J. Webster (ed.), *Wiley Encyclopedia of Electrical and Electronics Engineering*
Copyright © 1999 John Wiley & Sons, Inc.

IONOSPHERE CHARACTERISTICS

The ionosphere poses an interesting challenge for many radio systems that make use of signal transmission through all or some portion of that medium. Being a magnetoionic medium imbedded in a background neutral atmosphere, it exhibits very interesting refractive properties, including anisotropy, dispersion, and dissipation. The laminar ionosphere introduces an array of effects, which are related to the ionospheric component of radio refractivity. These include ray path bending, phase path increase, group path delay, absorption, Doppler shift, pulse dispersion, Faraday rotation, and magnetoionic path splitting. Inhomogeneities in the ionosphere give rise to temporal and spatial variations in the effects just cited. An understanding of the ionospheric personality provides information about a wide range of solar-terrestrial interactions, and it has significant space-weather implications. Space weather is a new discipline that includes a wide range of exoatmospheric phenomena of major importance to space systems and their operational effectiveness.

The main features of the ionosphere are well known, although details are subjects of continuing research. There are many excellent sources of information about the ionosphere, from both a theoretical and an experimental perspective. The books by Davies (1,2,3), Ratcliffe (4), and Giraud and Petit (5) should be consulted. Theoretical and plasma-physics aspects of the ionosphere have been discussed in a book by Kelley (6). A readable account of the basic physics of the ionosphere has been given by Rishbeth (7). Other useful references, which place the ionosphere within a larger context of the geospace weather system, include the Air Force *Handbook of Geophysics and the Space Environment* (8), and an *Introduction to the Space Environment* by Tascione (9). Various techniques for probing the ionosphere have been described in a monograph by Hunsucker (10). From a practical perspective, Goodman (11), Johnson et al. (12), and McNamara (13) have published expositions on the ionosphere in connection with radio system applications. There are also proceedings of topical conferences and workshops. The Ionospheric Effects Symposia (14) have chronicled ionospheric research activities and applications since 1975; and the Commission of the European Communities has published reports dealing with ionospheric prediction and modeling (15,16).

The purpose of this article is to provide a general understanding of the ionosphere. The emphasis is on those ionospheric processes and phenomena that are encountered by users of radio propagation systems. More complete descriptions of the underlying physical processes may be found in various references cited in the text. A final section on the current status of ionospheric research is provided as an aid to specialists and graduate students.

General Properties of the Ionosphere

Basic Structure. The ionosphere is an ionized region in the upper atmosphere that, by generally accepted convention, lies between an altitude range of 60 km to 1000 km. Nevertheless, the region above 1000 km but below 2000 km, called the protonosphere, is also ionized and may be important when considering the totality of ionization effects on radio systems. As a matter of convenience, some specialists have combined the ionosphere and protonosphere into a single region of ionization. For example, the integrated electron density from a ground station to a geosynchronous satellite (referenced to the vertical) is referred to as the total electron

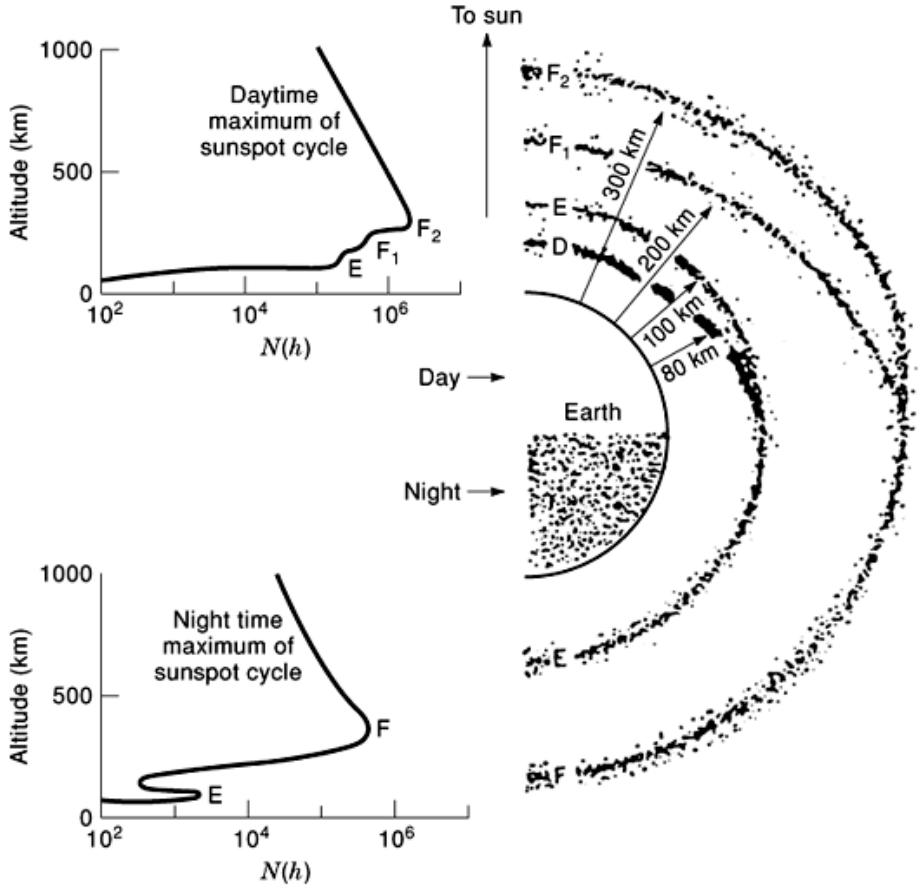

Fig. 1. Depiction of the ionospheric layers and the diurnal variation. [By permission of J. M. Goodman and Kluwer Academic Publishers, Norwell, MA (11).]

content of the ionosphere (*TEC*), even though both ionospheric and protonospheric electrons contribute to the integral. For the purpose of this article, we shall use the more restricted definition for the ionosphere, generally placing the upper limit at approximately 1000 km. While there are equal numbers of free electrons and positive ions within the ionosphere, it is the electron number density that characterizes the array of interesting phenomena associated with the region. The ionosphere is imbedded in the earth's magnetic field, and this situation influences the distribution of the ionized constituents. A clear indication of this may be seen in the worldwide distribution of electron density in the upper ionosphere, which tends to be described by geomagnetic rather than geographic coordinates. Moreover, being a magnetoionic medium, the ionosphere has a profound effect upon radiowaves that interact with the medium.

The ionospheric electron density distribution is logically evaluated first in terms of its height profile, followed by its geographical and temporal variabilities. Though there is abundant evidence suggesting a rather complex electron density profile comprised of several peaks and valleys, the basis for understanding fundamental properties of the ionosphere comes from a simple picture of an ionized medium dominated by a single region, or layer, having a distinct maximum in electron density. This is not without justification, since the highest and thickest component region, the so-called F layer, typically exhibits the greatest electron density. Moreover, in many radiowave applications, it is the F layer that exhibits the dominant interaction. Figure 1

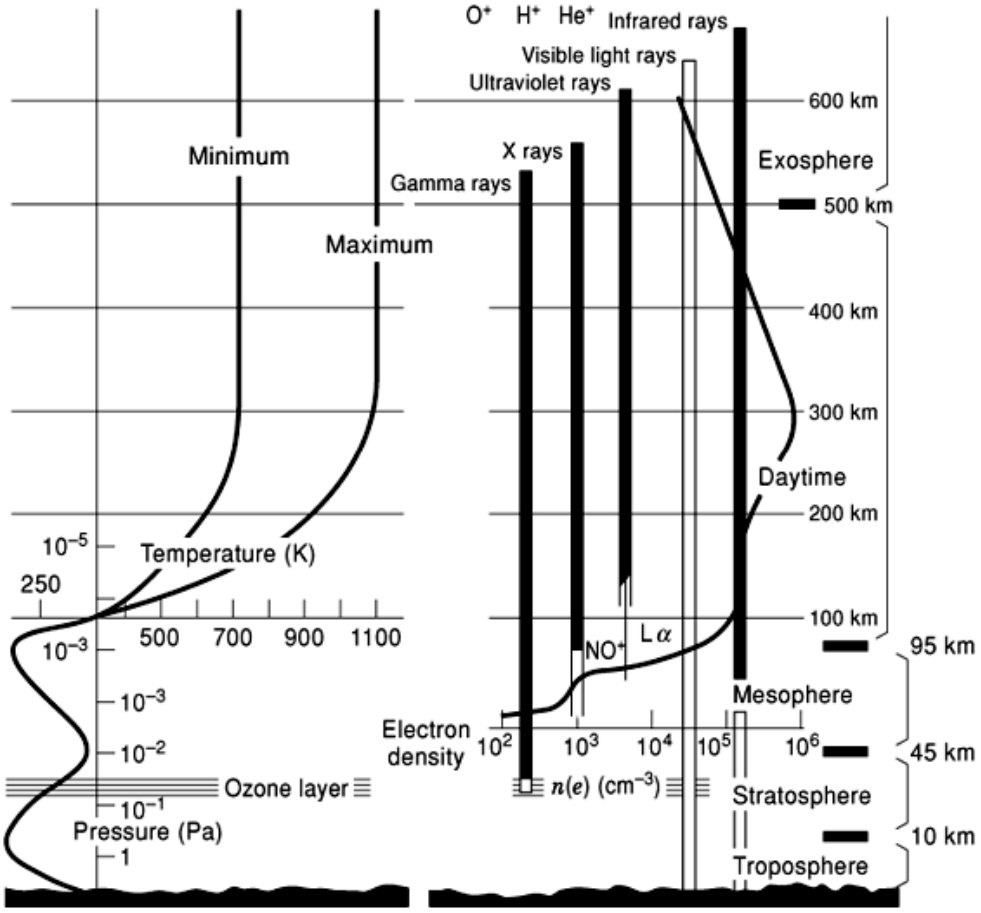

Fig. 2. Various atmospheric and ionospheric layers, the depth of penetration of solar radiation, and the thermospheric temperature profile. [From a National Research Council report (62).]

depicts the various regions or layers of the ionosphere in terms of the electron number density. It has been observed that the height profile varies diurnally, seasonally, and as a function of solar activity.

Formation of the Ionosphere. The sun exerts a number of influences on the upper atmosphere, but the interactions of most importance for our discussion are photodissociation and photoionization. Figure 2 depicts the neutral atmosphere, its various regions, and the depth of penetration of the various components of solar flux.

In the lower atmosphere, species such as N_2 and O_2 dominate the constituent population, though other species such as water vapor, carbon dioxide, nitric oxide, and trace element gases are influential in specific contexts. In the upper atmosphere, however, molecular forms are dissociated by incoming solar flux into separate atomic components. Formally the lowest portion of the ionosphere is the so-called D layer at an altitude of $\sim$ 60 km $\pm$ 20 km, but the free-electron and ion population rises dramatically at an altitude of $\sim$ 100 km, which is the median altitude of the E layer. Two things occur at this altitude. First, oxygen becomes dissociated as a result of solar UV radiation. Secondly, the mixing of the atmosphere, so efficient below 100 km, ceases rather dramatically, and the region where this occurs is called the turbopause.

The process of dissociation is so efficient that we treat the distribution of neutral species in a vast segment of the upper atmosphere (i.e., above 200 km) as that of a monatomic gas. In the lower atmosphere (i.e., below

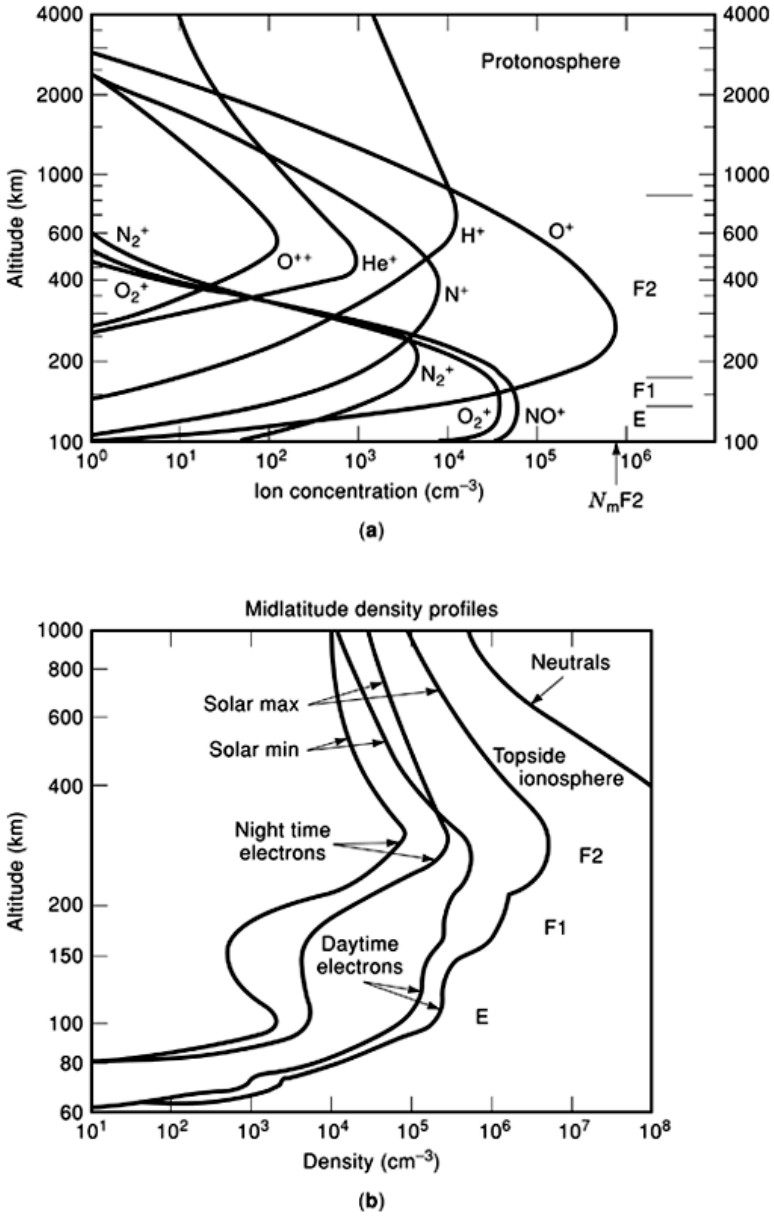

Fig. 3. (a) Profiles of ion concentrations, as a function of height, for daytime conditions. (b) Electron density distributions for day/night and solar maximum/minimum conditions. [From Jursa (8).]

roughly 200 km), the gas is largely polyatomic, although the transition between the two regimes is rather gradual between 100 km and 200 km. This has implications for the lifetime of ion–electron pairs created through photoionization. Also, in the altitude regime above about 200 km and well above the turbopause, collisions become a rarity, so that mixing of the various species becomes unimportant in comparison with diffusive forces. As a consequence, diffusive separation occurs, with constituents of the neutral gas seeking their own height distributions dictated by their atomic masses, the gas temperature, and the acceleration of gravity. Figure 3(a) shows height profiles of ionic species in the upper atmosphere, and Fig. 3(b) shows typical

distributions of midlatitude electron density for daytime and nighttime under solar maximum and minimum conditions.

It may be seen that ionized monatomic oxygen is the majority ion between roughly 180 km and 800 km, and is wholly dominant between about 200 km and 500 km. Atomic hydrogen ions become important above 500 km, and the region from about 800 km to 2000 km is called the protonosphere. It should also be noted that above 500 km (i.e., the base of the exosphere), the neutral atmosphere is virtually collisionless and particles tend to move about freely. On the other hand, electrons and ions in the exosphere are still influenced by the earth's magnetic field and electrodynamic forces.

The electron density distributions in the ionosphere and protonosphere are variable. Because of this, the boundary between the ionosphere and the protonosphere is not sharply defined, being dependent upon a number of factors including time of day, season, and solar activity. The protonosphere is often referred to as the plasmasphere, especially by magnetospheric scientists and those engaged in transionospheric TEC measurements.

Ionospheric Layering. Table 1 provides information about the various ionospheric layers, the altitude ranges of each, the principal ionic constituents, and the means of formation. A comment is appropriate here on the nature of ionospheric layering, with some emphasis on the historical distinctions made between the words *layer* and *region* as they pertain to the ionosphere. Often the terms are used interchangeably, and while neither is generally preferred, *region* is the more accurate description. This is because it does not convey the incorrect impression that sharp discontinuities in electron density exist at well-defined upper and lower boundaries. This is especially the case for the F region, and to a lesser extent for the D and E regions. From a historical perspective, the concept of layering derives from the appearance of the ionospheric regions on vertical-incidence ionospheric soundings, called ionograms (see the subsection "Sounder Measurement Method" below). Furthermore, the alphabetic designation of the ionospheric regions was also based upon the early sounding studies. On the other hand, there are certain situations for which the restrictive term *layer* is acceptable. For example, the normal E region may occasionally be characterized by an electron density profile displaying a degree of boundary sharpness. Aside from this, the most significant localized concentration of free electrons in the ionosphere, called sporadic E (or Es), exists as an isolated layer within the boundaries of the normal E region (see the section "Sporadic E" below). It is termed sporadic because it exhibits a generally unpredictable temporal and geographical distribution, and because of its limited geographical extent, it is sometimes referred to as a sporadic E patch.

As indicated above, the ionosphere is often described in terms of its component regions or layers. These were the so-called D, E, and F regions. These designations are largely based upon data obtained from crude sounder (i.e., ionogram) measurements undertaken in the 1920s and 1930s. These early measurements often exhibited evidence for an additional layer between regions E and F in the daytime ionosphere. This led to the notion that the F region is actually composed of two distinct regions (F_1 and F_2) having different properties.

The lowest region of the ionosphere, the D region, is important in the characterization of absorption losses for short-wave systems, and also as a reflecting layer for long-wave communication and navigation systems. There is also evidence for a bifurcation in the D region, with the upper portion (above 60 km) being produced by solar flux, and with the lower portion (below 60 km) being produced by galactic cosmic rays.

Ground-based vertical-incidence sounder measurements have provided the bulk of our current information about ionospheric structure (see the subsection "Sounder Measurement Method" below). Through application of ionogram inversion technology to allow for the radio-wave interaction effects, individual sounder stations provide information about the vertical distribution of ionization to the altitude of the F_2 maximum (i.e., 300 km to 400 km). In addition, the worldwide distribution of these systems has allowed a good geographical picture to be developed using sophisticated mapping algorithms. These measurements are somewhat limited in the characterization of certain features such as the so-called E–F valley, and they cannot evaluate ionization above the F_2 maximum. There is also a paucity of data over oceanic regions. Satellite measurements (viz., topside sounders and in situ probes) have been invaluable in the characterization of the F-region ionization density

Table 1: Properties of the Ionospheric Layers

Region	Height Range (km)	N_{max} Range (m^{-3})	f_p (MHz)	Major Ingredient(s)	Basis of Formation
D	70–90	10^8–10^9		NO^+, O_2^+	Lα X rays
E	90–130	$\sim 10^{11}$ (day)	$\sim$0.3 (night)	O_2^+,NO^+	Lβ X rays;
	$h_{max} \sim$ 110	$\sim 10^{10}$ (night)	$\sim$3.0 (day)		
		(smooth diurnal variation)			Chapman layer
Es	90–130	$\sim 10^{12}$		Metallic ions	Wind Shear & meteoric debris;
		(highly variable)			equatorial electrojet;
					Auroral electrojet and precipitation
F_1	130–210	$\sim 2 \times 10^{11}$ (day)	$\sim$ 3–6 (day)	O^+, NO^+	Helium II line;
	$h_{max} \sim$ 180	$\sim$ 0 (night)	(merges with F_2 layer at night)		UV radiation;
		(smooth diurnal variation)			Chapman layer
F_2	200–1000	$\sim 10^{12}$ (day)	$\sim$ 5–15 (day)	O^+	Upward
	$h_{max} \sim$ 300	$\sim 2 \times 10^{11}$ (night)	$\sim$ 3–6 (night)		diffusion from the F_1 layer;
		(asymmetric diurnal variation)	(presunrise minimum)		photoionization

over oceanic regions. Rocket probes and incoherent backscatter radar measurements, which provide a clearer representation of the true electron density profile, typically reveal a relatively featureless profile exhibiting a single F-region maximum with several underlying ledges or profile derivative discontinuities. Nevertheless, a valley of ionization may often be observed between the E and F regions. Ionization above the F_2 maximum may be deduced from satellite probes and Thomson-scattering radars, but a large amount of information has been derived from total electron content measurements using Faraday rotation or group path measurements of signals from geostationary satellites or Global Positioning System (*GPS*) satellites. Hunsucker (10) describes various ionospheric measurement techniques.

Simple layering occurs as the result of two factors. First, the atmospheric neutral density decreases exponentially with altitude, while the solar ionizing flux density increases with height above sea level. This leads to the formation of single region for which the ionization rate is maximized, and ultimately results in a layer having the so-called Chapman shape. This shape is based upon a simple theory advanced by Sidney Chapman in 1931 (Ref. 17; see Fig. 4). We observe nonetheless a degree of structure in the ionosphere, which suggests more than one layer. One cause for multilayer formation is the existence of a multicomponent atmosphere, each component of which possesses a separate height distribution at ionospheric altitudes. But there are other factors. Solar radiation is not monochromatic as suggested in the simple Chapman theory, and it has an energy density that is not evenly distributed in the wavelength domain. Furthermore, its penetration depth and ionization capability depend upon wavelength and atmospheric constitution. All of this results is a photoionization rate, and an associated electron density profile, that are structured functions of altitude. It has been shown that the Chapman model is valid for the D, E, and F_1 regions but is not generally valid for the F_2 region.

Chapman Layer Theory. One of the basic tenets of Chapman theory is that solar radiation will penetrate to an altitude for which the total number of atoms or molecules, P (populating a column of unit cross sectional area directed toward the sun) is equal to the reciprocal of the absorption (or interaction) cross

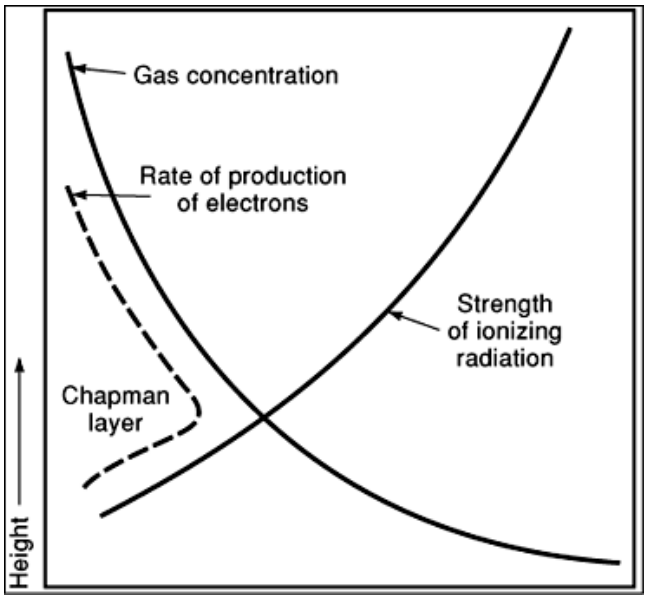

Fig. 4. An idealized representation of ionization production in the atmosphere as the solar radiation encounters a neutral gas with exponentially increasing density. [By permission of J. M. Goodman and Kluwer Academic Publishers, Norwell, MA (11).]

section a that is $P = 1/a$. The peak in ionization will be produced in the neighborhood of that altitude, and the concept is valid for oblique solar illumination as well as for the case in which the sun is directly overhead. It is convenient to look at the production rate in terms of its deviation from the peak (overhead) value at height h_0. For this it is useful to define a reduced height z, corresponding to the normalized departure of an arbitrary value of ionospheric height h from h_0.

$$z = \frac{h - h_0}{H} \tag{1}$$

where h_0 is the peak height for vertically incident radiation from the sun, and H is the neutral scale height given by the following expression:

$$H = kT/mg \tag{2}$$

where k is Boltzmann's constant, T is the absolute gas temperature, m is the atomic or molecular mass, and g is the acceleration of gravity. Within the thermosphere (with $h \gtrsim 100$ km), the gas temperature is monotonically increasing, reaching an asymptotic level near the base of the exosphere. The temperature rises from $\approx$ 180 K at the mesopause (and incidentally near the turbopause) to levels approaching a diurnal range of 600 K to 1100 K at solar minimum and 800 K to 1400 K at solar maximum. The heat sources include solar radiation, the dissipation of atmospheric gravity waves, and particle precipitation. The asymptotic levels of T are due to limits on the thermal conductivity of the gas. The scale height H is a convenient parameter, since it may be used as a measure of layer thickness for an equivalent fixed-density slab. More importantly, it has a physical meaning. If the atmosphere is in diffusive equilibrium governed by the force of gravity and the gas pressure

gradient, and N is the atomic or molecular gas density (as appropriate), we have

$$N = N_0 \exp(-h/H) \tag{3}$$

$$H = \left(\frac{1}{N} \left| \frac{dN}{dh} \right| \right)^{-1} \tag{4}$$

where N_0 is the atomic or molecular density at some reference height.

In a diffusively separated atmospheric environment, each constituent has its own unique scale height governed by its own molecular (atomic) mass. In an ionized gas in which the electrons and ions are coupled by electrostatic forces, the effective value of the mean molecular mass is $\approx \frac{1}{2}$ the mass of the positive ion. This is because the mass of the electron is essentially zero in comparison with the ion mass.

Figure 5 depicts the production-rate curves associated with an ideal Chapman-like production profile and a range of solar zenith angles χ. It is seen that there are a number of curves, parametrized in terms of χ , for which production rate maxima $q_{\max}$, may be observed. The largest $q_{\max}$ occurs for $\chi = 0$ (overhead case corresponding to $q = q_0$), and we see that other values for $q_{\max}$, corresponding to oblique geometries wherein $\chi \neq 0$, will decrease in magnitude and occur at increasing heights as χ becomes larger (i.e., the sun moves toward the horizon). Chapman theory yields the following rate-of-production formula:

$$q = q_0 \exp(1 - z - \sec \chi e^{-z}) \tag{5}$$

At altitudes well above the peak in q, the rate of electron production drops off in an exponential fashion imitating the exponential decrease in gas pressure with height. In order to relate Chapman production curves to actual electron density distributions, we must examine loss processes and certain dynamic factors.

The Continuity Equation and Equilibrium Processes

The equation that expresses the time rate of change of electron concentration, N_e, is the continuity equation:

$$\frac{d N_{\mathrm{e}}}{dt} = q - L(N_{\mathrm{e}}) - \mathrm{div}(N_{\mathrm{e}} \boldsymbol{V}) \tag{6}$$

where N_{e} is the electron density, $L(N_{\mathrm{e}})$ is the loss rate, which is dependent upon the electron density, div stands for the vector divergence operator, and $\boldsymbol{V}$ is the electron drift velocity.

The divergence of the vector in Eq. 6 is the transport term, sometimes conveniently called the *movement* term. The continuity equation says that the time derivative of the electron density within a unit volume is equal to the number of electrons that are generated within the volume (through photoionization processes) minus the number that are lost (through chemical recombination or attachment processes), and finally adjusted for those electrons that exit or enter the volume (as expressed by the transport term). To first order, the only derivatives of importance in the divergence term are in the vertical direction, since horizontal N_{e} gradients are generally smaller than vertical ones. In addition, there is a tendency for horizontal velocities to be small in comparison with vertical drift velocities. Consequently, we may replace div (N_{e} $\boldsymbol{V}$) by (d/dh) (N_{e} V_{h}), where V_{h} is the scalar

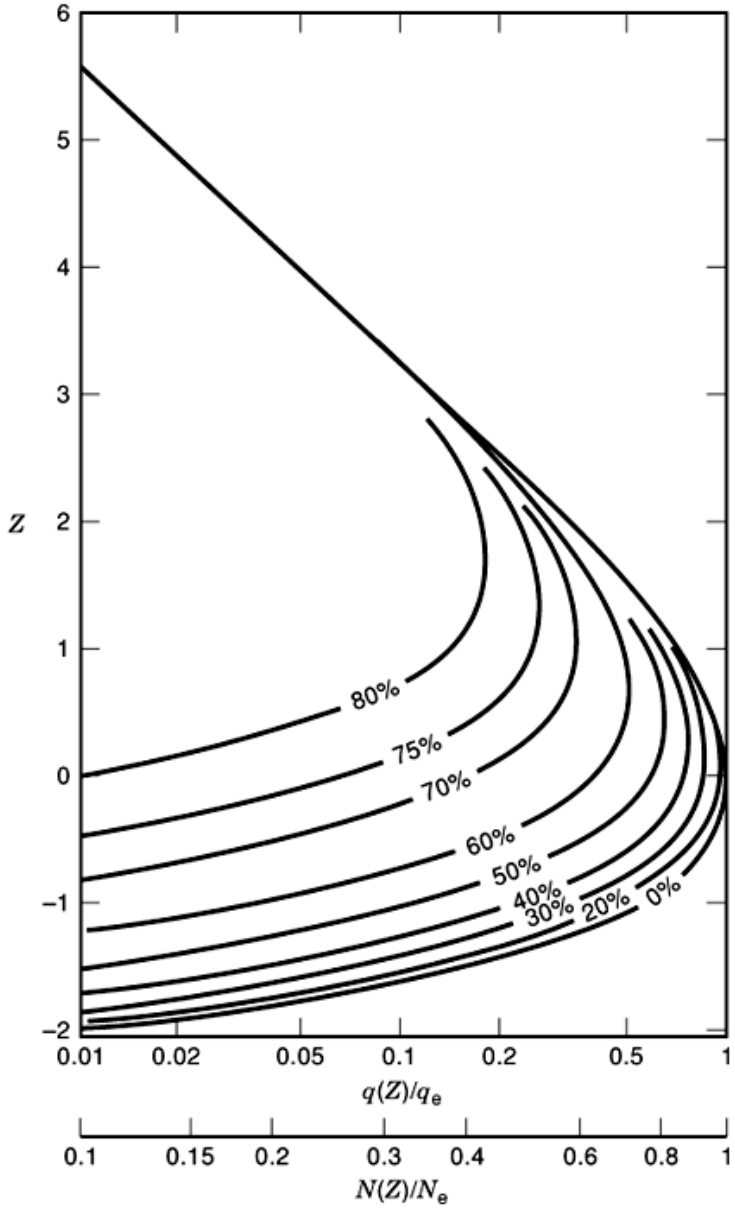

Fig. 5. Curve illustrating the rate of electron production as a function of reduced height ($h - h_0$) and for selected values of the solar zenith angle. [From Davies (1).]

velocity in the vertical direction. We rewrite Eq. 6 as follows:

$$\frac{d\,N_e}{dt} = q - L(N_e) - \frac{d}{dh}(N_e\,V_h) \tag{7}$$

Table 2: Types of Equilibrium Processes

Photochemical equilibrium (production balanced by loss):
$L(N_e) \gg \frac{d}{dh}(N_e V_h)$
Drift equilibrium (production balanced by drift):
$L(N_e) \ll \frac{d}{dh}(N_e V_h)$

Now let us look at some special cases. If $V_h = 0$ (no movement), then the time variation in electron concentration is controlled by a competition between production q and loss L. At nighttime, we may take $q = 0$, and this results in

$$\frac{dN_e}{dt} = -L(N_e) \tag{8}$$

In principle, there are two mechanisms to explain electron loss: attachment of electrons to neutral atoms (in the upper ionosphere), and recombination of electrons with positive ions (in the lower ionosphere). The attachment process is proportional to N_e alone, while recombination depends upon N_e with N_i, where N_i is the number of ions. Attachment involves radiative processes and has an extremely low cross section (probability of occurrence). We may ignore it in many practical situations and take recombination as the major source for electron loss. Since $N_e = N_i$, the recombination process obeys the equation $L = \alpha \cdot {N_e}^2$, where α is the recombination coefficient. Recombination is very rapid in the D and E regions, the process being accomplished in a time on the order of seconds to minutes. Attachment, the electron loss process for the upper ionosphere, has a time constant on the order of hours. This is the primary reason that the ionosphere does not entirely disappear overnight. Another reason is that there exists a second source of electrons associated with the plasmasphere. This reservoir of ionization is built up during the daytime through vertical drift, but bleeds into the ionosphere during nocturnal hours.

In the vicinity of local noon, $dN_e/dt = 0$ and we may analyze the quasiequilibrium conditions suggested by Eq. (7) when the left-hand side of the equation equals 0. The two main types of equilibrium processes are given in Table 2.

The equilibrium processes identified in Table 2 are the dominant possibilities during daytime when photoionization is significant. During nocturnal hours, equilibrium is seldom achieved at F-region heights, although it is approached in the period before sunrise.

While the continuity equation appears quite simple, the generic terms (i.e., production, loss, and transport) represent a host of complex photochemical and electrodynamic processes, which exhibit global variations and are influenced by nonstationary boundary conditions within the atmosphere and the overlying magnetosphere. Notwithstanding these complications, the equation provides a remarkably clear view of the basic processes that account for ionospheric behavior. In fact, the relative contributions of terms in the continuity equation will account for the majority of the anomalous ionospheric properties; that is, those ionospheric variations that depart from a Chapman-like characteristic. This is especially true for the F_2 layer, within which the movement term attains paramount status. In the E and F_1 regions, where the movement term is small compared with production and loss (through recombination), photochemical equilibrium exists in the neighborhood of midday. All of this has had a significant bearing on the development of ionospheric models and prediction methods.

Indeed, as it relates to the F region of the ionosphere, it may be said that the existence of a nonvanishing divergence term in the continuity equation has been the primary impetus for the development of statistical modeling approaches. Nevertheless, efforts to account for all terms in the continuity equation through physical modeling are ongoing.

The underlying assumptions used by Chapman in his theory of layer production are in substantial disagreement with observation. The Chapman layer was based upon an isothermal atmosphere, and it is well known that the atmosphere has a scale height, kT/mg, which varies with height. Moreover, the basic theory assumes a monochromatic source for photoionization and a single constituent gas. Corrections and extensions to the early Chapman theory have led to better agreement with observation, and to this day the Chapman layer provides a fundamental baseline for ionospheric profile modeling.

Description of the Ionospheric Layers

Sounder Measurement Method. In any discussion of the ionospheric electron density distribution, it is important to recognize that many experimental methods have been used to arrive at our current understanding. The major ones include ground-based vertical-incidence sounding (*VIS*), topside sounding using satellite platforms, incoherent backscatter radar, the Faraday rotation and signal delay of satellite signals, and in situ measurements using rocket probes and satellites. The VIS method, which employs the high-frequency (*HF*) band, was the earliest method and has provided the most comprehensive picture of the lower ionosphere and its worldwide distribution. The ionospheric D layer is an exception, and special methods are needed to determine the electron densities in that region. Much of the current nomenclature involving ionospheric structure and phenomena is a carryover from early VIS investigations. As a consequence we shall briefly examine the VIS method. For a discussion of other methods, the reader is referred to Hunsucker (10).

The plasma frequency associated with an electron gas, f_{p} (a natural resonant frequency), is proportional to the square root of the electron density of the gas:

$$f_{\mathrm{p}} = 8.98\, N_{\mathrm{e}}^{1/2} \tag{9}$$

where f_{p} is in hertz and N_e is in electrons per cubic meter.

It may be shown that a radio wave, propagating vertically upward into the ionosphere, will penetrate the region until it reaches a point at which the sounding frequency matches the plasma frequency. All frequencies less than this value will be reflected back to ground. An ionospheric sounder is essentially a radar, which maps out the height-dependent ionospheric electron concentration versus transmission frequency, where the probing frequency is typically a stepwise increasing function of time. A plot of signal echo time delay versus transmission frequency is called an ionogram. A typical ionogram and the corresponding ionospheric profile are given in Figure 6. If $N_{\max}$ is the maximum electron density of a layer, then we define a so-called critical frequency of reflection, f_{c}, which is the maximum plasma frequency within the layer. If the sounder transmission frequency exceeds f_{c}, then the signal is not reflected and penetrates the layer. There are as many critical frequencies in the ionosphere as there are layers or regions. A more complete treatment of the theory of radio propagation in the ionosphere shows that a magnetoionic medium supports two modes of propagation, the *ordinary* (O mode) and the *extraordinary* (X mode). These modes encounter slightly different indices of refraction and thus travel with slightly different velocities and directions. As a consequence, each ionogram consists of two traces, corresponding to O- and X-mode echoes. These traces may be closely aligned over a large portion of their respective propagation bands but can depart significantly at their respective critical frequencies, with the X mode supporting somewhat higher-frequency signal reflections. By convention, the O-mode trace is used for conversion of ionogram critical frequencies into maximum electron densities. The following convenient

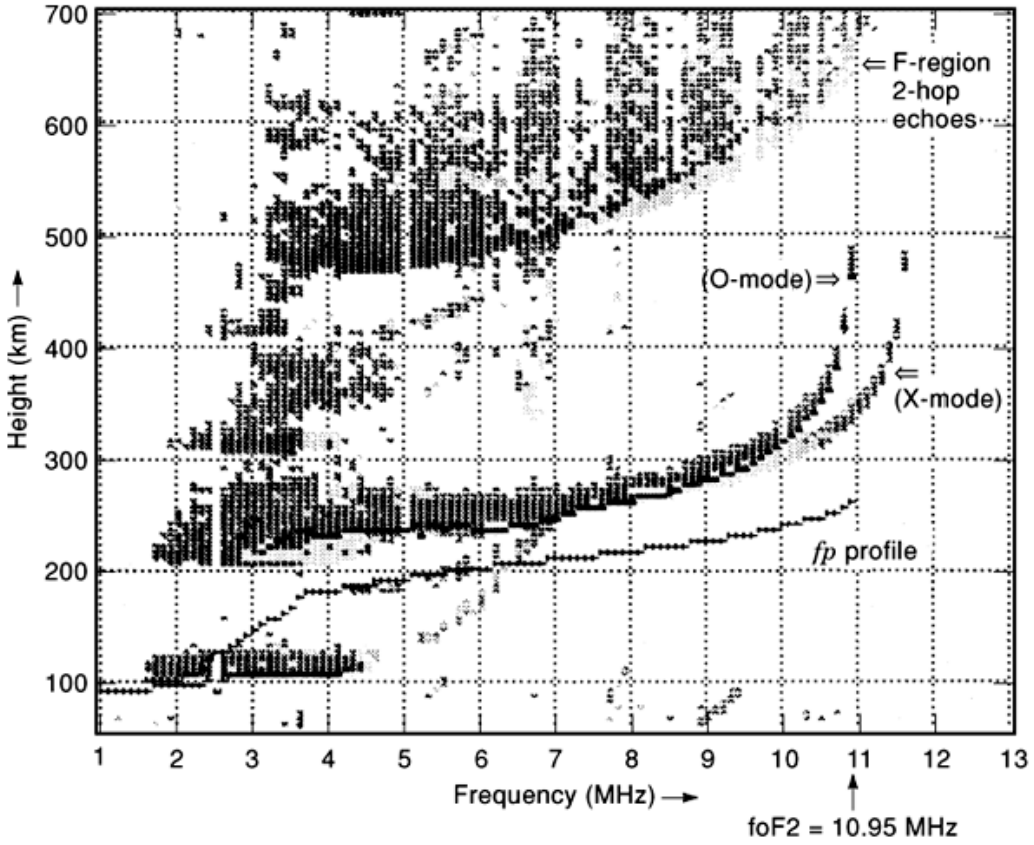

Fig. 6. Typical vertical-incidence ionosonde recording (i.e., ionogram) and the corresponding plasma frequency profile *fp(h)*. The electron density profile is related to the plasma frequency profile by the Eq. 9 in the text. [This ionogram was derived from U.S. Government web site (http://www.ngdc.noaa.gov/stp/), and the instrument was developed by University of Massachusetts-Lowell for the U.S. Air Force.]

expression is used:

$$N_{\max} = 1.24 \times 10^{10} f_O^2 \qquad (10)$$

where f_O is the ordinary-ray critical frequency (MHz) and $N_{\max}$ is the maximum electron density of the given layer (e/m^3). Equation (10) is equivalent to Eq. (9).

From a historical perspective, it is interesting to note that the concept of radar detection of aircraft derived from the early work of ionospheric specialists who were already using ionospheric sounders as a means to detect ionospheric layers.

The D Region. The D region is responsible for most of the absorption encountered by HF signals, which exploit the sky-wave mode. In most instances, D-region absorption is a primary factor in the determination of the lowest frequency, which is useful for communication over a fixed sky-wave circuit. In addition, the D region supports long-wave propagation at very low frequency (*VLF*) and low frequency (LF), and the medium is exploited in certain legacy navigation systems and strategic low-rate communication systems. The sounder method as described in the previous section is not useful for measurement of the D region, since the electron densities are relatively low. Details of D-region electron concentration are sketchy in comparison with

information available about the E and F regions, principally because of the difficulty in making diagnostic measurements. Moreover, analysis is hampered because many photochemical processes with poorly defined reaction rates take place in the D region. Over 100 reactions have been compiled.

Table 1 shows that the D region lies between 70 km and 90 km. In fact, the upper and lower levels are not precisely defined. It is evident that more than one source of ionization gives rise to the D-region electron density distribution. Sources include solar radiation at the upper levels and galactic cosmic rays at lower levels. In addition, relatively rare polar-cap absorption (*PCA*) events are characterized by highly energetic solar protons that provide an additional source for ionization of the lower D region within the polar cap. Some investigators place the lower boundary of the D region at 50 km to allow for the contribution of galactic cosmic rays in the neighborhood of 50 km to 70 km. This altitude regime, termed the C region, is not produced by solar radiation. It exhibits different characteristics from the region between 70 km and 90 km. Specifically, a minimum in electron concentration is observed during solar-maximum conditions for the lower portion (viz., region C), while the reverse is true in the upper portion (viz., region D). This can be explained if we assume that the galactic cosmic rays are partly diverted from the earth by an increase in the interplanetary magnetic field (*IMF*), which occurs during solar maximum conditions.

The E Region. In an (α) Chapman layer for which photochemical equilibrium has been established, the following equation represents the electron density distribution as a function of reduced height z:

$$N_e(z) = (q_0/\alpha)^{0.5} \exp(1 - z - e^{-z} \sec \chi) \qquad (11)$$

where α is the recombination coefficient, χ is the solar zenith angle, and q_0 is the maximum production rate in the layer. Recall that α is the recombination coefficient (see the section "The Continuity Equation and Equilibrium Processes" above). The quantity q_0/α is dependent upon the sunspot number and is specific to the region involved, in this case the E region. The maximum rate of electron production q_0 occurs only for the overhead sun. However, it may be shown that actual maxima for other zenith angles are simply related by this expression:

$$q_{\max} = q_0 \cos \chi \qquad (12)$$

It may be shown that the ordinary-ray critical frequency for the E region, which is directly related to the E-region maximum electron density through Eq. (10), may be found from Eq. (11), and is given by

$$\text{foE} = \kappa(\cos \chi)^n \approx (\cos \chi)^{0.25} \qquad (13)$$

where κ is a constant of proportionality, which is dependent upon the sunspot number. The exponent n tends to a value 0.25 for long-term seasonal behavior, and in compliance with Chapman theory, but some workers have found that a value for $n \approx 0.3$ better represents the diurnal dependence. The constant of proportionality κ ranges between about 3 MHz and 4 MHz, bearing in mind that Eq. (13) represents a climatological median value.

The solar-activity dependence of the ratio of peak production to the effective loss (recombination) coefficient has been studied by a number of workers, and the results enable values of foE to be deduced. There have also been direct measurements of foE using vertical incidence sounders. While there is some variability to be considered, it is possible to develop a relationship connecting the median value of foE, the solar zenith angle, and the 12-month running-mean sunspot number. A generally accepted candidate for the daytime E-region critical frequency is

$$\text{foE} = 3.3\,[(1 + 0.008R_{12})\cos\,\chi]^{0.25} \quad \text{MHz} \qquad (14)$$

where R_{12} is the running 12-month sunspot number, which may range between roughly 10 and 150.

Equation (14) provides excellent agreement with observation during the daytime, but alternative expressions are found to be more appropriate during the nighttime hours (3). Moreover, it has been found that

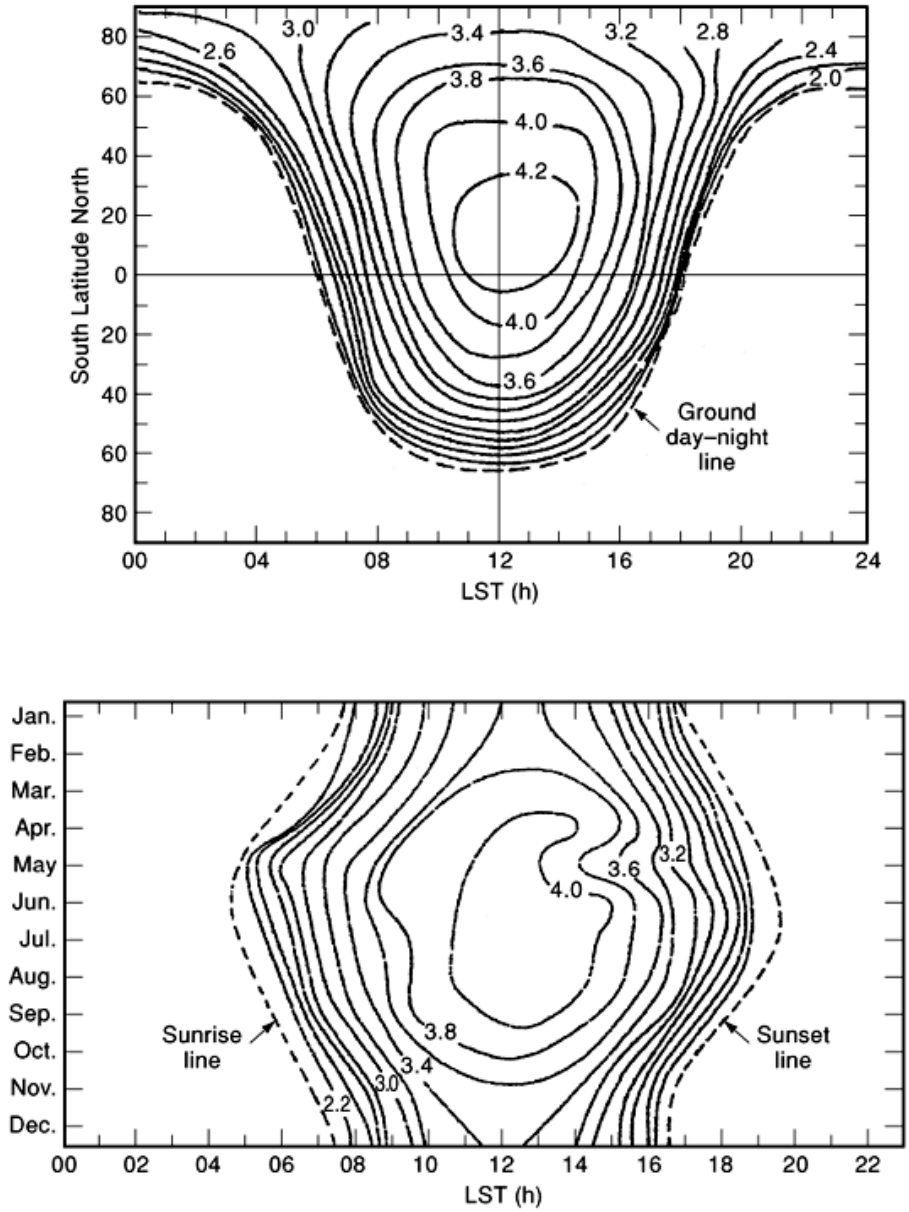

Fig. 7. (a) Depiction of the local-time (LST) and latitude dependence of foE for solar-maximum conditions in summer [from Davies (1)](b) Contours of foE at Fort Belvoir, Virginia, in 1958 (solar maximum), showing seasonal variations. The contours are in megahertz.

Eq. (14) is inaccurate in the very high latitudes, where other means of electron production become important, invalidating the Chapman hypothesis. Internationally adopted relations for monthly median foE are due to Muggleton (18); and an alternative relation, specific to the European region, has been published (15).

Figure 7(a) contains an E-region critical-frequency map for summer solstice conditions in 1958, a period of high solar activity (i.e., R_{12} large). The contours are representative of median conditions as a function of

geographic latitude and local time. It is seen that the E-region critical frequencies (and consequently the electron densities) are vanishingly small in regions devoid of solar illumination. This summer solstice behavior is consistent with Eq. (14), and other seasons have been shown to behave in conformance with $(\cos \chi)^{0.25}$ as well.

Figure 7(b) shows the monthly variation of foE for one station (Ft. Belvoir, Virginia) for the year 1958. The solar control is obvious in the median data plotted.

The F_1 Region. The F_1 region is not unlike the E region in the sense that it obeys many of the predictions of Chapman theory. We look for a relation for the ordinary-ray critical frequency that is formally similar to Eq. (13). A relation patterned after Chapman principles may be expressed as

$$\text{foF1} = f_s(\cos \chi)^n \tag{15}$$

Like the E region, the F_1 region exhibits more complicated behavior than that expressed by such a simple formula. Specifically, it has been found that the geomagnetic latitude tends to exhibit some control over the F_1-region electron densities. The function f_s in Eq. (15) depends upon sunspot number and magnetic latitude. It is also observed that the F_1 region disappears (i.e., merges with the F_2 region) at values of the solar zenith angle exceeding a certain maximum that itself depends upon both the sunspot number and the geomagnetic latitude. The Radio Sector of the International Telecommunications Union (*ITU-R*, previously the *CCIR*) has developed a method for computing foF1 taking all these factors into account (18). The internationally adopted monthly median foF1 formulation is based on the work of Ducharme et al. (19). The relation due to Davies (3) gives a convenient but approximate expression for the F_1-layer critical frequency:

$$\text{foF1} = (4.3 + 0.01R_{12}) \cos^{0.2} \chi \tag{16}$$

Figure 8 shows the solar-zenith-angle control of foF1 under sunspot maximum and minimum conditions.

The height of the F_1 ledge, hF1, is taken to be between 180 km and 210 km. From Chapman theory we anticipate that hF1 will be lower in summer than in winter and will be higher at midlatitudes than at low latitudes. Unfortunately, the reverse is true. Explanations for this behavior may be found in a detailed study of scale-height gradients, a nonvanishing movement term (as expressed in the continuity equation), or gradients in upper atmospheric chemistry.

The F_2 Region. The F_2 region is the most prominent layer in the ionosphere, and this significance arises as a result of its height (it is the highest of all the component layers) and of course its dominant electron density. It is also characterized by large ensembles of irregularity scales $\{\Delta L\}$ and temporal variations $\{\Delta T\}$. The F_2 region is a vast zone, which eludes prediction on the microscale ($\Delta L < 1$ km) and mesoscale (1 km $< \Delta L <$ 1000 km) levels, and even provides challenges to forecasters for global and macroscale ($\Delta L >$ 1000 km) variations. This is largely because of the elusive transport term in the continuity equation. There are also a host of so-called anomalous variations to consider, and these are the subjects of a succeeding section.

As in the E and F_1 regions, we may conveniently specify the behavior of the F_2 region in terms of equivalent plasma frequency rather than the electron density. For the peak of ionization we have

$$N_{\max}\text{F2} = 1.24 \times 10^{10}(\text{foF2})^2 \tag{17}$$

where foF2 is the ordinary-ray critical frequency.

While foF2 exhibits solar-zenith-angle, sunspot-number, and geomagnetic-latitude dependences, simple algebraic algorithms do not characterize these relationships. As a consequence, mapping methods are used to describe the F_2 region electron density patterns.

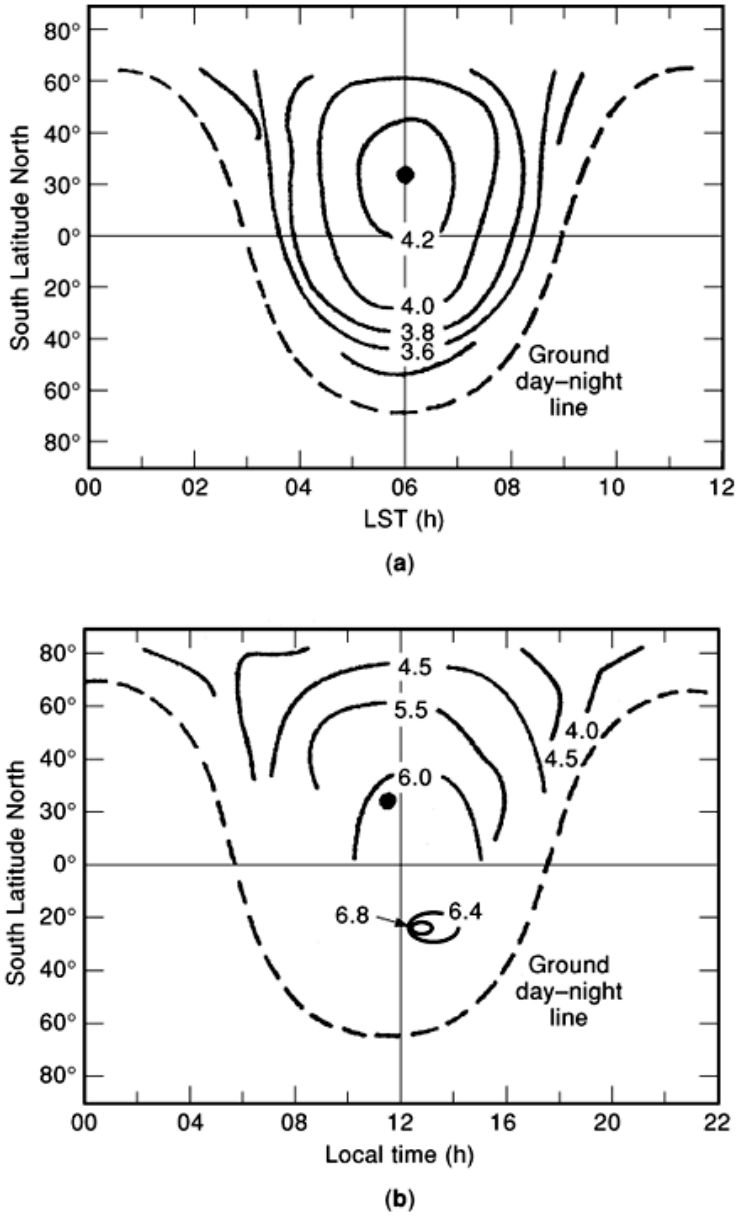

Fig. 8. Depiction of the local-time and latitudinal variation of foF1 for two different solar activity conditions: (a) solar minimum, June 1954; (b) solar maximum, June 1958. The contours are in megahertz. [From Jursa (8).]

The CCIR published its *CCIR Atlas of Ionospheric Characteristics*, which includes global maps of F_2-layer properties for sunspot numbers of 0 and 100, for every month, and for every even hour of Universal Time (20). Figure 9 is an illustration of the global distribution of foF2 for a sunspot number of 100. Such maps are derived from coefficients based upon data obtained from a number of ionosonde stations for the years 1954–1958 as well as for the year 1964. This set of coefficients is sometimes identified by an *ITS* prefix, but is known more

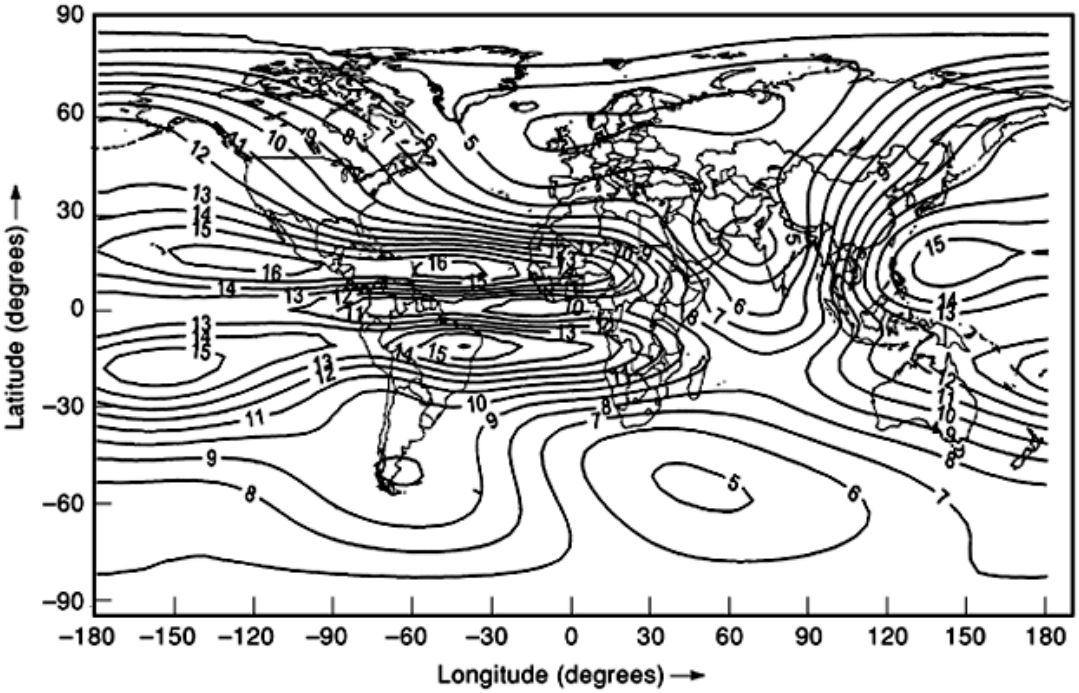

Fig. 9. Map of foF2 showing the worldwide distribution under the following conditions: 15 November, Sunspot Number = 135, Time = 0000 UTC. The countours of foF2 are developed using the URSI set of ionospheric coefficients. Curves similar to this are found in the *Atlas of Global Ionospheric Coefficients* (20). [By permission, Radio Propagation Services, November 2000.]

generally as the CCIR coefficients. Because of the paucity of data over oceanic areas, a method for improving the basic set of coefficients by adding theoretically derived data points was developed. As a result, a new set of coefficients has been sanctioned by International Union of Radio Science (URSI), and this is termed the URSI coefficient set. Many communication prediction codes, which require ionospheric submodels, allow selection of either set of ionospheric coefficients.

Anomalous Features of the Ionospheric F Region. The F_2 layer of the ionosphere is probably the most important region for many radio-wave systems. Unfortunately, the F_2 layer exhibits the greatest degree of unpredictable variability because of the transport term in the continuity equation. As indicated previously, this term represents the influences of ionospheric winds, diffusion, and dynamical forces. The Chapman description for ionospheric behavior depends critically upon the unimportance of the transport function. Consequently, many of the attractive, and intuitive, features of the Chapman model are not observed in the F_2 region. The differences between actual observations and predictions derived on the basis of a hypothetical Chapman description have been termed anomalies. In many instances, this non-Chapman-like behavior is not anomalous at all, but rather typical.

The following list represents the major forms of anomalous behavior in the F_2 layer: diurnal, Appleton, December, winter, and the F-region trough. A few comments are provided for each major form.

The Diurnal Anomaly. The diurnal anomaly refers to the situation in which the maximum value of ionization in the F_2 layer occurs at a time other than at local noon as predicted by Chapman theory. On a statistical basis, the actual maximum occurs typically in the temporal neighborhood of 1300 to 1500 LMT. Furthermore, there is a semidiurnal component that produces secondary maxima at approximately 1000 to

1100 LMT and 2200 to 2300 LMT. Two daytime maxima are sometimes observed (one near 1000 and the other near 1400), and these may cause the appearance of a minimum at local noon. This feature, when observed, is called the midday biteout.

Appleton Anomaly. This feature is symmetric about the geomagnetic equator and goes by a number of other names, including the geographic anomaly, the geomagnetic anomaly, and the equatorial anomaly. The Appleton anomaly is associated with the significant departure in the latitudinal distribution of the maximum electron concentration within 20° to 30° on either side of the geomagnetic equator. Early in the morning a single ionization peak is observed over the magnetic equator. However, after a few hours the equatorial F region is characterized by two distinct crests of ionization that increase in electron density as they migrate poleward. This phenomenon is described as an *equatorial fountain* initiated by an $\boldsymbol{E} \times \boldsymbol{B}$ plasma drift (termed a Hall drift), where $\boldsymbol{E}$ is the equatorial electrojet electric field and $\boldsymbol{B}$ is the geomagnetic field vector. This drift is upwards during the day, since the equatorial electric field $\boldsymbol{E}$ is eastward at that time. As the electrojet decays, the displaced plasma is now subject to downward diffusion when the atmosphere begins to cool. This diffusion is constrained along paths parallel to $\boldsymbol{B}$, which map to either side of the geomagnetic equator. The poleward extent of the anomaly crests is increased if the initial Hall-drift amplitude is large. This anomalous behavior accounts for the valley in the parameter foF2 (with peaks on either side) seen at the geomagnetic equator in Figure 9. There are significant day-to-day, seasonal, and solar-controlled variations in the onset, magnitude, and position of the anomaly. There are also asymmetries in the anomaly crest position and electron density. Asymmetries in the electron density in the anomaly crests appear to be the result of thermospheric winds that blow across the equator from the subsolar point. The effect of magnetic activity on the anomaly is to constrain the electron density and latitudinal separation of the crests. Magnetic activity is monitored worldwide, and the quasilogarithmic index Kp is used to represent the level of worldwide activity (21). When Kp$\geq$ 5 (on a scale from 0 to 9), the anomaly disappears.

The December Anomaly. This term refers to the fact that the electron density at the F_2 peak over the entire earth is 20% higher in December than in June, even though the solar-flux change due to earth eccentricity is only 5% (with the maximum in January).

The Winter (Seasonal) Anomaly. This is the effect in which the noontime peak electron densities are higher in the winter than in the summer despite the fact that solar zenith angle is smaller in the summer than it is in the winter. This effect is modulated by the 11-year solar cycle and virtually disappears at solar minimum.

The F-Region (High-Latitude) Trough. This is representative of a number of anomalous features that are associated with various circumpolar phenomena, including particle precipitation, the auroral arc formations, etc. The high-latitude trough is a depression in ionization, occurring mainly in the nighttime sector, and it is most evident in the upper F region (22). It extends from 2° to 10° equatorward of the auroral oval, an annular region of enhanced ionization associated with optical aurora (see the section "The High-Latitude Ionosphere"). The trough region is associated with a *mapping* of the plasmapause onto the ionosphere along geomagnetic field lines (see Fig. 17). The low electron density within the trough results from a lack of replenishment through candidate processes such as antisunward drift, particle precipitation, or the storage effect of closed field lines. The latitudinal boundaries of the trough may be sharp, especially the poleward boundary with the auroral oval. A model of the trough is due to Halcrow and Nisbet (23).

Irregularities in the Ionosphere. In addition to the various anomalous features, irregularities in the electron density distribution may be observed throughout the ionosphere. The size, intensity, and location of these irregular formations are dependent upon a number of factors, including geographical area, season, time of day, and levels of solar and magnetic activity. The *traveling ionospheric disturbance* (TID; see the subsection "Short-Term Variations" and the section "Ionospheric Predictions" below) belongs to a special class of irregular formations that are generally associated with significant changes in the electron density (more than a few percent) over large distances ($>$ 10 km). The remaining irregularities, loosely termed *ionospheric inhomogeneities*, typically develop as the result of ionospheric instability processes and are not directly associated

with TIDs. On the other hand, TIDs have been shown to be a possible catalyst in the formation of ionospheric inhomogeneities, especially in the vicinity of the Appleton anomaly. Relatively small-scale ionospheric inhomogeneities are important, since they are responsible for the rapid fading (scintillation) of radio signals from satellite communication and navigation systems. Such effects may introduce performance degradations or outages on systems operating at frequencies between 100 MHz and several gigahertz. Models of radiowave scintillation have been developed, and these are based upon a basic understanding of the global morphology of ionospheric inhomogeneities.

There are inhomogeneities in all regions of the ionosphere, but the equatorial and high-latitude regions are the most significant sources. Hunsucker and Greenwald (24) have reviewed irregularities in the high-latitude ionosphere, and Aarons (25) has examined the equatorial environment.

Equatorial inhomogeneities tend to develop following sunset and may persist throughout the evening, but with decreased intensity after local midnight. The irregularities are thought to be the result of an instability brought about by a dramatic change in F-region height at the magnetic equator following sunset. The scale lengths of the irregularities may range between roughly a meter and several kilometers, and the spectrum of the irregularities has been observed to exhibit a power-law distribution. There is a tendency for the irregularities to be field-aligned with an axial ratio of roughly 20 to 1. In addition, the irregularities are organized in distended patches. Though the situation is variable, the patch sizes range between $\sim$100 km and several thousand kilometers in the upper F region, and average $\sim$100 km in the lower F region. The equatorial irregularities tend to be more intense and widespread at the equinoxes and at solar maximum, but magnetic activity tends to suppress the growth of the irregularities.

High-latitude irregularities exist within the polar cap and the auroral zone, with the latter being primarily associated with the bottomside F region. The high-latitude F region is quite variable, and unlike midlatitudes, it may have an electron density less than the E-region during nocturnal hours. In the wintertime, structured auroral arcs may migrate within the polar cap, and the electron density enhancements within these formations may be several orders of magnitude greater than the normal background, especially during elevated solar activity. During disturbed geomagnetic conditions, structured electron density patches have been observed to travel across the polar cap in the antisunward direction. These irregularities may have a significant effect on communication systems. For both the auroral zone and the polar cap, increased geomagnetic activity has a dramatic influence on the growth of irregular ionospheric formations. Moreover, for large and sustained values of Kp, it has been observed that the high-latitude irregularity patterns tend to migrate equatorward, replacing the background midlatitude properties (see the section "The High-Latitude Ionosphere").

Diurnal Behavior of the Ionospheric Layers

Mean Variations. As indicated by the Chapman representation for $N_{\max}$, the respective critical frequencies for the layers D, E, F_1, and F_2 will generally peak during the daytime. Moreover, all layers, with the occasional exception of the F_2 region, closely follow the tendency for the existence of a peak in the mean value electron density in the neighborhood of local noon. Figure 10 shows the mean diurnal variation of the E, F_1 and F_2 critical frequencies at solar maximum for a midlatitude site. Day-to-day F-region variability is exhibited in Fig. 11 for a period of maximum solar activity and midlatitudes.

Short-Term Variations. Variations in layer critical frequencies will occur from hour to hour and from day to day, especially for the F_2 region. It appears that much of this variability owes its existence to the impact of geomagnetic storms, TIDs, and miscellaneous F-region dynamic effects. TIDs are one of the more fascinating features of the ionosphere. They are the ionospheric tracers of neutral atmospheric gravity waves, which derive from a number of sources in the upper atmosphere. These sources include localized heating effects, atmospheric explosions, enhanced auroral activity, and other atmospheric phenomena that are associated with rapid and

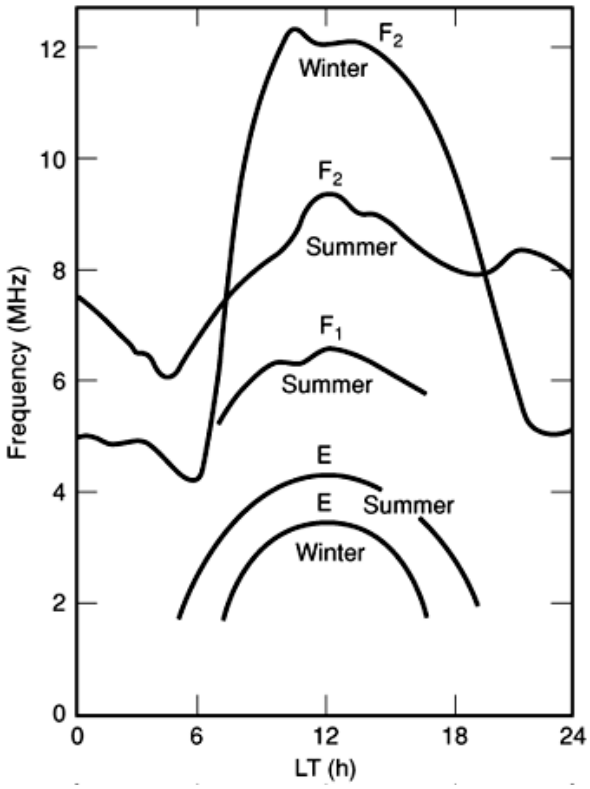

Fig. 10. Mean diurnal variation of foE, foF1, and foF2 for summer and winter under northern-hemisphere- and solar-maximum conditions. [From Jursa (8).]

nonuniform changes in atmospheric pressure. Figure 12 shows the variation of foF2 as a function of time, showing the impact of TIDs. Figure 13 shows the effect of a large geomagnetic storm.

Long-Term Solar-Activity Dependence of the Ionospheric Layers

There is a clear tendency for the ionospheric critical frequencies to increase with sunspot number. Figure 14 shows the long-term variation of R_{12}, foF2, and foE, and the D-layer absorption level (at 4 MHz), for noontime conditions. The D region is best characterized by the amount of absorption it introduces (see the subsection "Ionospheric Layering" above). A device for monitoring the D-region absorption is the *riometer*, which evaluates it as the product of D-region electron concentration and the electron collision frequency. From Figure 14, a slow 11-year modulation in the ionospheric parameters is evident. After smoothing, the results correlate well with sunspot number. Superimposed on this solar epochal variation is an annual variation, with D-region absorption and foE exhibiting summertime maxima, while foF2 exhibits a wintertime maximum (i.e., seasonal anomaly).

The slow but definite dependence upon mean sunspot number is illustrated in Fig. 15. This plot is unusual in that it presents running 12-month averages of the specified ionospheric parameters as well as of the sunspot number. This obscures the seasonal effects observed in Fig. 14.

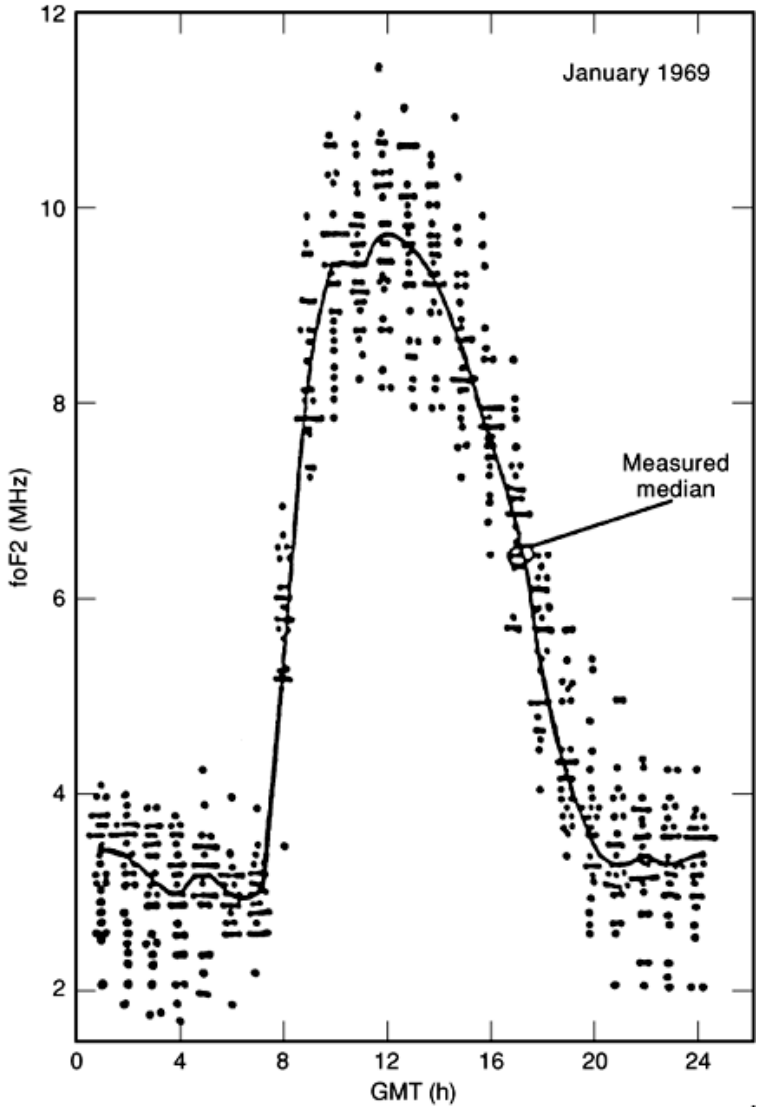

Fig. 11. Variations in the hourly values of foF2 as a function of the time of day, for January solar maximum conditions at a northern-hemisphere midlatitude site. The range of day-to-day variability in foF2 is $\sim \pm 10\%$, suggesting a variation in NmaxF2 of $\sim \pm 5\%$. [From Davies (1).]

Sporadic E

General Characteristics. Even though the normal E region is Chapman-like in nature, isolated forms of ionization are often observed in the E-region, having a variety of shapes and sizes. These ionization forms have been termed sporadic E, because they appear quasirandomly from day to day, and they generally defy deterministic prediction methods. Sporadic E (Es) ionization has been observed during rocket flights and with incoherent backscatter radar, and a layer thickness of the order of 2 km has been observed. It generally takes the form of large-scale structures, having horizontal dimensions of hundreds of kilometers at middle latitudes.

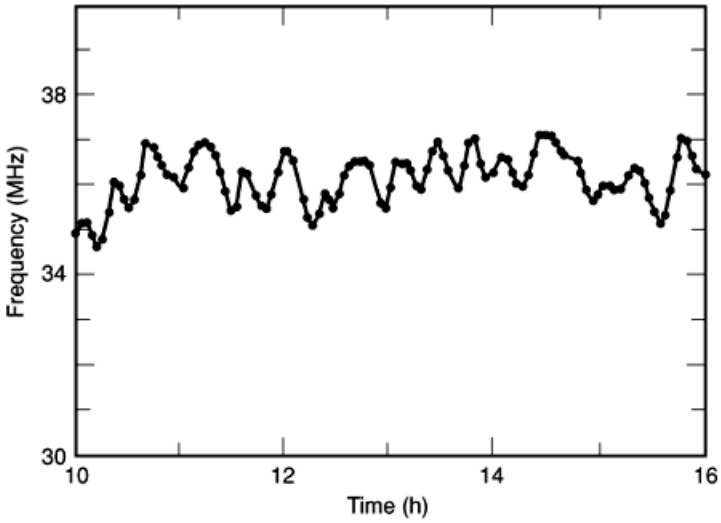

Fig. 12. Variations in the ionosphere thought to be associated with traveling ionospheric disturbances. The foF2 variations shown here are of the order of ±2% and have periods of ∼20 min. The NmaxF2 variations are ∼ ±1%. [From Paul (63).]

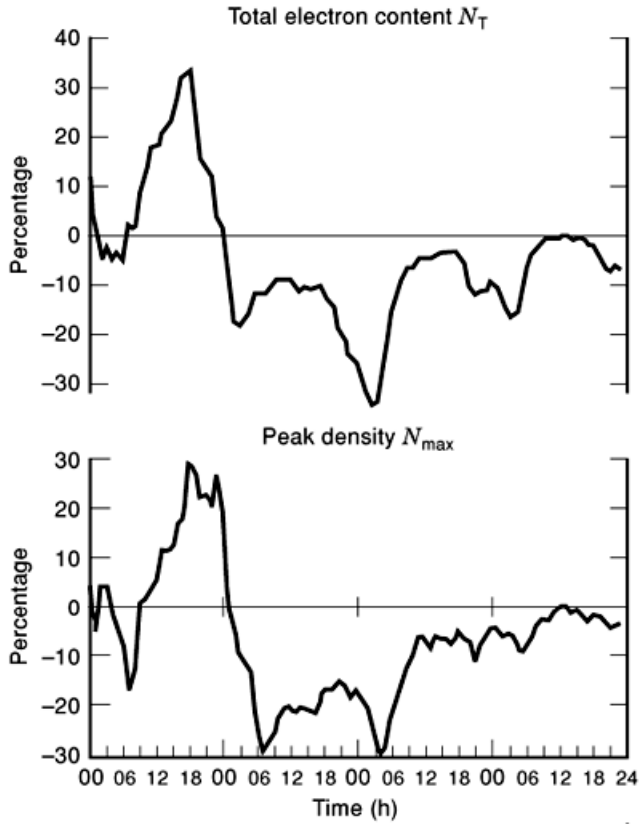

Fig. 13. Effect of a large geomagnetic storm on N_{max}. [By permission of J. M. Goodman and Kluwer Academic Publishers, Norwell, MA (11).]

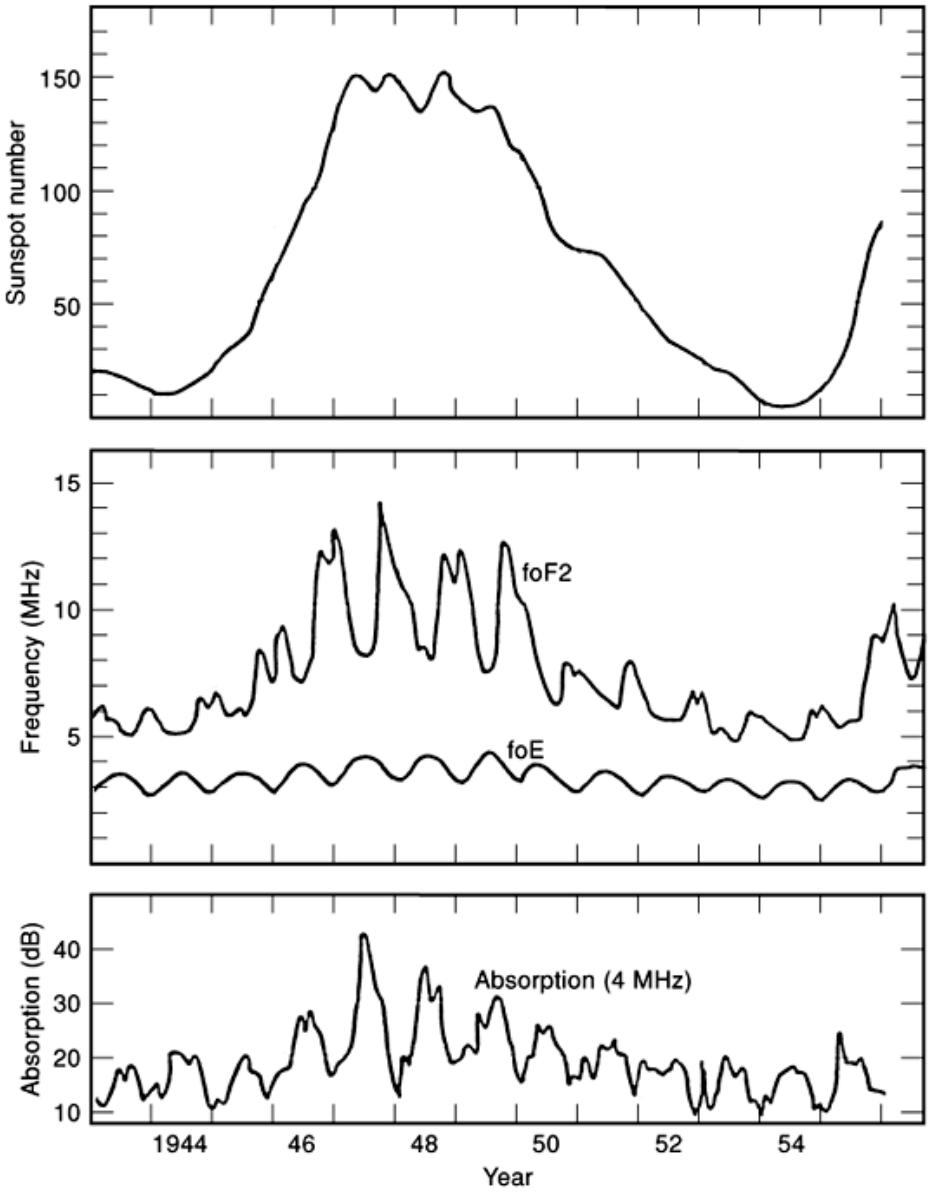

Fig. 14. Variation in R_{12}, foF2, foE, and 4 MHz absorption at noontime. The seasonal effects are clearly evident, the foE and D-layer variations being out of phase with the foF2 variations (i.e., seasonal anomaly). [By permission of J. M. Goodman and Kluwer Academic Publishers, Norwell, MA (11).]

Polar and equatorial forms have different structures and causal mechanisms. Although sporadic E consists of an excess of ionization (against the normal E-region background), it does not appear to be strongly tied to solar photoionization processes. Still, midlatitude Es occurs predominantly during summer days. Sporadic E does exhibit seasonal and diurnal tendencies, which have been examined statistically, and at least three different types of sporadic-E ionization have been discovered with distinct geographical regimes: low-latitude

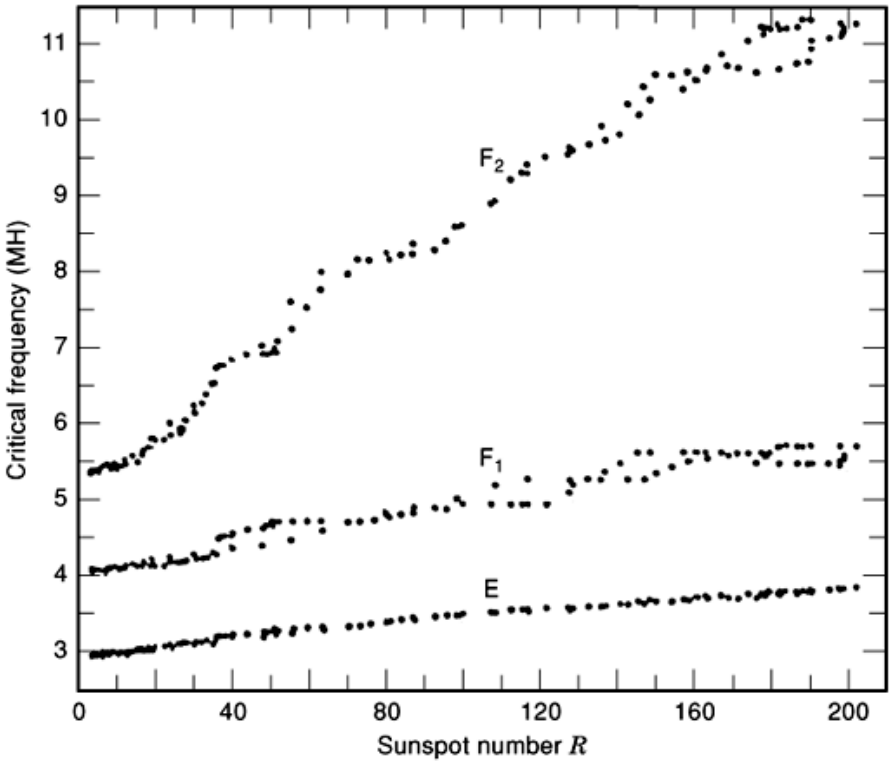

Fig. 15. Long-term variation in R_{12}, foF2, and foE at noontime. Since running 12-month averages were taken, the seasonal effects observed in Fig. 14 are smoothed out. [By permission of J. M. Goodman and Kluwer Academic Publishers, Norwell, MA (11).]

(or equatorial), midlatitude (or temperate), and high-latitude ionization. Figure 16 depicts the probability of Es occurrence.

Formation of Midlatitude Sporadic E. It has suggested that wind-shears in the upper atmosphere are responsible for the formation of sporadic E at midlatitudes. We shall review this process briefly.

It should be recalled from the examination of photochemistry in the ionosphere that molecular ions such as those that exist in the E region introduce rapid electron loss by recombination. At the same time it is recognized that an enormous number of meteors burn up in the E region. This meteoric debris is largely comprised of metallic ions, which are monatomic. Their presence has been confirmed by mass spectroscopy measurements using rockets, and they include iron, sodium, magnesium, etc. Since monatomic ions exhibit a small cross section for electron capture, the process by which atomic ions become concentrated in well-defined layers will lead to reduced loss rates for ambient free electrons in the interaction region.

The influx of this foreign mass of metallic ions, when distributed over the whole of the E region, would be insufficient to overwhelm the omnipresent molecular species (such as NO^+), which are in a state of photochemical equilibrium, were it not for a mechanism that preferentially concentrates the meteoric debris ions. Apparently wind shear is this mechanism. The basic wind shear theory was proposed by Whitehead (26), but it remained for Gossard and Hooke (27) to outline a process for meteoric ion concentration based upon the interaction of the meteoric debris with atmospheric gravity waves, the latter wave structures being responsible for the development of TIDs as well. The ultimate process involves a corkscrew propagation of atmospheric gravity waves and atmospheric tides, which results in a rotation of wind velocity as a function of altitude. This

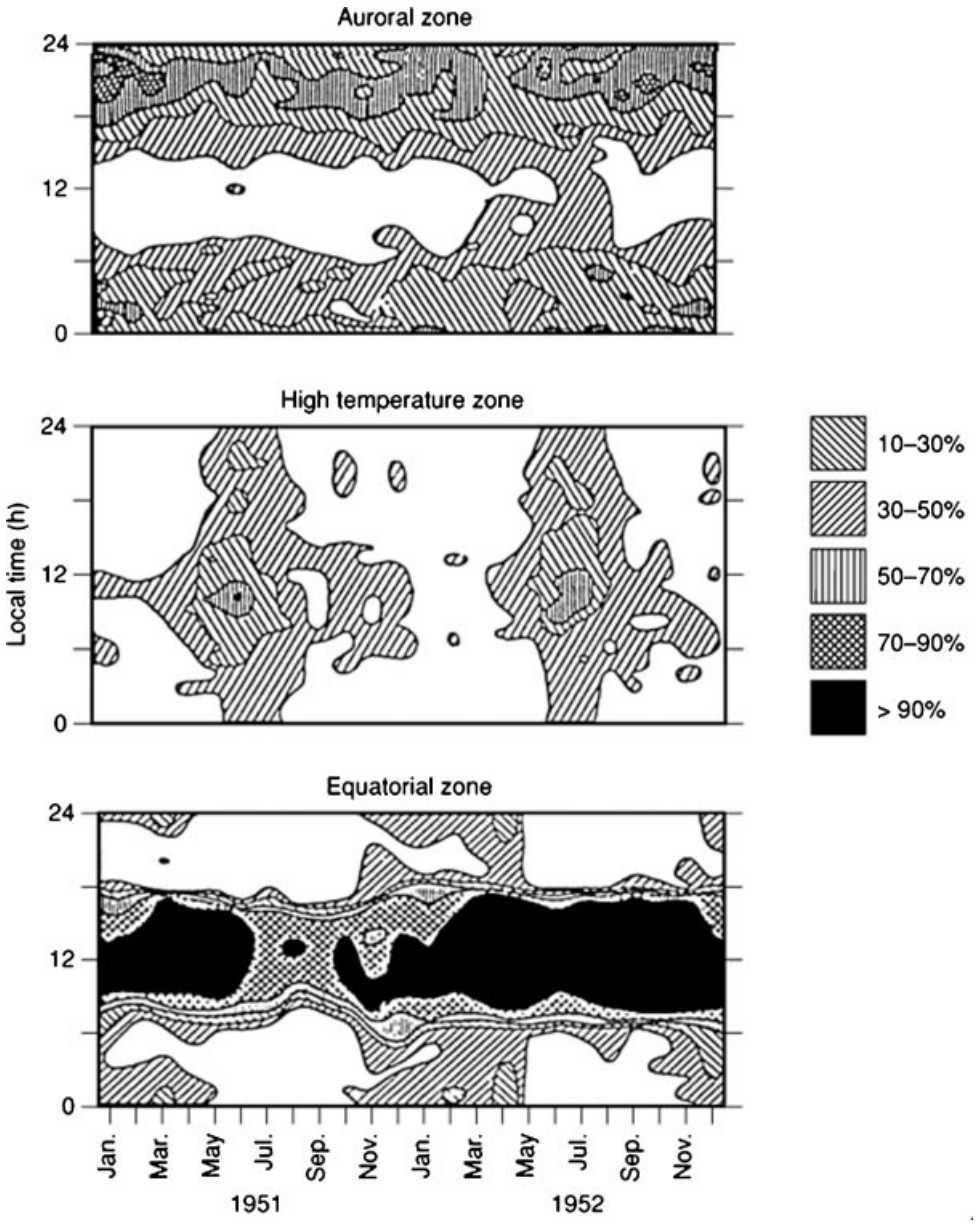

Fig. 16. Probability of Es occurrence as observed in the period 1951–1952. It is representative of the global, seasonal, and diurnal variation of sporadic-E ionization. [From Davies (1).]

effect can cause the wind to change direction over an altitude of only a kilometer or so, so as to trap meteoric ions at an intermediate point having zero velocity. 'This buildup in a narrow region is sufficient to generate an intense sporadic-E patch.

Sporadic E at Non temperate Latitudes. The high-latitude sources are evidently of two types, depending upon whether the observation is made in the neighborhood of the auroral oval or poleward of it

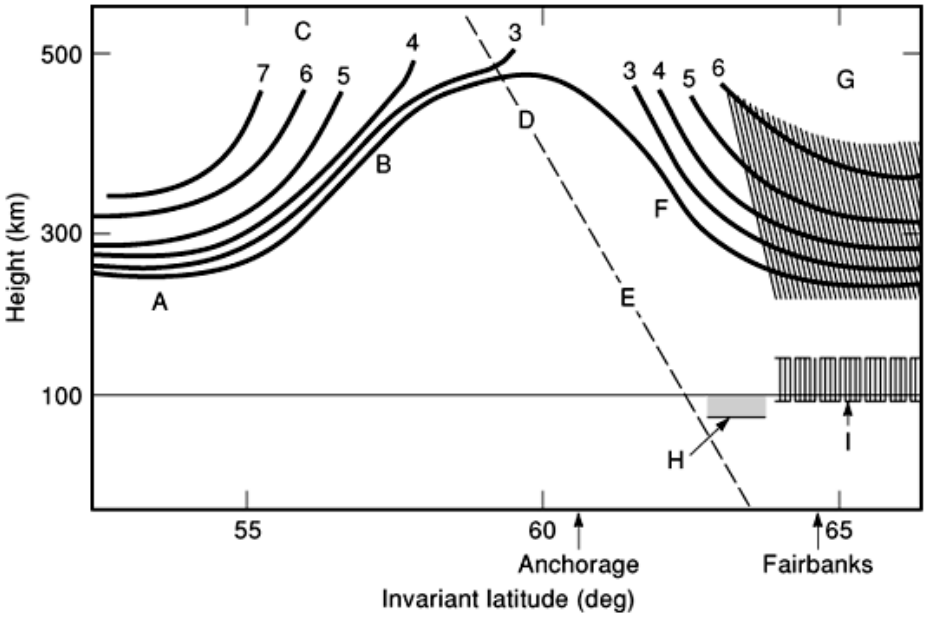

Fig. 17. Idealized picture of ionospheric plasma frequencies in a north–south plane through Fairbanks and Anchorage, Alaska. E, region equatorward of trough; B, equatorward edge of trough: C, plasma frequencies (MHz); D, trough minimum; E, plasmapause field line; F, poleward edge of trough; G, F-region blobs; H, enhanced D-region absorption; I, E-region irregularities. [By permission of J. M. Goodman and Kluwer Academic Publishers, Norwell, MA (11), after Hunsucker (28).]

(i.e., in the polar cap region). It has been found that auroral Es is basically a nocturnal phenomenon; it is associated with the optical aurora and is due to auroral electron precipitation. Because of its proximity to the seat of auroral substorm activity, it is not surprising to find some correlation between auroral Es and some appropriate magnetic index. Indeed, it has been found that auroral Es is positively correlated with magnetic activity. On the other hand, polar-cap Es may be relatively weak, and is negatively correlated with substorm activity.

Turning equatorward, it has been found that equatorial Es is most pronounced during daylight hours, and evidence points to the formation of ionization irregularities within the equatorial electrojet as the responsible agent at low latitudes.

The High-Latitude Ionosphere

From a morphological point of view, the high-latitude region is the most interesting part of the ionosphere. It has been said that the auroral zone and associated circumpolar features, are our windows to the distant magnetosphere, and the presence of visible aurorae has fascinated observers for centuries. The interplanetary magnetic field, which may be traced to its solar origins, has a significant impact on the geomorphology of the high-latitude ionosphere and its dynamics, including magnetic substorm development.

The high-latitude region of the ionosphere is characterized by a hierarchy of phenomena that are largely orchestrated by magnetospheric and interplanetary events (of a corpuscular nature) rather than solar (electromagnetic) flux variations. Hunsucker (28) has examined the salient features and they are depicted in Fig. 17, with particular emphasis on the high latitude trough. In Fig. 18, from Bishop et al. (29), many of the same features are depicted and compared with worldwide features.

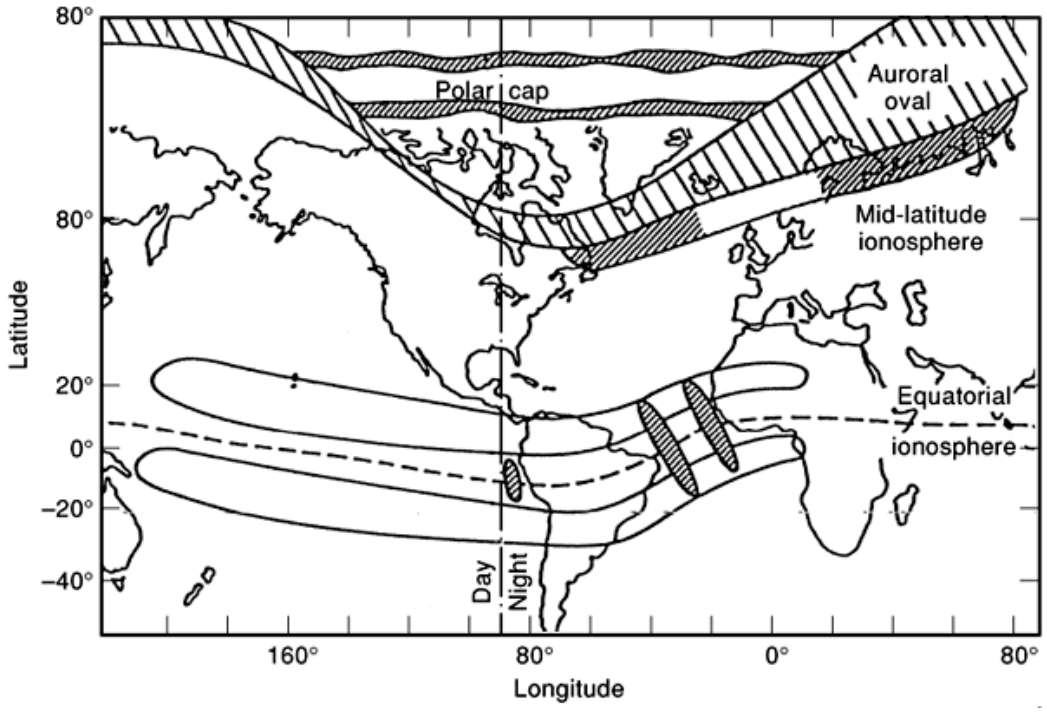

Fig. 18. Depiction of various ionospheric features at a given time such that the day–night terminator is passing through the middle of the United States (i.e., ≈ 2300 GMT). [By permission of J. M. Goodman and Kluwer Academic Publishers, Norwell, MA (11), from Bishop et al. (29).]

The magnetic activity index K_p is generally available and is typically used as the parameter of choice to determine the statistical position of the auroral zone. The concept of the auroral oval was developed by Feldstein and Starkov (30) on the basis of a set of all-sky camera photographs that were obtained during the International Geophysical Year. Other models exist, but the Feldstein picture is found in most models that attempt to include auroral effects in some way. The position of the oval is important, not only as an ionospheric feature in itself, but because it also represents a boundary between the decidedly different geophysical regimes that are poleward of it (the polar cap) and equatorward of it (the midlatitudes). Because the position of the auroral zone varies diurnally as well as with the index K_p, there are some sites that may be characterized by all four regimes at any given time: polar, auroral, trough, and midlatitude. Iceland is such a location.

One of the most fascinating properties of the various circumpolar features is their latitudinal motion as a function of magnetic activity. The ionospheric plasma is best organized in terms of some form of geomagnetic coordinates, but the high-latitude plasma patterns are not fixed in that frame of reference either. The equatorward boundary of the region of precipitating electrons has been deduced from DMSP satellite instruments, and it takes a form due to Gussenhoven et al. (31):

$$L(t) = L_0(t) + a(t)\,K_p \qquad (18)$$

where corrected geomagnetic coordinates are used, $L(t)$ and $L_0(t)$ are specified in degrees, and $L_0(t)$ is the equatorward boundary of the oval when $K_p = 0$. It is emphasized that $L_0(t)$ and $a(t)$ are functions of time in the magnetic local time (*MLT*) system.. Both functions are smoothly varying over the diurnal cycle: L_0 ranges between 65° at ≈0100 MLT and ≈72° at ≈1700 MLT, and $a(t)$ varies between ≈ −2 at 2400 MLT and

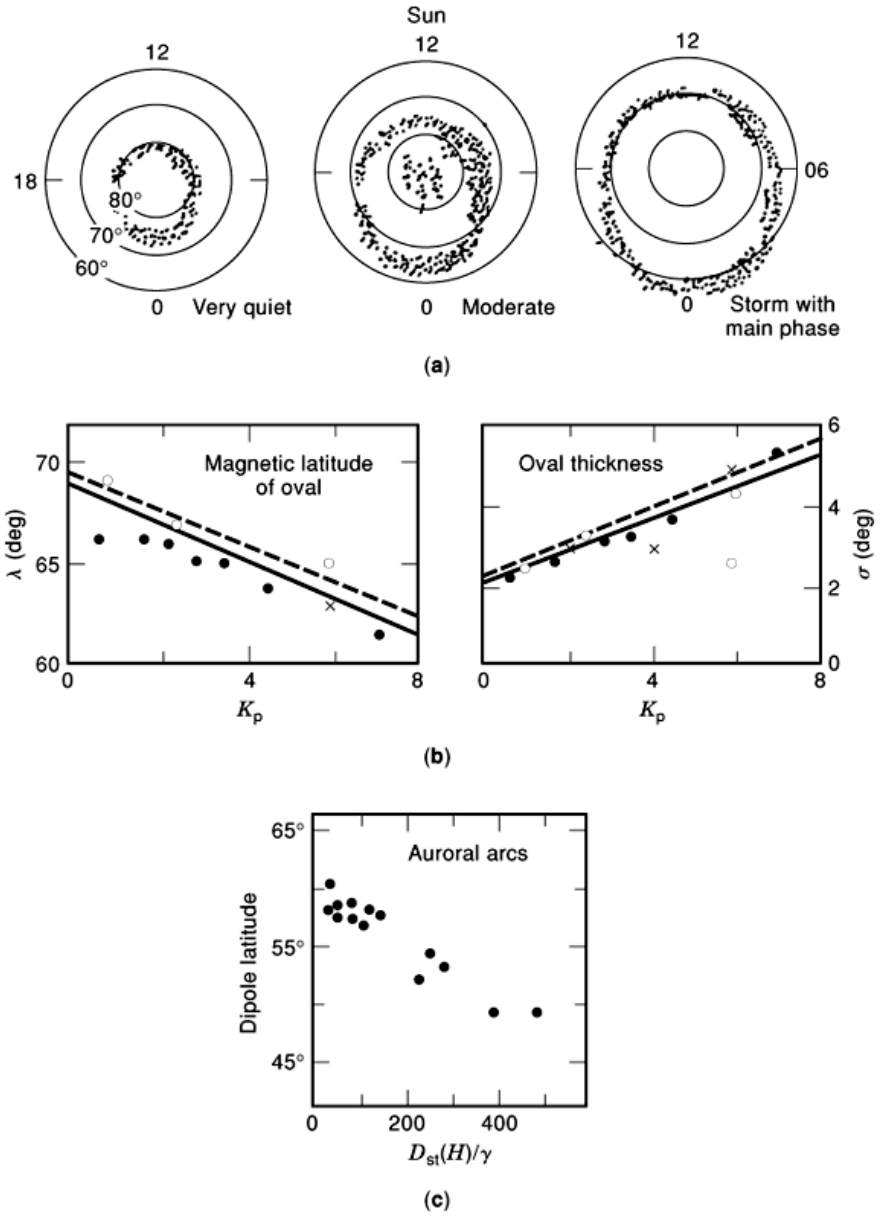

Fig. 19. (a) Descent of the auroral oval as a function of magnetic activity; (b) position of the auroral oval and its thickness versus the magnetic index K_p; (c) position of the auroral arc formations versus the magnetic activity index D_{st}. [By permission of J. M. Goodman and Kluwer Academic Publishers, Norwell, MA (11).]

-0.8 at $\approx$1500 MLT. Therefore the statistical representation of the oval has its greatest equatorward descent during nocturnal hours. Moreover this equatorward boundary is greatly influenced by magnetic activity. Chubb and Hicks (32) have found that the daytime aurora descends approximately 1.7 degrees per unit K_p, and the nighttime aurora descends at a rate of 1.3 degrees per unit K_p. The auroral oval and thickness are depicted in Fig. 19. Ultimately the auroral arcs, which reside within the auroral oval, are tied to interplanetary phenomena.

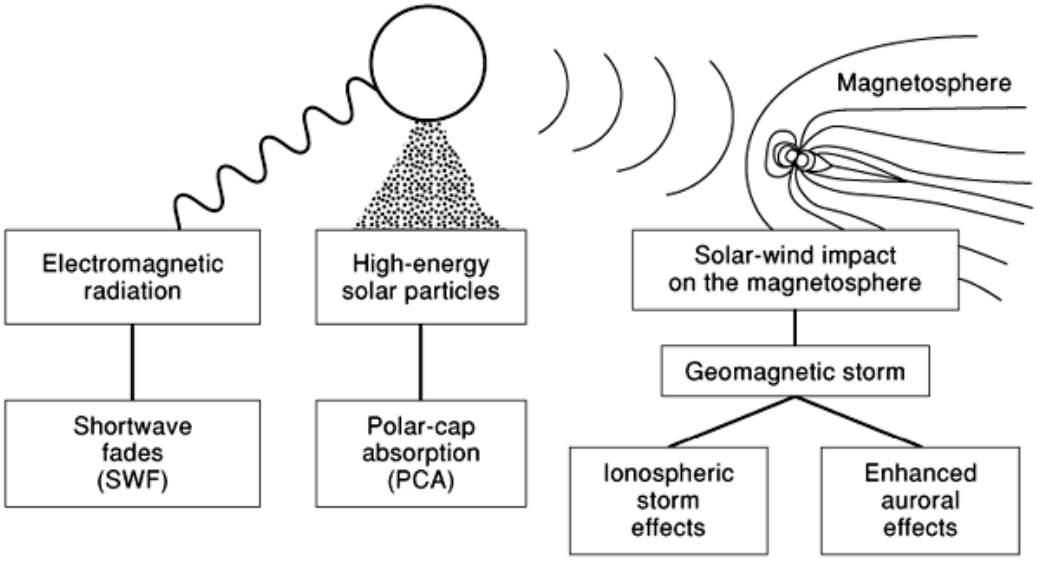

Fig. 20. Hierarchy of solar–terrestrial effects. [By permission of J. M. Goodman and Kluwer Academic Publishers, Norwell, MA (11).]

Workers have shown that the magnitude of the southward component of the interplanetary magnetic field is a key factor is the development of so-called geomagnetic substorms, wherein K_p exhibits large enhancements.

The US Air Force prepares daily summaries of an index Q in order to provide a basis for various analyses of the high-latitude ionosphere. The index Q ranges between 0 and 8, with larger values associated with a widening of the oval region and a general increase in intensity of activity within the oval. Moreover, the equatorward boundary of the auroral oval moves to lower latitudes as Q increases. Since Q, viewed as a parameter, defines the shape and location of the auroral zone, it is a convenient index for transmission to communication facilities and forecasting facilities. Its utility is dependent upon timeliness and accuracy. As originally designed by Feldstein, Q defines only a statistical relationship between the oval position and magnetic activity, the latter being parameterized by the planetary index K_p. Nevertheless, the Feldstein oval concept has been shown to have some utility under real-time circumstances. Satellite imagery is used to deduce an effective Q.

Auroral physics is an exceedingly rich and complex subject. Not all phenomena in the high-latitude region are understood, and insufficient data are available to fully characterize even those factors for which a general understanding exists.

Ionospheric Response to Solar Flares

Now we shall take note of a special class of effects called *sudden ionospheric disturbances* (*SID*s). These constitute those events that arise as a result of the atmospheric interaction with electromagnetic flux from solar flares. A book by Mitra (33) is an excellent treatise on the ionospheric effects of solar flares.

We recognize that the sun is the ultimate source for a large variety of ionospheric and magnetospheric effects. Fig. 20 exhibits the hierarchy of solar-induced ionospheric effects. There are many types of SID observed; one of the most important is the *short-wave fade* (*SWF*), which affects HF communication circuits on the sunlit side of the earth. The source of the enhanced D-region ionization responsible for the SWF is typically an impulse burst of X-ray energy from within an active region on the sun (generally a sunspot). An X-ray flare generates

a significant increase in D-layer ionization with a temporal pattern that mimics the flare itself. This results in an increase in the product of the electron density and the collision frequency. It is the growth of this product that accounts for the absorption of HF signals passing through the D region. Flares tend to be more prevalent during the peak in sunspot activity, and the individual-flare duration distribution ranges from a few seconds to roughly an hour.

The Ionospheric Storm

The magnetic storm is a fascinating geophysical phenomenon, which goes far beyond the visible evidence corresponding to auroral displays at high latitudes. It is central to the issues surrounding what is now referred to as space weather. A discourse on this subject is beyond the scope of this article, but the reader is referred to an excellent geophysical monograph edited by Tsurutani et al. (34).

The ionospheric storm is the ionosphere's response to a geomagnetic storm. While the ionospheric response to magnetic storms is varied, it has been shown that they may be conveniently classified as either *positive* or *negative* in nature. The main attribute of so-called negative storms is that they are generally associated with decreases in foF2. Positive storms exhibit the opposite behavior. At midlatitudes the ionospheric storm signature is typically commensurate with the main features of a negative storm, although variations may occur. Often the temporal (or storm-time) pattern is complex. For example, the midlatitude ionospheric response to a large magnetic storm is generally characterized by a short-lived increase in the F-region electron concentration in the dusk sector following storm commencement (*SC*), after which it decreases dramatically (see Fig. 13). The initial short-lived enhancement is observed in foF2 records, and it is correlated with the initial positive phase of the geomagnetic storm. The main phase of the geomagnetic storm is correlated with a concomitant foF2 diminution, and this reduction in foF2 may last for a day or longer. It is thought that the initial enhancement in foF2 is a result of electrodynamic forces, while the long-term reduction in foF2 is associated with changes in upper-atmospheric chemistry and modification of thermospheric wind patterns. A key factor in this process is ionospheric heating through dissipation of storm-induced atmospheric gravity waves. This heating effect will cause the thermosphere to expand, and ionospheric loss rates will increase.

Ionospheric Current Systems

Current systems are important in an understanding of ionospheric perturbations associated with the onset of geomagnetic storms and the progression of these events. There are four principal current systems in the ionosphere that give rise to relatively rapid fluctuations in the geomagnetic field: the ring current, the magnetopause current system, the atmospheric dynamo, and various high-latitude current systems. The first two are associated with magnetic storms and occur at magnetospheric distances. The atmospheric dynamo is important in an understanding of tidal-driven forces, which interact with the ionospheric plasma, causing a vertical drift of the F-region ionization. Excellent descriptions of ionospheric current systems and dynamo theory may be found in a monograph by Rishbeth and Garriott (35) and in a book by Ratcliffe (4). High-latitude (i.e., polar and auroral) currents and atmospheric dynamo currents are observed at lower ionospheric heights in the vicinity of the E layer. Brekke (36) provides a good treatment of relevant high-latitude current systems. There is also a current system within the neighborhood of the magnetic equator: the equatorial electrojet, a current that flows along the geomagnetic equator, eastward by day and westward by night. It is associated with a class of discrete ionospheric formations that are termed equatorial sporadic E.

Ionospheric Models

As in many areas of geophysical study, ionospheric modeling may assume a number of forms, ranging from the purely theoretical to the totally empirical. Approaches may also include a combination of these forms, although empirical models dominate the field. Recent developments include allowance for adaptivity within the models to accommodate exploitation in the near-real-time environment for special applications. While physical or theoretical principles are the inspiration for a number of models, in fact most models in use today are largely specified on the basis of semiempirical relationships derived from observational data.

Ionospheric models fulfill a variety of needs beyond basic research, with the most prominent application being the assessment and prediction of radio-system performance. For example, ionospheric models are the engines that drive HF system performance models such as IONCAP (37). Related models are supported by the U.S. Department of Commerce, including VOACAP, ICEPAC, and REC533 (38). Other applications include evaluation of transionospheric signal parameters and errors in ranging or geolocation introduced by the electron content of the ionosphere. A general discussion of the status of ionospheric modeling in the context of HF communication systems has been covered by Goodman (11), and recent information regarding telecommunication system planning has been published by the Commission of European Communities (16).

Ionospheric profile models are based upon the superposition of various submodels of the ionospheric layers or regions (i.e., D, E, Es, F_1, and F_2).The basic purpose of modeling is to represent the electron density profile under a variety of conditions [see Fig. 3(b)]. These submodels may represent the respective layers as thin horizontal sheets (e.g. sporadic E) or quasiparabolic regions in the vicinity of maximum ionization. The submodels are specified by the maximum electron density of the layer, the layer height, the layer thickness, and a functional representation of the layer shape. There are a number of models for the height profile, the main differences being the manner in which the component layers are combined. Figure 21 depicts the International Reference Ionosphere (39) and the ionospheric model in the computer program IONCAP. There are also geographical, seasonal, and solar epochal variations in the specified ionospheric profiles and the parameters upon which they are built. An example of the geographical variations in foF2 was shown in Fig. 9, and the *Global Atlas of Ionospheric Coefficients* was discussed in the section "Description of the Ionospheric Layers." Ionospheric coefficients used to produce maps similar to Fig. 9 are common to virtually all global models of the ionosphere. Currently there are two sets of ionospheric coefficients that may be specified: the original CCIR (or ITU-R) set, which is sanctioned by the ITU-R, and the newer URSI set (40).

A simplistic model of the ionosphere consists of a parabolic E layer, a linear increase in electron density in the F_1 layer a parabolic F_2 layer (41). At nighttime, the E and F_1 layers effectively disappear. A newer ITU-R recipe, consisting of multiple quasiparabolic layers to provide continuity of the overall profile and its height derivatives (42), has replaced this so-called Bradley–Dudeney profile model.

Significant improvements in empirical ionospheric modeling have been promoted by military agencies around the world, including the US Department of Defense, the UK Ministry of Defence, and others. This is not surprising in view of the large number of applications of ionospheric specification in radio-wave systems used by the military. The original ICED model was intended to be a northern-hemisphere ionospheric specification model to serve the requirements of the US Air Force. It was only a regional model, descriptive of midlatitude behavior but extending into the auroral zone. It was designed to allow for recovery of some of the dynamic features embodied in auroral climatology that are smeared out in most mapping procedures. The model, as described by Tascione et al. (43,44), is driven by an effective sunspot number and an index derived from auroral oval imagery. The effective sunspot number is not based on solar data at all, but is derived from ionospheric data extracted from the US Air Force real-time ionosonde network. This sunspot number is similar to an ionospheric T index developed by Australian workers, and to the pseudoflux concept used by the US Navy for HF predictions (11). The ICED model has been generalized to incorporate global considerations, while emphasizing near-real-time applications. Anderson et al. have developed a low-latitude ionospheric profile model, SLIM (45) and a fully analytic ionospheric model, FAIM (46) in order to eliminate the use of limiting

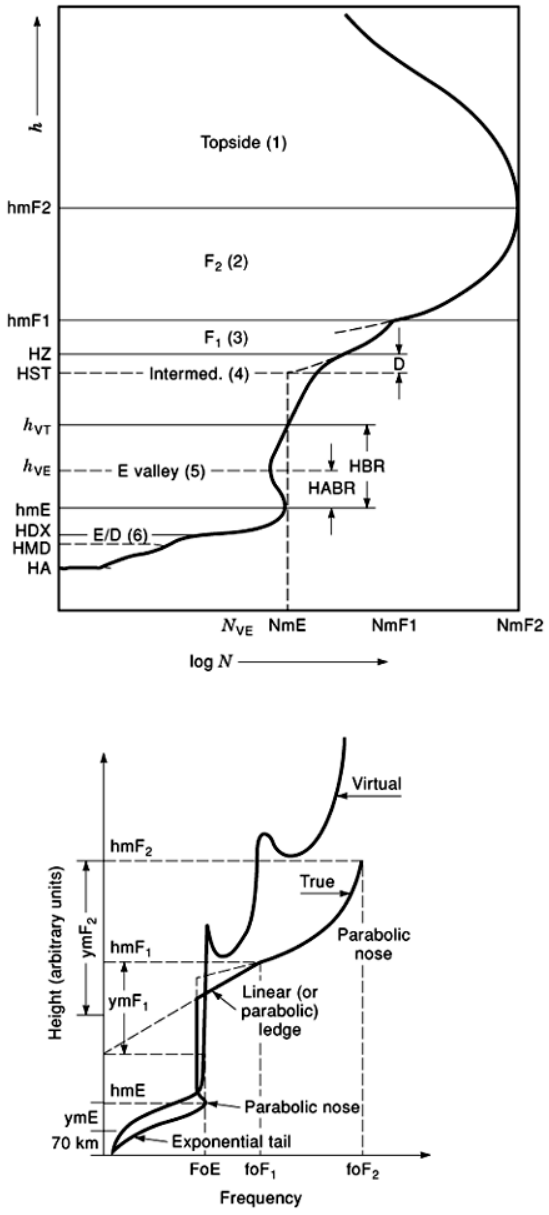

Fig. 21. (a) International Reference Ionosphere (*IRI*); (b) IONCAP ionospheric submodel. [By permission of J. M. Goodman and Kluwer Academic Publishers, Norwell, MA (11).]

simplifications in the driving parameters associated with prediction models. A discussion of SLIM and FAIM may be found in a paper by Bilitza (47). Other developments supporting Air Force requirements include PIM and PRISM. PIM (Parameterized Ionospheric Model) is a global model of theoretical and empirical climatology, which specifies the ionospheric electron and ion densities from 90 km to 25,000 km. PRISM (Parameterized Real-Time Ionospheric Specification Model) uses ground-based and space-based data available in real time

to modify PIM, thereby providing a near-real-time ionospheric specification. Another model, RIBG, developed by Reilly (48), combines ICED and several other models with a general ray-tracing utility. Current versions of these models and validation of PRISM are discussed by Doherty et al. (49). A survey of computer-based empirical models of the ionosphere has been published by Secan (50).

The IRI mentioned previously (see Fig. 21), is a global empirical model that specifies monthly averages of electron, ion, and neutral temperatures, in addition to electron and ion densities, from about 50 km to about 2000 km (39). The IRI development is a joint project of URSI and the international "Committee of Space Research" (COSPAR), and has proven to be a useful model for scientific research. Another model of primary interest to workers studying transionospheric propagation effects is the so-called Bent model, a profile model based upon topside and bottomside sounder data (51). Simplicity is not always important in the age of sophisticated computers, but the Ching–Chiu model (52) has found a number of scientific applications in cases in which detailed ionospheric specification is not paramount.

Aside from global modeling of the ionosphere, there have been attempts to model selected regions of the world more accurately. During the decade of the 1990s, European scientists affiliated with the COST program have taken a lead in regional modeling and mapping of the ionosphere (15,16)

Ionospheric Predictions

Ionospheric predictions influence several disciplines, including the prediction of radio system performance, a matter of some interest in planning as well as ultimate operations. Long-term predictions are generally based upon predictions of driving parameters such as the sunspot number, the 10.7 cm solar flux, and magnetic activity indices. Unfortunately, these parameters are not easy to predict. Moreover, the functions relating these parameters to the ionosphere are imprecise. Therefore, long-term predictions needed for system design are subject to a considerable amount of uncertainty. To first order the uncertainty in the median value of foF2 for a particular time and location is proportional to the uncertainty in the sunspot number.

In addition to the uncertainty in the mean parameters, we must allow for the fact that ionospheric parameters have real distributions, and with few exceptions the spread of these distributions is such that errors about the mean may be a dominant contribution. Short-term ionospheric predictions (or forecasts) generally refer to departures from the median behavior, the latter being well characterized by running averages of solar flux and related parameters (e.g., sunspot number). The short-term fluctuations may be specified in terms of hour-to-hour, day-to-day, and week-to-week variabilities. There are also second-to-second and minute-to-minute variations, they are generally unpredictable. Compensation for such fluctuations is quite difficult, but may be accommodated through use of system protocols that enable real-time channel evaluation (*RTCE*) measures to be initiated, such as channel sounding or probing. These very short-term forecasts are generally referred to as *nowcasts*.

There are four ITU-R documents that are pertinent to the investigation of the ionospheric forecasting problem. The first deals with the exchange of data forecasts (53), the second outlines various measures for forecasting of ionospheric parameters (54), the third deals specifically with solar-induced ionospheric effects (55), and the fourth outlines various real-time channel evaluation schemes (56). These reports should be consulted.

The distributions of parameters such as foF2, foEs, and hF2 are important, since these parameters depart significantly from fundamental intuition and from rules set forth by Sidney Chapman in his classic theory. Distributions of foF2 and foEs are available (57,58), but F_2-layer height distributions are not directly available. Ionospheric predictions in the short and intermediate terms provide the most exciting challenge for ionospheric researchers.

Observational data have shown that TIDs are the ionospheric tracers to a class of atmospheric gravity waves, and these disturbances are a major contribution to ionospheric variability, especially at F-region heights.

TIDs have a major impact on layer height as well as peak electron density, and possess a variety of scales, from kilometers to thousands of kilometers. The small- to intermediate-scale TIDs, having wavelengths of less than a few hundred kilometers and periods of the order of 10 min to 20 min, arise from local sources and have relatively small amplitudes away from the source region. The large-scale TIDs have sources that are located at great distances, and there is a strong correlation between this class of disturbances and geomagnetic storms. Evidence suggests that large-scale TIDs have an impact over global distances and originate within the auroral zone as a result of atmospheric modifications associated with precipitation and auroral arc formation. A survey of the effects of TIDs on radiowave systems may be found in a review paper by Hunsucker (59).

The field of ionospheric predictions is undergoing continuous evolution with the introduction of new scientific methods and instruments, which are providing fresh insight. The requirements for quasi-real-time products based upon current ionospheric specification has led to increased importance of so-called real-time ionospheric models. This class of models, in turn, is driven by a hierarchy of solar–terrestrial observations, which enables the analyst to examine the space-weather environment as an integrated complex of phenomena. This general approach is leading to an improvement in our understanding of ionospheric structure and it variations, if not better short-term forecasts. In the immediate future, it is anticipated that the primary ionospheric specification tools will consist of terrestrial sounding systems, including real-time networks of ionospheric sounders (60). Real-time data services based on these approaches are becoming available (61).

Fundamental Science Issues and Challenges in Ionospheric Research

There are a number of challenges facing ionospheric specialists and aeronomists. While theories explaining most facets of ionospheric behavior exist and are generally accepted, the theories do not always provide a good basis for prediction. This is because the driving forces and boundary conditions needed in a physical model are not always known, and estimates must be used. This has led to the development of semiempirical models for the purpose of system design, and these are used for operations as well. By and large these models exploit large ionospheric databases and yield only median representations of ionospheric parameters. To fix this problem, various *update* schemes have been developed to make the specification of the ionospheric state as current as possible. The physics is then used to let the system evolve. All of this can be very unsatisfactory unless an understanding of the nature of ionospheric variability (in both space and time) is established. There are many sources within the earth–sun system that contribute to the growth of ionospheric structure. While these have been characterized to some extent, the characterizations are not sufficient to provide predictions acceptable for many users of the ionospheric channel. Currently this is a major challenge facing the ionospheric research community.

The following topics require more attention from ionospheric specialists: (a) the driving forces of upper atmospheric winds and their impact on ionospheric structure and dynamics; (b) the hierarchy of energy sources within the earth–sun system that influence ionospheric behavior; (c) the development of geomagnetic storms and the impact that storms have on ionospheric behavior; (d) the development and evolution of ionospheric inhomogeneities; and (e) various methods for ionospheric prediction.

Finally, in the new millennium, the researcher is confronted with an enormous amount of data, both near-real-time and archived, that may be accessed via the Internet. Harnessing this information stream, and using the state-of-the-art computational assets, it should be possible to leverage ongoing science efforts, organize more efficient experimental campaigns, and enhance collaborative efforts, all resulting in a fuller understanding of ionospheric physics. Some examples include programs such as the Solar–Terrestrial Energy Program (*STEP*) and the STEP Results, Applications, and Modeling Phase (*S-RAMP*). These programs are international efforts run under the aegis of the Scientific Committee on Solar–Terrestrial Physics (*SCOSTEP*), in cooperation with the International Council of Scientific Unions (*ICSU*).

BIBLIOGRAPHY

1. K., Davies *Ionospheric Radio Propagation*, NBS Monograph 80, Washington: USGPO, 1965.
2. K., Davies *Ionospheric Radio Waves*, Waltham, MA: Blaisdel, 1969.
3. K. Davies *Ionospheric Radio*, IEE Electromagnetic Wave Series 31, London: Peregrinus, Ltd., IEE, 1990.
4. J. A. Ratcliffe *An Introduction to the Ionosphere and the Magnetosphere*, London: Cambridge Univ. Press, 1972.
5. A. Giraud M. Petit *Ionospheric Techniques and Phenomena*, Dordrecht: Reidel, 1978.
6. M. C., Kelley *The Earth's Ionosphere, Plasma Physics and Electrodynamics*, San Diego, CA: Academic Press, 1989.
7. H., Rishbeth Basic physics of the ionosphere, in M. P. M. Hall and L. Barclay, (eds.), *Radiowave Propagation* IEE Electromagnetic Series 30, Stevenage, Peregrinus UK, 1991.
8. A. S. Jursa (scientific ed.), *Handbook of Geophysics and the Space Environment*, Springfield, VA: Air Force Geophysics Laboratory, Air Force Systems Command, US Air Force, National Technical Information Service (NTIS), 1985.
9. T. Tascione *Introduction to the Space Environment*, Malabar, FL: Orbit, 1988.
10. R. D. Hunsucker *Radio Techniques for Probing the Terrestrial Ionosphere*, New York: Springer-Verlag, 1991.
11. J. M. Goodman *HF Communications: Science & Technology*, New York: Van Nostrand Reinhold, 1991.
12. E. E. Johnson *Advanced High Frequency Radio Communications*, Boston: Artech House, 1997.
13. L. McNamara *The Ionosphere: Communications, Surveillance, and Direction Finding*, Malabar, FL: Orbit, Kreiger, 1991.
14. J. M. Goodman (ed.) *Ionospheric Effects Symposia*, Springfield, VA: National Technical Information Service. 1975, 1978, 1981, 1984, 1987, 1990, 1993, 1996, 1999,
15. P. A. Bradley (chair, COST 238 Mgmt.Committee), PRIME: Prediction and retrospective ionospheric modeling over Europe, Final Report, Commission of European Communities, Chilton, Didcot, Oxfordshire, UK: Rutherford-Appleton Laboratory, 1999.
16. R. Hanbaba Improved quality of service in ionospheric telecommunication systems planning and operations, COST 251 Final Report, Warsaw: Space Research Centre, 1999.
17. S. Chapman The absorption and dissociative or ionizing effect of monochromatic radiation in an atmosphere on a rotating earth, *Proc. Phys. Soc*, **43**: 26, 1931.
18. L. M. Muggleton A method for predicting foE at any place and time, *Telecomm. J.*, **42** (7), 413–418, 1975.
19. E. D. Ducharme L. E. Petrie R. Eyfrig A method for predicting the F_1 layer critical frequency based upon zurich smoothed sunspot number, *Radio Sci.*, **8**: 837–839, 1973.
20. CCIR, *CCIR Atlas of Ionospheric Characteristics*, Report 340, General Assembly held in Oslo, Geneva: ITU, 1966.
21. P. N. Mayaud *Derivation, Meaning, and the Use of Geomagnetic Indices*, Washington: American Geophysical Union, 1980.
22. D. F. Muldrew F-layer ionization trough deduced from alouette data, *J.Geophys. Res.* **70**: 2635-2650, 1965.
23. B. W. Halcrow J. S. Nisbet A Model of F_2 peak electron densities in the main trough region of the ionosphere, *Radio Sci.*, **12**: 815-820, 1977.
24. R. D. Hunsucker R. A. Greenwald, (eds.) Special issue, radio probing of the high latitude ionosphere and atmosphere—new techniques and new results, *Radio Sci.*, **18**: 1983.
25. J. Aarons Equatorial scintillations: A review, *IEEE Trans. Antennas Propag.*, **25**: 729–736, 1977.
26. J. D. Whitehead Production and prediction of sporadic E, *Rev. Geophys. Space Phys.*, **8**: 65–144, 1970.
27. E. E. Gossard W. H. Hooke *Waves in the Atmosphere*, Amsterdam: Elsevier Science, 1975.
28. R. D. Hunsucker Anomalous propagation behavior of radio signals at high latitudes, in H. Soicher (ed.), *Propagation Aspects of Frequency Sharing, Interference and System Diversity*, AGARD-CF-332, Springfield, VA: NTIS, 1983.
29. G. J. Bishop *et al.* A modern trans-ionospheric propagation sensing system, in *Operational Decision Aids for Exploiting or Mitigating Electromagnetic Propagation Effects*, NATO-AGARD-CP-453, UK: Specialised Printing Services, 1989.
30. Y. I. Feldstein G. N. Starkov Dynamics of auroral belts and polar geomagnetic disturbances, *Planet. Space Sci.*, **No. 15**, p. 209, 1967.
31. M. S. Gussenhoven D. A. Hardy N. Heinemann Systematics of the equatorward diffuse auroral boundary, *J. Geophys. Res.*, **88**: 5692, 1983.
32. T. A. Chubb G. T. Hicks Observations of the aurora in the far ultraviolet from OGO 4, *J. Geophys. Res.*, **75**: 1290–1311, 1970.
33. A. P. Mitra *Ionospheric Effects of Solar Flares*, Dordrecht: Reidel, 1974.

34. B. T. Tsurutani et al. (ed.) *Magnetic Storms*, Geophysical Monograph 98, Washington: American Geophysical Union, 1997.
35. H. Rishbeth O. K. Garriott *Introduction to Ionospheric Physics*, New York: Academic Press, 1969.
36. A. Brekke Currents in the auroral zone ionosphere, in *The Physical Basis of the Ionosphere in the Solar–Terrestrial System*, AGARD-CP-295, London: Tech. Edit. & Reprod., pp. 13.1–13.9, 1980.
37. L. R. Teters, *et al.* Estimating the performance of telecommunication systems using the ionospheric transmission channel, IONCAP user's manual, NTIA Report 83-127, PB84-111210, Springfield, VA: U.S. Dept. of Commerce, 1983.
38. ITS, *HF Propagation Models*: [Online], 2000. Available www: http://elbert.its.bldrdoc.gov/hf.html
39. D. Bilitza International reference ionosphere, NASA, NSSDC 90-22, Greenbelt, MD: World Data Center A (Rockets and Satellites), 1990.
40. C. M. Rush *et al.* Ionospheric mapping—an update of fof2 coefficients, *Telecomm. J.*, **56**: 179–182, 1989.
41. P. A. Bradley J. R. Dudeney A simple model of the vertical distribution of electron concentration in the ionosphere, *J. Atmos. Terrest. Phys.*, **35**: 2131–2146, 1973.
42. M. I. Dick P. L. Bradley The RAL quasi-parabolic model ionosphere *Proc. COST–PRIME Workshop on Data Validation of Ionospheric Models and Maps*, COST 238TD (93) 001, 1992, pp. 67–83.
43. T. F. Tascione, *et al.* A technical description of the ionospheric conductivity and electron density profile model (ICED, Version 196-11), Scott AFB, IL: Hqrtrs Air Weather Service, US Air Force, 1987.
44. T. F. Tascione *et al.* New ionospheric and magnetospheric specification models, *Radio Sci.*, **23**: 211–222, 1988; also in Proc. IES'87.
45. D. N. Anderson M. Mendillo B. Herniter A semi-empirical low-latitude ionospheric model *Radio Sci.*, **22**: 292, 1987.
46. D. N. Anderson J. M. Forbes M. Codrescu A fully analytic, low latitude ionospheric model, *J. Geophys. Res.*, **94**: 1520, 1989.
47. D. Bilitza Solar terrestrial models and application software, *Planet. Space. Sci.*, **40**(4): 541–579, 1992.
48. M. H. Reilly M. Singh A transionospheric radio propagation model, in J. M. Goodman, (ed.), *IES93 Proc.* Springfield, VA: NTIS, 1993.
49. P. H Doherty *et al.* Validation of PRISM: The climatology, in J. M. Goodman, (ed.), *Proc. IES99* Springfield, VA: NTIS, 1999.
50. J. A. Secan A survey of computer-based empirical models of ionospheric electron density, Report NWRA CR-89-11038, Bellevue, WA: Northwest Research Associates, 1989.
51. R. B. Bent *et al.* The development of a highly-successful worldwide empirical ionospheric model and its use in certain aspects of space communications and worldwide total electron content investigations, in J. M. Goodman (ed.), *Effect of the Ionosphere on Space Systems and Communications, IES'75*, USGPO, Springfield VA: NTIS, 1976, pp. 13–28.
52. B. K. Ching Y. T. Chiu A phenomenological model of global ionospheric electron density in the E, F_1, and F_2-regions, *Atmos. Terrest. Phys.*, **35**: 1615–1630, 1973.
53. ITU-R, Exchange of information for short-term forecasts and transmission of ionospheric disturbance warnings, ITU-R Recommendation P.313, Geneva: International Telecommunications Union, 1995.
54. ITU-R, Short-term forecasting of critical frequencies, operational maximum useable frequencies and total electron content, ITU-R P.888, Geneva: International Telecommunications Union, 1994.
55. ITU-R, Short-term prediction of solar induced variations of operational parameters for ionospheric propagation, ITU-R P. 727, Geneva: International Telecommunications Union, 1994.
56. CCIR, Real-time channel evaluation of Hf ionospheric radio circuits, Report 889-2, Geneva: International Telecommunications Union, 1990.
57. D. Lucas G. W. Haydon Predicting statistical performance indices for high frequency telecommunication systems, ITSA-1, Boulder, CO: US Dept. of Commerce, 1966.
58. M. Leftin S. M. Ostrow C. Preston Numerical maps of foEs for solar cycle minimum and maximum, ERL 73-ITS 63, Boulder CO: US Dept. of Commerce, 1968.
59. R. Hunsucker Atmospheric gravity waves and traveling ionospheric disturbances: thirty years of research, in J. M. Goodman (ed.), *Effect of the Ionosphere on Radiowave Signals and System Performance, IES90*, USGPO, Springfield, VA: NTIS, 1990.
60. I. A. Galkin B. W. Reinisch D. Kitrosser Advances in digisonde networking, in *Proc. IES99*, J. M. Goodman (ed.), Springfield, VA: NTIS, 1999.

61. J. M. Goodman J. W. Ballard Dynacast-assisted frequency management for HF communication and broadcasting systems, in J. M. Goodman (ed.), *Proc. IES99*, Springfield, VA: NTIS, 1999.
62. NRC, *Solar–Terrestrial Research for the 1980s*, Washington: National Research Council, National Academy Press, 1981.
63. A. Paul Ionospheric Variability, Technical Report 1277, San Diego, CA: Naval Ocean Systems Center, 1989.
64. F. Lied (ed.) *High Frequency Radio Communications with Emphasis on Polar Problems*, AGARDograph 104, NATO, Maidenhead, UK: Technivision, 1967.

JOHN M. GOODMAN
Radio Propagation Services, Inc.