Chapter Three

Principles Of the ionosphere at middle and low latitudes

Chapter Three
Principles of the ionosphere at middle and low latitudes

Introduction

1. The ionized part of the atmosphere, the ionosphere, contains significant numbers of free electrons and positive ions. There are also some negative ions at the lower altitudes.
2. The medium as a whole is electrically neutral, there being equal numbers of positive and negative charges within a given volume.
3. Although the charged particles may be only a minority amongst the neutral ones they exert a great influence on the medium's electrical properties, and herein lies their importance.
4. The first suggestions of electrified layers within the higher levels of the terrestrial atmosphere go back to the 19th century, but interest was regenerated with Marconi's well known experiments to transmit a radio signal from Cornwall in England to Newfoundland in Canada in 1901, and with the subsequent suggestions by Kennelly and by Heaviside (independently) that, because of the Earth's curvature, the waves must have been reflected from an ionized layer. The name *ionosphere* was coined by R. Watson-Watt in 1926, and came into common use about 1932.
5. The main regions are designated D, E, Fl and F2, with the following daytime characteristics:
    D region, 60-90 km: 10^8-10^{10} m^{-3} (10^2-10^4 cm^{-3});
    E region, 105-160 km: several 10^{11} m^{-3} (10^5 cm"3);
    Fl region, 160-180 km: several 10^{11}-10^{12} m^{-3} (10^5-10^6 cm^{-3});
    F2 region, maximum variable around 300 km: up to several 10^{12} m^{-3} (10^6 cm^{-3}).
6. The D, E and Fl regions vanish at night, and the F2 region, however, tends to persist though at reduced intensity.

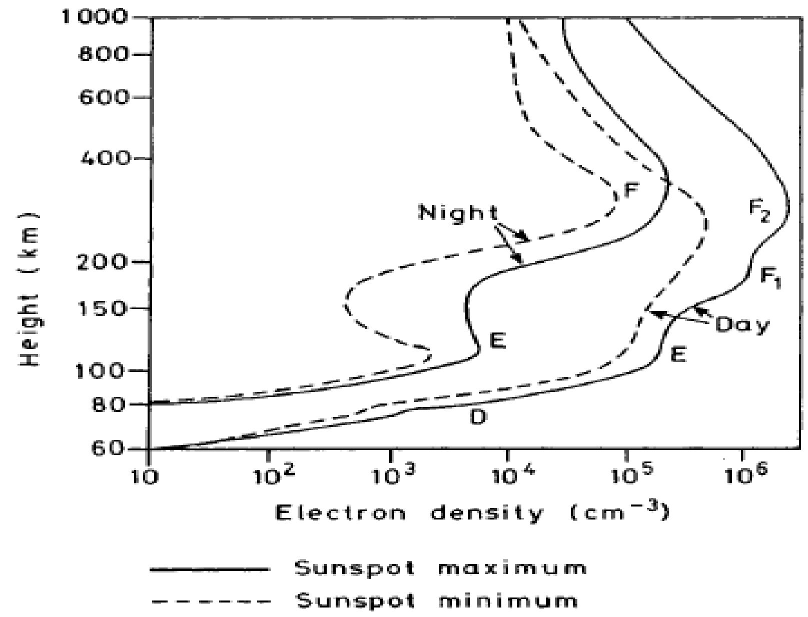


Fig. 1 Typical vertical profiles of electron density in the mid-latitude ionosphere.

Physical Aeronomy

Principles

1. The ionosphere is formed by the ionization of atmospheric gases such as N_2, O_2 and O.
2. At middle and low latitude the energy required comes from solar radiation in the extreme ultra-violet (EUV) and X-ray parts of the spectrum.
3. Once formed, the ions and electrons tend to recombine and to react with other gaseous species to produce other ions. Thus there is a dynamic equilibrium in which the net concentration of free electrons depends on the relative speed of the production and loss processes.
4. In general terms the rate of change of electron density is expressed by a continuity equation:

$$\frac{dN}{dt} = q - L - div\,(NV)$$

   Where q is the production rate, L the loss rate by recombination and attachment, and div (NV) expresses the loss of electrons by movement, V being their mean drift velocity.

5. Following the 'law of mass action', if we consider an ionization and recombination reaction, then at equilibrium

$$[X][h\nu] = const. * [X+][e]$$

   Thus, since [e] = [X+] for electrical neutrality,

$$[e]^2 = const. *[X]\,[h\nu].$$

   During the day the intensity of ionizing radiation varies with the elevation of the Sun, and the electron density responds.

6. Since the concentration of ionizable gas (X) reduces with increasing height while the intensity of ionizing radiation increases. It is reasonable to anticipate from this that the electron density will pass through a maximum at some altitude.

The Chapman production function

1. The rate of production of ion-electron pairs can be expressed as the product of four terms:

$$q = \eta \sigma n I$$

Where,

   A. I is the intensity of ionizing radiation at some level of the atmosphere (J/m^2.s)
   B. n is the concentration of atoms or molecules capable of being ionized by that radiation. For an atom or molecule to be ionized it must first absorb radiation,
   C. The amount absorbed is expressed by the absorption cross-section σ.
   D. However, not all this energy will go into the ionization process, and the ionization efficiency,η, takes that into account, being the fraction of the absorbed radiation that goes into producing ionization.

2. From this simple beginning S. Chapman, in 1931, developed a formula which predicts the form of a simple ionospheric layer and how it varies during the day. At this stage we deal only with **the rate of production of ionization (q), and the formula expressing this is the Chapman production function**. The derivation makes some assumptions:

   a. The atmosphere is composed of a single species, exponentially distributed with constant scale height.
   b. The atmosphere is plane stratified: there are no variations in the horizontal Plane.
   c. Solar radiation is absorbed in proportion to the concentration of gas particles.
   d. The absorption coefficient is constant: this is equivalent to assuming monochromatic radiation.

3. The Chapman production function is usually written in a normalized form as:

$$q = q_{mo}\, exp(1 - z - \sec(\chi)\ \exp(-z))$$

Where,

z is the reduced height for the neutral gas, $z = (h — h_{mo})/H$,
H is the scale height.
χ is the solar zenith angle,
h_{mo} is the height of maximum production rate when the Sun is overhead ($\chi = 0$),
q_{mo} is the production rate at h_{mo}, also when the Sun is overhead.

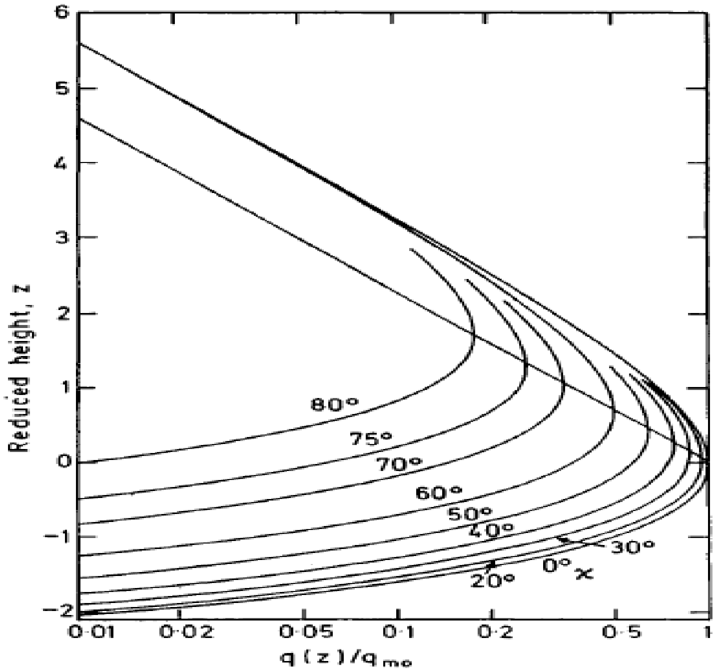


Figure (2) Chapman production function

4. The latest equation can be written in the following form

$$q_m = q_{mo}\cos(\chi)$$

Where q_m if the maximum production, q_{mo} is the maximum production at solar zenith angle equals zero.

Ionization by energetic particles

1. The other source of ionization in the terrestrial ionosphere is energetic particles, which are not entirely absent at middle latitudes but are much more important at high latitudes where they frequently provide the dominant source of ionization. Electrons, Bremsstrahlung X-rays, Protons.

Principles of chemical recombination

1. Working out the rate of electron production is just the first step in calculating the electron density in an ionized layer, and the next step is to reckon the rates at which electrons are removed from the volume under consideration. These are represented in the continuity equation by two further terms,
    a. <u>one for the recombination between ions and electrons to reform neutral particles, and</u>
    b. <u>The other to account for movement of plasma into or out of the volume.</u>
2. We deal first with the principles of chemical recombination. The question of which individual reactions are most important in different parts of the ionosphere will be addressed.
3. First we assume that the electrons recombine directly with positive ions and that no negative ions are present:

$$X^+ + e \rightarrow X.$$

Then the rate of electron loss is

$$L = \alpha[X^+]\, N = \alpha N^2$$

Where N is the electron density (equal to the ion density $[X^+]$) and α is the recombination coefficient. At equilibrium, therefore,

$$q = \alpha N^2$$

Taking the production rate q from the Chapman production function 3, we obtain

$$N = N_{mo}\, \exp\frac{1}{2}(\,1 - z - \sec(\chi)\exp(-z))$$

in which $z = (h - h_o)/H$. And from Equation 4 it is seen that the electron density at the peak of the layer varies as $cos^{1/2}\ (\chi)$

$$N_m = N_{mo}\, cos^{1/2}(\chi)$$

a layer with these properties is called an α -Chapman layer.

4. If one is concerned particularly with electron loss, **then attachment to neutral particles to form negative ions can itself be regarded as another type of electron loss process.** We can see that the attachment type of reaction can be written

$$M + e \rightarrow M^-$$

And the rate of electron loss (due to attachment) is

$$L = \beta N$$

Where β is the attachment coefficient. At equilibrium,

$$q = \beta N$$

Taking q from the Chapman production function, as before, gives

$$N = N_{mo}\, \exp(1 - z - \sec(\chi)\exp(-z)).$$

The peak electron density now varies as:

$$N_m = N_{mo}\, cos\ \ (\chi)$$

Such a layer is a β-Chapman layer.

In fact β is expected to vary with height because it depends on the concentration of the neutral molecules (M).

5. It is known that electron loss in the F region occurs in a two stage process:

$$X^+ + A_2 \rightarrow AX^+ + A \qquad \beta[X^+] \qquad (1)$$

$$AX^+ + e \rightarrow A + X \qquad \alpha[AX^+]N \qquad (2)$$

In which A_2 is one of the common molecular species like O_2 and N_2.

a. The first step moves the positive charge from X to AX (<u>charge exchange</u>),

b. The second one dissociates the charged molecule through recombination with an electron, <u>a dissociative recombination reaction.</u>

c. At low altitude β is large, equation (1) goes quickly and all X^+ is rapidly converted to AX^+; the overall rate is then governed by the rate of equation (2), giving an α -type process because $[AX^+] = N$ for neutrality.

d. At a high altitude β is small and equation (1) is slow and controls the overall rate. Then $[X^+] = N$ and the overall process appears to be β type.

e. As height increases, the reaction type therefore alters from α -type to β-type. Figurer (3).

f. The change from α to β -behavior occurs at height h_t where

$$\beta(h_t) = \alpha N$$

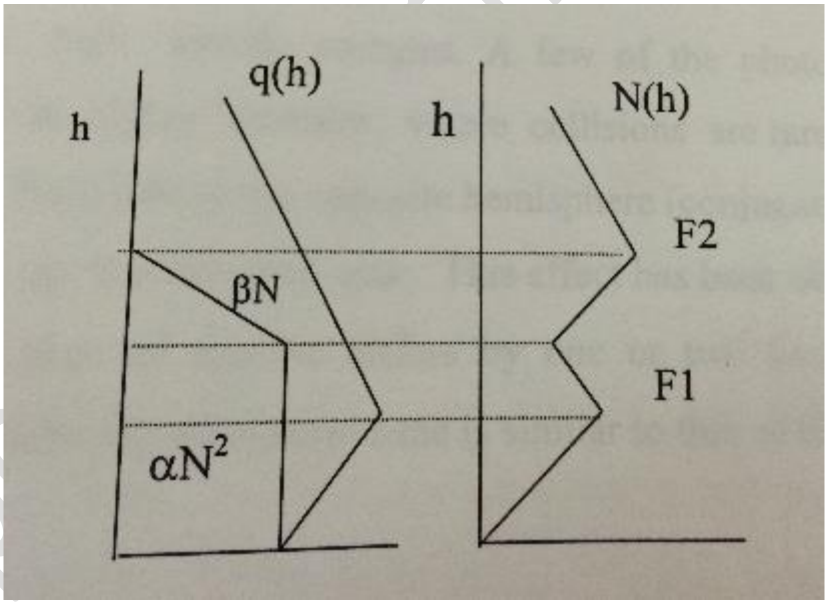


Figure (3) The formation of the maximum of the ionosphere at the height of the most favorable balance between the ionizing radiation I and the particle density of the atmosphere N_n.

6. *The relaxation time t_r* is the time in which the electron density world reduce to half if there was no more production of electrons. Setting q=0 in the continuity equation and neglecting attachment and diffusion we have

$$\frac{dN}{dt} = \alpha N^2$$

Solving this simple differential equation, we find that the time t_r in which N (t_r) becomes $N_o/2$

$$t_r = \frac{1}{\alpha N}$$

Vertical transport

1. The final term of the continuity equation is to account for changes of electron and ion density at a given location due to bulk movement of the plasma.
2. Such movements can have various causes and can occur in the horizontal and the vertical planes in general,
3. We shall concentrate here on the vertical component of movements in the F region.
4. Assuming that photochemical production and loss are negligible by comparison with the effect of movements, the continuity equation gives

$$\frac{\partial N}{\partial t} = -\nabla.(Nv)$$

Where v is the drift velocity. Taking vertical movement only,

$$\frac{dN}{dt} = -\frac{\partial}{\partial h}(Nw)$$

Where w is the vertical drift speed and h is the height. Supposing that this drift is due to diffusion we can

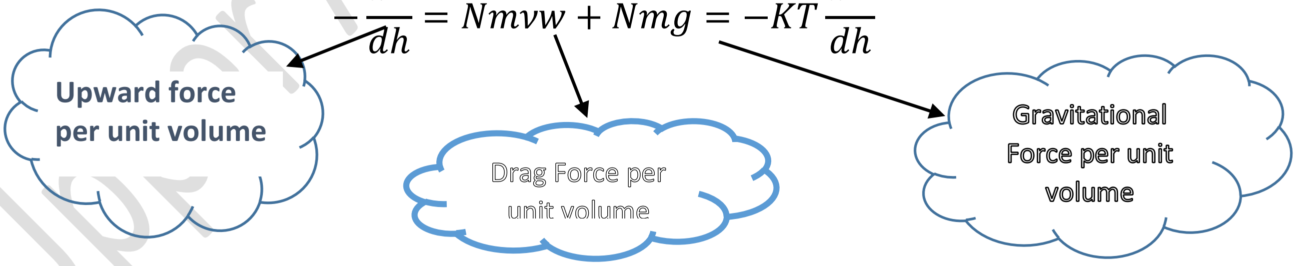
$$w = -\frac{D}{N}\frac{\partial N}{\partial h}$$

D being the diffusion coefficient, given in its simplest form by

$$D = \frac{KT}{mv}$$

5. In the present case the **minority gas** is the plasma composed of ions and electrons, and the **majority gas** is the neutral air. Moreover, for drift in the vertical direction the force of gravity also acts on each particle, giving

$$-\frac{dP}{dh} = Nmvw + Nmg = -KT\frac{dN}{dh}$$



Where we have used P = NkT. Since $D = \frac{KT}{mv}$ and $H_N = \frac{KT}{mg}$ (H_N being the scale height of the minority gas), this may be rearranged to give

$$\frac{dN}{dt} = \frac{\partial}{\partial h}\left[D\left(\frac{\partial N}{\partial h} + \frac{N}{H_N}\right)\right]$$

This is the equation that has to be satisfied by the time and height variations of the upper F region and the protonosphere where **ion production and recombination are both sufficiently small.**

6. A plasma is composed of two minority species, ions and electrons. Initially the ions, being heavier, tend to settle away from the electrons, but the resulting separation of opposite charges produces an electric field, E, and a restoring force eE on each charged particle.
7. To find how this affects the drift of the plasma, we write separate equations for each species and include the electrostatic force:

Electrons

$$-\frac{dP_e}{dh} = Nm_e g + EeN + Nm_e \nu_e w = -KT_e \frac{dN}{dh}$$

Ions

$$-\frac{dP_i}{dh} = Nm_i g - EeN + Nm_i \nu_i w = -KT_i \frac{dN}{dh}$$

Summing

$$Nm_i g + N\nu_i m_i \omega = -K(T_i + T_e)\frac{dN}{dh}$$

it being assumed that $m_e << m_i$, $m_e\nu_e << m_i\nu_i$, $N_e = N_i$ (= N), and $w_e = w_i$ (= w). Hence, for a plasma,

$$Nw = -D_p\left(\frac{dN}{dh} + \frac{N}{H_p}\right)$$

Where

$$D_p = K\ \frac{(T_e + T_i)}{m_i\,\nu_i}$$

And

$$H_P = \frac{K\ (T_i + T_e)}{g\ m_i}$$

Respectively known as the ambipolar or plasma diffusion coefficient and the plasma scale height.

The continuity equation for a plasma is:

$$\frac{dN}{dt} = \frac{\partial}{\partial h}\left[D_p\left(\frac{\partial N}{\partial h} + \frac{N}{H_p}\right)\right]$$

Chemical aeronomy

Introduction

1. To work out what the actual ionosphere should be like on Earth or any other planet requires (يتطلب) a detailed consideration (تعابير مفصله) of many factors.

A. We need to know the composition of the neutral atmosphere and its physical conditions such as density and temperature.

B. We also need full information on the solar spectrum and any energetic particle fluxes able to ionize the constituents (مركبات) of the atmosphere.

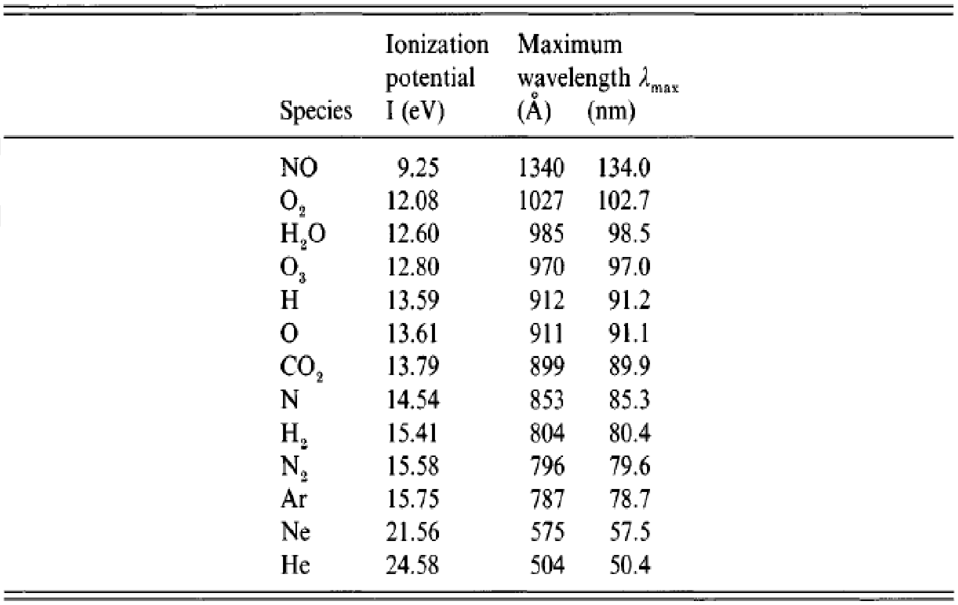
C. We would have to determine what gases could be ionized by the radiation incident, and then determine the ionization rate of each species, summing over all wavelengths and all gases to get the total production rate in a given volume.

D. Then the loss processes, both Chemical and by transport.

2. The composition of the neutral atmosphere is composed mainly of the nitrogen/oxygen mixture generally known as 'air' up to about 100 km, and above that the atomic species O, He, and H progressively enter the scene due to diffusive separation.
3. To be ionized, a species must absorb a quantum of radiation whose energy exceeds the ionization potential. Table 1 lists the ionization potential and the corresponding maximum wavelength λ_{max} of radiation able to effect an ionization for various atmospheric gases.
4. Only radiation of wavelength less than λ_{max} can produce ionization.
5. These values of λ_{max} immediately identify the relevant parts of the solar spectrum as the:

X-ray (0.1-17 nm, 1-170 A) and Extreme ultra-violet, EUV, (17-175 nm, 170-1750 A), emissions which come from the solar chromosphere and corona. Some are enhance during flares.

Table (1) Ionization potentials



Species	Ionization potential I (eV)	Maximum wavelength λ_{max} (Å)	(nm)
NO	9.25	1340	134.0
O_2	12.08	1027	102.7
H_2O	12.60	985	98.5
O_3	12.80	970	97.0
H	13.59	912	91.2
O	13.61	911	91.1
CO_2	13.79	899	89.9
N	14.54	853	85.3
H_2	15.41	804	80.4
N_2	15.58	796	79.6
Ar	15.75	787	78.7
Ne	21.56	575	57.5
He	24.58	504	50.4

6. However, not every photon with sufficient energy will actually produce an ionization. The chance that the photon will be absorbed is given by the absorption cross section (σ sigma) and the chance that absorbed photon produces an ionization is given by the ionization efficiency (η Eta). The product ($\sigma \eta$) is called the ionization cross section (σ_i).
7. In the Chapman formulation the value of the absorption cross section, σ, generally increases with increasing wavelength up to λ_{max} and then falls rapidly to zero. Figure 4 shows how (σ) depends on wavelength for some common gases.

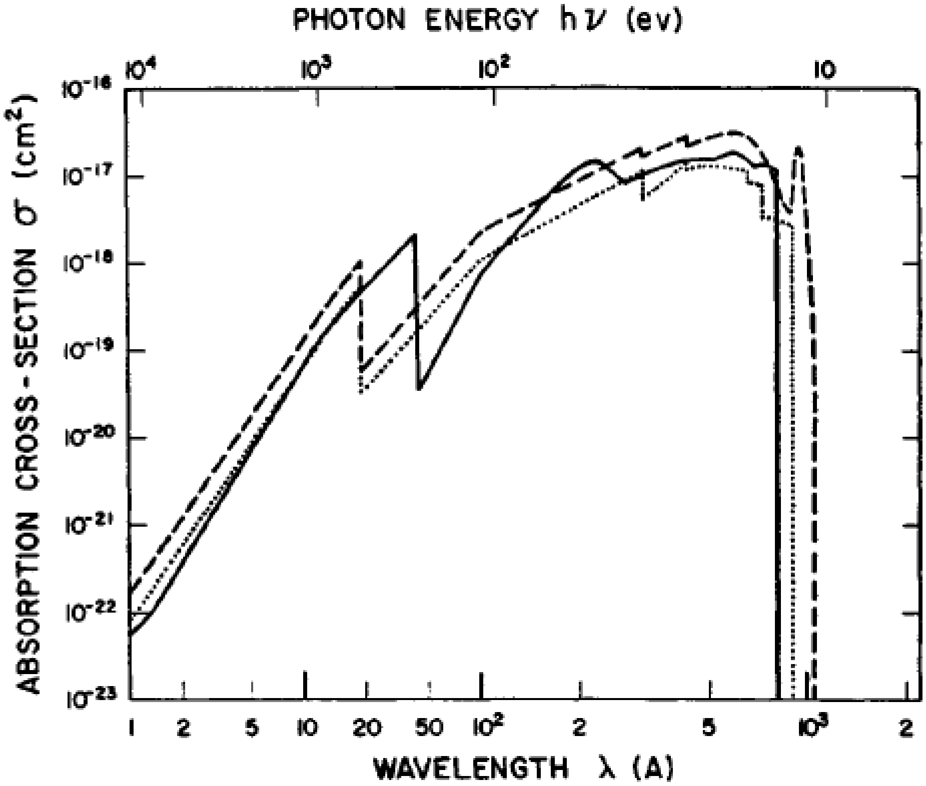


Fig. 4 Absorption cross-sections for O_2, O and N_2

8. The efficiency with which the absorbed radiation leads to ionization is expressed by the ionization efficiency, η. With atomic species, all the absorbed energy goes into ion-electron production, $\eta=1$, at the rate of one ion-electron pair for every 34 eV of energy.
9. For molecules $\eta<1$,

$$\eta = \frac{360}{\lambda(A^o)}$$

10. The intensity of solar radiation decreases due to absorption as it passes through the atmosphere. The peak of ionization production occurs at that altitude where the intensity has been reduced to 1/e times the intensity at the top of the atmosphere ($I_m = I/e$). The intensity is then said to reach one optical depth at the peak of production.

11. The Chapman theory shows that the production rate is a maximum at the level where the optical depth ($\sigma nH \sec \chi$), is unity. If the absorption at given wavelength is due to several species, then the condition for maximum production is

$$\sum \sigma_i n_i H_i \sec \chi = 1$$

At that height the rate is given by,

$$q_m = \eta I_\infty \cos \chi / eH$$

Summed over several species if necessary.

The height of unit optical depth in a model terrestrial atmosphere is given as a function of wavelength in Figure 5.

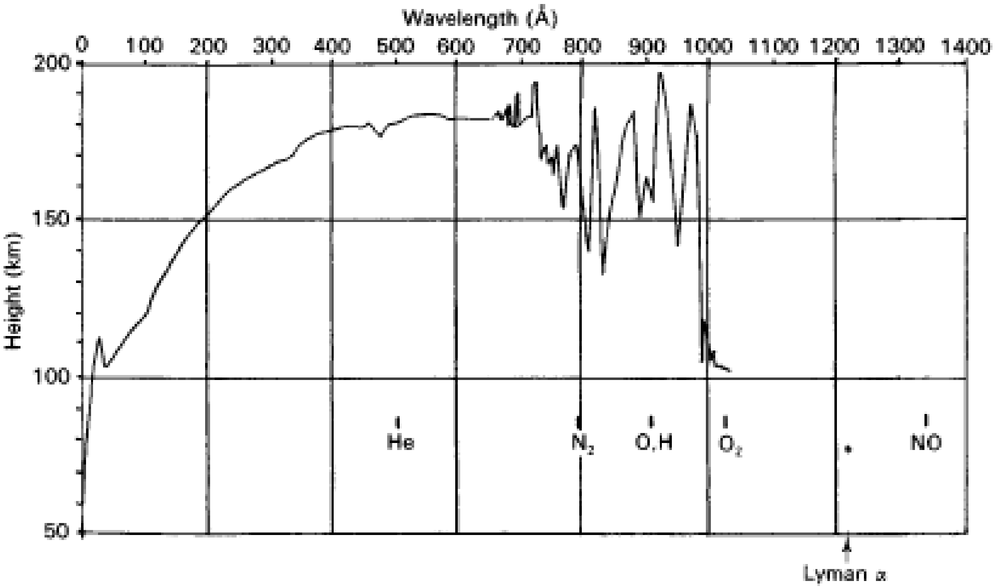


Figure (5) Height where the optical depth reaches unity for radiation vertically incident on a model atmosphere. Ionization limits for common gases are marked.

The transfer of solar radiation in the atmosphere does not involve emission sources or scattering into the original beam, and the process is well described by the Lambert-Beer exponential absorption law,

$$I(\lambda) = I_\infty(\lambda) \exp[-\tau(\lambda)]$$

At any point in the atmosphere the intensity at wavelength λ is $I(\lambda)$ and $\tau(\lambda)$ is the optical depth.

The optical depth, $\tau(\lambda)$, a parameter that appears in the Lambert-Beer law specifies the attenuation of the solar irradiance by the atmosphere.

For a vertical sun, and at altitude z_0 it is, simply,

$$\tau^{v}_{z_0}(\lambda) = \sum_j \sigma^{a}_j(\lambda) \int_{z_0}^{\infty} n_j(z)\, dz$$

Where $\sigma_j^a(\lambda)s$ are wavelength dependent absorption cross section of species j and the $n_j(z)$s are the height profiles of their concentrations.

For slant path

$$\tau(\lambda, \chi_0, z_0) = \sum_j \sigma_j^a(\lambda) \int_{z_0}^{\infty} n_j(z) \left[1 - \left(\frac{R + z_0}{R + z} \right)^2 \sin^2 \chi_0 \right]^{-1/2} dz$$

E and Fl regions

E-region

1. The E region, peaks at 90-110 km.

2. EUV radiation between 800 and 1027 A° (the ionization limit of O2) is absorbed by molecular oxygen to form O_2^+.

3. At the short wavelength end X-rays of 10-100 A° (1-10 nm) ionize all the atmospheric constituents.

4. The main primary ions are $N_2^+, O_2^+ \ and \ O^+$, but the most numerous are again observed to be $NO^+ and \ O_2^+$.

5. The intensity of solar X-rays varies over the solar cycle and they probably make little contribution to the E region at solar minimum.

6. Direct radiative recombination of the type

$$e + X^+ \rightarrow X + h\nu$$

Is slow relative to other reactions and is not significant in the E and F regions. Dissociative recombination, as

$$e + XY^+ \rightarrow X + Y,$$

Is 10^5 times faster with a reaction coefficient of (10^{-13} m^3/s) and in both E and F regions <u>the electron and ion loss proceeds via molecular ions.</u>

7. The main recombination reactions of the E region are therefore

$$e + O_2^+ \rightarrow O + O$$
$$e + N_2^+ \rightarrow N + N$$
$$e + NO^+ \rightarrow N + O.$$

8. Metallic ions of meteoric origin, such as Fe^+, Mg^+, Ca^+ and Si^+, are also present in the E region. They cannot recombine dissociatively and therefore have recombination coefficients typical of the radiative process (10^{-18} m^3/s), which gives them relatively long lifetimes.

F1 region

1. The Fl region at 160-180 km are fairly well understood. The Fl region is attributed to the most heavily absorbed part of the solar spectrum, between about 200 and 900 A°.
2. The ionization limit of atomic oxygen is at 911 A° for which the optical depth reaches unity from about 140 to 170 km.
3. The primary reaction products are O_2^+, N_2^+, O^+, He^+, and N^+, but subsequent reactions leave NO^+ and O^+ as the most abundant.
4. In the F region the principal primary ion is O^+, which is first converted to a molecular ion by a charge exchange reaction:

$$\left.\begin{aligned} O^+ + O_2 &\rightarrow O_2^+ + O \\ O^+ + N_2 &\rightarrow NO^+ + N. \end{aligned}\right\}$$

The molecular ion then reacts with an electron as above, to give as the net result. Or, the atomic oxygen ions recombine with electrons with a third particle

$$\left.\begin{array}{l} e+O^{+}+O_2 \rightarrow O+O+O \\ e+O^{+}+N_2 \rightarrow O+N+N. \end{array}\right\}$$

5. In the Fl region the overall reaction is controlled by the rate of the dissociative recombination.

F2-region

1. It extends from 200 to roughly 1000 Km and has a daytime maximum near 250 Km of about $5*10^5$ el/cm^3.
2. As we move to higher altitudes the rate at which electrons and ions recombines decreases rapidly with height and the relaxation time at higher altitudes is much longer.
3. The F2 peak is formed because the loss rate decreases rapidly as the dominant term in L changes from αN^2 to βN in the F2 region.
4. The most ionized gas presents in these heights F2-region is O^+ which is lost in the following two stages reactions as shown earlier

$O^+ + O_2 = O + O_2^+$ (Charge transfer)
(1)
Or
$O^+ + N_2 = N + NO^+$ (ion- atom interchange)
(2)

And then the new ions are dissociated in neutralizing collisions with the free electrons,

$e^- + O_2^+ = O + O$
$e^- + NO^+ = N + O$ (Dissociative Recombination)

5. In the F2 region, where the number density of neutral molecules is considerable smaller, equations (1 & 2) becomes bottleneck reactions.

The Upper Ionosphere

1. At altitudes above the F2 peak both the production and the loss of electrons tend to zero, and the only final term of continuity equation $div\ (NV)$.
2. Electrons and ions tend to diffuse as a body because their separation builds up an electric field which brings them back together. This phenomena called ambipolar diffusion.
3. Around 1000 Km O^+ is replaced by He^+ as the predominant ion, and at higher altitudes (≈ 2500 Km) He^+ is replaced by H^+, i.e., by free protons.
4. The layer where helium ions dominate is often called heliosphere and the region above it is called the protonosphere.

D region

1. The D region of the ionosphere does not include a maximum.

2. The D region is also the most complex part of the ionosphere from the chemical point of view. This is due, first, to the relatively high pressure, which causes minor as well as major species to be important in the photochemical reactions, and, second, because several different sources of ionization contribute significantly to ion production.

Sources

To treat the whole of the D region from 60 to 95 km over all latitudes one has to reckon with 6 sources of ion production:

a. The Lyman-α line of the solar spectrum at 1215 A° penetrates below 95 km and ionizes the minor species nitric oxide (NO) whose ionization limit is at 1340 A.

b. The EUV spectrum between 1027 and 1118 A° ionizes another minor constituent, excited oxygen in the state $O_2\,(^1\Delta_g)$.

c. EUV radiation also ionizes O_2 and N_2, as in the E region.

d. Hard X-rays of 2-8 A° ionize all constituents, most effect being therefore from the major species O_2 and N_2.

e. Galactic cosmic rays, which affect the whole atmosphere down to the ground, become a major ionization source in the lower D region. At this level the production rate increases downward in proportion to the total air density.

f. Energetic particles from the Sun or of auroral origin ionize the D region at high latitudes, where at times they form the main source.

3. The relative contributions of these different sources vary with **latitude, time of day, and level of solar activity**.
4. At greater solar zenith angles the contributions from Lyman-α and X-rays are reduced, cosmic rays becoming relatively more important below about 70 km.
5. The X-ray flux varies strongly with solar activity (by a factor of a hundred to a thousand) and is probably not significant in the D region at sunspot minimum.

Principles of airglow

1. Many of the photochemical reactions within the upper atmosphere are accompanied by the emission of radiation in the visible, infra-red or ultra-violet parts of the electromagnetic spectrum. These emissions comprise the airglow.
2. Aurora and airglow are often treated together, but there are some important differences. They are, for instance, energized from different sources.
3. The aurora is a high-latitude phenomenon, occurring most intensely during and after solar disturbances, which is energized by charged particles entering the atmosphere from the magnetosphere. It shows much structure in both space and time.
4. Airglow, is always present, covers all latitudes, and is virtually unstructured. On a moonless night (محاق) the airglow contributes the major part of the light arriving from the sky, exceeding starlight in total intensity

5. R. J. Strutt, fourth Baron Rayleigh and son of the great Lord Rayleigh (the third Baron), pioneered airglow studies in the 1920s, and in his honour the unit of airglow emission rate is called the Rayleigh:

$$1R = 10^6 \text{ photons/cm}^2\,\text{s}$$

   This measures the height-integrated emission rate, such as would be observed by a photometer situated at the ground.
6. The name 'airglow' actually dates from 1950. Airglow (and auroral) emissions are generated when an excited species returns to the ground state. There are various immediate causes, including :
   (a) Radiative recombination reactions which emit a photon.
   (b) Reaction products in an excited state, which then radiate the excess energy in returning to the ground state
   (c) Excitation by hot electrons (following ionization) and by electric fields
   (d) Excitation by solar radiation, giving resonance emissions at the same wavelength.
7. Airglow measurements have several applications to upper-atmosphere science.
   1. Neutral-air winds can be determined from the Doppler shift of an airglow line,
   2. Derive the height profile of atomic oxygen.

Figure (6) Images of Airglow

Figure (7) Images of Aurora

Ionospheric phenomena at middle and low latitudes

Introduction

The geospace environment is complex and subtle. To a large extent the ionosphere varies in a regular and predictable manner, but these regularities may not always accord with simple theory. In addition, major perturbations called storms occur from time to time, and the spatial structure includes irregularities of various sizes. Indeed, it appears that the structure of the ionosphere includes all the scales of space and time that are accessible to observation.

1. Observed behavior of the mid-latitude ionosphere

1. We will begin with the regularities of the ionosphere - those variations with altitude, time of day, latitude, season, and solar cycle which are repeatable and therefore predictable, and also with the ionosphere's response to solar flares and eclipses.
2. All of these relate to the larger scales of variation: vertical distances measured in tens and hundreds of kilometers, horizontal distances in hundreds and thousands of kilometers, times in hours to years.

3. Historically, ionospheric observations were compared with the Chapman theory of the production of an ionospheric layer, and major departures from the theory were christened' anomalies'.

1.1 E -region and sporadic-E

1. The behavior of the E region is close to that of an α-Chapman layer.
2. On average its critical frequency, foE, varies with the solar zenith angle as $cos^{1/4}(\chi)$, which means that the peak electron density, N_m(E), varies as $cos^{1/2}(\chi)$.
3. The E layer does not quite vanish at night, but a weakly ionized layer remains with electron density about 5 x 10^9 m^{-3} (against 10^{11} m^{-3} by day). One possible cause is meteoric ionization.
4. The most remarkable anomaly of the E region is *sporadic-E*. On ionograms sporadic-E is seen as an echo at constant height which extends to a higher frequency than is usual for the E layer; for example to above 5 MHz less than a kilometer across.
5. The principal cause of sporadic-E at middle latitude is a variation of wind speed with height, a *wind shear*, which, in the presence of the geomagnetic field, acts to compress the ionization.
6. The time scale of the process needs ions of relatively long life, and it is thought that these are metallic ions of meteoric origin.
7. At high latitude, sporadic-E is attributed to ionization by incoming energetic particles. Sporadic-E is significant in radio propagation since it may reflect signals that would otherwise penetrate to the F region.

1.2 Fl region

1. The Fl region is another well behaved layer as viewed from the Chapman theory, though it seldom exists as a distinct peak and is more correctly called the *F1 ledge*.
2. Its critical frequency, foFl, varies as $cos^{1/4}(\chi)$ indicating α-Chapman behavior.
3. However the ledge is not always present, being more pronounced in summer and at sunspot minimum. (Fl is never seen in winter at sunspot maximum.)
4. The explanation is to be found by comparing h_t, the transition height between α -type and β-type recombination and h_m, the height of maximum electron production rate. The Fl ledge appears only if $h_t > h_m$, and since h_t depends on the electron density, the ledge vanishes when the electron density is greatest.

1.3 F2 region and its anomalies

1. Since the F2 region has the greatest concentration of electrons, it is also the region of greatest interest in radio propagation, it is also the most variable, the most anomalous, and the most difficult to predict. The anomalies may be summarized as follows:

A. The diurnal variation may be asymmetrical about noon. There can be rapid changes at sunrise but perhaps little or no change at sunset. The daily peak may occur either before or after local noon in the summer, though it is likely to be near noon in the winter. On some days a secondary minimum appears near noon between the morning and evening maxima.

B. The daily pattern of variation may not repeat from day to day. (If it did, the next day could at least be predicted from the previous one.) Figure 8 illustrates points (A) and (B).

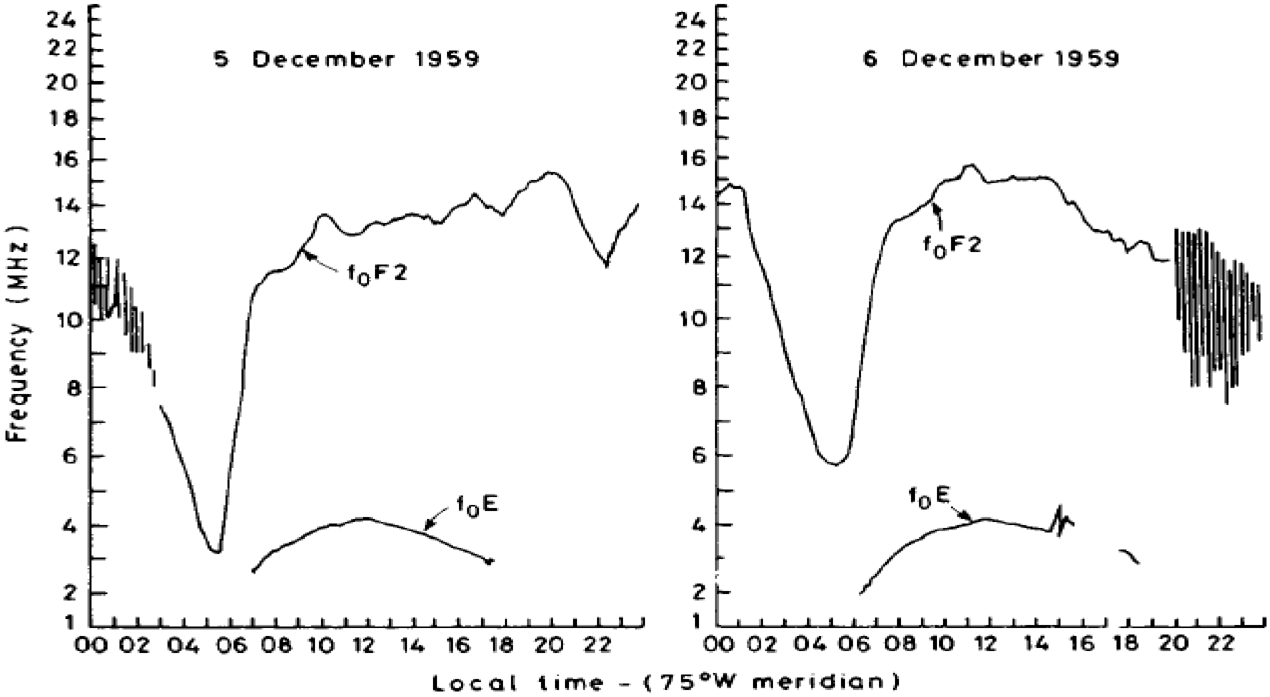


Fig. 8 Diurnal change of foF2 on successive days in December 1959 at a low latitude station, Talara, Peru. Note, by contrast, the regularity of the E layer.

C. There are several anomalous features in the seasonal variation. Noon values of foF2 are usually greater in winter than in summer, whereas the Chapman theory leads us to expect the opposite. This is the *seasonal anomaly,* which is shown in Figure 9.

D. There is also an annual anomaly, such that, if averaged between hemispheres, the electron densities are 20 % greater in December than in June. The changing distance between Sun and Earth would account for only a 6% difference. Also, the electron content is abnormally large at the equinoxes, giving the semi-annual anomaly

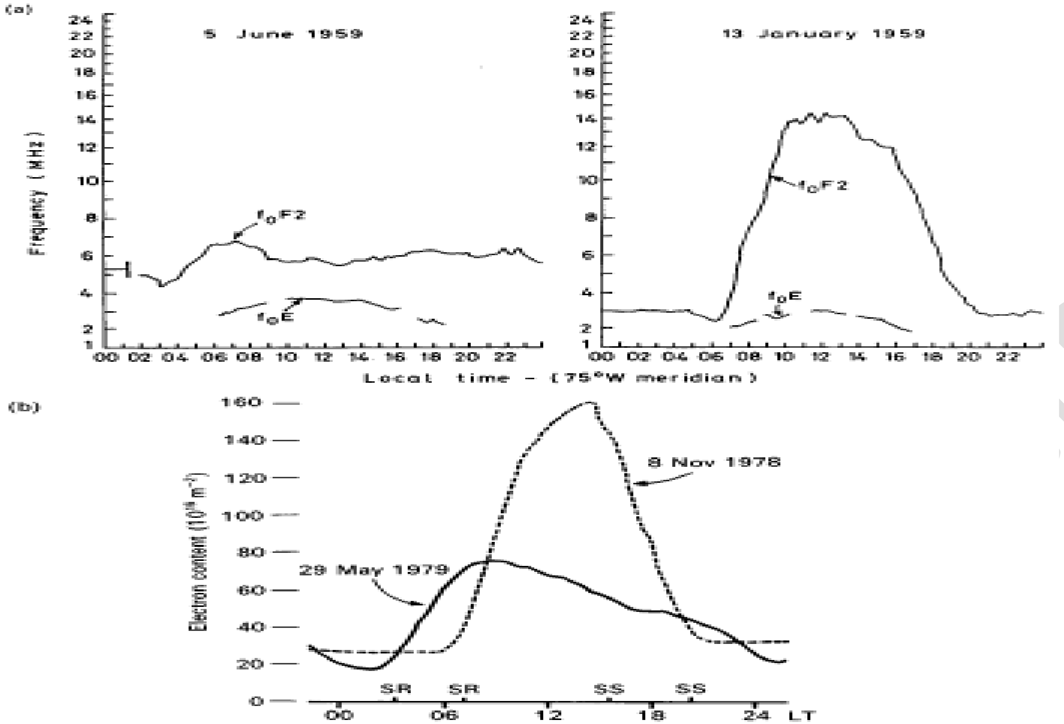


Fig. 9 (a) Diurnal change of foF2 in summer and winter at a high-latitude station in the northern hemisphere, Alaska. The F region is anomalous while the E layer behaves as expected. (b) Summer and winter electron content measured at Fairbanks, Alaska.

There are four main causes of seemingly anomalous behavior:

a) Reaction rates are sensitive to temperature,
b) The chemical composition varies.
c) There are winds in the neutral air which lift or depress the layer by a complicated mechanism.
d) The ionosphere is influenced by the protonosphere and the conjugate hemisphere.

Reaction rates are generally temperature sensitive. The rate for the reaction

$$O^+ + N_2 \longrightarrow N + NO^+$$

The reaction rate constant k is given by

$$k = 4.5 \times 10^{-20} (T/300)^2 \quad m^3/s \quad \text{if} \quad T > 1000\ K,$$

And

$$k = 5 \times 10^{-19} \quad m^3/s \quad \text{if} \quad T < 1000\ K.$$

This property obviously contributes to both the persistence (بقاء) of the night F region and to the seasonal anomaly.

2. The electron production rate depends on the concentration of atomic oxygen, O, whereas the loss rate is controlled by the molecular species N_2 and O_2. Thus, increases in the ratios $[O]/[O_2]$ and $[O]/[N_2]$ will increase the equilibrium electron density. (Fig.10).

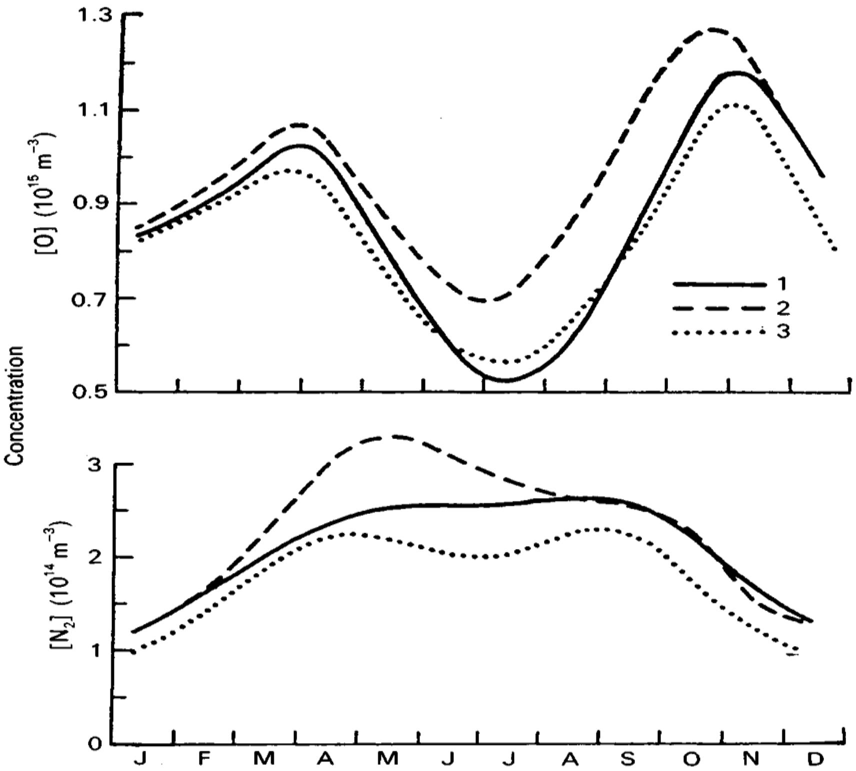


Fig. 10 Variations of O and N2 concentration at 300 km at latitude 45° and 1400 LT, according to three models, of which model 2 is MSIS (see Section 4.4).

3. The ratio [O]/[N2] at 250-300 km is measured as about 6 in winter and about 2 in summer, a seasonal change amounting to a factor of three.
4. The change of composition is attributed to the pattern of global circulation in the thermosphere.
5. In winter months, representative quantities such as N_m (the peak electron density in the layer) and N_T (the electron content) show a single daily maximum at or near to local noon.
6. In summer such quantities vary more gradually during the middle of the day when there might even be a local minimum or 'biteout'; peaks are more likely to appear in the morning or in the evening, or sometimes in both.
7. Meridional component of the thermospheric neutral wind, which acts to depress the ionosphere when flowing poleward and elevates it when flowing equatorward, can exert a major influence both on electron densities and on electron content.
8. At 300 km the neutral wind flows poleward by day and equatorward by night at speeds ranging between tens and hundreds of meters per second. Thus its effect is usually to depress the ionosphere and thereby increase the rate of loss by day, but to lift the region and reduce its rate of decay at night. These effects can be modelled mathematically (Figure 11) and they are found to be major.

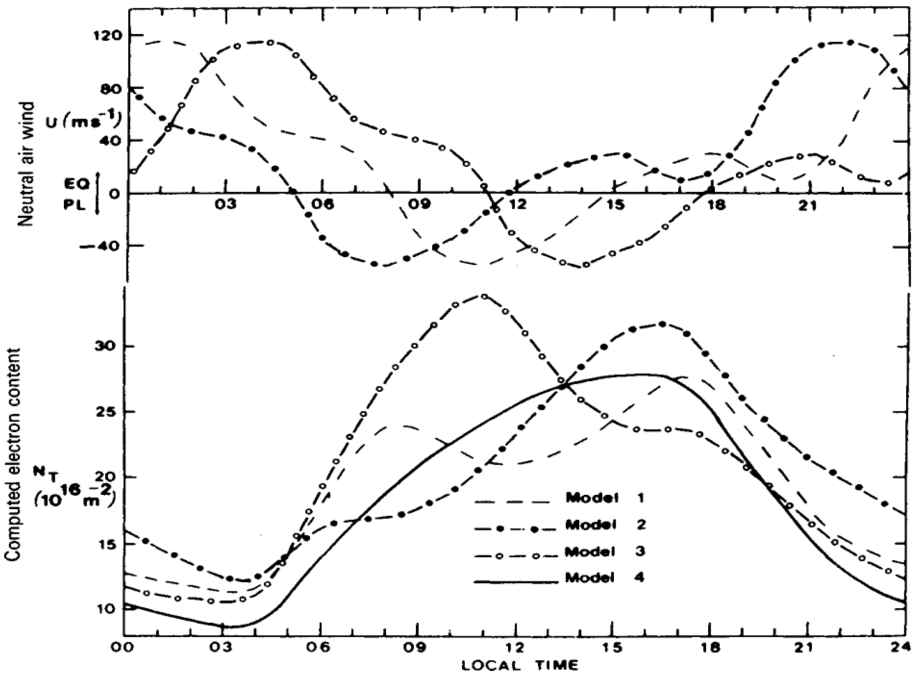


Fig. 11 Daily patterns of the neutral-air wind for four models, and the consequential daily variations of electron content. In models 2 and 3 the wind pattern is shifted by 3 hours with respect to model 1. The wind is zero in model 4.

9. The amount by which the layer is raised or lowered may be estimated by applying Equation

$$\Delta h_m \sim \mathbf{WH^2/D \sim W/\beta}$$

Where **D** is the diffusion coefficient and **H** the scale height, β is the recombination coefficient and vertical drift, **W**, due to the wind.

Taking H = 60 km for the neutral scale height, D = 2 x 10^6 m^2/s for the diffusion coefficient, and W = 30 m/s as a typical vertical drift due to the daytime poleward wind, the layer peak is lowered by about 50 km.

---

11. It is not possible to understand the F2 region without taking account of its temperature variations. The photoelectrons produced by the ionization processes are hotter than the neutral atoms from which they were formed. This excess energy is gradually shared with the positive ions, though transfer to the neutrals is less efficient.
12. Consequently the plasma is hotter than the neutral air, and within the plasma the electrons are hotter than the ions ($T_e > T_i$). The electron temperature can be two or three times the ion temperature by day, though by night their temperatures are more nearly equal.
13. Diurnal temperature variations, measured by incoherent scatter radar, are illustrated in Figure 12. These temperature changes strongly affect the

distribution of F2-region plasma. When hotter, the plasma has a greater scale height

$$H_p = k(T_e + T_i)/m_i g,$$

And so spreads to greater altitudes, where it tends to persist for longer because the loss rate is smaller.

14. At the greater altitudes the positive ions are protons, and the ionosphere and the protonosphere are strongly coupled through the charge exchange reaction between protons and atomic oxygen ions

$$H + O^+ \rightleftharpoons H^+ + O$$

As the F region builds up and is also heated during the hours after sunrise, plasma moves to higher altitudes where protons are created, and these then flow up along the field lines to populate the protonosphere. In the evening the proton population flows back to lower levels, where it charge exchanges back to oxygen ions and so helps to maintain the F region at night.

15. Via the protonosphere the magnetically conjugate ionosphere may also have an effect, since protonospheric plasma, coming mainly from the summer ionosphere, is equally available to replenish (سد النقص) the winter ionosphere.
16. Overall, the winter ionosphere benefits from the conjugate region in the summer hemisphere. The inter-hemispheric coupling provided by the geomagnetic field has another consequence since photoelectrons arriving from the conjugate region may help to heat the local ionosphere. This is most noticeable at sunrise when electron densities are low and thus a given amount of heat causes the greatest temperature rise.
17. A marked increase in the slab thickness is given by:

$$r = I/N_m$$

(This is the thickness of a hypothetical ionosphere with uniform electron density N_m and content I. The slab thickness is a useful parameter of the ionosphere and it depends physically on the temperature and the ion composition) is observed at sunrise, which is plainly due to the arrival of photoelectrons.

18. One of the most remarkable things about the F region is its variability from one day to the next (Figure 12).
19. In the Polar Regions this might not be surprising because of the sporadic nature of solar and auroral activity. But these are not dominant influences inside the plasmasphere, and the magnitude of the day-to-day changes poses a riddle whose solution so far lacks even good experimental clues. Presumably the origin must be a source in the terrestrial atmosphere or in the solar wind.

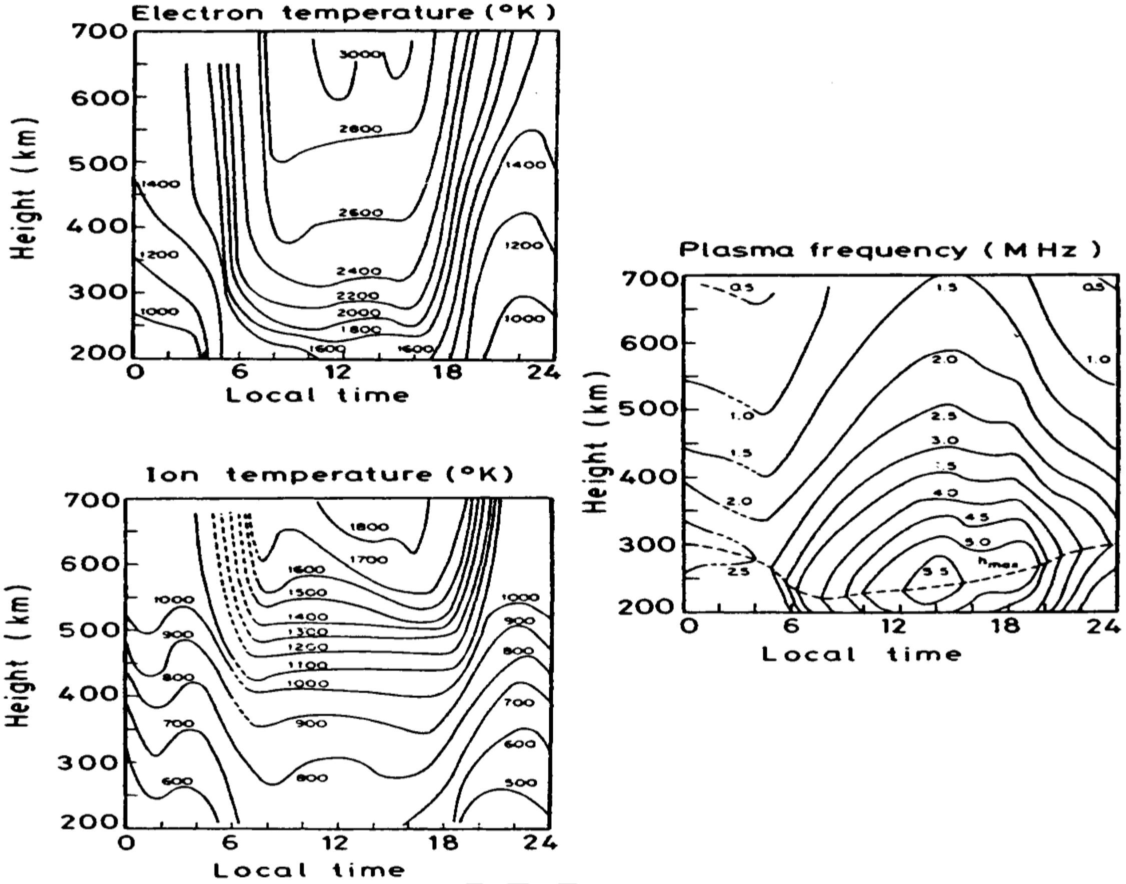


Fig. 12 Diurnal variation of electron temperature (Te), ion temperature (T_i), and plasma frequency (f_N) between 200 and 700 km, measured by incoherent scatter at a mid-latitude station, April 1964.

20. Although correlations between properties of the F region and sunspot numbers are significant over the long term, from day to day the correlation is generally poor and, by elimination, it seems most likely that the origin must lie in the neutral-air wind in the thermosphere. Even if thermospheric wind variations are accepted as a plausible hypothesis, another lurking problem is that the source of that variation is itself not known.
21. It appears likely that the various classical anomalies of the F2 region arise from combinations of the factors outlined above, though the details may not be clear in any particular case. With the mechanisms established in essence, further progress depends on acquiring better data on the thermosphere and on more precise modelling of their effects using computer techniques.

1.4 D region

1. The mid-latitude D region is complex chemically, but observationally its behavior may be deceptively simple. The region is under strong solar control and it vanishes at night.
2. We know that, to a first approximation, VLF (f <30 kHz) radio waves are reflected as at a sharp boundary in the D region. For VLF waves incident on the ionosphere at steep incidence, the reflection height, h, appears to vary as

$$h = h_o + H \ln \sec(\chi)$$

Where χ is the solar zenith angle. h_0 comes to about 72 km, and H to about 5 km which happens to be the scale height of the neutral gas in the mesosphere.

3. At oblique incidence, when transmitter and receiver are more than about 300 km apart, the height variation follows a quite different pattern. The reflection level falls sharply before ground sunrise, remains almost constant during the day, and then recovers fairly rapidly following ground sunset. The reason has to do with the formation and detachment of negative ions at sunset and sunrise, coupled with electron production by cosmic ray ionization - a source with no diurnal variation. This lower part of the D region is sometimes called a *C layer*. These patterns of height variation are illustrated in Figure 13.
4. The D region is the principal seat of radio absorption, and absorption measurements are one way of monitoring the region. The absorption per unit height depends on both the electron density and the electron-neutral collision frequency

$$\kappa = \frac{\omega}{c} \cdot \frac{1}{2\mu} \cdot \frac{XZ}{1+Z^2} = \frac{e^2}{2\varepsilon_0 mc} \cdot \frac{1}{\mu} \cdot \frac{N\nu}{\omega^2+\nu^2}.$$

   And the measurement gives the integrated absorption up to the reflection level.
5. Generally, the absorption varies with solar zenith angle as $(\cos \chi)^n$ with n in the range 0.7-1.0.
6. However, the seasonal variation contains an intriguing (مثيره للاهتمام) anomaly, which is that during the winter months the absorption exceeds by a factor of two or three the amount that would be expected by extrapolation from the summer. Moreover, the absorption is much more variable from day to day in the winter. This phenomenon is the *winter anomaly of ionospheric radio absorption.*

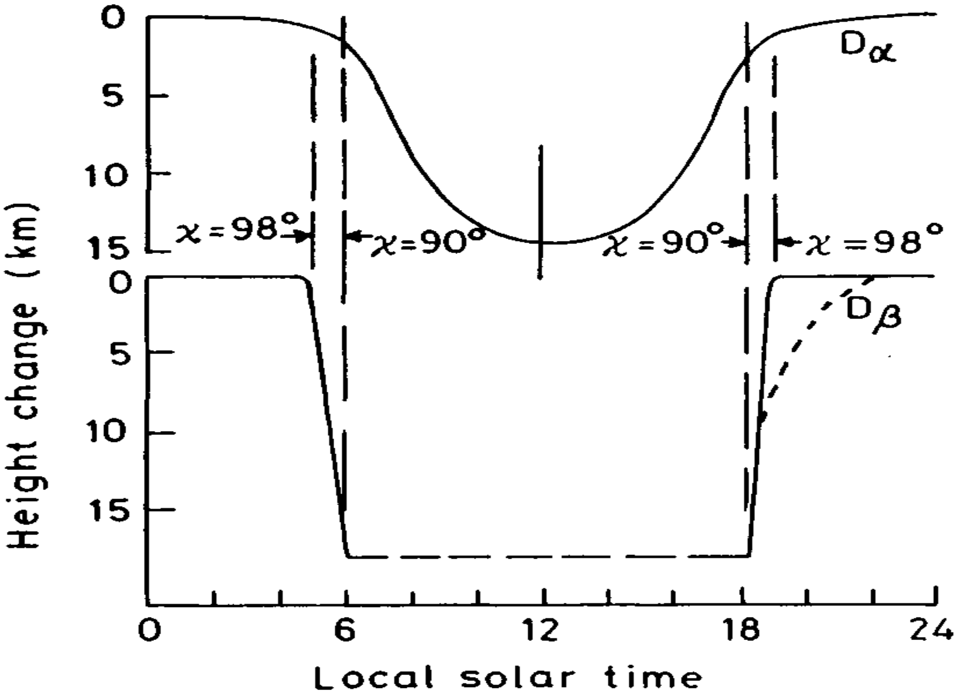


Fig. 13 Two kinds of diurnal behavior of the D region inferred from VLF radio propagation at vertical and oblique incidence. The regions originally called Dα and D_β are now more usually known as D and C. The evening recovery at oblique incidence tends to be more gradual than in a simple D_β pattern and similar to the dashed curve.

1.5 Effects of solar flares

1. The solar flare serves as a useful indicator of solar activity since more occur when the Sun is more active. It also has direct consequences in the ionosphere.
2. In 1937, J. H. Dellinger recognized that *fadeouts* in high-frequency radio propagation were the result of abnormally strong absorption in the ionosphere occurring at the same time as a solar flare.
3. The fadeout had a rapid onset and a typical duration of tens of minutes, like the visible flare. Because they begin suddenly, all the immediate effects of a solar flare are known as *sudden ionospheric disturbances (SID).* The absorption effect discovered by Dellinger was for long called the *Bellinger fade* but is now generally termed a *shortwave fadeout (SWF).*
4. The nature and timing of the SWF immediately provide two clues to its nature. The simultaneity between the fadeout and the visible flare shows that the cause is electromagnetic; and the occurrence of radio absorption indicates that the electron density in the D region has been increased.
5. The enhancement is most likely to be in the Lyman- α line or in the X-ray flux. Lyman- α is enhanced by a few per cent during a flare, and for many years this was thought to be the cause of the SWF. But when it became possible to measure hard X-rays from rockets it was observed that they intensified by several powers of ten during a flare; thus the SWF is now attributed to the X-rays emitted from the flare.

6. There is in fact a wide range of ionospheric disturbances, as summarized in Table 2. Though the effects are most marked in the D region, E and F region effects can also be detected. The electron content, governed mainly by the F region, is increased by a few per cent.
7. The range of effects shows that a considerable band of the spectrum is enhanced. Moreover, there is some difference of timing: the EUV radiation and the hardest X-rays tend to be enhanced early in the flare (at the same time as the 10 cm radio noise), whereas the softer X-rays last longer and correspond more closely with the optical flare. These phases are shown schematically in Figure 14.
8. Comparing the predicted and observed effects of flares is a useful way to verify models of the ionosphere and its production and loss mechanisms. The computation would take X-ray and EUV fluxes measured directly on a satellite, include an atmospheric model and assume the photochemistry.
9. Figure 15 shows increases of electron density calculated from the enhancement of flux in six bands. Note that most of the F-region increase comes from the middle bands, 260-796 A, whereas radiation at longer and shorter wavelength contributes more at the lower altitudes. The D-region effect is due to the hardest X-rays.
10. Figure 16 shows how the electron density profile changed during an actual flare. The F-region effect was by no means negligible, though it was smaller than the D- and E-region effects in percentage terms.
11. All SID effects cover the whole of the Earth's sunlit hemisphere and are essentially uniform except for a dependence on the solar zenith angle. Even flares that are not observed visually can be detected by their ionospheric effects and some properties of the flare deduced. One good reason for studying the effects of solar flares is that nuclear explosions also create dramatic effects in the ionosphere and it is important not to confuse the two!

Table 2 *SID phenomena*

		Technique	Effect	Region	Radiation
SWF	Shortwave fadeout	HF radio propagation	Absorption	D	Hard X-rays 0.5–8 Å
SCNA	Sudden cosmic noise absorption	Riometer	Absorption		
SPA	Sudden phase anomaly	VLF radio propagation	Reflection height reduced		
SEA	Sudden enhancement of atmospherics	VLF atmospherics	Intensity enhanced		
SFE	(Magnetic) solar flare effect	Magnetometer	Enhanced ionospheric conductivity	E	EUV and soft X-rays
SFD	Sudden frequency deviation	HF Doppler	Reflection height reduced	E + F	EUV
—	Electron content enhancement	Faraday effect	Content enhanced	F	EUV

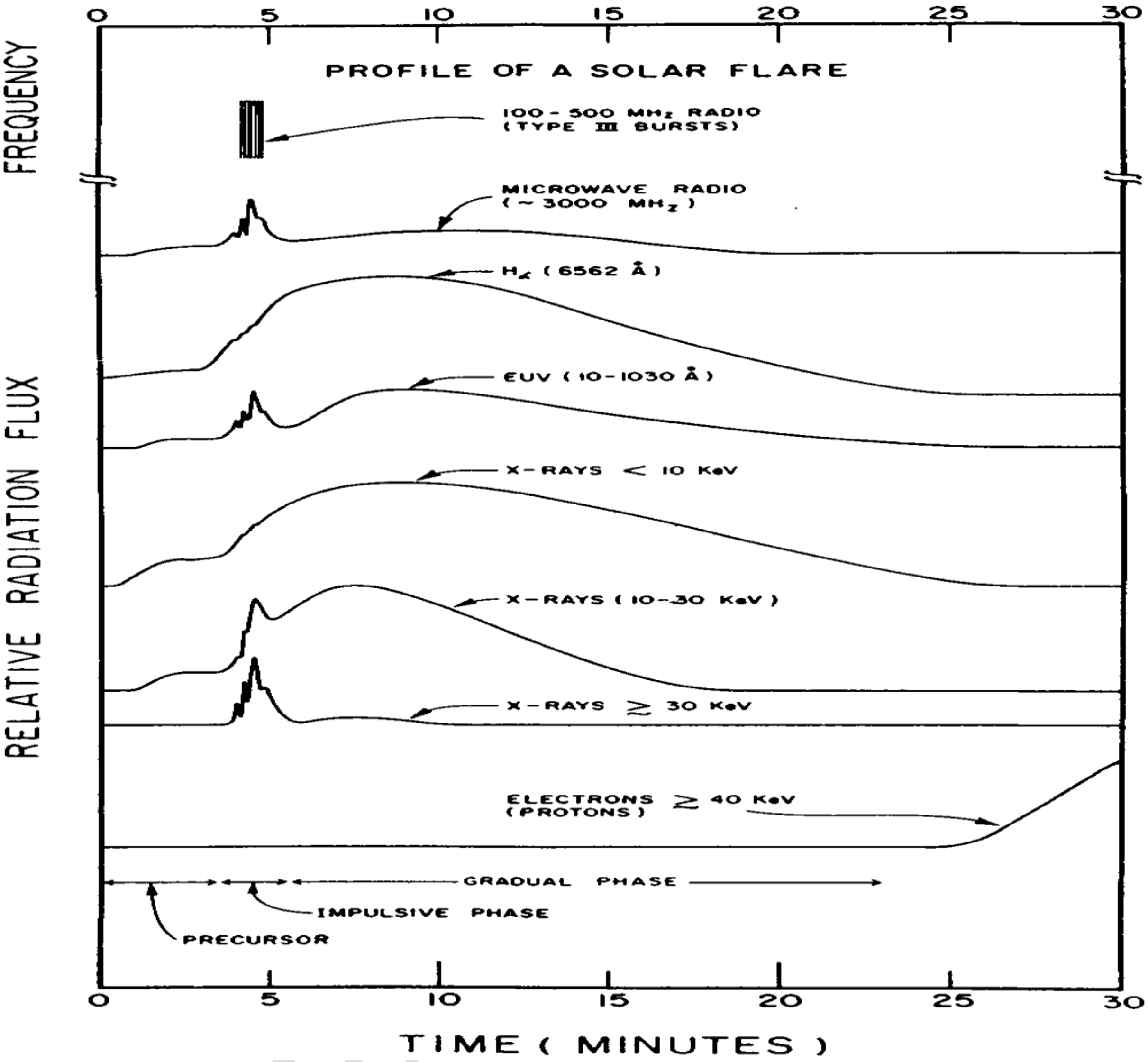
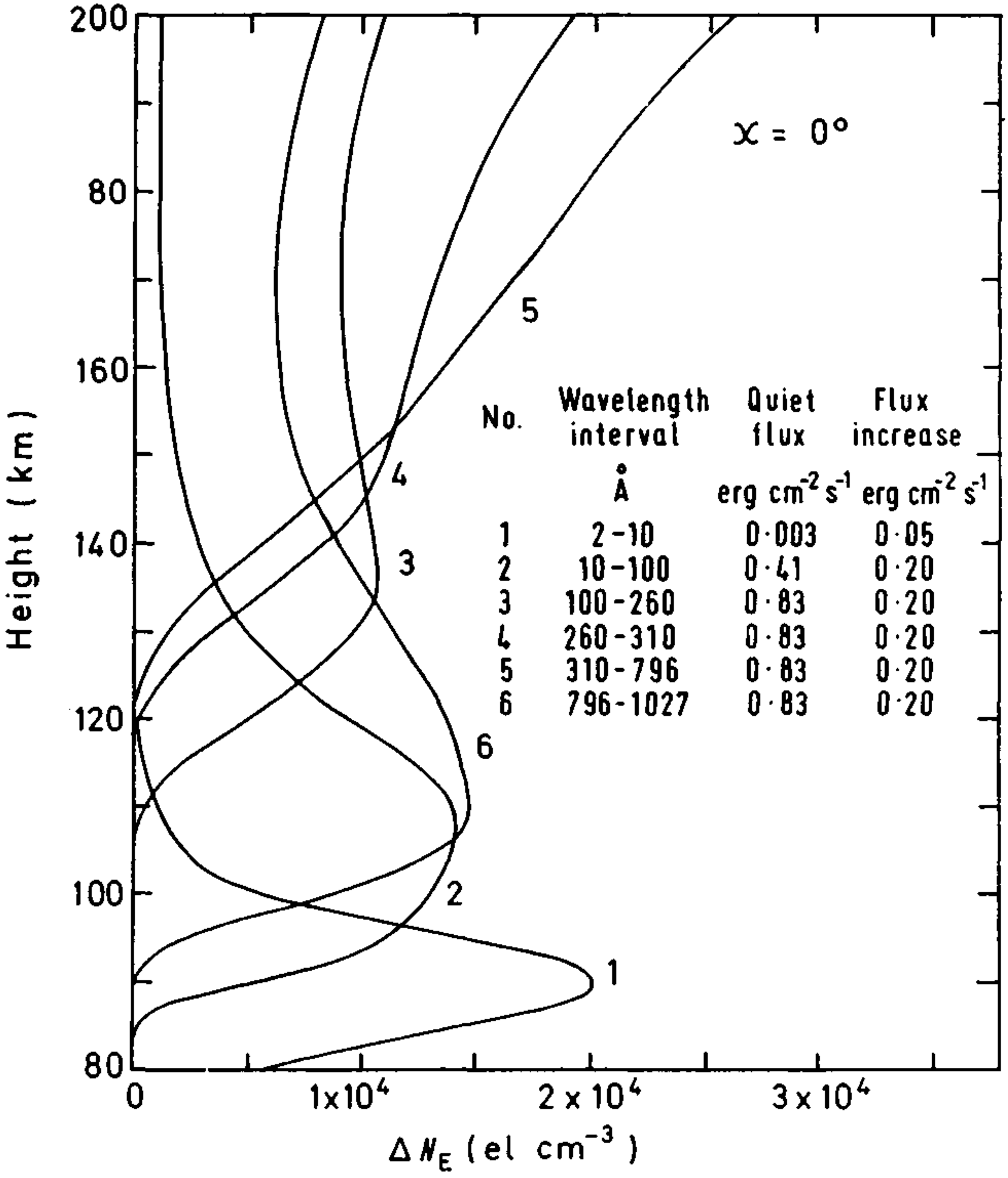


Fig. 14 Typical time scale of emissions from a flare. (D. M. Rust, in *Solar System Plasma Physics* (eds. Parker *et al.)*. North-Holland, 1979. Elsevier Science Publishers)



No.	Wavelength interval Å	Quiet flux erg cm^{-2} s^{-1}	Flux increase erg cm^{-2} s^{-1}
1	2-10	0·003	0·05
2	10-100	0·41	0·20
3	100-260	0·83	0·20
4	260-310	0·83	0·20
5	310-796	0·83	0·20
6	796-1027	0·83	0·20

Fig. 15 The computed increase of electron density due to flux increases in the stated wavelength intervals. (A. D. Richmond, private communication)

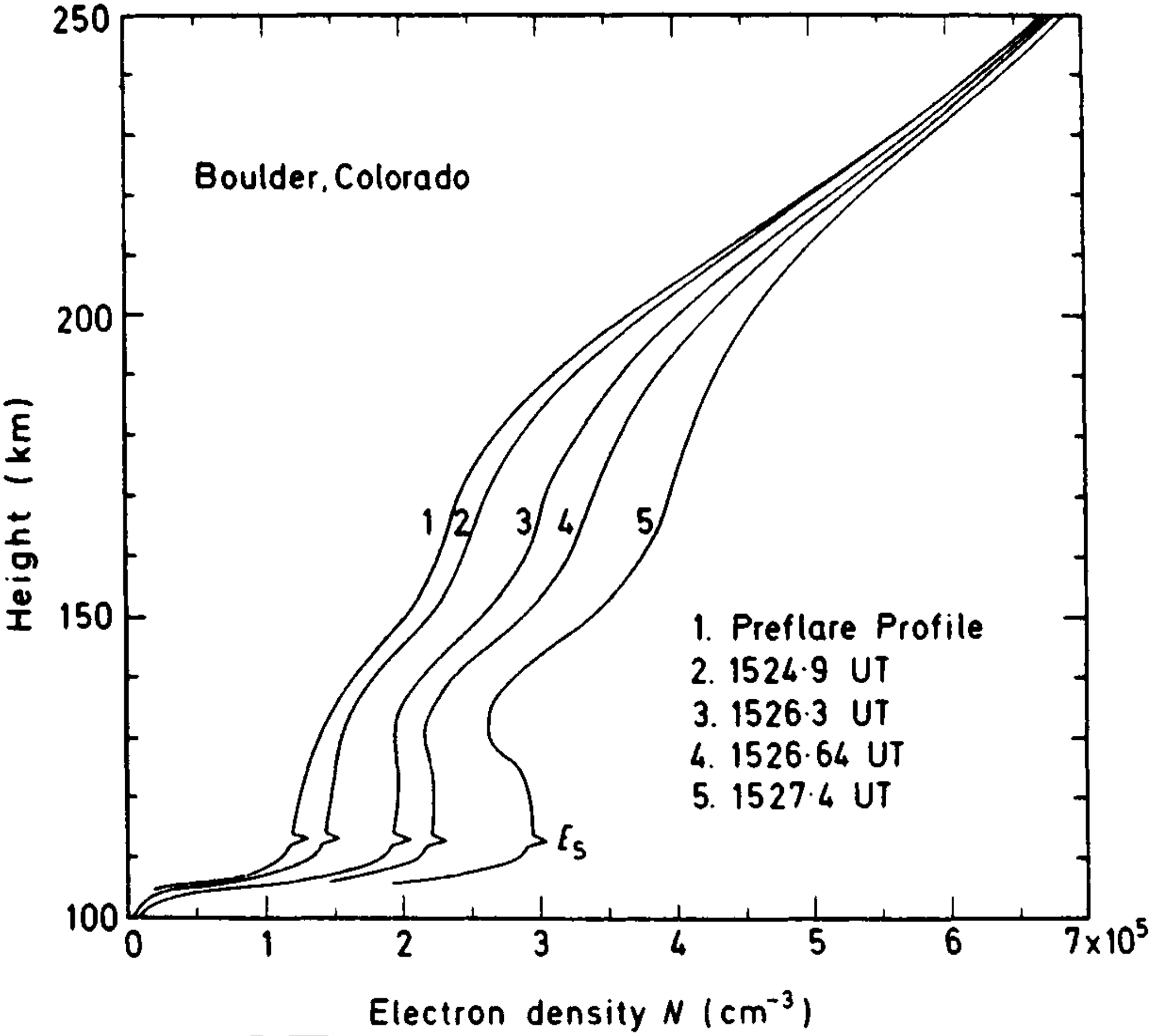


Fig. 16 Change of ionospheric profile during a flare, deduced from SFD observations. (R. F. Donnelly, *Report ERL 92-SDL6,* Environmental Research Laboratories, Boulder, Colorado, 1968)

Home works

1. Explain the continuity equation in detail.
2. What are the main mechanism for producing the maximum production and ionization in the ionosphere?
3. The rate of production of ion-electron pairs depend on four important parameters. What are these?
4. What are the Chapman production function in a normalized form?
5. Prove that α -Chapman layer is given by the following electron density formula.
$$N_m = N_{mo}\, cos^{1/2}(\chi)$$
6. Prove that β-Chapman layer is given by the following electron density formula.
$$N_m = N_{mo}\, cos\ \ (\chi)$$
7. Explain the optical depth in details?
8. Derive the relaxation time for β-Chapman layer?
9. What are the main causes for airglow and auroral emissions?
10. The D region is the most complex part of the ionosphere. Why?
11. Define 1. Ambipolar diffusion. 2. Absorption cross section 3. ionization efficiency
12. If you have two ions species with recombination coefficients equal (10^{-18} and 10^{-13} m^3/s), which one will remain long lifetime? And why?
13. Estimate the ionization efficiency for O_2 where the maximum wavelength is 102.7 nm?
14. Explain in details the loss mechanism in F2 region?