SEL-71-048

The Coupling between the Ionosphere and the Protonosphere and Its Implications on the Long-Term Variations of Ionospheric Electron Content

by

Helio Waldman

September 1971

Technical Report No. 13

Prepared under

National Aeronautics and Space Administration
Research Grant No. NGR 05-020-001

RADIOSCIENCE LABORATORY

STANFORD ELECTRONICS LABORATORIES

STANFORD UNIVERSITY · STANFORD, CALIFORNIA

SEL-71-048

THE COUPLING BETWEEN THE IONOSPHERE AND THE PROTONOSPHERE AND ITS IMPLICATIONS ON THE LONG-TERM VARIATIONS OF IONOSPHERIC ELECTRON CONTENT

by

Helio Waldman

September 1971

Technical Report No. 13

Prepared under

National Aeronautics and Space Administration
Research Grant No. NGR 05-020-001

Radioscience Laboratory
Stanford Electronics Laboratories
Stanford University Stanford, California

ABSTRACT

Recent experimental determinations of the flux of ionization between the ionosphere and the protonosphere have indicated that, during the day, the plasma will often flow upwards at a rate that by far exceeds the previously calculated theoretical limit of that flux. The theory of the coupling between the two layers is reexamined in the light of the new evidence, and the discrepancy mentioned above is shown to be probably due to a previous underestimation of the neutral hydrogen concentrations in the thermosphere.

In the upper F-region, where diffusion dominates the behavior of the ionospheric oxygen ions, there is a maximum velocity with which such ions can flow upwards. This fact, combined with the condition of photochemical equilibrium prevailing in the lower ionosphere, gives rise to the existence of a limiting upward flux of oxygen ions from the ionosphere. In the topside ionosphere, a similar situation occurs with the protons: up to a certain height they are under chemical equilibrium, and above, they can escape upwards at a rate which is limited by collisions with the oxygen ions, the dominant charged species. Consequently, a maximum proton flux also exists, which limits the rate at which ionization can be transferred from the ionosphere to the protonosphere. Such flux is controlled by the abundance of oxygen ions in the topside ionosphere, which in turn is strongly dependent on the upgoing flux of this species. For continuity reasons, the flux of oxygen ions coming from the bulk of the F region must match the flux of protons escaping the ionosphere.

It is shown that, on account of these simultaneous relationships, the limiting exchange flux cannot be properly determined from a fixed model of the F-region (as has been attempted in the past) particularly when

the resulting flux value is comparable with the limiting upward flux of oxygen ions in the ionosphere. In order to circumvent this difficulty, a self-consistent limiting proton flux is derived.

The effect of the exchange flux on the daytime ionosphere is studied. It is found that fluxes as large as have been observed can cause significant depletion in the daytime ionosphere, particularly in the topside. On the other hand, it is argued that the daytime proton flux normally approaches its limiting value because of low pressures prevailing in the protonosphere; some experimental support to this contention is presented.

The long-term variations in the daytime exchange flux are estimated with the use of model hydrogen concentrations based on the well-known inverse relationship between the abundance of neutral hydrogen and the neutral temperature in the thermosphere. The results are found to be quite compatible with the observed long-term behavior of the ionospheric electron content at a midlatitude location, as revealed by Faraday observations using geostationary satellites. It is found that the difference in behavior between the solar-cycle variations in the protonosphere and in the daytime ionosphere can be understood as resulting from the solar-cycle variation in hydrogen concentration. Moreover, the semiannual variation in electron content, peak electron concentration and the height of the peak can be linked to the well-known semiannual variation in the thermospheric neutral temperature through the effects of the exchange flux.

The present work also includes a review of the basic processes occurring in the ionosphere, with emphasis on the concepts of limiting velocity and limiting flux. In addition, an approach to the problem of numerical simulation of the ionosphere is presented and discussed.

ACKNOWLEDGEMENTS

I am very grateful to Dr. A. V. da Rosa for his invaluable advice and support throughout the course of this work.

Mr. S. C. Hall was responsible for the collection of the data used in this study, many of which were reduced and processed under supervision of Mr. O. G. Almeida. I would like to thank Miss M. Budney for typing this dissertation. This work was supported by NASA Research Grant No. NGR 05-020-001.

TABLE of CONTENTS

TABLES

ILLUSTRATIONS

Chapter 1

INTRODUCTION

Most of the plasma content of the ionosphere is associated with the F region, a layer consisting mainly of O^+ ions and electrons and extending from 200 km to at least 600 km, and occasionally up to 2000 km, with the higher altitudes occurring during the day. This upper boundary is defined by the existence of another layer on top of the ionosphere, consisting mainly of protons and electrons. This layer, which is called the protonosphere, extends to several earth's radii up to a region where the geomagnetic lines are not normally closed throughout the course of a day.

One of the most puzzling features of ionospheric behavior is the striking disparity between the ways in which the plasma content varies with time in the F region and in the protonosphere. Within both regions, the plasma is forced to move along the geomagnetic lines, except for occasional electromagnetic drifts set up by transverse electric fields. Each protonospheric tube of force is connected to the F region at both ends. A flow of protons to or from the ionosphere is continually cutting across the transition region of each end, either draining out or supplying the plasma content of the protonospheric tube. In the upper part of the F2 region, this flow of protons is converted into a flow of O^+ ions, or vice-versa, through the reactions:

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

The midlatitude ionosphere undergoes large daily variations in its columnar electron content, due to the incidence of solar ionizing

radiation during the day and its disappearance during the night. As ions and electrons are produced by photoionization during the day, their number will build up until the rate in which they are neutralized (through chemical reactions) matches the rate of their generation. At night, photoionization essentially ceases to exist, and the number of charged particles will decay as the ions and electrons are recombined into neutral particles. The resulting ratio between daytime and nighttime ionospheric electron content will depend on the season or the degree of solar activity, but will generally fall between 3 and 10.

If the exchange of ionization between the ionosphere and the protonosphere were rapid enough, one could expect to observe the same kind of diurnal variation in the content of a protonospheric tube of magnetic lines connected with the midlatitude ionosphere. Actually, however, this is not the case; large diurnal changes of a regular nature do not seem to occur in the protonospheric tubular content. Since the beginning of the last decade, this apparent discrepancy has prompted closer investigations into the nature and extent of the coupling between the two layers. Hanson and Ortenburger [1961] have shown that the nighttime flow of protons into the ionosphere is not high enough to effectively deplete the content of a tube during the course of one night. On the other hand, Geisler [1967] has derived theoretically a limiting value for the daytime flux of protons into the protonosphere which was so small that the protonospheric tubular content could not possibly follow the diurnal variation in the ionospheric content. The limitation in the flow of protons is not due to slowness in the charge-exchange process itself, but rather to the limited ability of the protons to flow through a majority of O^+ ions between the protonosphere (defined here as starting

where the concentration of protons exceeds that of atomic oxygen ions) and the region in the F2-layer where most of the charge exchange takes place.

In the protonosphere, unlike in the ionosphere, there is practically no local production or loss of ionization. For this reason, the plasma content of a protonospheric tube at any time is determined by the past history of the exchange of protons between the ionosphere and the protonosphere. In contrast, the ionospheric content is determined (in steady-state) by an equilibrium between local losses, a topside flux to or from the protonosphere and, during the day, local production through photoionization. During the night, the topside flux has to be negative (downward) in order to maintain a steady-state ionosphere without local ionization production. The existence of a nighttime steady-state ionosphere is shown by the fact that after sunset the electron content often does not decay to zero but rather to a finite value which can only be sustained by a flux of protonospheric origin in the absence of nighttime ionization sources.

The impact of the coupling flux on the daytime ionosphere is still unclear at present. The limiting daytime flux from the ionosphere to the protonosphere depends on a number of parameters, including the poorly known concentration of neutral hydrogen in the charge-exchange region (roughly between 500 and 800 km). Using a value of 10^4 cm^{-3} for this quantity at 400 km, Geisler [1967] derived a limiting flux of 1.5×10^7 cm^{-2} sec^{-1}, which can hardly be expected to have any sizeable effect on the ionosphere. More recently, however, evidence is accumulating which shows that the daytime flux from the ionosphere to the

protonosphere is actually in the range $4 \times 10^7 - 4 \times 10^8$ cm^{-2} sec^{-1}. This does not contradict Geisler's theory, because the concentration of neutral hydrogen has also been found to be one to two orders of magnitude higher than the value assumed in his calculations, although this result is still being disputed [Reber et al., 1967; Meier and Prinz, 1970]. In Chapter 4 of this work, the results of a theory by Geisler and Bowhill [1965] are used to show that fluxes of the magnitude mentioned above do have a powerful effect on the ionospheric content, by depleting the ionic concentrations, especially in the topside. Therefore, any seasonal or solar-cycle variations in these fluxes should reflect in a corresponding variation in the daytime ionospheric content. This is not by any means a straightforward result, since these fluxes are still a very minor fraction of the total integrated production of electrons.

Even granting that the coupling between the ionosphere and the protonosphere is sufficiently weak to keep large diurnal changes from occurring in the protonosphere, one might still expect that very slow variations in the ionospheric content (such as seasonal or solar-cycle effects) are reproduced in the protonospheric tubular content. Here again, however, nature denies our first expectations. While both nighttime and daytime ionospheric electron content are strongly affected by the level of solar activity, the protonospheric tubular content does not present any appreciable solar-cycle variation at all. Seasonal variations in the ionosphere and in the protonosphere are also apparently unrelated.

In the present work, it is shown that some of these disparities can be tentatively explained by slow variations in the coupling between

the two layers. These variations are attributed to changes in the concentration of neutral hydrogen which, due to its strong sensitivity to the global level of neutral exospheric temperatures, is thought to have strong seasonal and solar-cycle variations quite different from those of the other species present in the thermosphere. Our main purpose, however, is to explore the effect of these changes on the daytime ionospheric F-layer, and especially on its columnar content. This is done in Chapters 4 and 5.

In Chapter 4, the present state of the theory concerning the coupling between the ionosphere and the protonosphere is reviewed and discussed, and an extension to this theory is proposed. The extension is deemed necessary at the present stage because upward fluxes of the order of 10^8 cm^{-2} sec^{-1}, which have been measured during the day, are found to be able to deplete the O^+ concentrations on the topside F-layer. Since this quantity is one of the factors that determine the limiting proton flux itself, a more self-consistent treatment than that of Geisler's (who assumed a model ionosphere) is indicated. The extended theory seems able to explain the observed semiannual variation in the midlatitude daytime ionospheric electron content and its evolution throughout the solar cycle.

In Chapter 5, a large amount of ionospheric electron content data is studied; the discussion is focused on the long-term ionospheric variations as revealed by these data and other independent sources. Several components of these variations are distinguished: The solar-cycle variation follows the well-known variation of solar activity throughout the 11-year period; the annual and seasonal variations have both a 12-month period, but while the former has the same phase at two

opposite latitudes, the latter has opposite phases at opposite latitudes; and the semiannual variation, with a 6-month period. No attempt was made to distinguish between seasonal and non-seasonal components in the observed semiannual electron content variation. The comparative morphology of the semiannual variations in the ionospheric electron content, in the peak electron concentration and in the height of the peak for middle latitudes is shown to be fairly compatible with the idea that they are caused by changes in the daytime flux from the ionosphere to the protonosphere. The solar cycle variation in ionospheric electron content is also discussed and shown to be consistent with the assumed evolution of the coupling between the ionosphere and the protonosphere.

Chapters 2 and 3 have tutorial value only, as far as the main purpose of this dissertation is concerned. The former is essentially a review of the basic processes that control the ionospheric plasma. Special emphasis is given to the existence of limiting velocities in regions where diffusion dominates the plasma under the action of gravity. This concept is very important for the understanding of the limiting flux that arises whenever such a region exists on top of a chemical equilibrium region, as discussed in Chapter 4.

Chapter 3 discusses the numerical simulation of the ionosphere, an increasingly important tool for ionospheric research. An approach to this problem is proposed in which a single program can handle different kinds of boundary conditions. It is hoped that this scheme will add new flexibility to the use of numerical techniques in the study of ionospheric processes.

Chapter 2

THE DYNAMIC AND THERMAL BEHAVIOR OF THE IONOSPHERIC PLASMA

The last decades have witnessed a tremendous progress in our knowledge of the ionospheric behavior, in particular of the F-layer. The extensive and continued use of sounding techniques in many localities around the world for some decades has allowed the formation of a global picture of the behavior of the critical frequency of the layer. More recently, the development of radio techniques and the use of rockets and satellites have provided the opportunity to examine the detailed structure of the ionosphere, both in its bottomside (below the peak of the F-layer) and in the topside. Today, the increasing number of geostationary satellites in orbit is creating the proper conditions for continuous worldwide measurement of ionospheric columnar electron content.

The extensive use of old and new experimental techniques of measuring ionospheric parameters has generated a very large amount of data. As more and more information was brought about by these data, more and more "anomalies" were revealed in the ionospheric behavior, reflecting the relative poorness of the physical models being used to describe the ionosphere. These anomalies or departures from a very simple model of behavior may be related to many different features of ionospheric morphology, such as spatial gradients (e.g., the equatorial anomaly), annual variation (e.g., the seasonal anomaly), and persistence of the nighttime ionosphere. The prevalence of anomalous behavior of the ionospheric F-layer has given rise to a general effort aimed at creating a physical picture of the ionosphere that would be more consistent with observed behavior.

The situation is complicated by the fact that the ionospheric plasma is subject to many different processes that act on it simultaneously; each region of the ionosphere is controlled by a different kind of balance between these processes, but all the regions are in some degree coupled to each other either through diffusive or conductive processes. During the day, for instance, ions and electrons are being constantly created through ionization of the ambient neutral particles by incident solar radiation. Basically, this photoionization mechanism will produce O^+, N_2^+ and O_2^+ ions in comparable amount, but the last two species of ions are relatively short-lived and rapidly lose their charge either by direct recombination with electrons or through a chain of reactions involving the ambient neutral particles. For this reason O^+ is the dominant ion in most of the F layer, with the exception of the lower altitude region sometimes called the F1-layer, where the electron concentration is so low and the concentration of neutral molecular species (which react with the O^+ ion) is so high that there the NO^+ and O_2^+ ions have a longer life than the O^+ ions.

Partly through the action of Coulomb collisions with the thermal electrons, the energetic photoelectrons created by the photoionization are thermalized to the ambient electron temperature. This results in heating of the ionospheric plasma, which, in turn, transfers heat to the neutral atmosphere mainly through elastic collisions between ions and neutral particles. If the temperatures that would result from a simple local equilibrium between heating and cooling processes were allowed to exist, they would set up very strong conductive heat fluxes due to their high gradients. Actually, the presence of conductive heat fluxes along the magnetic lines in the ionosphere allows for a considerable

imbalance between heat production and loss on a local basis. The resulting ion and electron temperatures will influence the rate at which the plasma will diffuse along the magnetic lines. This diffusion is limited by collisions between the ionized and neutral particles, thus creating a mechanical link between these species. The mechanical coupling between the ionospheric plasma and the neutral atmosphere is also manifested in the action of neutral winds dragging the ions along the magnetic lines, and conversely by the limiting of these winds by the retarding action of the ions.

The plasma is also chemically coupled to the neutral particles, since the O^+ ions react with the N_2 and O_2 molecules, creating NO^+ and O_2^+ ions that are rapidly lost through recombination with electrons. Since the concentration of neutral particles decreases exponentially with height, this recombination action creates a plasma loss region in the bottomside of the ionosphere, in which the concentration of ionized particles is determined by a balance between production and loss processes. Above this region, diffusion is fast enough to give the ionized particles a chance to redistribute themselves before recombining, with the result that the peak and the topside of the F layer are located in a diffusion region.

In the topside, a chemical balance exists between protons and oxygen ions as a result of their exchange of charges with oxygen and hydrogen atoms, respectively. The ambient hydrogen ions may escape to the protonosphere, thereby depleting the F region. In order to escape upwards, however, they must diffuse through the O^+ ions (which, among all background gases, are the one that presents the largest frictional resistance to a moving gas of protons in the upper F- region)

and they can do so only with a limited velocity. The depletion of the ionosphere caused by the outflow of hydrogen ions is then limited by the presence of a diffusive barrier in the topside of the ionosphere. The characteristics of this barrier are such that the maximum possible amount of escape flux crossing it is determined by many different factors, in particular the poorly known abundance of neutral hydrogen atoms.

It is clear that the behavior of the ionospheric F-layer is determined by a very complex collection of factors that mutually influence each other, making it quite difficult to predict the ionospheric behavior that will result from a given set of conditions in the neutral atmosphere and a given flux of solar ionizing radiation. The temptation to neglect certain features of the problem is great, and, as one might expect, this was indeed the approach adopted in early ionospheric work. Initially, much attention was given to production and loss processes, and the assumption of balance between these two mechanisms was found to explain the ion concentrations observed in the lower part of the layer. This assumption, however, did not fit at all the observed profiles of the F-layer around and above its peak. The introduction of diffusive processes into the physical model of the ionosphere, which led to the Chapman-Ferraro equation, was probably the largest step ever made in bringing ionospheric theory closer to observed behavior. As a result of this step, a satisfactory explanation was achieved for the general shape of the observed profiles of electron concentration. More recently, it has become clear that electric fields and neutral winds in the thermosphere play a major role in controlling the F-layer. A visible

gap between theory and observation still remains, however, with many observed features of ionospheric behavior still lacking a final interpretation.

The variety of mechanisms that control the behavior of the F-layer is not exhausted by the oversimplified picture presented above. Some of these mechanisms cannot be suitably accounted for, either because some relevant physical constant cannot be accurately measured in the laboratory (e.g., many reaction rates) or because some important physical parameter of the neutral atmosphere is poorly known (e.g., the rate of production of neutral hydrogen, the velocity of neutral winds). Moreover, the simultaneous action of many different interrelated processes on the plasma makes it difficult to determine the final behavior of the F-layer that results from all these mechanisms.

The present effort to create a better understanding of the ionosphere is being developed along many different lines. The need to know with best possible accuracy the physical constants that are relevant to individual ionospheric processes has motivated the study of the dynamic response of the ionosphere to specific events such as eclipses, as well as prompted the performance of a series of laboratory measurements. On the other hand, the need to understand how the different physical processes act on the ionospheric plasma in the presence of one another has spurred a sequence of works in the field of numerical simulation of the ionosphere.

The purpose of this chapter is to review the basic processes that control the behavior of ions and electrons under ionospheric conditions. This review will provide the groundwork for later discussion of

specific problems raised by ionospheric observations, and at the same time will present the derivation of the equations that are to be solved in an ionospheric simulation program.

The momentum transport equation will be applied to ionospheric conditions, and the resulting expressions will be combined with the continuity equation to derive an equation that describes the time-varying behavior of the ionic concentration. Special emphasis will be given to the concepts of full flow mode and limiting velocities, since these concepts are basic for the understanding of the coupling between the ionosphere and the protonosphere, a subject that is to be further developed in a later part of this study.

At some point in the review, the treatment has to be specialized for the case of the major ion and the minor ions. The treatment given to the major ion is mainly applicable to the O^+ ions, which are dominant over a wide range of altitudes in the ionosphere. The differences in behavior between the major ion and a minor ion are discussed and the reasons for these differences are analyzed.

The distribution of ionization in the ionosphere is largely determined by the electron and ion temperatures, which are determined by a balance between heat production, cooling and conduction that is closely related to the plasma concentration. For this reason, the heat flow equation is coupled to the momentum transport and continuity equations, and a realistic numerical simulation of the ionosphere will have to solve these equations simultaneously. Accordingly, a discussion of the heat equations for the ionospheric plasma is also presented in this review.

Since the present research is mainly concerned with the ionosphere at midlatitudes, many assumptions will be adopted in the derivation of the equations that would not be reasonable under equatorial or polar conditions.

A. The Momentum Transport Equation for the Major Ion

The momentun transport equation describes the rate of change of the momentum carried by an elementary cell of a species of a given gaseous mixture under the influence of pressure gradient, friction with other species and external forces. It can be rigorously derived from Boltzmann's equation by taking its moment in velocity space [Spitzer, 1962]. However, a somewhat less general approach will be used here, so that the physical origin of each term of the equation can be properly identified. In addition, the relative importance of each term will be evaluated for ionospheric conditions.

Let us consider a species with particle mass m, concentration n, and bulk (average) velocity $\vec{V}$. Then, a unit volume cell carries a momentum given by $nm\vec{V}$. The time derivative of this quantity must then equal the sum of all forces applied to the particles of the cell, either by its own species (pressure gradient), or by collision with particles of other species, or by external fields (electric, magnetic and gravitational). In addition, the total momentum of the cell will also be affected by the momenta brought in by newly created particles within the cell, or taken out by particles that disappear through recombination or other loss processes.

Summing up all contributions, and neglecting the Coriolis term due to rotation of the Earth:

$$\frac{d}{dt}(nm\vec{V}) = n\vec{F} - \vec{\nabla}p + \sum_i (\vec{V}_i - \vec{V})\, mn\nu_i - Lm\vec{V} + Qm\vec{V}_q \qquad (1)$$

where: $\vec{F}$ = sum of all external forces acting on a particle of the species

p = partial isotropic pressure of the species

$\sum_i$ = summation over all species except the one under consideration

ν_i = collison frequency for momentum exchange with species "i"

L = rate of loss of particles per unit volume

Q = rate of creation of new particles per unit volume

$\vec{V}_q$ = average velocity of the newly created particles

Assuming a perfect gas

$$p = nkT, \qquad (2)$$

where T is the temperature of the species under consideration, and k is Boltzmann's constant.

The continuity equation is

$$\frac{\partial n}{\partial t} = Q - L - \nabla \cdot (n\vec{V}) \qquad (3)$$

Rewriting the left-hand side of (1) with the use of the continuity equation:

$$\frac{d}{dt}(nm\vec{V}) = nm\frac{d\vec{V}}{dt} + m\vec{V}\left[\frac{\partial n}{\partial t} + \vec{V} \cdot \nabla n\right] =$$

$$= nm\frac{d\vec{V}}{dt} + m\vec{V}\left[Q - L - n(\nabla \cdot \vec{V})\right] =$$

$$= nm\frac{\partial \vec{V}}{\partial t} + nm\left[(\vec{V} \cdot \nabla)\vec{V} - (\nabla \cdot \vec{V})\,\vec{V}\right] + m\vec{V}\,(Q - L)$$

Substituting in (1):

$$nm \frac{\partial \vec{V}}{\partial t} + nm \left[(\vec{V} \cdot \nabla) \vec{V} - (\nabla \cdot \vec{V}) \vec{V} \right] =$$

$$= n\vec{F} - \vec{\nabla} (nkT) + \sum_i (\vec{V}_i - \vec{V}) \, nm\upsilon_i + Qm(\vec{V}_q - \vec{V})$$

If the macroscopic quantities such as concentration, velocity and temperature vary in space over a characteristic length H, then the term $nm[(\vec{V} \cdot \nabla)\vec{V} - (\nabla \cdot \vec{V}) \vec{V}]$ will have a magnitude of the order of $nm \frac{V^2}{H}$, or much smaller. By comparison, the pressure gradient term will be approximately.

$$\vec{\nabla} (nkT) \sim \frac{nmv_T^2}{H}$$

where v_T is the mean thermal velocity of the particles of the species.

If the bulk velocity $\vec{V}$ of the particles is negligible when compared with the thermal velocities, one can then make the following approximation:

$$nm \frac{\partial \vec{V}}{\partial t} = n\vec{F} - \vec{\nabla}(nkT) + \sum_i (\vec{V}_i - \vec{V}) nm\upsilon_i + Qm (\vec{V}_q - \vec{V})$$

In the ionosphere, this approximation is generally valid since the plasma velocity rarely exceeds 100 m/sec., while the ionic thermal velocities in the F region are mostly above 1 km/sec. For the electrons, the approximation is even much better.

In order to lump the effects of all other species into the equivalent of a single background species, it is convenient to define the following quantities:

$$\upsilon \equiv \sum_i \upsilon_i$$

$$\vec{U} \equiv \frac{1}{\upsilon} \sum_i \vec{V}_i \, \upsilon_i$$

While ν represents the total frequency of momentum-removing collisions for the species in question, $\vec{U}$ represents the bulk velocity of the "equivalent" single background species. Making use of these definitions, and neglecting momentum change due to production of new particles:

$$nm[\frac{\partial \vec{V}}{\partial t} + \nu\vec{V}] = n\vec{F} - \vec{\nabla}(nkT) + nm\nu\vec{U}$$

At any time then, the velocity will be approaching its steady-state value given by:

$$\vec{V}_{ss} = \frac{\vec{F}}{m\nu} - \frac{1}{mn\nu}\vec{\nabla}(nkT) + \vec{U}$$

If there is a sudden change in $\vec{V}_{ss}$, the velocity $\vec{V}$ will approach its new steady-state value with time constant ν^{-1}. If, however, $\vec{V}_{ss}$ varies with time scale much larger than ν^{-1}, then there will be time enough for the velocity to be always close to $\vec{V}_{ss}$. In the case of the O^+ ion, the dominant ion in the F_2-layer, one may safely assume $\nu^{-1} < 10$ min. for all altitudes. Assuming then that all macroscopic quantities, as well as the external forces driving the plasma, vary over time scales much longer than 10 minutes, one can say that $\vec{V} = \vec{V}_{ss}$ for all t, and so

$$nm\nu\vec{V} = n\vec{F} - \vec{\nabla}(nkT) + nm\nu\vec{U} \qquad (4)$$

This approximation is not justified when the application of Eq. (4) produces time constants shorter than ν^{-1}, as it happens at high altitudes in the F layer. In this region both production and loss processes have negligible influence and the application of (4) implicitly assumes that the plasma is controlled by diffusion through the neutral species. In this regime, the time required for the plasma distribution to reach steady-state (diffusion time) is proportional to ν: frequent

collisions produce fast response, and vice-versa. Above a certain altitude, the diffusion time will then be shorter than ν^{-1}, and consequently a diffusive time-varying response is not possible. Under these conditions the equations which follow will yield a response that is faster than the actual ionospheric response, since velocity is being replaced by its instantaneous final value at each time. In the steady-state case, Eq. (4) can be properly applied over all heights, since $\frac{\partial \vec{V}}{\partial t}$ is then zero by definition.

Let us now assume that spatial gradients exist only in the vertical direction. Calling h the height, and considering a gravitational, an electric and a magnetic field:

$$nm\nu\vec{V} = nm\vec{g} + nq\vec{E} + nq\,(\vec{V} \times \vec{B}) - k\,\frac{\partial}{\partial h}\,(nT)\,\hat{h} + nm\nu\vec{U}$$

where q is the charge of the particles and the symbol $\hat{}$ indicates a unit vector.

Letting $\hat{z}$ be in the direction of $\vec{B}$, $\hat{y}$ the direction perpendicular to $\vec{B}$ upwards in the plane $(\vec{B}, \hat{h})$, and $\hat{x} = \hat{y} \times \hat{z}$, one has the following three components of the equation above:

$$z: \quad nm\nu V_z = nmg \sin I + nqE_z - k\,\frac{\partial}{\partial h}\,(nT) \sin I + nm\nu U_z \qquad (5a)$$

$$y: \quad nm\nu V_y = nmg \cos I + nqE_y - nqV_xB - k\,\frac{\partial}{\partial h}\,(nT) \cos I + nm\nu U_y \qquad (5b)$$

$$x: \quad nm\nu V_x = nqE_x + nqV_yB + nm\nu U_x \qquad (5c)$$

where I is the magnetic dip angle.

Both for ions and electrons, the gyrofrequencies $\left|qB/m\right|$ by far exceed the momentum transfer collision frequencies ν over the range of interest in the ionosphere. Therefore, provided that $(V_x - U_x)$ is not exceedingly large, the collision terms in (5c) may be neglected when

compared with the magnetic term, resulting

$$V_y = \frac{E_x}{B}$$

The vertical component of the velocity is

$$v = V_z \sin I + V_y \cos I = V_z \sin I - \frac{E_x}{B} \cos I$$

Introducing (5a) to eliminate V_z:

$$nm\upsilon v = nm\upsilon \left[U_z \sin I - \frac{E_x}{B} \cos I\right] + nmg \sin^2 I + nqE_z \sin I - k \frac{\partial}{\partial h} (nT) \sin^2 I \tag{6}$$

Inspection of this equation reveals that the vertical component v of the plasma velocity is defined with respect to a velocity u specified by the external electric and magnetic fields, as well as the "background" velocity $\vec{U}$, in the following manner:

$$u \equiv U_z \sin I - \frac{E_x}{B} \cos I$$

The velocity u represents the sum of the vertical components of the drifts forced on the plasma by the background wind $\vec{U}$ and the electric field. Writing Eq. (6) for one ion species and for electrons, using indexes i and e respectively:

$$\begin{cases} n_i m_i \upsilon_i v_i = n_i m_i \upsilon_i u_i + n_i m_i g \sin^2 I + n_i eE_z \sin I - k \frac{\partial}{\partial h} (n_i T_i) \sin^2 I & (7a) \\ n_e m_e \upsilon_e v_e = n_e m_e \upsilon_e u_e + n_e m_e g \sin^2 I - n_e eE_z \sin I - k \frac{\partial}{\partial h} (n_e T_e) \sin^2 I & (7b) \end{cases}$$

Multiplying (7b) by $\left(n_i/n_e\right)$:

$$n_i m_e \nu_e v_e = n_i m_e \nu_e u_e + n_i m_e g \sin^2 I - n_i e E_z \sin I -$$

$$- k \frac{n_i}{n_e} \frac{\partial}{\partial h} (n_e T_e) \sin^2 I \qquad (7c)$$

Since $m_e \ll m_i$, and ν_e is less than 2 orders of magnitude above ν_i [Rishbeth and Garriott, 1969], one may approximate:

$$\begin{cases} m_e + m_i \cong m_i \\ m_e \nu_e v_e + m_i \nu_i v_i \cong m_i \nu_i v_i \\ m_e \nu_e u_e + m_i \nu_i u_i \cong m_i \nu_i u_i, \end{cases}$$

provided that v_e and v_i, as well as u_e and u_i, do not differ too much.

Then adding (7a) and (7c) in order to eliminate explicit reference to the electric field:

$$n_i m_i \nu_i v_i = n_i m_i \nu_i u_i + n_i m_i g \sin^2 I - k \sin^2 I \left\{ \frac{\partial}{\partial h} (n_i T_i) + \frac{n_i}{n_e} \frac{\partial}{\partial h} (n_e T_e) \right\} \qquad (8)$$

In the case of a single ionic species, one can assume $n_i = n_e$ due to the charge neutrality, thereby simplifying the equation. Although the ionosphere is not formed by a single ionic species, this assumption is still relevant for the study of ionospheric plasma, since at most altitudes of interest the mixture is strongly dominated by one of the ions. Accordingly, the discussion will now be specialized for the case of the major ion, for which $n_i = n_e$. In a later section, the minor ions will be examined. Dropping the subscripts in (8):

$$nm\nu v = nm\nu u + nmg \sin^2 I - 2k \sin^2 I \frac{\partial}{\partial h} (nT),$$

where: $T \equiv \frac{T_e + T_i}{2}$

Let us call $D \equiv D_a \sin^2 I$, where D_a is the ambipolar diffusion coefficient

$$D_a \equiv \frac{2kT}{m\upsilon}$$

Then:

$$v = u + \frac{g}{\upsilon} \sin^2 I - D \left[\frac{1}{n}\frac{\partial n}{\partial h} + \frac{1}{T}\frac{\partial T}{\partial h}\right] \tag{9a}$$

Or:

$$v - u = D \left[\frac{mg}{2kT} - \frac{1}{n}\frac{\partial n}{\partial h} - \frac{1}{T}\frac{\partial T}{\partial h}\right] \tag{9b}$$

Except for the inclusion of the vertical drift u, the equation above has now been known for a long time [Ratcliffe and Weekes, 1960]. This is an equation that relates density, velocity and temperature for the major ion. One can see from it that if at any time the temperature and velocity profiles are known, the density profile at that time can be determined. This is not true in general of course, but rather it is a consequence of the quasi-equilibrium condition ($^{dv}/_{dt} \cong 0$) assumed previously. In general the accelerations would also have to be known in order to construct the density profile.

From inspection of (9a), one can see that the vertical velocity v of the plasma is the sum of the three following terms:

(a) A "background" velocity

$$u = U_z \sin I - \frac{E_x}{B} \cos I \tag{10}$$

which combines the effects of the electric field and of neutral motions on the plasma dynamics.

(b) $(g/\nu) \sin^2 I$, which is the average downward velocity gained by an isolated ion between collisions with neutrals; thus, this velocity measures the effectiveness of the gravitational field in bringing an ion particle down in the presence of neutrals.

(c) A velocity given by:

$$-\frac{D}{nT}\frac{\partial}{\partial h}(nT) = \frac{-\frac{\partial}{\partial h}(2knT)}{nm\nu} \sin^2 I$$

Since (2knT) is the plasma pressure, $-\frac{1}{nm}\frac{\partial}{\partial h}(2knT)$ is the acceleration of the particles due to gradient in the plasma pressure. Therefore, this is the velocity gained by a particle with this acceleration, after a mean free time ν^{-1}.

All vertical forces are of course multiplied by $\sin^2 I$. Indeed, if F is a vertical force, only its component F sin I along the magnetic field is effective, and F $\sin^2$ I is the vertical projection of the effective component.

B. The Velocity Profile

In this section Equation (9) is applied to some specific cases. Since (9) is valid for the major ion, for which $n = n_e$, we can think of these equations as relating to the plasma distribution.

Let us start with the case when $(v - u) = 0$. This will occur, for instance, if there are no external electric fields and the plasma is stationary with respect to the neutral background material. Then, from (9b):

$$\frac{1}{nT}\frac{\partial}{\partial h}(nT) = -\frac{mg}{2kT} \equiv \frac{1}{2H}$$

where H is the scale height of a neutral gas with atomic mass m and temperature T.

Integrating:

$$n = \frac{n_o T_o}{T} \exp\left[-\int_{h_o}^{h} \frac{d\xi}{2H(\xi)}\right] \tag{11a}$$

Assuming $\frac{\partial T}{\partial h} = \frac{\partial g}{\partial h} = 0$:

$$n = n_o \exp\left[-\frac{(h - h_o)}{2H}\right] \tag{11b}$$

The density profile described by Equation (11) is called a hydrostatic equilibrium distribution.

It is also clear that if $\frac{|v - u|}{D} \ll \frac{1}{2H}$, then the expression in brackets in Equation (9b) is approximately zero, so that the distribution will still be nearly one of hydrostatic equilibrium described by (11), even though $(v - u) \neq 0$.

In order for this to happen, it is sufficient that:

$$|v - u| \ll \frac{D}{2H} = \frac{|g|}{\nu} \cdot \sin^2 I \tag{12}$$

As seen previously, the velocity (D/2H) is the average downward velocity gained by an ion from the gravitational field between two successive collisions. Therefore, (12) implies that the plasma is under control of gravity and its own pressure gradient, and is not controlled by collisions.

Since the collision frequency ν is proportional to the neutral concentration, it decreases exponentially with height. Then, at a sufficiently high altitude there will frequently be a region where (12) is satisfied and the hydrostatic equilibrium distribution is approached. In this region, n(h, t) will approach the value given by Equation (11).

When (12) does not hold:

$$\frac{1}{nT}\frac{\partial}{\partial h}(nT) \equiv -H_p^{-1} = \frac{1}{2H} - \frac{v-u}{D} \tag{13}$$

where H_p is the effective scale height of the plasma pressure at any given height.

Some interesting features of the interrelation between the velocity and concentration profiles are revealed by an inspection of (13):

(a) At the peak of the ionization profile, $H_p = \infty$. Therefore:

$$(v-u)_{peak} = -\frac{D}{2H} = \frac{g}{\nu}\sin^2 I$$

This is the average velocity gained by a particle between consecutive collisions, under the action of a vertical acceleration $g \sin^2 I$. Due to the absence of vertical pressure gradient at the peak of the F2-layer, the plasma vertical velocity there is under exclusive control of the gravitational field, collisions and the externally imposed vertical drift u.

This result makes it possible to express the flux Ψ_{peak} at the peak of the F2 layer as a function of the concentration N_m and the collision frequency ν_m at that point:

$$\Psi_{peak} = N_m \left[u + \frac{g}{\nu_m}\sin^2 I\right]$$

Under steady-state conditions, this flux must equal the excess of production over loss integrated up to the peak. This is an important condition in determining the height of the peak. During the night, for example, Ψ_{peak} must be negative, because of the absence of ionization

sources in the ionosphere. Consequently:

$$\nu_m < -\frac{g}{u} \sin^2 I$$

which places a lower boundary for the location of the nighttime peak of the F2-layer. It is of interest to notice the control exerted by the externally imposed vertical drift u on this condition.

(b) Below the peak, $H_p < 0$, and therefore:

$$v - u < \frac{g}{\nu} \sin^2 I$$

In other words, the plasma falls faster than it would under the exclusive control of gravity and collisions, because the pressure gradient exerts an additional downward force.

(c) Above the peak, since $H_p > 0$:

$$v - u > \frac{g}{\nu} \sin^2 I$$

which is explained by the upward force exerted on the plasma by its own pressure gradient.

Figure 1 depicts schematically the different equilibrium conditions normally existing at various altitudes in the ionosphere.

Owing to the fact that the diffusion constant D increases exponentially with height, Equation (13) suggests that at sufficiently high altitudes, the O^+ concentration falls off with a scale height $2H = -\frac{2kT}{mg}$ characteristic of a collisionless ($D = \infty$) situation. This is, in fact what frequently happens, except in the special case in which the vertical plasma velocity v grows so rapidly that the ratio v/D does not tend to zero with increasing heights.

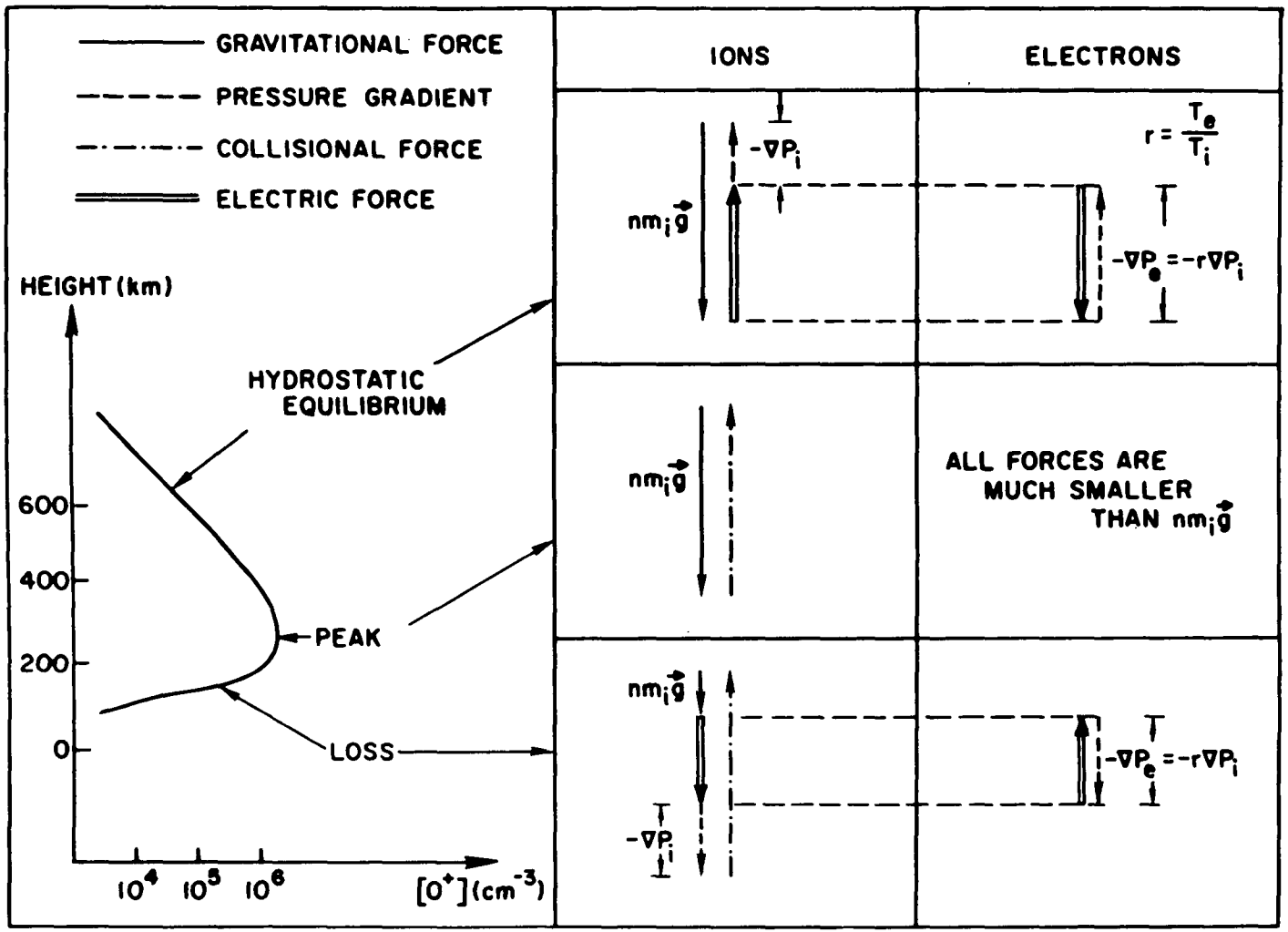


Fig. 1. EQUILIBRIUM CONDITIONS IN DIFFERENT REGIONS OF THE IONOSPHERE. The above relation between electronic and ionic pressure gradients is exact in the isothermal case, but is also a good approximation in the real ionosphere.

Rewriting Equation (13) for $|v| \gg |u|$:

$$\frac{1}{H_p} = \frac{-m}{2kT}\left(g - \frac{v\upsilon}{\sin^2 I}\right) \tag{13a}$$

This is just the expression for the plasma scale height in the collisionless case, with g replaced by $\left(g - \frac{v\upsilon}{\sin^2 I}\right)$. Clearly, when $|v\upsilon| \gg |g| \cdot |\sin^2 I|$, the plasma scale height will depart very little from its value under collisionless conditions.

If one assumes a steady-state condition, absence of production and loss processes ($Q = L = 0$), and negligible divergence in the magnetic lines, the flux nv must be height-independent, which means that v has to vary as $\exp\left(\frac{h}{H_p}\right)$. The collision frequency will vary with height with the scale height H_f of the species that is producing the friction. Therefore, the product $v\upsilon$ varies as $\exp\left[\left(\frac{1}{H_p} - \frac{1}{H_f}\right) h\right]$.

If $H_p > H_f$, $v\upsilon$ will tend to zero with increasing altitudes, and consequently H_p will approach 2H, according with Equation (13a). Therefore, if $H_p > H_f$, H_p has to approach 2H at high altitudes. However, if $H_p = H_f$ this need not happen for the product $v\upsilon$ will then keep constant over all altitudes. The velocity needed for this to occur may be derived by equating H_f to the plasma scale height as given by Eq. (13a).

$$H_f^{-1} = -\frac{m}{2kT} \cdot \left(g - \frac{v\upsilon}{\sin^2 I}\right)$$

$$\therefore \quad v = \frac{g \sin^2 I}{\upsilon} + \frac{2kT \sin^2 I}{m\upsilon H_f} = \frac{g \sin^2 I}{\upsilon} + \frac{D}{H_f} \equiv v_L$$

This is clearly a limiting value, since any velocity smaller than v_L would result in $H_p > H_f$, thus making the collisional force negligible and $H_p \cong 2H$ at high altitudes. On the other hand, if $v = v_L$, the collisional force will be height-independent and the plasma will be distributed with the scale height H_f of the friction-producing medium. In this case, the flow of plasma is being limited by collisions, since the produce $v\upsilon$ does not fall off with increasing height.

Thus, it is clear that the plasma scale height may approach a value other than 2H with increasing heights. By multiplying Eq. (13a) by n and rearranging, one obtains an equation which shows the existence of only two possible values that can be approached by H_p:

$$\frac{\partial n}{\partial h} - \frac{mg}{2kT} n = - \frac{mnv}{2kT \cdot \sin^2 I} \cdot \upsilon$$

Under the conditions that normally prevail at the topside of the F-layer (steady-state, $Q = L = 0$), the flux nv is constant, and therefore the equation above is an inhomogeneous linear first-order differential equation. The homogeneous solution is given by $n_1 \cdot \exp\left(\frac{mgh}{2kT}\right)$, corresponding to the normal collisionless ($\upsilon = 0$) or hydrostatic ($v = 0$) profile. Since the independent term is proportional to υ, the particular solution is an exponential with scale height H_f; this solution is called a full flow mode [Banks et al., 1969]. The general solution will be:

$$n = n_1 \cdot e^{-h/2H} + n_2 \cdot e^{-h/H_f} \qquad (14)$$

where n_1 and n_2 are arbitrary constants.

In the case of the 0^+ ions in the F region, the main colliding species is the monoatonic oxygen for which the scale height is $H_f = -\frac{kT_n}{mg} = \frac{T_n}{2T} \cdot 2H$, where T_n is the temperature of the neutral species. Since T_n is always smaller than 2T, it is clear that 2H is always larger than H_f. For this reason, the mode $n_1 \cdot \exp(-\frac{h}{2H})$ will eventually dominate at a sufficiently high altitude, unless $n_1 = 0$. Consequently, n_1 can never be negative, for the ionic concentration is obviously a positive quantity.

By applying this concentration profile to Eq. (9), one can derive the corresponding velocity profile, which for $u = \frac{\partial T}{\partial h} = 0$ is:

$$v = \frac{n_2}{n} \cdot [\frac{g}{\nu} \sin^2 I + \frac{D}{H_f}] \cdot e^{-h/H_f}$$

Because n_1 cannot be negative, the ratio between n_2 and n cannot exceed $\exp(\frac{h}{H_f})$ at any point. A consequence of this is that a maximum possible value exists for the velocity at each point, given by:

$$v_L = \frac{g}{\nu} \sin^2 I + \frac{D}{H_f} \qquad (15)$$

As might be expected, the limiting velocity is the characteristic velocity of the full flow mode. A pure full flow mode will then appear whenever the ions are escaping upwards at the maximum possible rate.

Specializing Eq. (15) for the case of 0^+ ions diffusing through neutral monoatomic oxygen particles:

$$v_L = \frac{g}{\nu} \sin^2 I \cdot (1 - \frac{2T}{T_n})$$

Since this velocity increases rapidly with height, it may reach a point where it exceeds the thermal velocity of the ions. At this point, the inertia term $nmv \frac{\partial v}{\partial h}$ is no longer negligible when compared to the pressure gradient, and the approximations leading to these equations are no longer valid. Since the plasma velocity will then be limited by its inertia as well as collisions, the limiting velocity as given by Eq. (15) will then be an overestimate. The limiting velocity under these conditions is difficult to calculate because the steady-state differential equation for n is then non-linear, and the concentration profile cannot be thought of as a superposition of independent modes. The thermal velocity of the 0^+ ions in the ionosphere is about 10^3 m/sec, so that for $T = 1.5T_n$ this transition will occur when $\nu \cong 2.10^{-2}$ sec^{-1}, i.e., around 600 km. Therefore, above this height a pure full flow mode for the 0^+ ions is impossible; a superposition of two modes as in Eq. (14) may still exist, as long as the ion velocity is small enough for the low-speed condition ($v < v_T$) to prevail.

However, this distinction is seldom important because the 0^+ ion velocities are kept well below the limiting value by the relatively small upward flux that can be supported by the H^+ ions immediately above the region of chemical equilibrium between H^+ and 0^+ [Hanson et al., 1961; Geisler, 1967; Banks et al., 1969]. The question of the limiting H^+ flux is also related to the full flow mode concept, but applied to a minor ion; this point will be covered later in this study.

Another restriction to the existence of a pure full flow mode arises at high altitudes for the 0^+ ion, and it is due to the decreasing concentration of the friction-producing species. As this concentration decreases, the mean free path for collisions increases, and a point is

reached where it becomes longer than the scale height of the friction-producing medium. At this point, the inhomogeneity of the background species cannot be overlooked any more in the collision mechanism, and the very concept of a local collison frequency loses its significance. Since $\nu = \frac{v_T}{\lambda}$, where λ is the mean free path, this condition will arise for $\nu < \frac{v_T}{H_f} \cong 2 \cdot 10^{-2}$ sec^{-1}, considering monatomic oxygen as the only colliding species with H_f = 50 km. For the same reasons as stated above, this restriction is also of little practical significance.

In middle latitudes, the magnetic lines that cut across the ionosphere remain closed all the time and intersect the ionosphere at a conjugate point in the opposing hemisphere. For this reason, large outflows from the ionosphere cannot ordinarily be accomodated, and the full flow mode is then seldom large enough to have a noticeable effect on the scale height of the O^+ ion well above the peak. In the polar ionosphere, however, this is not necessarily true, because the magnetic lines will then be open for some time during each day; this effect is discussed in detail by Banks et al. [1969].

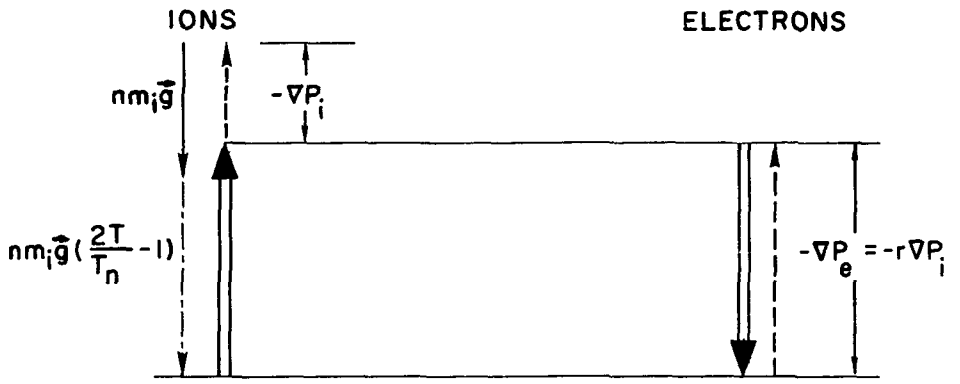


Fig. 1a. BALANCE OF FORCES UNDER LIMITING VELOCITY CONDITIONS.

C. The Minor Ions

As mentioned above, Eq. (8) describes the balance of forces acting on a cell of any ionic species, whatever its relative abundance in the mixture. This equation contains a term that depends on $\frac{1}{n_e} \frac{\partial}{\partial h} (n_e T_e)$, where n_e is the electron concentration. Because macroscopic charge neutrality has to be assumed (otherwise very large electric fields would soon force this condition to be restored), n_e is the sum of the concentrations of all ions present in the ionosphere. Therefore, the n_e-dependent term in the transport equation of each ion effectively couples it to the corresponding equations for all other ions, complicating their solution. In the case of the major ion, this difficulty is avoided by replacing n_e with n_i in the equation, thereby ignoring the marginal control on electron concentration that is exercised by the minor ions.

In the case of the minor ions, another approach has to be used. When the major ion concentration overwhelmingly exceeds the sum of all other ionic concentrations, as is the case in many important circumstances, one may assume $n_e \cong \bar{n}_i$, where $\bar{n}_i$ is the concentration of the major ion. One may then solve first for $\bar{n}_i$, and use the resulting concentration profile to write an uncoupled transport equation for each minor ion. Since the profile of $\bar{n}_i$ can take many different shapes, there is no general treatment that could describe the dynamic behavior of a minor ion.

There is, however, a more specialized treatment that can be given to the minor ion whenever the distribution of the major ion can be assumed to approach that of hydrostatic equilibrium. This treatment will be relevant in the case of H^+ ions in a majority of O^+ ions above the peak of the F-layer, as well as O^+ ions in the protonosphere.

Under this assumption:

$$\frac{1}{\bar{n}_i T_i} \frac{\partial}{\partial h} (\bar{n}_i T_i) = \frac{\bar{m}_i g}{2kT}$$

where $\bar{m}_i$ is the mass of the major ion.

Assuming further that the ratio between electron and ion temperatures is constant with height:

$$\frac{1}{n_e} \frac{\partial}{\partial h} (n_e T_e) = \frac{T_e}{2T} \frac{\bar{m}_i g}{k} = \frac{\bar{m}_i}{1 + \frac{T_i}{T_e}} \frac{g}{k}$$

Substituting in Eq. (8), one has for the minor ion:

$$n_i m_i \nu_i v_i = n_i m_i \nu_i u_i + n_i m_i g \sin^2 I - k \sin^2 I \frac{\partial}{\partial h} (n_i T_i) - \frac{n_i \bar{m}_i g}{1 + \frac{T_i}{T_e}} \sin^2 I$$

Dropping some of the indexes and rearranging:

$$v = u + \left[1 - \frac{\frac{\bar{m}_i}{m_i}}{1 + \frac{T_i}{T_e}}\right] \frac{g}{\nu} \sin^2 I - D_i \left[\frac{1}{n} \frac{\partial n}{\partial h} + \frac{1}{T_i} \frac{\partial T_i}{\partial h}\right]$$

where: $D_i = \frac{kT_i}{m_i \nu} \sin^2 I$ = ionic diffusion coefficient.

Comparison of this relation with Eq. (9), which is the corresponding equation for the major ion, brings to light the existence of two differences in behavior. Firstly, in the case of the minor ion the diffusion coefficient depends on T_i instead of $(T_e + T_i) = 2T$. This merely reflects the fact that the diffusion of the minor ions is not ambipolar, for unlike the major ion they are not constrained to follow the same distribution as the electrons.

The second difference is that the minor ions behave as if they were acted upon by a gravitational field given by

$$g' = g \left[1 - \frac{\frac{\bar{m}_i}{m_i}}{1 + \frac{T_i}{T_e}} \right]$$

The modified gravitational field is actually due to the presence of the electric field that is set up to make the electrons and the ions of the majority species follow the same distribution, thus preserving neutrality of the plasma. This electric field may be derived from Eq. (7b):

$$e E_z \sin I = - \frac{k}{n_e} \frac{\partial}{\partial h} (n_e T_e) \sin^2 I =$$

$$= - kT_e \frac{\bar{m}_i g}{2kT} \sin^2 I =$$

$$= - \frac{\bar{m}_i}{1 + \frac{T_i}{T_e}} g \sin^2 I$$

The sum of the vertical electric and gravitational forces on a minor ion of mass m_i will be:

$$m_i g \sin^2 I + e E_z \sin I = m_i \left[1 - \frac{\frac{\bar{m}_i}{m_i}}{1 + \frac{T_i}{T_e}} \right] g \sin^2 I =$$

$$= m_i g' \sin^2 I$$

Therefore, the modified gravitational field g' effectively simulates the combined effects of the real electric and gravitational fields.

It is clear that Eq. (9) can be applied for both major and minor ions under the present assumptions, provided that in the latter case g be replaced by g' and 2T by T_i. The discussion of the behavior of the major ion presented in the previous subsections is thus pertinent to the case of a minor ion. The concept of the full flow mode, which was developed for the major ion, is just as applicable for a minor one.

Assuming Q = L = u = 0, the steady-state concentration profile will have the same form as in Eq. (14), provided the appropriate replacements are made for g and T:

$$n = n_1 e^{\frac{m_i g' h}{kT_i}} + n_2 e^{-\frac{h}{H_f}}$$

Specializing this discussion for the important case of protons diffusing through O^+ ions in the topside of the F-layer, one has:

$$g' = \left[1 - \frac{16}{1 + \frac{T_i}{T_e}}\right] g$$

At high altitudes, the thermal coupling between ions and electrons is much stronger than the coupling between ions and the neutrals because of the low concentrations of the latter. Under these conditions, $T_e = T_i$ is a good assumption, and therefore g' = -7g. The scale height of the hydrostatic equilibrium mode will be $\frac{kT_i}{7m_i g}$; since this is a negative quantity, this mode will correspond to concentrations rapidly increasing with height. On the other hand, the full flow mode is characterized by

the scale height of the friction-producing species, which in this case is the 0^+ ion. Therefore, $H_f = \frac{-2kT}{16m_i g}$ and the hydrostatic equilibrium mode will always dominate the concentration profile unless n_1 is zero, i.e., when the protons are escaping with the maximum possible velocity. The limiting velocity may be derived from Eq.(15), just by replacing g by g' and D by D_i:

$$v_L = \frac{g'}{\nu} \sin^2 I + \frac{D_i}{H_f} =$$

$$= - \frac{7g}{\nu} \sin^2 I + \frac{kT_i}{m_i \nu} \sin^2 I \left(- \frac{16 m_i g}{2kT_i} \right) =$$

$$= - 15 \frac{g}{\nu} \sin^2 I \qquad (16)$$

The H^+ ions, when controlled by diffusion as a minor ion in the middle of a majority of 0^+ ions, may assume two different scale heights. When diffusing downwards, or when diffusing upwards with velocities lower than v_L, the concentration will increase with height. If, however, the protons are escaping with the limiting velocity v_L, their concentration will have the same scale height as the 0^+ ions, therefore, decreasing with height. This fact is important, since it permits one to identify a situation in which the protons are escaping at the maximum possible velocity just by inspection of the ionic concentration profile.

D. The Concentration Equation

By combining Eq. (9) with the continuity equation, one can eliminate the velocity variable v, and obtain a second-order partial differential equation for n(h, t).

The continuity equation is:

$$\frac{\partial n}{\partial t} = Q - L - \frac{\partial}{\partial h}(nv)$$

Substituting v as given by Eq. (9), and assuming $\frac{\partial u}{\partial t} = 0$:

$$\frac{\partial n}{\partial t} = Q - L - u\frac{\partial n}{\partial h} - \frac{\partial}{\partial h}\left\{D\left[\frac{nmg}{2kT} - \frac{\partial n}{\partial h} - \frac{n}{T}\frac{\partial T}{\partial h}\right]\right\}$$

After expanding the last derivative, applying the definition of D, rearranging and grouping the terms:

$$\frac{\partial n}{\partial t} = Q - L - u\frac{\partial n}{\partial h} + D\left[\frac{\partial^2 n}{\partial h^2} + A\frac{\partial n}{\partial h} + Bn\right] \qquad (17)$$

where:

$$A = \frac{\nu}{T}\frac{\partial}{\partial h}\left(\frac{T}{\nu}\right) - \frac{mg}{2kT} + \frac{1}{T}\frac{\partial T}{\partial h}$$

$$B = \frac{\nu}{T}\frac{\partial}{\partial h}\left(\frac{T}{\nu}\right)\left[-\frac{mg}{2kT} + \frac{1}{T}\frac{\partial T}{\partial h}\right] - \frac{m}{2kT}\frac{\partial g}{\partial h} + \frac{mg}{2kT^2}\frac{\partial T}{\partial h} - \frac{1}{T^2}\left(\frac{\partial T}{\partial h}\right)^2 + \frac{1}{T}\frac{\partial^2 T}{\partial h^2}$$

In this form, this equation is valid for the major ion of the mixture. However, under the conditions specified in the previous section and with the substitutions that were derived there, it can also be applied to a minor ion.

The complexity of Eq. (17), and in particular the dependence of the temperature T on the plasma concentration preclude any attempt to solve this equation analytically unless drastic simplifying assumptions are made. In the next chapter, a description is given of an ionospheric simulation program that solves Eq. (17) given an initial profile n(h, 0) and two boundary conditions at any time. All the functions that determine the coefficients of the equation are generated from given models, except the temperatures T_e and T_i, upon which the coefficients A and B depend. These temperatures are obtained through simultaneous solution of the heat equation for ions and electrons.

The production function for O^+ ions is given by the rate of photoionization of oxygen atoms at each point. This rate is calculated by assuming an incident solar ultraviolet flux, dividing this radiation into several bands and taking into account the absorption of each band by the atmosphere between the point in question and the sun:

$$q_O(h) = [O](h) \sum_{i=1}^{m} \sigma_{Oi}^{(i)} \Phi_{\infty i} \exp\left\{ - \int_h^{\infty} (\sigma_{Oi}^{(a)}[O] + \sigma_{O_2 i}^{(a)}[O_2] + \sigma_{N_2 i}^{(a)}[N_2]) \sec \chi \, dh \right\}$$

where:

m = number of bands in which the solar EUV radiation is divided

$\Phi_{\omega i}$ = exospheric photon flux at band i

$\sigma_{M_i}^{(i)}$, $\sigma_{M_i}^{(a)}$ = ionization and absorption cross-sections of species M in the band i

The photoionization rates of N_2 and O_2 molecules are calculated in the same way, and the total electron production is given by $q_e = q_O + q_{O_2} + q_{N_2}$. Figure 2 presents a profile of the production functions calculated for 1100 LT at approximately equinoctial conditions. In this calculation, the marginal contribution of ionization by collisions between the neutral particles and very energetic photoelectrons (secondary ionization) is also taken into account.

The O^+ ions are lost through the following reactions.

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

$$O^+ + O_2 \rightarrow O_2^+ + O$$

The loss rate of O^+ ions may then be expressed as $L = \beta[O^+]$, where $\beta = \gamma_{N_2}[N_2] + \gamma_{O_2}[O_2]$, γ_{N_2} and γ_{O_2} being the rates of the reactions described above. Because the loss rate is proportional to $[O^+]$, this type of loss is called a linear loss. The linear loss coefficient β decreases exponentially with height, following the scale height of the N_2 or O_2 molecules.

When applied to O^+ ions in the F-layer, Eq. (17) becomes

$$\frac{\partial}{\partial t}[O^+] = q_O - \beta[O^+] - u\frac{\partial}{\partial h}[O^+] + D\left\{\frac{\partial^2}{\partial h^2}[O^+] + A\frac{\partial}{\partial h}[O^+] + B[O^+]\right\}$$

In the isothermal case, $\frac{\partial T}{\partial h} = 0$, and assuming that $\frac{\partial g}{\partial h} = 0$, one has

$$A = -\frac{1}{\nu}\frac{\partial \nu}{\partial h} - \frac{mg}{2kT} = \frac{1}{H_f} + \frac{1}{2H}$$

$$B = -\frac{1}{\nu}\frac{\partial \nu}{\partial h}\left(\frac{-mg}{2kT}\right) = \frac{1}{H_f}\frac{1}{2H}$$

Under these conditions and in the absence of neutral winds or external electric fields (u = 0), the steady-state equation for the 0^+ ions will be:

$$q_0 - \beta[0^+] + D \left\{ \frac{\partial^2[0^+]}{\partial h^2} + \left(\frac{1}{H_f} + \frac{1}{2H} \right) \frac{\partial}{\partial h} [0^+] + \frac{1}{H_f} \frac{1}{2H} [0^+] \right\} = 0$$

Being inversely proportional to ν, the diffusion coefficient D increases exponentially with height. Below about 200 km, D is small enough for $[0^+]$ to be given by photochemical equilibrium, i.e., $[0^+] = \frac{q_0}{\beta}$.

At high altitudes, on the other hand, both q_0 and β become very small, so that:

$$\frac{\partial^2}{\partial h^2} [0^+] + \left(\frac{1}{H_f} + \frac{1}{2H} \right) \frac{\partial}{\partial h} [0^+] + \frac{1}{H_f} \frac{1}{2H} [0^+] = 0$$

and the profile will be given by Eq. (14).

Between 200 and 400 km, a transition between loss and diffusion-dominated conditions occurs. The peak of the layer is normally located in this region.

E. The Heat Equations

The coefficients of Eq. (17) (from which the concentrations are calculated) depend on the ion and electron temperatures. Therefore, if the dynamic behavior of the ionospheric plasma is to be simulated, one must either adopt arbitrary models for the temperatures or solve the heat equations for ions and electrons. Since these last equations depend on the electron concentration, a system of coupled partial differential equations has to be solved.

The heat equations for electrons and ions can be obtained from Boltzmann's equation as:

$$\frac{\partial}{\partial t}\left(\frac{3}{2}\,nkT\right) = \frac{\partial_c}{\partial t}\left(\frac{3}{2}\,nkT\right) - \nabla\cdot\left(\frac{3}{2}\,nk\vec{V}\right) + \nabla\cdot\left(B_c T^{5/2}\nabla T\right) - nkT\left(\nabla\cdot\vec{V}\right)$$

where n, T and $\vec{V}$ are the concentration, temperature and bulk velocity of the species (electrons or ions), k is the Boltzmann's constant, $B_c T^{5/2}$ is the conductivity of a gas of charged particles [Spitzer, 1962] and $\frac{\partial_c}{\partial t}$ indicates time variation due to interaction of the species with external sources or sinks of energy (production and loss of particles, thermal contact with other species, or absorption and emission of radiative energy).

Since $\frac{3}{2}kT$ is the average thermal energy of the particles, $\frac{3}{2}nkT$ is the thermal energy contained in a unit volume of the gas. Bearing this fact in mind, one can make an intuitive interpretation of the terms of the heat equation as written above.

The left-hand side of the equation is the time variation of the thermal energy contained in a unit volume fixed in space. On the right-hand side, each term represents one or more physical mechanisms that contribute to this time variation. The first term is the time variation due to interaction with external sources, and it may be represented as:

$$\frac{\partial_c}{\partial t}\left(\frac{3}{2}\,nkT\right) = \frac{\partial_c n}{\partial t}\,\frac{3}{2}\,kT + \frac{3}{2}\,kn\,\frac{\partial_c T}{\partial t} = (Q - L)\,\frac{3}{2}\,kT + \bar{Q} - \bar{L}$$

where Q and L are the rates of production and loss of the species under consideration (as discussed in the previous sections), $\bar{Q}$ is the heat

energy input and $\bar{L}$ the energy loss due to collisions with particles of other species.

The second term represents the contribution due to the divergence of the heat flux caused by the flow of material. The third term expresses the divergence of the heat flux due to conduction caused by a temperature gradient. The last term is due to the power dissipated in compression or released in the expansion of the gas.

Expanding some derivatives in the heat equation:

$$\frac{\partial n}{\partial t}\frac{3}{2}kT + \frac{3}{2}kn\frac{\partial T}{\partial t} = (Q - L)\frac{3}{2}kT + \bar{Q} - \bar{L} - \frac{3}{2}kT\ \nabla.(n\vec{V}) -$$

$$- \frac{3}{2}kn\vec{V}.(\nabla T) + \nabla.(B_c T^{5/2}\nabla T) - nkT(\nabla.\vec{V})$$

From the continuity equation, one has:

$$\frac{\partial n}{\partial t} = Q - L - \nabla.(n\vec{V})$$

Substituting into the expanded heat equation, and dividing by $\frac{3}{2}kn$:

$$\frac{\partial T}{\partial t} = \frac{\bar{Q} - \bar{L}}{\frac{3}{2}kn} + \frac{\nabla.(B_c T^{5/2}\nabla T)}{\frac{3}{2}kn} - \vec{V}.(\nabla T) - \frac{2}{3}T(\nabla.\vec{V})$$

and under the assumption of horizontal stratification:

$$\frac{\partial T}{\partial t} = \frac{\bar{Q} - \bar{L}}{\frac{3}{2}kn} + \frac{1}{\frac{3}{2}kn}\frac{\partial}{\partial h}\left[\frac{B_c}{\sin^2 I}T^{5/2}\frac{\partial T}{\partial h}\right] - v\frac{\partial T}{\partial h} - \frac{2}{3}T\frac{\partial v}{\partial h} \qquad (19)$$

This is the equation to be used for ions and electrons in the simulation program discussed on the next chapter. Some drastic simplifications will be made for the case of the ions, but for the electrons all terms will be retained in the equation.

It is instructive at this point to investigate the relative importance of each term in Equation (19). One way of doing this is to isolate selected terms in the equation, and derive the temperature behavior that would result from ignoring all other terms. In the following discussion, this approach will be used to study three hypothetical situations: local thermal equilibrium, purely conductive flow and the adiabatic case. Finally, an appraisal is given of how much the real ionosphere departs from these idealized situations.

Local Thermal Equilibrium. Historically, the problem of computing ionospheric plasma temperatures was first tackled by ignoring all terms other than the ones due to scattering of energy into ($\bar{Q}$) or out of ($\bar{L}$) the species under consideration. This greatly simplifies the mathematical treatment, because it frees Equation (19) of all spatial derivatives. This approach, by ignoring the important effects of heat conduction in the electron gas, often leads to completely unrealistic results, as shown further on.

The major source of energy during the day is the photoionization of neutrals. When a photon is absorbed in this process, part of its energy is used in the ionization of the particle. Because of total momentum conservation, the remaining energy is almost entirely transformed into kinetic energy of the ejected electron (photoelectron), while the resulting ion stays with almost the same kinetic energy as

the original neutral particle. However, the heat input from the photoelectrons to the ambient electron gas is not equal to the rate of generation of photoelectron energy at each point because two complications arise:

a) particularly at low altitudes (below 300 km) part of the energy of the photoelectrons is lost in inelastic collisions with neutral particles, and only the remaining part is transferred to the ambient electrons through elastic Coulomb collisions. Some of these collisions will produce secondary ionization, so that some of the energy lost by the primary photoelectron may still revert to the electron gas. However, most inelastic collisions will simply excite the neutral particle without ionizing it. The excited particles, in returning to their ground state, will emit photons in the optical and infrared, which have too little energy to produce further ionization;

b) at high altitudes (above 300 km) the mean free path for energy loss of the photoelectrons is so long that they are not thermalized locally; their suprathermal energy is lost over an extended trajectory, and some of them even escape the ionosphere before relaxing to the thermal energies [Geisler and Bowhill, 1965].

The handling of these complications in a numerical simulation program is described in the next chapter; for the purpose of the present discussion, $\bar{Q}$ is assumed to be a known function of height. Figure 2

shows the electron production q_e and the electron heat input $\bar{Q}$ as a function of height, calculated for 1100 LT at solar cycle maximum. Around 350 km, the ratio $(\frac{\bar{Q}}{q_e})$ is about 10 eV, which is the average energy of the newly created photoelectrons. Therefore, at this height the effects of inelastic collisions and non-local heating compensate each other. Above this height, non-local heating enhances the ration $\frac{\bar{Q}}{q_e}$, while below it the ratio falls due to increasingly frequent inelastic collisions between neutrals and photoelectrons.

The energy gained by the electron gas from the photoionization of neutral particles will keep it at a substantially higher temperature than the neutrals. Consequently, the ions will be heated by the electron gas and cooled by the neutrals, thus remaining at an intermediate temperature between these two species. For the sake of simplicity, let us now consider only the 0^+ ion in the presence of monoatomic oxygen; in the simulation program, the contributions of other species are obtained just by summing the transfer rates over all ionic and neutral species.

The rate of energy transfer from electrons to ions is given by [Banks, 1966]:

$$\frac{L_{ei}}{\frac{3}{2}kn} = 3.7 \times 10^{-3}\, n\, \frac{(T_e - T_i)}{T_e^{3/2}} \quad °K/sec$$

The rate of energy transfer from 0^+ ions to monoatomic oxygen is given by:

$$\frac{L_{in}}{\frac{3}{2}kn} = 7.7 \times 10^{-10}\, [0]\, (T_i - T_n)\ °K/sec$$

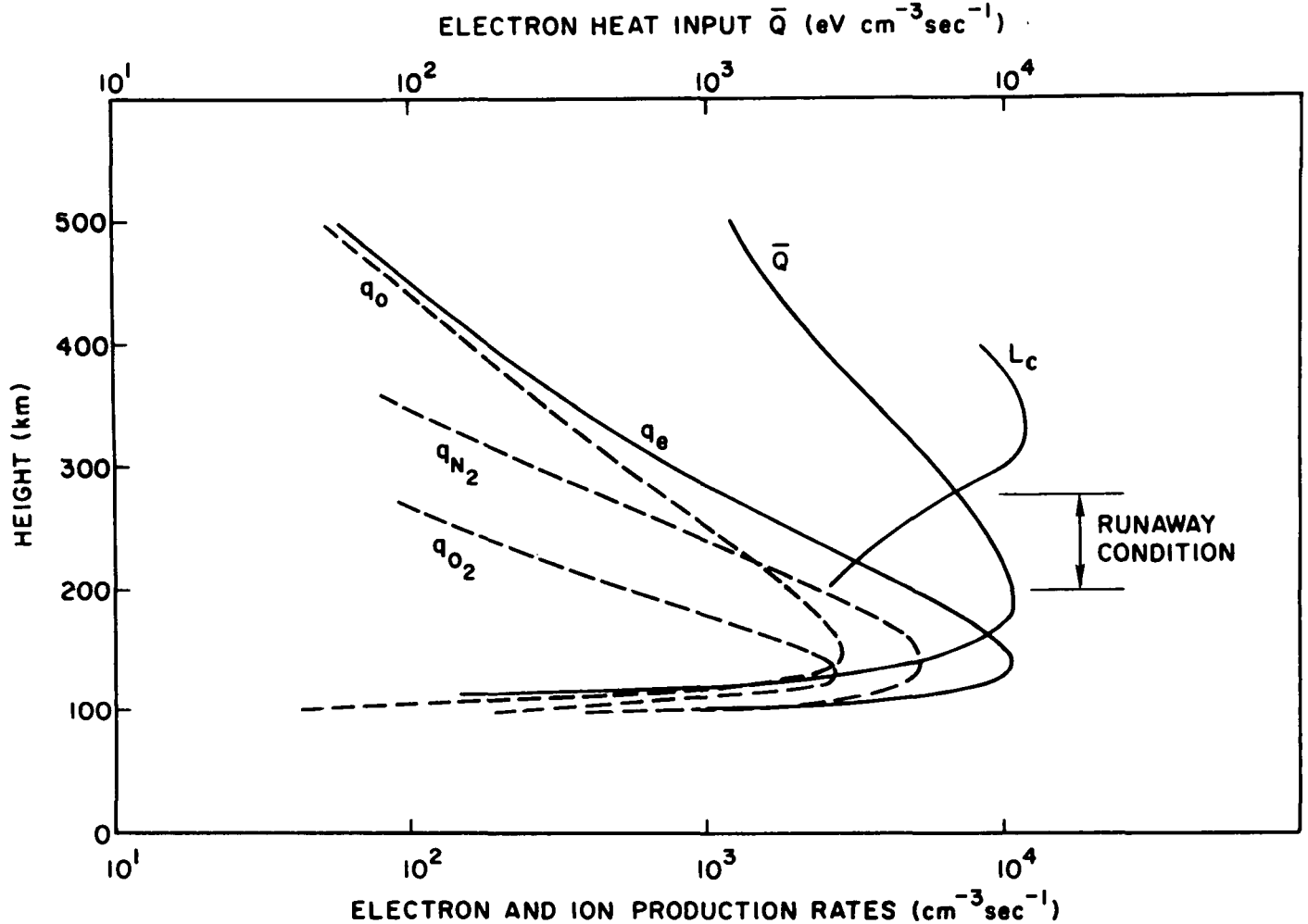


Fig. 2. CALCULATED PROFILES OF PRODUCTION RATES. The profiles shown above were calculated with the ionospheric simulation program for 1100 LT under midlatitude, solar-cycle maximum conditions. When $\bar{Q} > L_c$, the ions cannot dissipate the heat gained by the electrons (runaway condition); the resulting peak in T_e is shown on Figure 3.

assuming $T_i + T_n \cong 2500°K$ in the expression given by Banks [1966] which has a slow dependence on $(T_i + T_n)$.

The steady-state ionic temperature will be given by $L_{ei} = L_{in}$, or:

$$X = \frac{T_i - T_n}{T_e - T_i} = 4.8 \times 10^6 \frac{n}{[O]\, T_e^{3/2}}$$

When T_i is closer to T_e than to T_n, the parameter X will be larger than unity; as T_i approaches T_e, X will approach infinity. Conversely, when T_i approaches T_n, X will approach zero. At 300 km, typical daytime values are $n = 10^6$ cm^{-3}, $[O] = 4.5 \times 10^8$ cm^{-3} and $T_e = 2000°K$, so that $X(300\text{ km}) \cong 0.005$. Therefore, at this height the ionic temperature is very close to T_n. With increasing altitudes, however, the ratio $n/[O]$ increases rapidly. At about 550 km, X approaches unity, which means that T_i is about halfway between T_e and T_n. Above this height, T_i continues to approach T_e. Figure 3 illustrates this transition between neutral and electronic control of the ionic temperature.

The electron temperature will be determined by the heat input $\bar{Q}$ and the rates of energy transfer L_{ei} to the ions and L_{en} to the neutral particles, which is given by [Dalgarno et al., 1963]:

$$\frac{L_{en}}{\frac{3}{2}kn} = 10^{-14}\, T_e^{1/2}\, [O]\, (T_e - T_n)\ °K/\text{sec}$$

The steady-state electron temperature will be given by $\bar{Q} = L_{ei} + L_{en}$, or:

$$\frac{\bar{Q}}{\frac{3}{2}kn} = 3.7 \times 10^{-3}\, n\, \frac{(T_e - T_i)}{T_e^{3/2}} + 10^{-14}\, T_e^{1/2}\, [O]\, (T_e - T_n)$$

From the definition of X, it follows that $(T_e - T_i) = (T_e - T_n)/(1 + X)$, and therefore:

$$\frac{\bar{Q}}{\frac{3}{2}kn} = \left[3.7 \times 10^{-3} \frac{n}{(1+X)T_e^{3/2}} + 10^{-14}\, T_e^{1/2}\, [O]\right]\, (T_e - T_n)$$

During the night, $\bar{Q} = 0$ and therefore $T_e = T_n$ in steady-state. The following discussion is pertinent to the daytime ionosphere.

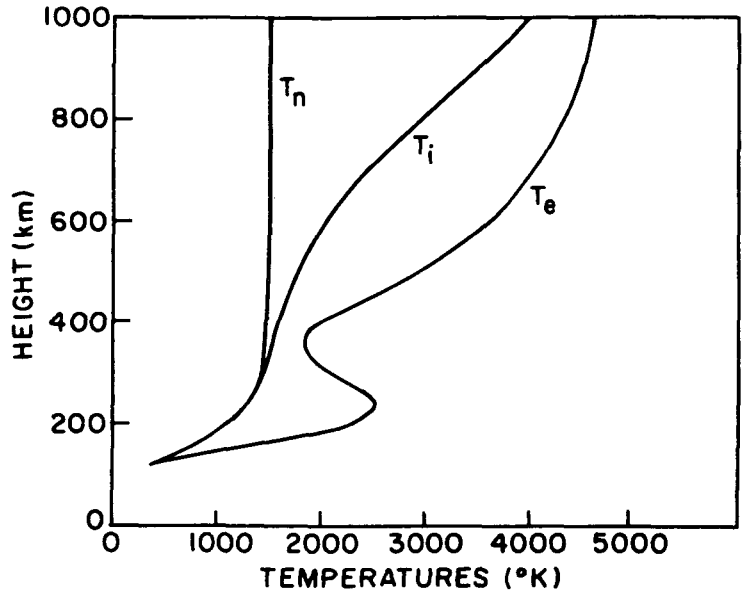
At low altitudes (below 550 km), $X \ll 1$, and therefore one can make the approximation $1 + X \cong 1$, with the following result:

$$\frac{\bar{Q}}{\frac{3}{2}kn} = \left[\, 3.7 \times 10^{-3} \frac{n}{T_e^{3/2}} + 10^{-14}\, T_e^{1/2}\, [O]\right](T_e - T_n)$$

The first term in the brackets is due to thermal coupling with the ions, and the second to coupling with the neutral particles. Figure 3 shows the relative importance of each term as a function of T_e for the values of n, [O] and T_n corresponding to a selected point. Naturally, the magnitude of each term will depend very strongly on height, because of the strong vertical gradients of n and [O].

It is interesting to notice that the energy coupling between ions and electrons has a maximum in T_e at a temperature T_{em} given by:

$$\frac{\partial}{\partial T_e} \quad \frac{(T_e - T_n)}{T_e^{3/2}} \quad = 0$$



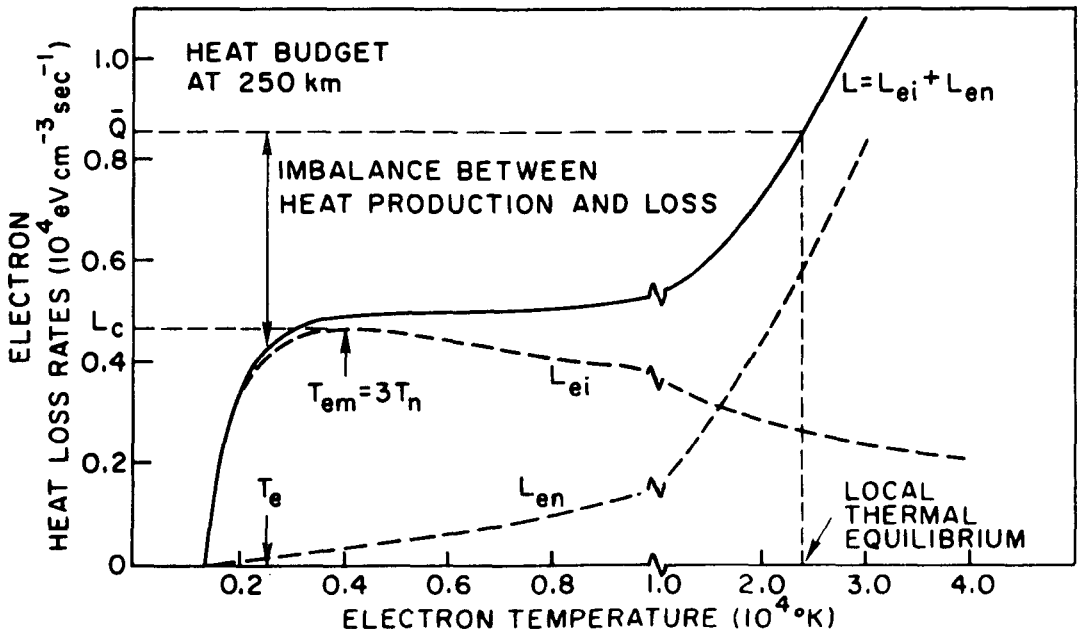


Fig. 3. CALCULATED TEMPERATURE PROFILES. The upper boundary condition for the heat equation was a downward heat conduction flux of 5 x 10^9 eV cm^{-2} sec^{-1} at 1000 km. At 250 km, local thermal equilibrium would give T_e = 24000°K, while the temperature calculated with conduction is only 2500°K.

$$\therefore \quad T_{em} = 3T_n$$

At low altitudes (e.g., below 180 km), where the neutral concentration is high, the electron-neutral coupling term is strong enough to mask this maximum in the loss function, so that the total electron heat loss will be a monotonically increasing function of T_e. Above 200 km, however, the electron-ion loss term will generally dominate, provided that T_e is as low as the electron temperatures normally found in the ionosphere. At these heights, the total loss function has both a maximum, and a minimum, as shown on Figure 3. The heat loss corresponding to $T_e = T_{em}$ is given by the following critical value:

$$L_c = 1.83 \times 10^{-7} \frac{n^2}{T_n^{1/2}} \text{ eV/cm}^3 \times \text{sec}$$

If the heat input $\bar{Q}$ is less than L_c, thermal equilibrium is possible at an electron temperature lower than $T_{em} = 3T_n$. If, however, $\bar{Q}$ exceeds this critical value, there is no temperature at which the ionic loss alone can balance the heat input from the photoelectrons to the electron gas. In the absence of other cooling processes, the electron temperature would then have to rise until the excess heat input can be dissipated by thermal coupling to the neutral gas. Since this situation would cause extremely high temperatures, it is called a runaway temperature condition [Hanson et al., 1961]. In a hypothetical conductionless ionosphere such a runaway condition might occur in a height interval between, say, 200 and 300 km. On the other hand, depending on the value of heat input and the amount of ionization present, the condition may never occur at all. However, even when it does occur, the high temperatures

implied by the above mechanism can never be reached in the real ionosphere because they would give rise to a strong conductive heat flow that would promptly remove from the region in question the excess heat that the ions are unable to dissipate. If the electron temperature in this region stays at about $T_{em} = 3T_n$, the amount of heat generation that cannot be dissipated locally is about $\bar{Q} - L_c$, which can be of the order of 100 eV cm^{-3} sec^{-1}. If the region is 100 km wide, one would have a total imbalance of 10^9 eV cm^{-2} sec^{-1} between heat production and loss in the runaway region. A heat flux of this magnitude would then have to exist in order to remove the excess heat generation from this region; at the temperatures prevailing in the ionosphere, moderate temperature gradients are enough to set up such a flux, and for this reason the electron temperature is generally kept below $T_{em} = 3T_n$ at low altitudes (below 500 km). The resulting temperature gradients, however, may be enough to cause the appearance of a peak in electron temperature between 200 and 300 km, which sometimes shows up in measured electron temperature profiles [Wand, 1969]. An example of a runaway condition and the resulting temperature profiles is given on Figure 3, where a peak in T_e is quite visible.

The fact that the ions are sometimes unable to dissipate locally all the heat that is being received by the electrons in an extended region will raise the general level of the electron temperature in the ionosphere. This effect, however, when compared with a downward conductive heat flux of about 5×10^9 eV cm^{-2} sec^{-1} from the protonosphere to the F-layer during daytime [Nagy et al., 1969], does not seem to be of primary importance.

At high altitudes (above 600 km), $X \gg 1$ and therefore one may approximate $1 + X \cong X$ in the steady-state equation, with the result:

$$\frac{\bar{Q}}{\frac{3}{2}kn} = \left[7.7 \times 10^{-10} + 10^{-14} T_e^{1/2} \right] [0] (T_e - T_n) \cong$$

$$\cong 7.7 \times 10^{-10} [0] (T_e - T_n)$$

Under this condition, the electron and ion temperatures are being controlled by the heat transfer from the ions to the neutral particles. In the absence of heat conduction, the electron temperature would then increase exponentially with height. Actually, however, the effectiveness of a temperature gradient in setting up a conductive heat flux increases with $T_e^{5/2}$. For this reason, very strong heat fluxes are generated in this region, causing severe attenuation of the temperature gradients.

As seen above, a simple conductionless model will normally imply extremely high electron temperature gradients that are not observed in any measurement of the ionospheric temperature profile. The conclusion may be drawn that, although photoionization and heat exchange between different species have a major role in the thermal budget of the ionospheric plasma, these processes alone do not accurately represent the thermal behavior of the electrons in the ionosphere. As for the ions, however, a simple model based on equilibrium between local heating and cooling will yield a much more accurate representation of their temperature profile because the thermal conductivity of the ions is more than one order of magnitude smaller than that of the electrons.

Heat Conduction. For the electron temperatures found in the ionosphere, relatively strong conductive heat fluxes can be set up by small gradients in electron temperature; the divergence of the heat flux will tend to neutralize the imbalance between heat production and loss at each point. If one ignores all external heating and cooling processes ($\bar{Q} = \bar{L} = 0$), as well as the transport terms in the heat flow equation, and if a constant heat flux Φ is considered, then:

$$BT_e^{5/2} \frac{\partial T_e}{\partial h} = - \Phi$$

$$\therefore \quad T_e = \left\{ T_{eo}^{7/2} - \frac{7\Phi}{2B} (h-h_o) \right\}^{2/7}$$

A negative (downward) heat flux will cause the temperature to increase slowly with height. For electrons, $B = 7.7 \times 10^5$ eV/sec cm $^\circ K^{7/2}$ in the absence of neutral species; the presence of neutral particles in the F-region attenuates the conductivity of the electron gas [Herman and Chandra, 1969], but their concentration in this region is not enough to make this effect important. For this reason, the influence of the neutral particles on the electron conductivity is disregarded in this discussion, but is taken into account in the simulation program described in the next chapter.

The temperature gradient needed to maintain a heat flux Φ will be:

$$\frac{\partial T_e}{\partial h} = - \frac{\Phi}{7.7 \times 10^5 \, T_e^{5/2}}$$

During the day, the photoelectrons that escape the ionosphere are partly thermalized in the protonosphere, raising the plasma temperature there. As a result, a conductive heat flux is set up from the protonosphere to the ionosphere. Downward heat fluxes in excess of 10^9 eV cm^{-2} sec^{-1} coming from the protonosphere are predicted from calculation of the escape flux of photoelectrons [Nisbet, 1968]. For $\bar{\Phi} = -5 \times 10^9$ eV cm^{-2} sec^{-1}, a temperature gradient will exist of about 15° K/km for $T_e = 1000°$ K, 2.5°K/km for $T_e = 2000°$ K and 1° K/km for $T_e = 3000°$ K. Temperature gradients of this order have been observed [Nagy et al., 1969; Wand, 1967], thus confirming Nisbet's calculations.

Due to their much lower thermal velocities (for comparable temperatures), the ions are much less effective than the electrons in transporting heat via conduction. The conductivity is inversely proportional to the square root of the mass of the particles, so that the same heat fluxes with the same temperatures would require the O^+ ions to have temperature gradients 170 times greater than the electrons. Since such gradients have never been observed, heat fluxes conducted by electrons must be much greater than those carried by the ions under all normal circumstances.

Transport Heating. A third instructive way of isolating selected terms in Eq. (19) is to ignore the presence of conductive heat flow, as well as energy scattering into or out of the species ($\bar{Q} = \bar{L} = 0$). If this is done, only the convective term $-v \frac{\partial T}{\partial h}$ and the compression term $-\frac{2}{3} T \frac{\partial v}{\partial h}$ remain; heating and cooling will depend only on the energy involved in the transport motion of the gas. This is the adiabatic case, since the energy contained in a cell of gas

remains constant. Considering steady-state in this case:

$$\frac{\partial T}{\partial h} = -\frac{2}{3} T \frac{\partial v}{\partial h}$$

If the additional assumption is made that no particles are being created (Q = 0) or lost (L = 0), the flux of particles will have to be height-independent and therefore $\frac{1}{v}\frac{\partial v}{\partial h} = -\frac{1}{n}\frac{\partial n}{\partial h}$. Consequently:

$$\frac{1}{T}\frac{\partial T}{\partial h} = \frac{2}{3}\frac{1}{n}\frac{\partial n}{\partial h}$$

$$T = T_o \left(\frac{n}{n_o}\right)^{2/3}$$

This relation is used in a large number of situations in which an adiabatic regime may be reasonably assumed to exist. Given its great simplicity, its use would be quite convenient, but unfortunately it is not justified by the physical conditions in the ionosphere. Both for ions and electrons, the transport terms in Eq. (19) are always of secondary importance when compared with the conduction term.

The relative importance of each term in Eq. (19) was reviewed above. For electrons, it was stated that heating by collisions with photoelectrons, cooling by thermal coupling with ions and neutral particles, and heat conduction are all essential components of the equation, while the contribution of the transport terms are generally unimportant. Because of the difficulty in obtaining values for the gas velocity, the latter terms are then often ignored. In the simulation program presented in this work, these velocities are available

from the simultaneous solution of the momentum transport equation, so that inclusion of the transport terms introduces no significant further complication. For this reason, they are kept in the equation to be solved in the numerical program.

As for the ions, their thermal behavior is largely controlled by the neutral air at lower altitudes and by the electrons at higher altitudes. An assumed absence of ionic conduction during the day implies a relatively high ionic temperature gradient in the transition region between electron and neutral control of the ion temperature, since the electrons are at a much higher temperature than the neutral gas. Banks [1967] has shown that the ionic conductivity is efficient enough to smooth this transitional behavior. The effect of ionic thermal conduction is therefore not always negligible, but it is not a dominant one under any circumstance, and its inclusion increases the order of the equation, complicating its numerical solution. As for the transport terms, Banks [1967] has calculated that velocities of about 70 m/sec at 600 km are needed for them to be comparable to the ionic conduction term. Velocities of this magnitude have been measured [Evans et al., 1970], but only for some time during the early morning. For these reasons, the ionic temperature is calculated in the simulation program through a simple equilibrium between heating by electrons and cooling by the neutral gas.

The analysis of ionic thermal behavior is further complicated by the existence of different species of ions. Although these are strongly coupled to each other, there may be a small difference between the temperatures of the H^+ and O^+ ions in the topside. These

differences have only a minor effect in the overall energy balance of the plasma [Banks, 1967], and they will be ignored in the present work. A common ionic temperature is assumed at every point, but separate energy transfer rates are considered for each ionic species.

Chapter 3

NUMERICAL SIMULATION OF THE IONOSPHERE

The complexity of Eqs. (17) and (19) clearly precludes any attempt to solve them simultaneously in a simple analytical fashion. An analytical solution of the continuity-momentum transport equation can be obtained for assumed values of electron and ion temperatures, but the resulting concentration profile will not in general be compatible with the assumed temperatures. As for the heat-flow equation, its non-linear nature prevents even this kind of solution, except when the conduction term is ignored, which would yield results in complete disagreement with the observations.

Many analytical studies of the continuity equation have been made in the past with different simplifying assumptions. Although these studies have revealed many important features of ionospheric behavior, it soon became clear that numerical solutions had to be used if one were to take into account the non-isothermal nature of the temperature profiles. Accordingly, some numerical solutions of this equation were obtained, but with the use of model temperatures for the electrons, ions and neutrals. Conversely, studies were made of the temperature variations of the ionospheric plasma, in which model concentration profiles were involved. In these studies, the coupling between the dynamic and thermal variations in the ionospheric plasma was ignored.

As more and more unexplained features of ionospheric behavior were identified, the need for a self-consistent solution of all equations describing the variations in the ionospheric plasma temperatures

and concentrations became clear. Not only is it necessary to couple the equations describing the temperatures and concentrations of the charged particles, but also a simultaneous solution has to be obtained for the neutral winds, since these, while having a powerful influence on the plasma distribution in the F-layer, are themselves heavily influenced by the amount of ions present. At present, the trend for self-consistency in numerical simulation of the ionosphere has found its most advanced point in the work of P. Stubbe (1970), in which time-varying coupled simultaneous solutions are given for the individual temperatures and concentrations of the O^+ and H^+ ions, the electrons and neutral particles, the concentrations of NO^+ and O_2^+ ions, as well as for the two components of a horizontal wind.

Self-consistent solutions are valuable because they reveal certain variation patterns that would normally remain hidden because of the difficulty in visualizing the coupling between the equations. From a simulation point of view, however, the inclusion of one more equation in the problem is advantageous only when the corresponding boundary values and coefficients are well known; otherwise, it is just as good and much cheaper to use a model. Stubbe, for example, recommends the use of model neutral temperatures at present. In practice, these factors have to be carefully considered because of the increased costs of "complete" solutions.

Assuming a good working scheme for solution of the equations, the quality of the results will depend on the adoption of realistic boundary conditions. For the O^+ concentration equation, for example, it is reasonable to assume local photochemical equilibrium at 100 km,

and this assumption provides a lower boundary conditon. The conditions in the upper F-region, however, are not as well known; for this reason, many numerical solutions of this equation in the past have used a "zero flux" upper boundary condition. Such a boundary condition is clearly not appropriate to describe a steady-state nighttime ionosphere which cannot exist without a flux from the protonosphere unless a nighttime source of ionization is postulated. Moreover, as shown on the remaining chapters of this work, this "zero flux" condition is not adequate for the daytime ionosphere either.

Many ionospheric parameters, such as the total columnar electron content and the peak electron concentration, are much better known than the concentration or flux at any fixed height of the upper F-region. In order to make the solution follow prescribed values of these quantities, however, one would normally have to try different values of the top concentration at each step in an iterative procedure. In an already lengthy "complete" solution, this would be quite costly.

In the program described in this chapter, a scheme is included that allows the use of boundary conditions other than a given concentration at the top of the solution interval in a non-iterative manner. The additional computation time brought about by this scheme is minimal.

The proposed program solves simultaneously the coupled differential equations describing the concentration of the O^+ ions, their temperature, the electron temperature, and the concentration of O_2^+, NO^+, and N_2^+ ions. The self-consistency of this approach is basically limited by the fact that the neutral wind is entered as an input to the program, and not solved for in a closed form. The effect of the neutral winds on the F-layer, however, has to be superimposed on electromagnetic drifts caused

mainly by East-West electric fields, as shown in Eq. (10). Therefore, it seems immaterial to include the calculation of these winds in an ionospheric simulation program unless the fields are also calculated. The calculation of the fields, however, is much more difficult than that of the winds, since it involves processes that occur in the E region, which has a completely different morphology from the F layer.

The proposed program can be operated both in a steady-state and in a time-varying mode. Because the use of a strictly time varying mode at high altitudes would require very short timesteps, a hybrid mode was formulated to allow use of timesteps of the order of 5 minutes while essentially simulating the time-varying behavior of the ionosphere.

A. Stability of the Solution

Equation (17) is a second-order parabolic partial differential equation. In order to solve it, one must know two boundary conditions at any time t, and one initial concentration profile n(h, 0).

Provided that these boundary conditions, as well as the coefficients of the equation, are known or can be reasonably estimated, Equation (17) can be solved with a computer. To do so, a difference equation is constructed so that its solution converges to the solution of the differential equation when the time and height steps tend to zero in a well defined way.

Among all possible schemes for construction of such difference equation, the most straightforward is based on an explicit method. Each derivative in Equation (17) is replaced by the corresponding finite-difference approximation, resulting in the following:

$$\frac{n_{i,j+1} - n_{i,j}}{\Delta t} = D_{i,j}\left[\frac{n_{i+1,j} - 2n_{i,j} + n_{i-1,j}}{(\Delta h)^2} + A_{i,j}\frac{n_{i+1,j} - n_{i-1,j}}{2\Delta h} + B_{i,j}n_{i,j}\right] + Q_{i,j} - L_{i,j} + u_j \frac{n_{i+1,j} - n_{i-1,j}}{2\Delta h}$$

where: $$n_{i,j} = n(i\Delta h,\ j\Delta t)$$

In this simple scheme, each new concentration value $n_{i,j+1}$ is calculated directly from the concentrations $n_{i,j}$ of the previous time step, which are all known.

In spite of its attractive simplicity, the explicit method cannot be used because of severe limitations imposed by the stability requirements. The basic cause of instability is the second-order term. Let us then study the stability conditions in the explicit difference equation for the problem:

$$\frac{\partial n}{\partial t} = D \frac{\partial^2 n}{\partial h^2}$$

The corresponding difference expression is

$$\frac{n_{i,j+1} - n_{i,j}}{\Delta t} = D_{i,j} \frac{n_{i+1,j} - 2n_{i,j} + n_{i-1,j}}{(\Delta h)^2}$$

For the purpose of discussing the stability question, let us assume that $D_{i,j} = D$ = constant for all i and j, which, although not true in the case of the ionosphere, will not restrict the nature of our conclusion regarding the stability problem. Rearranging the expression above:

$$n_{i,j+1} = (1 - 2\theta)\ n_{i,j} + \theta\ (n_{i+1,j} + n_{i-1,j})$$

where: $$\theta = \frac{(\Delta t)D}{(\Delta h)^2}$$

The equation describing the propagation of errors will then be:

$$\epsilon_{i,j+1} = (1 - 2\theta)\ \epsilon_{i,j} + \theta\ (\epsilon_{i+1,j} + \epsilon_{i-1,j})$$

where $\varepsilon_{i,j}$ is the error associated with $n_{i,j}$.

If $\theta < 1/2$, $\varepsilon_{i,j+1}$ will be a weighted average of three errors of the j-th level, with all the weights between 0 and 1. Therefore, the maximum error at any time cannot be greater than the maximum error at the preceding level. A sufficient condition for stability is then:

$$\theta = \frac{D(\Delta t)}{(\Delta h)^2} < \frac{1}{2}$$

This condition can also be proven to be necessary for stability. For a height step of 10 km and $D = 10^{11}$ cm^2 sec^{-1}, one should then have $\Delta t < 5$ sec in order to obtain stability. Therefore, the requirement of stability would put a rather severe limitation on the timestep if one were to use an explicit method.

The use of implicit methods must then be considered. In these, each space derivative is replaced by a weighted sum of finite difference expressions for the two consecutive times. One would then have for this problem:

$$\frac{x_i - n_i}{\Delta t} = \frac{D}{(\Delta h)^2}\left[a(x_{i+1} - 2x_i + x_{i-1}) + (1-a)(n_{i+1} - 2n_i + n_{i-1})\right]$$

where:

$$x_i = n_{i,j+1}$$

$$n_i = n_{i,j}$$

Rearranging terms:

$$-a\theta x_{i+1} + (1 + 2a\theta)\, x_i - a\theta x_{i-1} = (1 - a)\,\theta n_{i+1}$$

$$+ [1 - 2(1 - a)\theta]\, n_i + (1 - a)\,\theta\, n_{i-1}$$

Assuming then that the error is kept zero at the boundaries ($i = 1$ and $i = m$), the equation that describes the propagation of errors can be put in the following matrix form:

$$A\varepsilon_{j+1} = B\varepsilon_j$$

where:

$$A = I + a\theta C$$

$$B = I - (1 - a)\theta C$$

$I = (m - 2) \times (m - 2)$ identity matrix

$$C = \begin{vmatrix} 2 & -1 & 0 & 0 \ldots \\ -1 & 2 & -1 & 0 \ldots \\ 0 & -1 & 2 & -1 \ldots \\ 0 & 0 & -1 & 2 \ldots \end{vmatrix}$$

$$\varepsilon_j = \begin{vmatrix} \varepsilon_{2,j} \\ \varepsilon_{3,j} \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_{m-1,j} \end{vmatrix}$$

Let us now try $\varepsilon_j = \varepsilon_o \lambda^j$ as a solution of this matrix equation. If this is a solution for the equation, it is clear that stability can only be achieved if $\lambda \leq 1$ for all possible values of λ. Since for

this solution one would have $\epsilon_{j+1} = \lambda\epsilon_j$, substitution into the equation would yield:

$$(\lambda A - B)\ \epsilon_j = 0 \text{ for arbitrary } \epsilon_j$$

This expression represents a linear homogeneous system of equations with an infinite number of solutions. Then:

$$\det\ (\lambda A - B) = 0$$

The solutions of this equation for λ must all satisfy the stability condition $\lambda \leq 1$ in order for the method to be stable. Let us then solve for λ; substituting the expressions for A and B:

$$\det\ [\lambda I + \lambda a\theta C - I + (1 - a)\ \theta C] = 0$$

$$\det \left\{ C - \frac{(1 - \lambda)}{[1 - (1 - \lambda)\ a]\theta} \right\} = 0$$

Therefore, all possible values for λ are given by

$$\frac{1 - \lambda}{[1 - (1 - \lambda)\ a]\ \theta} = \mu$$

where μ is any eigenvalue of the matrix C.

From Fox [1962], we know that all eigenvalues of C are between zero and four, and so all possible values of λ must lie between 1 and the solution of:

$$\frac{1 - \lambda}{[1 - (1 - \lambda)\ a]\ \theta} = 4$$

Solving this equation, one has for the minimum of the interval allowed for λ:

$$\lambda_{min} = 1 - \frac{4\theta}{1 + 4a\theta}$$

Therefore, a necessary and sufficient condition for stability is

$$\frac{4\theta}{1 + 4a\theta} \leq 2$$

$$\theta\ (1 - 2a) \leq \frac{1}{2}$$

Then, if $a \geq 1/2$, stability is guaranteed for any θ, while for $a < 1/2$ the method is stable only if

$$\theta \leq \frac{1/2}{1 - 2a}$$

which checks with the result already obtained for the case $a = 0$.

The case $a = 1/2$ corresponds to the method introduced by Crank and Nicolson [1947], which has unrestricted stability for the simple problem discussed above. Richtmyer [1957] discussed the effect of lower order terms on stability, and concluded that there is practically no effect, at least for the case of constant coefficients. Fox [1962] called attention to the fact that, although implicit methods may be unconditionally stable, they still may give rise to a large truncation error when θ is exceedingly large. Since the diffusion coefficient D increases exponentially with height, we will be faced with exactly this situation at large altitudes. Alternate forms for the difference equation will then be used in order to deal with this problem.

B. The Equation of the O^+ Ion

The ionosphere is formed by a mixture of many different species of ions; therefore, one has to solve an equation for each ion in order to calculate the plasma dynamics. Due to complexity of the equation and the number of ionic species involved, it is then convenient to ignore certain terms of the equation that represent unimportant physical processes for the ion under consideration. Accordingly, diffusion will not be considered for the ions O_2^+, NO^+ and N_2^+, since these ions exist mainly at regions where photochemical processes are dominant. For the same reason, vertical transport of ionization due to presence of neutral winds or transverse electric fields will also be ignored for these ions. As a result, the corresponding equations are of zero-order in h, thus not requiring any boundary condition; only an initial condition is necessary.

As for the O^+ ion, diffusion is dominant at one altitude range while photochemical processes are dominant at another; there are transition regions where these processes coexist in competition. Consequently, the complete Eq. (17) will be solved for this ion.

Since the loss function for O^+ depends on the concentration of this ion, one must explicitly include this dependence in the equation. The loss of O^+ ions is due mainly to the following reactions:

$$O^+ + O_2 \rightarrow O_2^+ + O$$

$$O^+ + N_2 \rightarrow NO^+ + O$$

Therefore, the loss rate is given by:

$$L = \gamma_{O_2} [O_2]n + \gamma_{N_2} [N_2]\, n$$

where γ_{O_2} and γ_{N_2} are the corresponding reaction rates, and n is the O^+ concentration.

Letting $\beta = \gamma_{O_2} [O_2] + \gamma_{N_2} [N_2]$, one has:

$$L = \beta n,$$

where β is called the linear loss coefficient.

With this kind of loss rate, the corresponding difference equation for Eq. (17) is:

$$\frac{x_i - n_i}{\Delta t} = Q_i - \frac{1}{2}\beta_i (x_i + n_i) - \frac{u}{4\Delta h}(x_{i+1} + n_{i+1} - x_{i-1} - n_{i-1}) +$$

$$+ D_i \frac{1}{2(\Delta h)^2}[x_{i+1} + n_{i+1} - 2x_i - 2n_i + x_{i-1} + n_{i-1}] +$$

$$+ \frac{A_i}{4\Delta h}(x_{i+1} + n_{i+1} - x_{i-1} - n_{i-1}) + \frac{B_i}{2}(x_i + n_i)$$

where:

$$x_i = n_{i,j+1}$$

$$n_i = n_{i,j}$$

$$Q_i = Q_{i,j+1/2}$$

$$\beta_i = \beta_{i,j+1/2}$$

$$u = u_{j+1/2}$$

$$D_i = D_{i,j+1/2}$$

$$A_i = A_{i,j+1/2}$$

$$B_i = B_{i,j+1/2}$$

Then, rearranging all the terms, one has:

$$a_i\, x_{i+1} + \bar{b}_i\, x_i + c_i\, x_{i-1} = \bar{d}_i \qquad (20)$$

where:

$$a_i = \mu_i \left[\frac{1}{2\Delta h} + \frac{A_i}{4}\right] - \frac{\Delta t}{4\Delta h} \cdot u$$

$$\bar{b}_i = -1 - \frac{\mu_i}{\Delta h} + \frac{\Delta t}{2}\,(B_i D_i - \beta_i)$$

$$c_i = \mu_i \left[\frac{1}{2\Delta h} - \frac{A_i}{4}\right] + \frac{\Delta t}{4\Delta h}\, u$$

$$\bar{d}_i = -\, a_i\, n_{i+1} - (\bar{b}_i + 2)\, n_i - c_i\, n_{i-1} - \Delta t\, Q_i$$

$$\mu_i = \frac{\Delta t}{\Delta h}\, D_i$$

Equation (20), plus two boundary conditions at any time, yields the numerical solution of the problem. The procedure used to solve (20) will naturally depend on the nature of the boundary conditions; these procedures will be discussed at the end of this chapter.

C. The Steady-State Mode

The differential equation for the steady-state concentration profile is derived by making $\frac{\partial n}{\partial t} = 0$ in Equation (17):

$$0 = Q - \beta n - u\,\frac{\partial n}{\partial h} + D\left[\frac{\partial^2 n}{\partial h^2} + A\,\frac{\partial n}{\partial h} + Bn\right]$$

The corresponding difference equation can be derived from (20) by making $n_i = x_i$ for all i:

$$a_i x_{i+1} + \bar{b}_i x_i + c_i x_{i-1} = - a_i x_{i+1} - (\bar{b}_i + 2) x_i - c_i x_{i-1} - \Delta t\, Q_i$$

$$a_i x_{i+1} + (\bar{b}_i + 1) x_i + c_i x_{i-1} = - \frac{1}{2} \Delta t\, Q_i \qquad (21)$$

It is worthwhile to notice that Equation (21) gives the steady-state solution based on given temperatures of ions and electrons. One must then iterate it with the heat equation in order to get steady-state concentration and temperature profiles that are mutually compatible. Therefore, unless one is not concerned with the feedback between temperature and concentration, there is little practical advantage in using Equation (21) (steady-state mode) over Equation (20) (time-varying mode), since the latter permits the observation of real-time convergence to steady-state.

However, the steady-state mode will have important application in a different context, as explained in the next section.

D. The Hybrid Mode

As seen in the previous chapter, different regions of the ionosphere are dominated by different physical processes. Over a wide range of altitudes in the topside, diffusion is the dominant process for the O^+ ion, while in the bottomside the same ion species is largely controlled by chemical processes. Since both the diffusion constant and the recombination coefficient have exponential altitude variations with scale heights of less than 50 km, the time variations of electron concentration are characterized by time constants that take a wide range of values in the ionosphere. This is very important from a numerical standpoint, for in order to solve the equation strictly on

a time-varying basis, the timestep to be used should be smaller than the smallest time constant present in the height interval in which the solution is desired; otherwise, consecutive overshoots will cause the solution of the difference equation to present large oscillations.

Using some typical ionospheric values, the following values are obtained for the recombination time β_r^{-1} at 200 km and the diffusion time τ_d at 700 km:

$$\beta_r^{-1} \text{ (200 km)} \cong 100 \text{ sec}$$

$$\tau_d \text{ (700 km)} \cong 10 \text{ sec}$$

Therefore, in order to avoid large oscillations in the time-varying mode between 200 and 700 km, one would have to use a timestep smaller than some 10 seconds, which would be prohibitively expensive. Besides, as far as the global behavior of the layer is concerned, such a small timestep would be unnecessary. In order to overcome this difficulty, the program may be run in a hybrid mode. In this mode of operation, a test is made at each height to compare the timestep Δt with the smallest time constant τ present at that height. If $\Delta t < \tau$, Equation (20) is used, corresponding to the time-varying procedure. If $\Delta t > \tau$, Equation (21) is used, so that the linearization error is greatly reduced.

Figure 4 illustrates the situation. In both graphs, the curves are assumed real continuous time-variations of n. The solutions for n_{i+1} given by the steady-state and the time-varying procedures are indicated in each case. It is clear that when $\tau < \Delta t$, the steady-state solution is a better approximation to the real time varying continuous

solution than the one given by Equation (20).

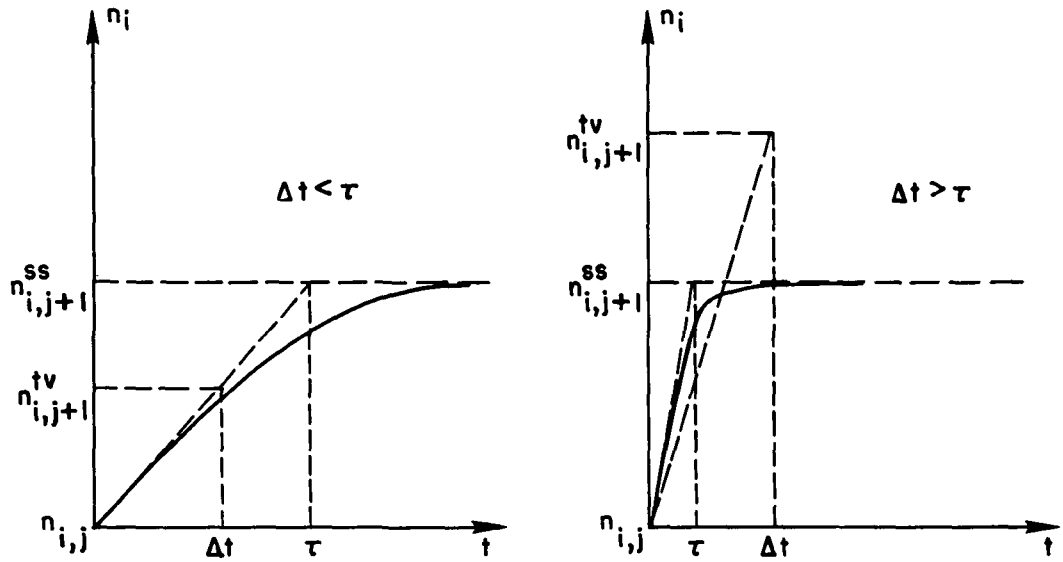


Fig. 4. COMPARISON BETWEEN TIME-VARYING AND STEADY-STATE MODES. Two assumed true variations of n_i with time, both with initial value given by $n_{i,j}$, are presented; $n^{tv}_{i,j+1}$ and $n^{ss}_{i,j+1}$ are approximations to $n_{i,j+1}$ given by the time-varying and steady-state modes respectively.

In a hybrid mode, it is then possible to use any convenient timestep.

E. The Boundary Conditions

Eqs. (20) and (21) may be written as:

$$a_i\, x_{i+1} + b_i\, x_i + c_i\, x_{i-1} = d_i \qquad i = 2, 3, \ldots, m-1 \qquad (22)$$

where the definition of b_i and d_i will depend on whether the program is run in the steady-state or time-varying mode, or, in the case of the hybrid mode, on how the timestep compares with the local time constant.

At each step, one must then solve for m unknowns x_1, x_2, . . ., x_m, and for this purpose (22) provides us with (m - 2) linear equations. In order to solve the system, two more equations involving the x_i's have to be added; these are provided by two boundary conditions. At the lower boundary, which in this program is located at 100 km, it is reasonable to assume that $[O^+]$ is given by photochemical equilibrium, since the recombination rates are so high there.

Therefore:

$$x_1 = \left.\frac{q_o}{\beta}\right|_{i=1} \qquad (23)$$

The specification of x_1 allows the transformation of (22) into another system of the form:

$$x_i = E_i\, x_{i+1} + F_i \qquad i = 1, 2, 3, \ldots, m-1 \qquad (24)$$

The coefficients E_i and F_i can be obtained by substituting the expression above into Equation (22):

$$a_i\, x_{i+1} + b_i\, x_i + c_i\, (E_{i-1}\, x_i + F_{i-1}) = d_i$$

$$\therefore \quad x_i = \frac{-a_i}{b_i + c_i E_{i-1}} x_{i+1} + \frac{d_i - c_i F_{i-1}}{b_i + c_i E_{i-1}}$$

Identifying this equation with Equation (24):

$$E_i = \frac{-a_i}{b_i + c_i E_{i-1}}$$

$$F_i = \frac{d_i - c_i F_{i-1}}{b_i + c_i E_{i-1}} \qquad (25)$$

All coefficients E_i and F_i can then be calculated from E_1 and F_1, which are specified by the lower boundary condition given in Eq. (23) as:

$$E_1 = 0$$

$$F_1 = \left.\frac{q_o}{\beta}\right|_{i=1}$$

If the top concentration x_m is now specified by the upper boundary condition, the x_i's can be all immediately calculated by the recurrent application of Equation (24). It is often desired, however, to specify conditions other than a given concentration value at a fixed height in the top of the ionosphere. Let us consider a general condition given in the following form:

$$r_i x_i + r_2 x_2 + \ldots + r_m x_m = R \qquad (26)$$

where all the r_i's and R are specified constants. The linear system to be solved would then be the following:

$$\begin{bmatrix} 1 & -E_1 & 0 & 0 & & 0 \\ 0 & 1 & -E_2 & 0 & & 0 \\ 0 & 0 & 1 & -E_3 & & 0 \\ \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot \\ 0 & 0 & 0 & 0 & 1 & -E_{m-1} \\ r_1 & r_2 & r_3 & r_4 & r_{m-1} & r_m \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ x_{m-1} \\ x_m \end{bmatrix} = \begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ F_{m-1} \\ R \end{bmatrix}$$

If the first row of coefficients is multiplied by r_1 on both sides and subtracted from the last row, the solution of the system will be preserved. The new system, however, will have a zero replacing r_1, while r_2 and R will be replaced by r_2' and R' given by:

$$r_2' = r_2 + E_1 r_1$$

$$R' = R - F_1 r_1$$

and all other coefficients remain unaltered. If now the second row is multiplied by r_2' and subtracted from the last row, r_2' will be replaced by a zero, while r_3 and R' will be replaced by r_3' and R'' given by:

$$r'_3 = r_3 + E_2\, r'_2$$

$$R'' = R' - F_2\, r'_2$$

If this procedure is repeated (m - 1) times, the only non-zero coefficients remaining in the last row of the system will be r'_m and the independent term $R^{(m-1)}$. The top concentration x_m that is needed for the solution to satisfy Eq. (26) will therefore be given by:

$$x_m = \frac{R^{(m-1)}}{r'_m}$$

Once x_m is determined, all other unknowns are calculated through recurrent application of Eq. (24). In this scheme, a condition in the form of Eq. (26) can be imposed on the solution in a non-iterative way, thereby saving precious computer time. There are many relevant ionospheric parameters that can be specified in the form of Eq. (26). A discussion is given below of some of the applications.

a) Integrated Columnar Electron Content. A quantity that is often well-known is the integrated electron content n_T of the ionosphere. From knowledge of this quantity and calculations of the columnar contents of other ions, it is possible to estimate $[O^+]_T$, the content of O^+ ions in the ionosphere. The problem is then to make the solution present the specified content value. Using a trapezoidal approximation, this condition is satisfied when:

$$\frac{1}{2}\,(x_1 + x_m) + \sum_{i=2}^{m-1} x_i = \frac{[O^+]_T}{\Delta h}$$

This condition is clearly in the form of Eq. (26). Therefore, it can be imposed on the solution through the scheme described above if the coefficients of that equation are specified as:

$$r_1 = r_m = .5$$

$$r_i = 1 \qquad i = 2, 3, \ldots, m-1$$

$$R = \frac{[O^+]_T}{\Delta h}$$

Figure 5 shows a series of electron concentration profiles obtained with the application of this boundary condition. The simulation program was run in the hybrid mode, with Δt = 6 minutes, Δh = 10 km, and u = 0. The values of columnar electron content used as boundary conditions are 31-day averages around 24 March 1967, at Stanford.

Naturally, if it is so desired, the content can be specified at any particular region, just by making $r_i = 0$ corresponding to heights h_i outside the interval in question.

b) Height, h_{max}, of the Peak Ion Concentration. If the peak of the profile is required to be at a certain height, h_{max}, between h_p and h_{p+1}, the condition may be approached by imposing that $x_{p+1} - x_p = 0$. This can be obtained if the following coefficients are specified on Eq. (26):

$$r_p = - r_{p+1} = 1$$

$$r_i = 0 \qquad i = 1, 2, \ldots, p-1, p+2, \ldots, m$$

$$R = 0$$

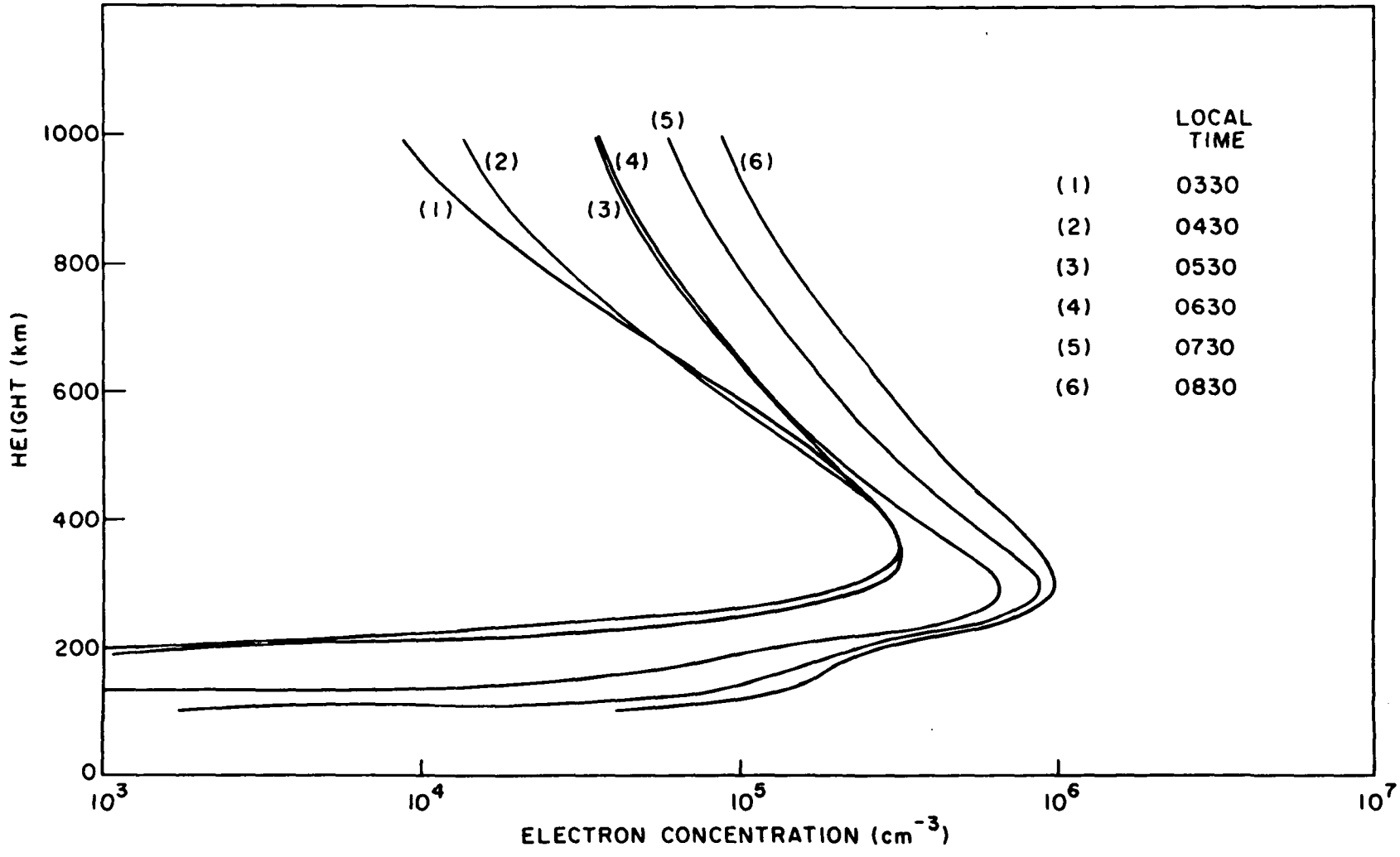


Fig. 5. ELECTRON CONCENTRATION PROFILES CALCULATED WITH SIMULATION PROGRAM. The values of electron content used as boundary values in this simulation are given on Figure 6.

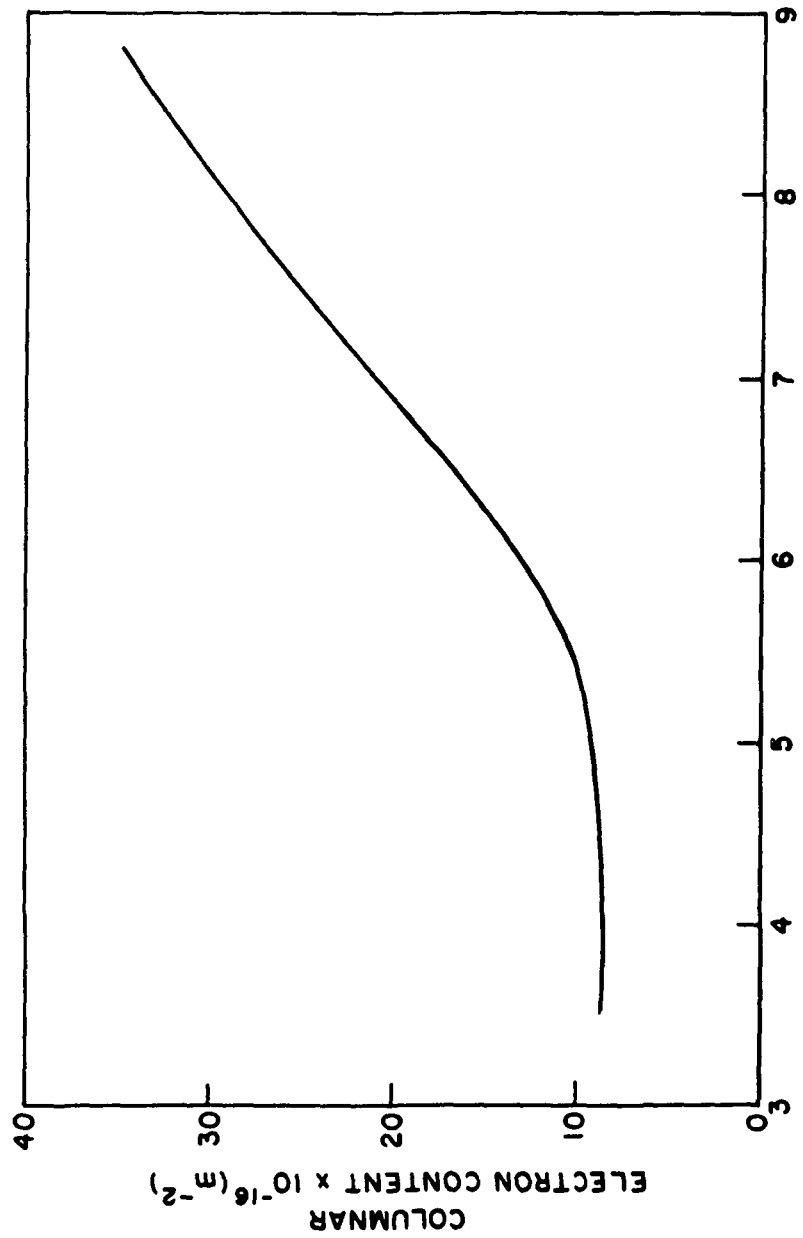


Fig. 6. TIME-VARYING BOUNDARY VALUES USED IN A SIMULATION. The electron content values presented above are 31-day averages around 24 March 1967, at Stanford.

It should be noticed at this point that all that can be insured by this procedure is that the solution will present a zero height derivative around h_p. The presence of a peak corresponds to an additional requirement of negative second derivative, which cannot be imposed on the solution. Moreover, some choices of h_p may produce negative values for the concentration. Therefore, even though the method provides a means for imposing the height of the peak, the specification of this height must be within a certain range of plausible choices, outside of which the resulting solutions would be physically impossible.

This limitation becomes particularly critical for the nighttime steady-state ionosphere, which, in the absence of nighttime ionization sources, is a purely accretive layer, for which the shape does not depend on the magnitude of the flux that is sustaining it. In this case, the height of the peak cannot be used as a boundary condition.

F. The Ions of the Lower F-Region

Among the ions that form the lower part of the F-region, the most important are the O_2^+, NO^+, and N_2^+ ions, and only these will be considered in the present work. The program solves the equations describing the time-varying behavior of the concentrations of these ions. Since they are produced mainly below 200 km during daytime, these ions do not have as much opportunity as the O^+ ions to diffuse through the neutral species, which are very dense in this region. This makes it possible to neglect the diffusion term in the equations of these ions. Above 250 km, this will cause considerable departure between the simulated profiles and the true ones, but there the contribution of these ions in the total mixture is minimal.

Besides photoionization, the ions of the lower F-region will be controlled by a number of different reactions, eight of which are shown on Table 1. Although it is difficult to predetermine the relative importance of each reaction under the most general conditions, the use of only these eight reactions in the program yields ion concentration profiles in reasonable agreement with the experimental results for the major ions of the lower F-region (e.g., Johnson, 1967). However, other reactions might also be important under special circumstances.

TABLE 1

Reaction	Reaction Rates		
	Symbol	Numerical Value	Reference
$O_2^+ + e \rightarrow O + O$	α_{O_2}	$7 \times 10^{-8}(\frac{1000}{T_e})$	Whitten et al. (1964)
$N_2^+ + e \rightarrow N + N$	α_{N_2}	$2.8 \times 10^{-7}(\frac{1000}{T_e})^{3/4}$	Whitten et al. (1964)
$NO^+ + e \rightarrow N + O$	α_{NO}	$4.7 \times 10^{-8}(\frac{1000}{T_e})^{3/2}$	Whitten et al. (1964)
$N_2 + O^+ \rightarrow NO^+ + N$	γ_{N_2}	5×10^{-13}	Stubbe (1969)
$O_2 + O^+ \rightarrow O_2^+ + O$	γ_{O_2}	1.6×10^{-11}	Stubbe (1969)
$O_2 + N_2^+ \rightarrow N_2 + O_2^+$	δ_{O_2}	10^{-10}	Fehsenfeld et al. (1965)
$O + N_2^+ \rightarrow NO^+ + N$	δ_O	2.5×10^{-10}	Fehsenfeld et al. (1965)
$O_2^+ + NO \rightarrow NO^+ + O_2$	ϵ	8×10^{-10}	Mitra (1968)

Considering only the reactions in Table 2, the concentrations of O_2^+, N_2^+ and NO^+ are given by the following differential equations:

$$\frac{d}{dt}[N_2^+] = q_{N_2} - \alpha_{N_2} n_e [N_2^+] - \delta_{O_2} [O_2] [N_2^+] - \delta_O [O] [N_2^+]$$

$$\frac{d}{dt}[O_2^+] = q_{O_2} + \gamma_{O_2} [O_2] [O^+] + \delta_{O_2} [O_2] [N_2^+] - \alpha_{O_2} [O_2^+] n_e - \epsilon[NO][O_2^+]$$

$$\frac{d}{dt}[NO^+] = \delta_O [O] [N_2^+] + \gamma_{N_2} [N_2] [O^+] + \epsilon [NO] [O_2^+] - \alpha_{NO} [NO^+] n_e \quad (27)$$

where:

$$n_e = [O^+] + [N_2^+] + [NO^+] + [O_2^+] = \text{electron concentrations}$$

The value of $[O^+]$ is derived from the complete continuity equation in the manner described in the previous sections, and is therefore regarded as a known quantity in these equations. The 3 differential equations are coupled together, and there are quadratic terms contained wherever n_e appears. Consequently, the system does not have a straightforward solution, and some simplifying procedures are appropriate at this point.

In the scheme used in this program, $\frac{d}{dt}$, is initially set equal to zero in Eqs. (27), resulting in a system of 3 coupled quadratic equations for each reference point. The solution of this system produces the steady-state concentration values $[N_2^+]_s$, $[O_2^+]_s$ and $[NO^+]_s$ corresponding to each point. If the program is being run in the steady-state mode, the calculation of the ionic concentrations is then complete. If the time-varying mode is being used, the computer now calculates how

close to their steady-state values the concentrations of O_2^+, NO^+ and N_2^+ ions were brought since the previous timestep. The differential equations (27) permit the calculations of the instantaneous time derivatives at the previous time step. If $[M]_j$ is a generic ion concentration at time step j, $[M]_s$ is the steady-state value calculated at the next step and $\frac{d}{dt}[M]$ is the corresponding initial time variation given by Eq. (27), one can then derive an effective time constant τ_M given by:

$$\tau_M = - \frac{[M]_j - [M]_s}{\frac{d}{dt}[M]}$$

With this time constant, the concentration for the next time step (j + 1) is calculated as:

$$[M]_{j+1} = [M]_s + \left\{ [M]_j - [M]_s \right\} \exp\left[- \frac{\Delta t}{\tau_M} \right]$$

This procedure is applied individually for each of the ions in question. Figure 7 is an example of a resulting ionic concentration profile.

G. The Heat-Flow Equation and the Simultaneous Solution

The heat flow equation (19) differs basically from the continuity-momentum transport equation in that its second-order term (which is the term that limits the stability of the solution in an explicit scheme) is non-linear. For this reason, it is difficult to apply to the heat-flow equation the discussion about stability that was presented in a

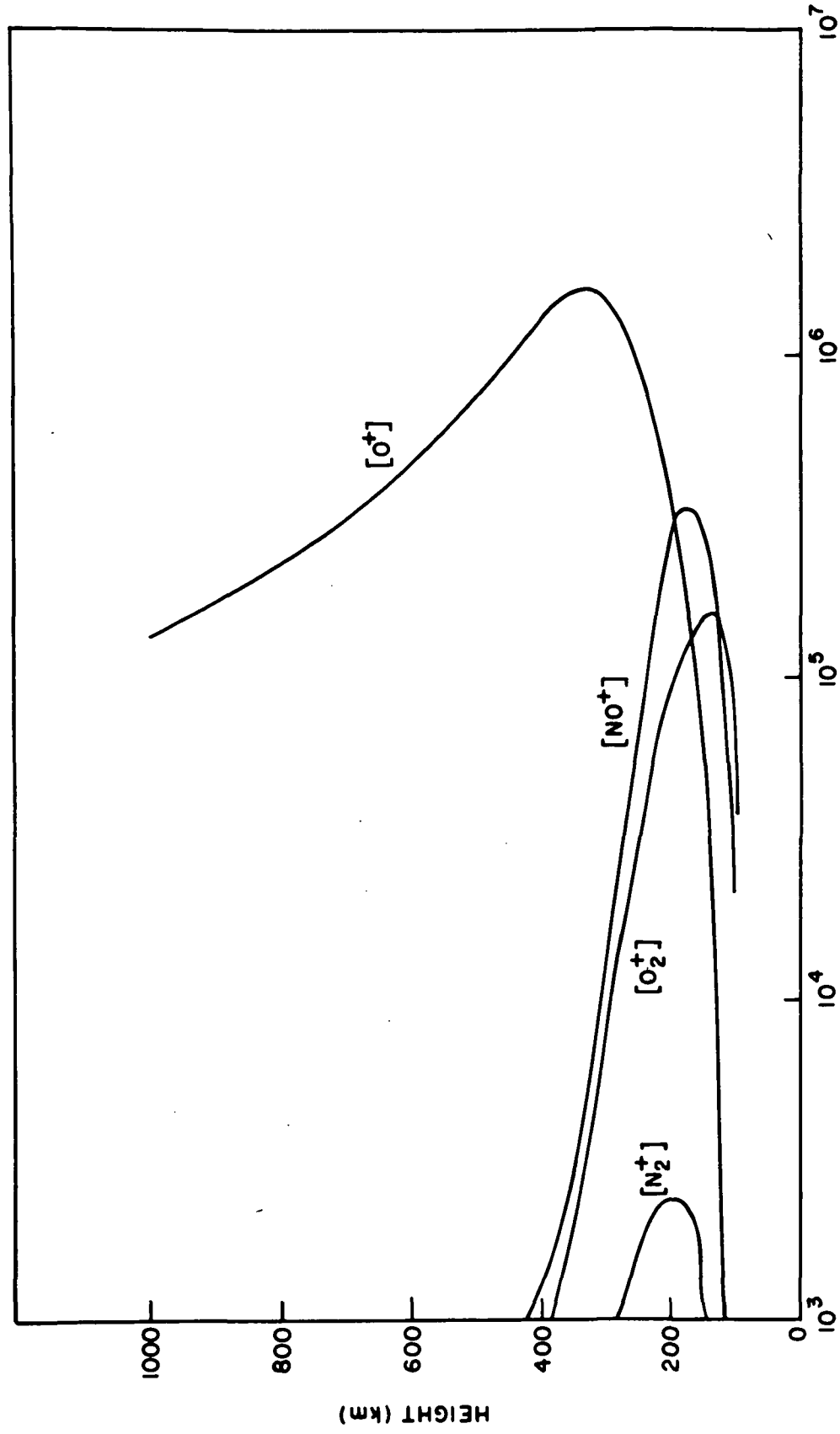


Fig. 7. CALCULATED IONIC CONCENTRATION PROFILES.

section of this chapter, and it might be expected that the stability conditions are even more restrictive in this case. In spite of these difficulties, da Rosa (1965) has formulated an implicit scheme that gives stable solutions of the heat-flow equation. This scheme is used in the present work.

In order to solve simultaneously the heat-flow and continuity-momentum transport equations for the ionospheric plasma, the following steps are taken at each time step:

a) the neutral temperatures and concentrations are calculated for the time of the present step, using an interpolation between values given for every 2 hours by the CIRA model atmosphere appropriate for the assumed degree of solar activity;

b) The production functions q_0, q_{O_2} and q_{N_2} are calculated at all heights where the solution is to be calculated. For this purpose, the solar EUV and x-ray fluxes are divided into 15 bands, and the intensity of the incident radiation in each of these bands is specified as an input;

c) the average energy of the photoelectrons produced by photoionization of each of the 3 main neutral species by each of the 15 radiation bands is calculated by averaging over the different possible resulting ionic states. The probability factors assigned to these states were taken from Shea et al. [1968]. Using the electron concentrations from the previous step, the computer then calculates the energy transferred by each group of photo-

electrons to the electron gas before it loses its remaining energy in an inelastic collision. These energies are then multiplied by the corresponding photoelectron production rate and summed over the 3 neutral species and 15 radiation bands to yield the local electron heat production at each point. A correction is then introduced to take into account the non-local thermalization of the photoelectrons, in the manner suggested by Geisler and Bowhill [1965];

d) the concentration of O_2^+, NO^+ and N_2^+ ions are calculated in the manner described in the previous section, using values of $[O^+]$ and T_e given by the previous step;

e) the concentration of O^+ ions is calculated by the scheme previously described, using values of T_e and T_i given by the previous time step;

f) electron temperature values are calculated through da Rosa's scheme, using electron and ion concentration values given in the previous time step;

g) ion temperature values are calculated from the new values of T_e and T_n, assuming ionic thermal equilibrium at every point.

Chapter 4

THE COUPLING BETWEEN THE F-REGION AND THE PROTONOSPHERE

In Chapter 2, a discussion was made of the conditions encountered by protons when they exist as a minor species among oxygen ions in the ionosphere. The concepts presented in that discussion should aid in understanding the mechanisms by which charge is transferred from the oxygen ions of the F-region to the neutral hydrogen atoms, and by which these newly created protons may or may not reach the protonosphere, where they will constitute the major species. The purpose of this chapter is to make an inquiry into these mechanisms, and in so doing develop concepts and ideas that will clarify some aspects of the long-term variations of some ionospheric parameters, especially the ionospheric columnar electron content, of which a large amount of observations is discussed later in this work.

A. The Diffusive Barrier

Above some 400 km and up to a height that depends on a number of parameters, the distribution of hydrogen ions follows closely a condition of chemical equilibrium. Under this condition, the proton concentration is determined by the balance between the reactions:

$$O^+ + H \underset{k_2}{\overset{k_1}{\rightleftharpoons}} O + H^+ \qquad (36)$$

The reaction rates k_1 and k_2 adopted in the present work will be approximations to the expressions given by Holzer and Banks [1969] for low-speed conditions:

$$k_1 = 4.3 \times 10^{-11} T_n^{1/2}$$

$$k_2 = 3.8 \times 10^{-11} T_i^{1/2} \qquad (37)$$

In the absence of diffusion processes, the equation describing the time-varying behavior of $[H^+]$ will be:

$$\frac{\partial}{\partial t}[H^+] + k_2 [O][H^+] = k_1 [O^+][H] \qquad (38)$$

Given the concentration of neutral hydrogen, neutral oxygen and O^+ ions, the concentration of protons will approach its corresponding steady-state value with time constant $\{k_2[O]\}^{-1}$. Below 100 km, this time constant is everywhere shorter than 10 minutes. Since the concentrations of the other species involved in (38) never vary significantly in such a short time, it is safe to assume that, wherever the charge exchange is rapid enough to hinder the action of diffusive processes on the hydrogen ion distribution, the concentration of these ions will be given by its steady-state value:

$$[H^+] = \frac{k_1}{k_2}\frac{[H]}{[O]}[O^+] = \frac{9}{8}\frac{[H]}{[O]}\left(\frac{T_n}{T_i}\right)^{1/2}[O^+] \qquad (39)$$

The concentration of neutral hydrogen is thought not to vary strongly with height between 400 and 1000 km; consequently, since T_n and T_i vary much more slowly with height than do the concentrations, $[H^+]$ will vary approximately as $[O^+]/[O]$ in the chemical equilibrium region, i.e., it will increase exponentially with height. As the neutral oxygen concentration decreases with increasing altitudes, the distance traveled by a proton between its creation and loss will become longer and longer with

higher altitudes, until it becomes longer than the neutral oxygen scale height. When this happens, the distribution of hydrogen ions is no longer determined by a chemical equilibrium condition, because most of the newly created protons are not lost locally.

The height at which the transition between chemical and diffusive control of the protons occurs is called the critical level, and it was first identified and estimated by Hanson and Ortenburger [1961]. These authors pointed out that the condition for non-local loss of the protons is not simply that the mean free path for charge exchange must be longer than the scale height of neutral oxygen, because the protons are also suffering elastic (Coulomb) collisions with the much heavier O^+ ions. These collisions have the effect of partially confining the protons, thereby extending the chemical equilibrium region. If λ is the mean free path for elastic scattering by O^+ ions, the proton will perform a random walk with steps of mean length λ; therefore, in order to move a distance equal to the neutral oxygen scale height H_n, it will have to move on the average $\left(\frac{H_n}{\lambda}\right)^2$ steps of length λ without undergoing a charge exchange. The critical level will then be located where:

$$H_n^2 = \lambda \rho \tag{40}$$

where ρ is the mean free path for charge exchange.

For the elastic scattering and the charge exchange mean free path, the following expressions will be used [Hanson and Ortenburger, 1961]:

$$\lambda = 1.3 \times 10^4 \quad \frac{T_i^2}{[O^+]}$$

$$\rho = \frac{4.15 \times 10^{14}}{[O]}$$

Substituting into Equation (40), the following condition is found for the critical height:

$$[O^+]_c \, [O]_c \left(\frac{T_n}{T_i}\right)^2 = 1.3 \times 10^{11} \qquad (41)$$

where the subscript c refers the subscripted quantity to the critical level.

Above the critical level, the protons are free to travel large distances before they are neutralized. The distribution of hydrogen ions will then be determined by an equilibrium between pressure gradient and electrical, gravitational and collisional forces. As seen on Chapter 2, the electric field E in the region where H^+ is still a minor species is upwards and eE is comparable with the weight of an oxygen ion. Since the hydrogen ions are much lighter but have the same charge as the O^+ ions, they will be rapidly accelerated upwards, provided that the pressure at higher altitudes is low enough. On the other hand, it was shown in Chapter 2 that there is a maximum upward bulk velocity with which the protons can travel through this region. For this reason, the region that is bounded from above by the protonosphere (which is defined as starting where $[H^+] \geq [O^+]$) and from below by the critical level is called a diffusive barrier region.

The existence of the diffusive barrier is naturally dependent upon the condition that the chemical equilibrium region terminates at a point (critical level) where the hydrogen ions are still a minor species. If one assumes that the chemical equilibrium values are followed by the

proton concentration up to the critical level, the ratio $[H^+]/[O^+]$ at that point will be:

$$\frac{[H^+]_c}{[O^+]_c} = \frac{9}{8}\frac{[H]_c}{[O]_c}\left(\frac{T_n}{T_i}\right)^{1/2} = \frac{9}{8}\left(\frac{T_n}{T_i}\right)^{1/2}\frac{[H]_{500}}{[O]_{500}}\, e^{\frac{15}{16}z_c} \qquad (42)$$

where the subscript 500 refers to the 500 km level (taken here as a reference point), $z_c = (h_c - 500)/H_n$ where h_c and H_n are in km and neutral hydrogen is assumed to follow a diffusive equilibrium distribution above 500 km.

The reduced height z_c of the critical level can be calculated from Eq. (41), which describes the condition that defines that point. Assuming that the diffusive equilibrium mode dominates the distribution of O^+ ions above 500 km, z_c will be given by

$$[O^+]_{500}\,[O]_{500}\left(\frac{T_n}{T_i}\right)^2 e^{-\frac{r+1}{r}z_c} = 1.3 \times 10^{11}$$

where $r = \dfrac{T_e + T_i}{T_n}$

Letting $s = \dfrac{15r}{16(r+1)}$ and eliminating z_c in Eq. (42):

$$\frac{[H^+]_c}{[O^+]_c} = \frac{1.125}{(1.3 \times 10^{11})^s}\,\frac{[H]_{500}\,[O^+]_{500}^{\,s}}{[O]_{500}^{1-s}}\left(\frac{T_n}{T_i}\right)^{2s+\frac{1}{2}} \qquad (43)$$

Since $r \geq 2$, the parameter s may take values between 5/8 and 1, although it will probably never exceed 0.9. For a typical value like $r = 4$, s will be 3/4 and then:

$$\frac{[H^+]_c}{[O^+]_c} = 5.1 \times 10^{-9} \frac{[H]_{500} [O^+]_{500}^{3/4}}{[O]_{500}^{1/4}} \left(\frac{T_n}{T_i}\right)^2 \qquad (43a)$$

The substitution of typical values for the concentrations and temperatures in the expression above will generally yield values of $[H^+]_c/[O^+]_c$ well below unity. This fact corroborates the idea that a diffusive barrier will generally exist above the chemical equilibrium region.

Inside the diffusive barrier, as the hydrogen ions are more and more controlled by diffusive processes alone, their distribution will approach the superposition of two exponential modes. The diffusive equilibrium mode, which would be the only one present in the hydrostatic ($v = 0$) or in the collisionless ($\nu = 0$) case, is characterized by a scale height given by

$$H_d = -\frac{kT_i}{m_H g'} = -\frac{kT_i}{m_H g}\left(\frac{\frac{T_i}{T_e} + 1}{\frac{T_i}{T_e} - 15}\right)$$

where m_H is the proton mass.

When $T_i = T_e$, H_d will be given by $-\frac{8}{7} H_p$, where H_p is the plasma scale height; the proton concentration will then increase exponentially with about the same growth rate as that below the critical level. For this reason, it is difficult to recognize the critical level by merely inspecting the ionic concentration profiles when the diffusive equilibrium mode dominates the proton distribution in the diffusive barrier.

The second mode is the full flow mode, which is the only one present when the protons are flowing upwards with the maximum possible velocity

This velocity is called the limiting velocity and is defined by Equation (16) in Chapter 2. The full flow mode is characterized by the scale height of the friction-producing species, which in this case is formed by the 0^+ ions.

Assuming $T_e = T_i$, the concentration of hydrogen ions in the diffusive barrier will be given by

$$[H^+] = n_1 e^{\frac{7(h-h_c)}{8H_p}} + n_2 e^{-\frac{(h-h_c)}{H_p}} \tag{44}$$

The velocity corresponding to this profile can be calculated from Eq. (9), where both the diffusion coefficient and the gravitational field have to be modified in the way that was specified for minor ions in Chapter 2. For an isothermal case, the velocity profile will be given by:

$$v(H^+) = \frac{g'}{\nu} \sin^2 I - \frac{D}{[H^+]} \frac{\partial}{\partial h} [H^+]$$

where:

$$D = \frac{kT_i}{m_p \nu} \sin^2 I$$

$$g' = g \left[1 - \frac{16}{1 + \frac{T_i}{T_e}}\right]$$

and ν is the collision frequency for momentum transfer from protons to 0^+ ions.

Assuming $T_i = T_e$, the proton flux $\Psi = [H^+]\ v\ (H^+)$ will then be given by:

$$\Psi = -\ 15\ n_2\ \frac{g}{\upsilon}\ \sin^2 I\ e^{-\frac{h - h_c}{H_p}} \qquad (45)$$

The total flux depends only on the full flow mode of the hydrogen ions distribution. Although the expression given above for the flux of protons through the diffusive barrier appears to be height-dependent, it actually is not, because υ varies exponentially with scale height H_p. The flux may then be expressed as

$$\Psi = -\ \frac{15\ n_2\ g\ \sin^2 I}{\upsilon_c}$$

Since the flux Ψ is determined by the amplitude of the full flow mode, it should be possible to determine its magnitude from the observation of ionic concentration profiles. Unfortunately, however, it is not easy to do so in the general case, because the rapid exponential growth of the diffusive equilibrium mode will tend to mask the full flow mode. Figure 8, which was reproduced from a paper by Geisler [1967], makes this point very clear. One can see that the flux will not alter drastically the distribution unless $n_2 \gg n_1$. When this happens, the velocity will approach the characteristic value for the full flow mode, which is the limiting velocity given by Eq. (16):

$$v_L = -\ 15\ \frac{g}{\upsilon}\ \sin^2 I$$

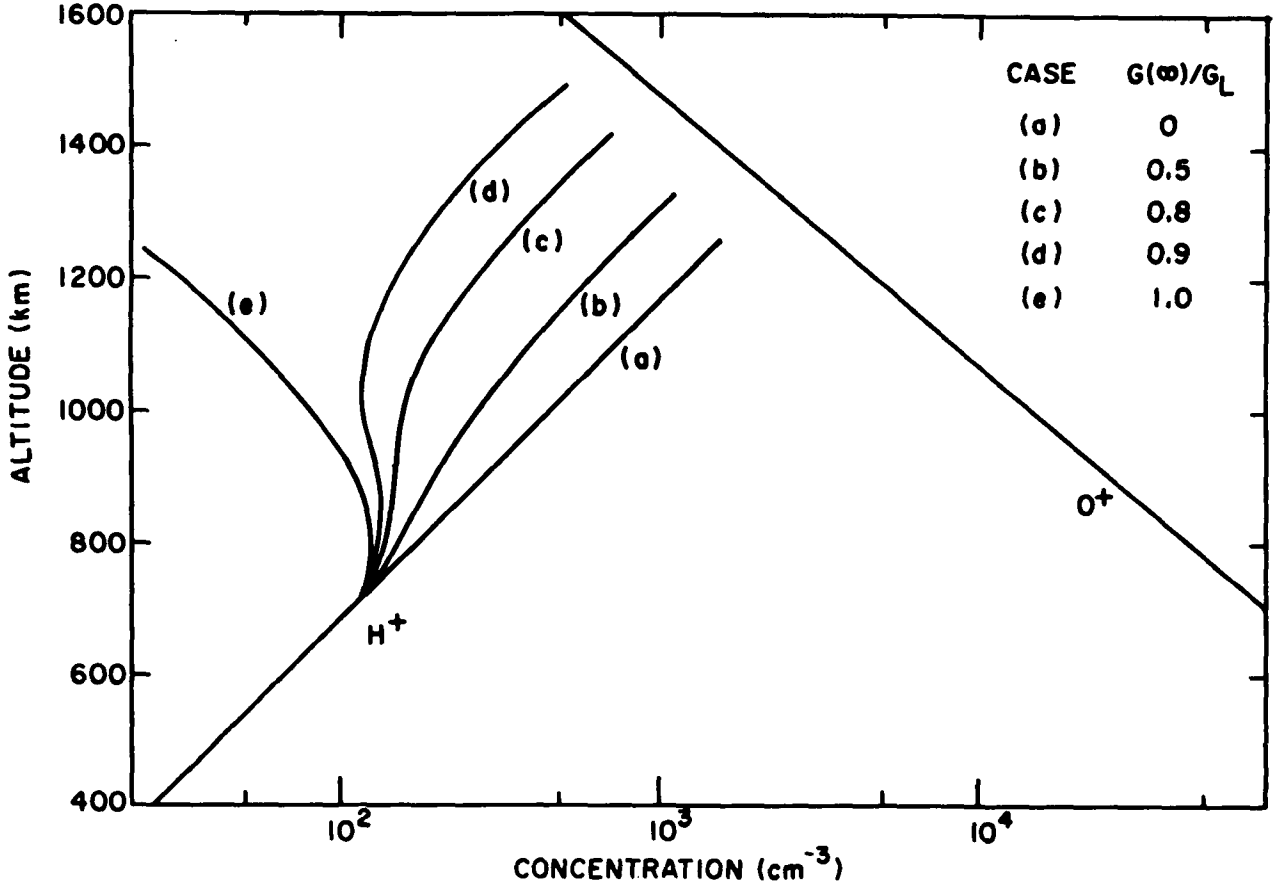


Fig. 8. PROTON CONCENTRATION PROFILES. This figure, which was reproduced from a paper by Geisler [1967], shows how little impact the upward flow of protons has on their distribution except when it approaches the limiting value, Ψ_D, which is called G_L in Geisler's paper.

The limitation on the velocity of protons imposed by collisions with O^+ ions in the diffusive barrier, when combined with the fact that $[H^+]$ is specified at the base of the diffusive barrier by a chemical equilibrium condition, implies a limitation in the total upward flux that can flow across the diffusive barrier and penetrate the protonosphere. The maximum upward flux that can be supported by diffusion through the diffusive barrier is called the limiting flux Ψ_D.

Under the limiting flow condition, $[H^+]$ in the diffusive barrier will be given by a pure full flow mode, so that the limiting flux will be:

$$\Psi_D \simeq - 15 \, [H^+] \, \frac{g}{\nu} \sin^2 I$$

Using for ν the values given by Banks and Holzer [1969] for low velocities:

$$\Psi_D = 1.1 \times 10^4 \, T_i^{3/2} \, \frac{[H^+]}{[O^+]} \sin^2 I \qquad (46)$$

Naturally, the ratio $[H^+]/[O^+]$ is independent of height in the diffusive barrier under these conditions, since only the full flow mode is present. The value of this ratio will determine the limiting flux. An estimate of the limiting flux can be obtained by assuming that $[H^+]$ follows the chemical equilibrium condition up to the critical level and then changes abruptly to a distribution given by the full flow mode. Substituting in (46) the $[H^+]/[O^+]$ ratio at the critical level given by Eq. (43):

$$\Psi_D = \frac{1.24 \times 10^4}{(1.3 \times 10^{11})^s} \, \frac{[H]_{500} \, [O^+]_{500}^s}{[O]_{500}^{1-s}} \, T_n^{2s+1/2} \, T_i^{1-2s} \sin^2 I \qquad (47)$$

The error affecting this estimate is due to the assumption of an abrupt change in the concentration profile at the critical level. Figure 8, however, suggests that this error is small. The critical level that results from the reference concentrations adopted by Geisler in his paper is at 780 km; one can see on the figure that at this height the value of $[H^+]/[O^+]$ that would be given by chemical equilibrium is very close to the value of this ratio in the diffusive barrier under the limiting flow condition.

A more accurate way of calculating the limiting flux is to consider an equation that contains both the diffusion and the charge-exchange terms, and determine the maximum value that can be approached by the flux with increasing altitudes so that this equation will still yield physically plausible solutions. This was done by Geisler [1967], but only for r=2. However, since the plasma temperature is generally higher than T_n, Eq. (47) will be adopted for the limiting flux. The error that can be expected from the procedure used to obtain this expression is by far smaller than the uncertainties that affect the concentrations involved in it. For r=4, one has:

$$\Psi_D = 5.6 \times 10^{-5} \frac{[H]_{500} [O^+]_{500}^{3/4}}{[O]_{500}^{1/4}} T_n^2 T_i^{-1/2} \sin^2 I \qquad (47a)$$

The limiting flux across the diffusive barrier, as expressed in this equation, will depend on a number of concentrations and temperatures that present spatial, temporal and solar variations. The knowledge of these quantities is generally limited by considerable uncertainties. Among all the relevant parameters, however, the most poorly known is by

far the concentration of neutral hydrogen. Figure 9 shows the variation of the ratio between Ψ_D and $[H]_{500}$, calculated with model atmospheres, and solar activity. While this ratio increases eight times from solar cycle minimum to maximum, $[H]_{500}$ is expected to fall 10 to 50 times in the same interval. Therefore, the limiting flux will probably decrease with solar activity, due to the sensitivity of [H] to neutral temperature. Moreover, Ψ_D will depend directly on the magnitude of [H], which is very poorly known.

In the next section, a brief discussion is presented of the present state of knowledge about the abundance of neutral hydrogen in the thermosphere, its variations, and the possible impact that a proton flux from the F-region to the protonosphere with the limiting value may have on its global circulation and balance.

B. The Concentration of Neutral Hydrogen

Hydrogen atoms are produced below 100 km by dissociation of water vapor and methane. The total amount of production per unit area is thought to be fairly constant throughout the year (except in the polar regions) and throughout the solar cycle, but its magnitude is very poorly known. A value of 2.5×10^7 cm^{-2} sec^{-1} was used by Hanson and Patterson [1963] for the source hydrogen flux; more recently, however fluxes of the order of 10^8 cm^{-2} sec^{-1} are being favored.

Neutral temperatures in the thermosphere are of the order of 1000°K; at this temperature, the mean thermal velocity of a hydrogen atom is 5 km/sec, while that of an oxygen atom is 1.25 km/sec. In order for any object to escape the gravitational field of the earth, it must have an outward radial velocity (at the surface) of at least 11.4 km/sec. Therefore,

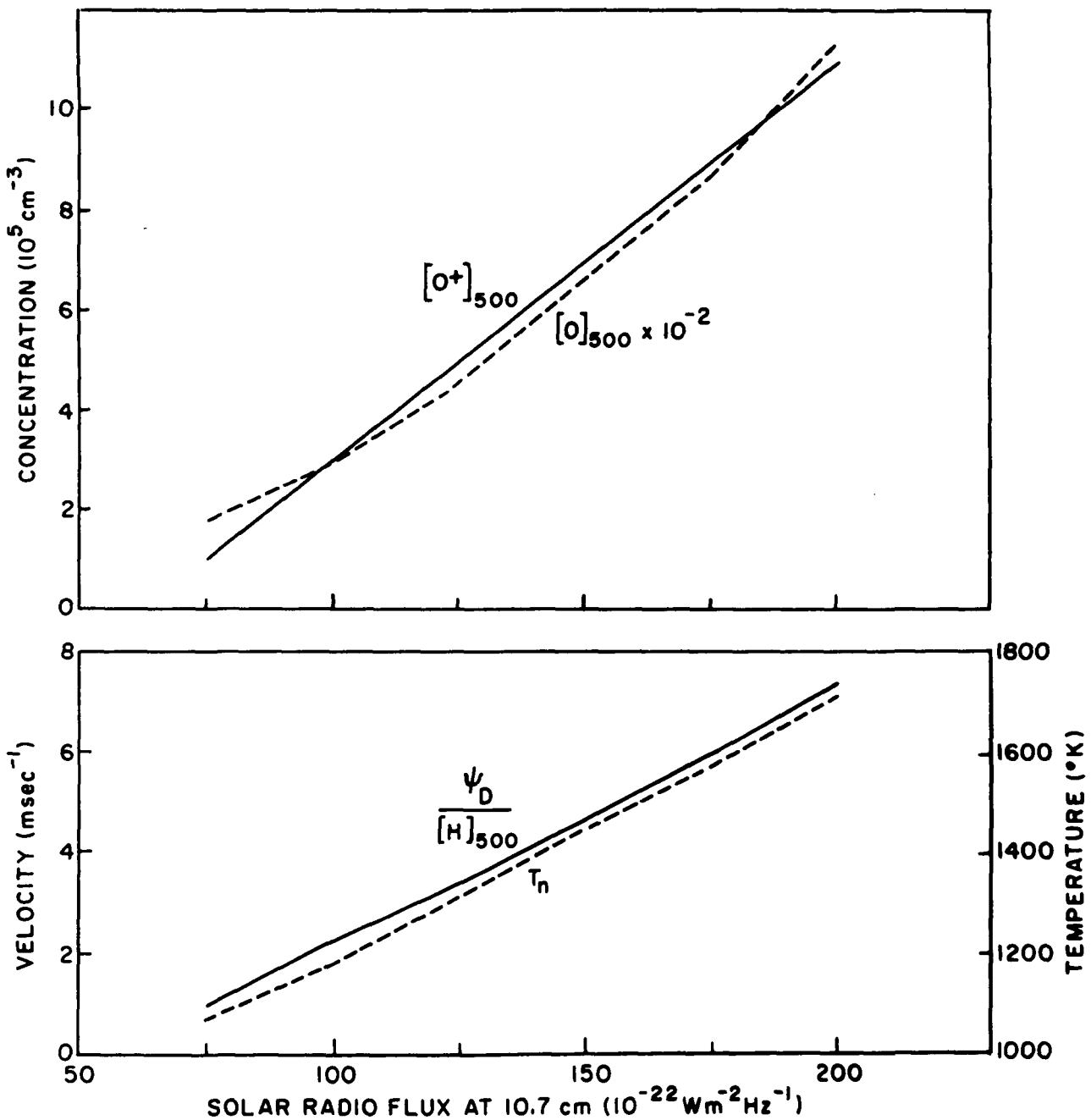


Fig. 9. SOLAR-CYCLE VARIATION OF THE PROTON ESCAPE FLUX. The curves for T_n, $[O^+]_{500}$ and $[O]_{500}$ are assumed model variations.

once the hydrogen atoms penetrate the region where they are no longer confined by collisions with heavier particles (exosphere, above about 500 km), a significant fraction of them will be able to escape the earth. This gives rise to a thermal escape flux; in the cases of the heavier gases, this effect is unimportant because of the lower thermal velocities.

Considerable theoretical difficulties are involved in the evaluation of the rate of thermal escape, unless the effect of the escape on the particle velocity distribution is ignored. If this is done, one has:

$$\frac{\Psi_e}{[H]} = 2.97 \times 10^7 \frac{\exp\left(-\frac{7010}{T_n}\right)}{\sqrt{T_n}} \quad cm^{-2}\ sec^{-1} \qquad (48)$$

where Ψ_e is the flux of hydrogen atoms due to thermal escape. Figure 10 shows the variation of this thermal escape velocity with solar activity, assuming a neutral temperature variation given by the CIRA model. One can see that the escape rate is very sensitive to the neutral temperature. Since the abundance of neutral hydrogen is roughly determined by a balance between a constant production and the thermal escape loss, there will be a strong variation of [H] throughout the solar cycle, as shown on Figure 12.

If a proton flux with the limiting value exists in the diffusive barrier, this flux must be fed by an equal flux of hydrogen atoms and O^+ ions. Therefore, when the F-region is supplying the protonosphere with a limiting flux of protons during daytime, this results in an additional loss of hydrogen atoms in the dayside of the earth. It is important to notice that, although this additional flux increases the total loss of neutral hydrogen on a local basis, it actually tends to retain

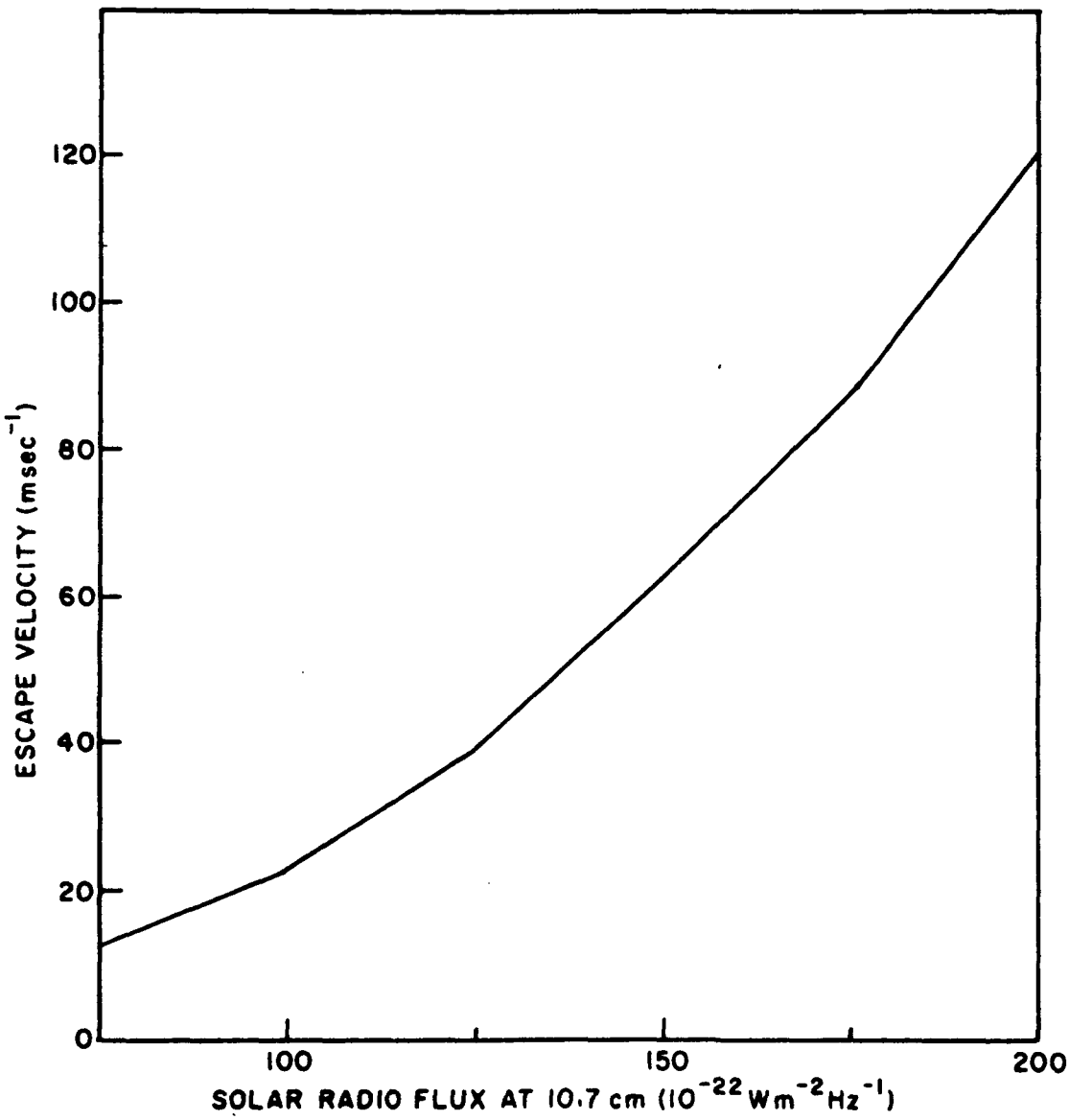


Fig. 10. THERMAL ESCAPE VELOCITY OF HYDROGEN ATOMS. The velocities were calculated for the daytime neutral temperatures of the CIRA reference atmosphere.

neutral hydrogen on a planetary scale, since the hydrogen atoms are stored in the protonosphere in the form of protons. These protons, when dumped back to the ionosphere, are neutralized and will tend to maintain the general level of hydrogen abundance in the ionosphere. The processes by which the ions of the protonosphere are returned to the F-region are not clear, but they seem to occur both in a slow, steady fashion in the night-side, and in an irregular way during stormy events such as geomagnetic storms. Figure 11 shows, in a schematic fashion, the role that the protonosphere, as a reservoir of protons, plays on the general circulation of atomic hydrogen in the exosphere. By giving some hydrogen atoms on the day-side a temporary alternative to thermal escape, the transfer of ionization to the protonosphere will actually delay the global loss of hydrogen, increasing its abundance.

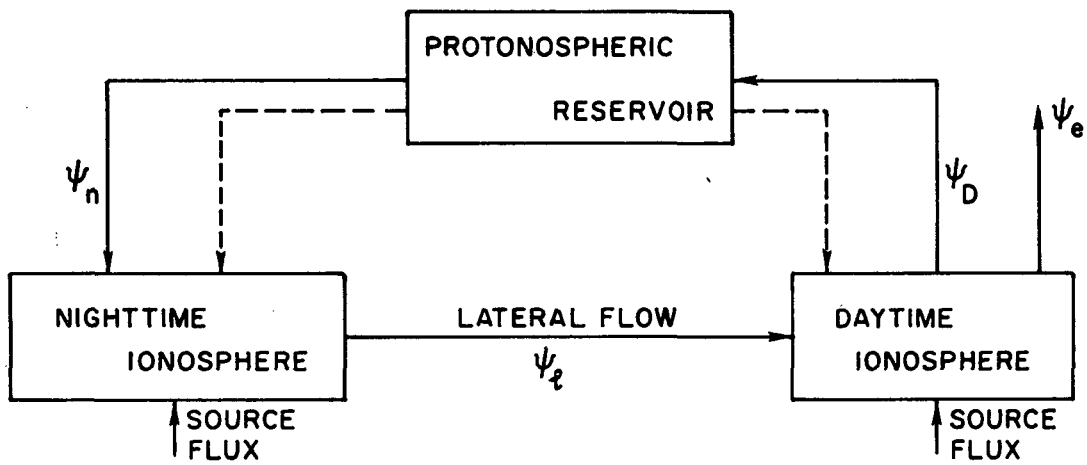


Fig. 11. SCHEMATIC DIAGRAM OF NEUTRAL HYDROGEN GLOBAL CIRCULATION IN THE EXOSPHERE. Dashed lines represent dumping of thermal protons on the ionosphere during stormy events.

Even if the F-region is transferring protons to the protonosphere at or near the limiting rate Ψ_D, the importance of this mechanism to the total neutral hydrogen balance will depend on how Ψ_D compares with the thermal escape Ψ_e. Comparison between Figures 9 and 10 indicate that Ψ_D is one order of magnitude lower than Ψ_e as given by Equation (48). Therefore, if the thermal escape is taken to be given by this equation, the circulation and escape of atomic hydrogen must not be strongly affected by coupling between the F-region and protonosphere.

Since the escaping particles are among the most energetic of the hydrogen atoms, there will be a tendency for the high energy tail of the velocity distribution to be depressed with respect to a Maxwell distribution. Whether this effect will significantly attenuate the escape rate or not is still a matter for controversy. Liwshitz and Singer [1966] found that because of this effect the true escape rate may be only one-fourth of the escape rate corresponding to a normal Maxwell distribution. The lower escape rates would naturally result in higher concentrations of neutral hydrogen. If this is true, however, the thermal escape rate will be comparable with the limiting flux of protons through the diffusive barrier. Therefore, it is possible that the exchange of ionization between the F-region and the protonosphere is after all an important factor in the global circulation of neutral hydrogen.

Due to the lower temperatures that prevail during the night, thermal escape on the night-side is many times slower than at the day-side of the earth. Hydrogen will therefore build up during the night, giving rise to a large day-night asymmetry in its concentration. Hanson and Patterson [1963] have shown that under these conditions a horizontal flow is set up that is very efficient in attenuating the asymmetry.

If protons are removed from the F-region during daytime at the maximum possible rate, and if this rate is comparable or faster than the thermal escape rate, an additional hydrogen shortage will be caused on the day-side by transfer of ionization to the protonosphere. Unless a lateral flow can eliminate this effect, the limiting flux itself will be reduced due to lack of hydrogen. Banks [1969b] has mentioned this problem in connection with the polar wind, and suggested that in that case a lateral flow would be fast enough to eliminate possible shortages of hydrogen in the polar regions due to the polar wind. While this seems plausible for the polar caps, it will certainly be more difficult to occur for the entire day-side of the earth. It is possible that in many days lateral flow from the night-side will maintain the level of hydrogen abundance only on the fringes of the day-side. In this case, high limiting fluxes Ψ_D would be possible only in the middle of the norming and afternoon, but not around noon.

From the standpoint of the global circulation and escape of neutral hydrogen, the discussion above is quite incomplete. It has been shown, however, that the coupling between the F-region and the protonosphere can only be thoroughly analyzed when considered in its connection with the whole morphology of that species in the exosphere.

In the present work, we wish to address the problem of ionization exchange between F-region and the protonosphere in connection with its possible effects on the F-region. In order to make this possible, a model will be adopted for the neutral hydrogen concentration in the exosphere. The use of model concentrations for the neutral species is a common procedure in the study of ionospheric processes. This is generally a good procedure, because the abundance of neutral gases are

seldom affected by their interaction with the ions. In the case of neutral hydrogen, this is not necessarily true, so that the use of a model is somewhat questionable. The complexity of the mechanisms that control the behavior of neutral hydrogen, however, would make a complete study of this problem extremely difficult, so that the use of a hydrogen model seems appropriate at this stage. However, in interpreting the results of this study, one must be aware of the possible impact that the resulting proton fluxes may have on the assumed neutral hydrogen concentrations.

The hydrogen concentrations used in this work were suggested by Stubbe [1970], and are given by the following analytical expression:

$$[H]_{500} = 1.5 \times 10^{7} \exp\left[-\frac{T_n - 700}{178}\right] \text{ cm}^{-3} \qquad (49)$$

The resulting variation of $[H]_{500}$ with solar activity, for the neutral temperatures of the CIRA model, is shown on Figure 12. The exponential nature of the variation with temperature is in general agreement with the other existing models, although disagreement exists as to its intensity and the magnitude of the day-night asymmetry that it implies. The main differences, however, concern the absolute values of [H], and are caused by discrepancies between observations made with different experimental methods. While spectroscopic measurements tend to give values of 10^4 cm^{-3} for $[H]_{500}$ during daytime at medium solar activity, Reber et al. [1967] reported $[H]_{500} = 10^6$ cm^{-3} within a factor of 3 from a direct measurement made with a mass spectrometer in the Explorer 32 satellite during May 1966. A few other mass spectrometer measurements have generally yielded [H] values one to two orders of

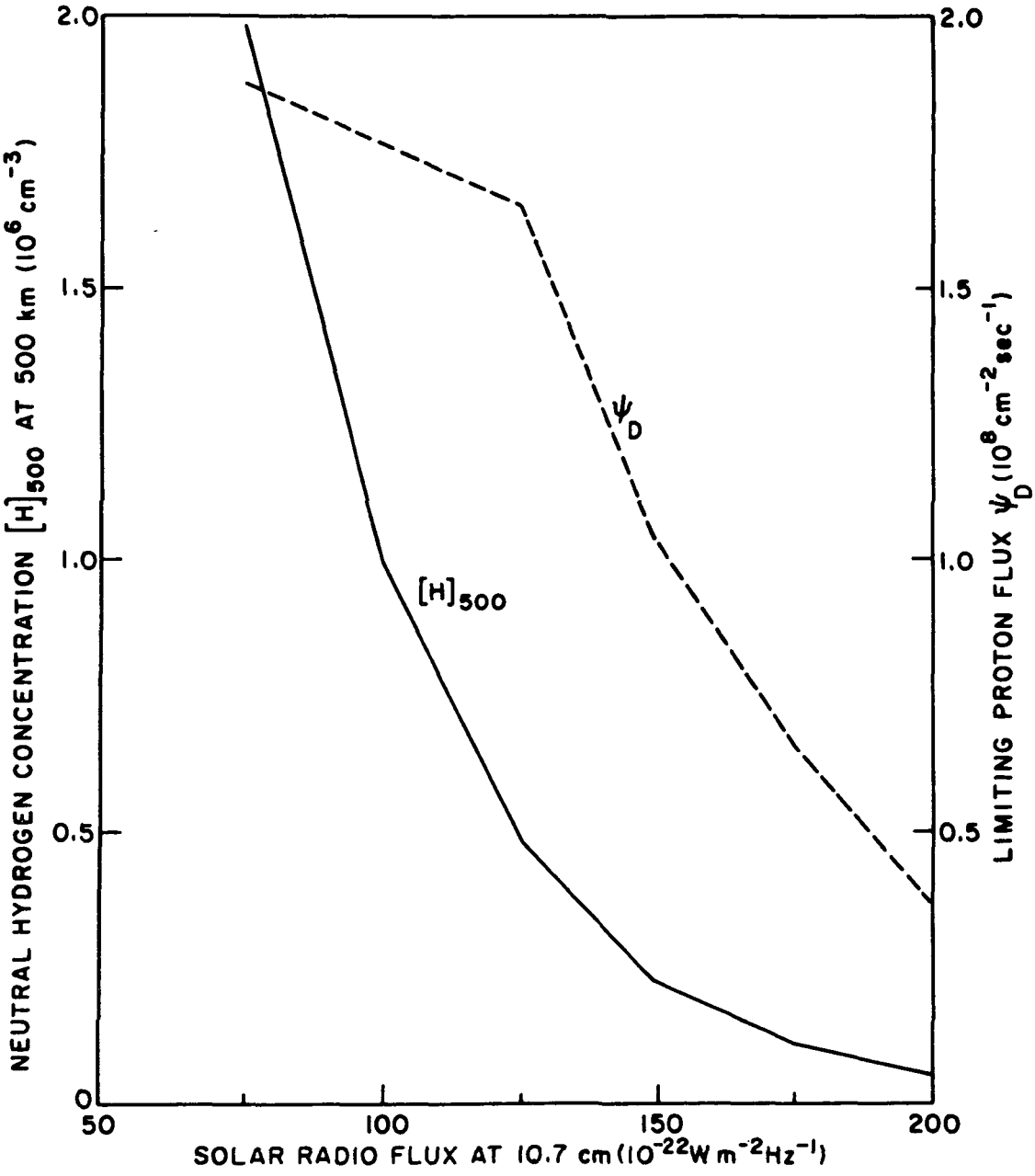


Fig. 12. VARIATION OF HYDROGEN CONCENTRATION WITH SOLAR ACTIVITY. Also shown is the corresponding variation in Ψ_D for the model quantities given in Figure 9.

magnitude above the optical determinations. While the larger number of spectroscopic measurements would tend to give them higher reliability, hydrogen concentration values can be inferred from these measurements only when some restrictive assumptions are made [Stubbe, 1970].

The values of [H] given by Equation (49) tend to agree much better with the direct mass spectrometer measurements. The reason for the choice of this particular model, however is that the corresponding values of $[H]_{500}$, when substituted in Equation (47a), yield values of the limiting flux Ψ_D that agree fairly well with values of this quantity which can be inferred from ionic composition measurements. These fluxes are discussed in the next section.

C. The Occurrence of the Limiting Flux

Much of the discussion presented in the previous sections is referred to a situation in which the F-region during the day is transferring ionization to the protonosphere at or near the maximum possible rate. When this happens, the F-region will be said to be under the limiting outflow condition. In the next chapter, it will be shown that many annual and solar variations in ionospheric columnar electron content, as well as other ionospheric parameters, can be better understood if the limiting flux is assumed to represent the common state of the ionization exchange between the daytime F-region and the protonosphere. In this section, we wish to discuss some independent observations that tend to support this assumption.

Hoffman et al. [1969] have reported daytime midlatitude ion composition measurements taken simultaneously with a rocket up to 700 km, and with the mass spectrometer on board the Explorer 31 satellite

from 700 km to almost 1100 km. In the height range covered by the satellite, the measured hydrogen ion and oxygen ion concentrations presented the same scale heights, with a ratio $[H^+]/[O^+]$ of about 0.034. This situation can only be explained by an upward proton flux at the limiting value. For an ionic temperature of 2700° K (which is suggested by the observed scale height), this flux can be calculated from Equation (46) to be 4.5×10^7 cm^{-2} x sec^{-1}. For this flux, the value of $[O^+]_{500}$ that was simultaneously measured (10^5 cm^{-3}) would tend to indicate a hydrogen concentration of the order of 5×10^5 cm^{-3} at the exobase, while our model hydrogen concentration would give $[H]_{500}$ in excess of 10^6 cm^{-3} on that day. Although this would suggest that the adopted model hydrogen concentrations are overestimates, another daytime limiting flux determination made by Brinton et al. [1969] for 2 March 1966 resulted in a much higher flux value. These authors have studied ion composition measurements made with Geoprobe rocket up to 630 km. At the higher altitudes covered by the experiment, the $[H^+]$ profile clearly departed from diffusive equilibrium, and the $[H^+]/[O^+]$ ratio was much higher than observed in Hoffman's experiment. By fitting the experimental data with theoretical $[H^+]$ and $[O^+]$ profiles generated through solution of the continuity equation, Brinton et al. derived an upward flux of protons of at least 1.5×10^8 cm^{-2}. The abundance of neutral hydrogen needed for fluxes of this magnitude approaches the value $[H]_{500} = 10^6$ cm^{-3}, which was given by direct mass spectrometer measurements for May 66 [Reber et al, 1968] and is in general agreement with Equation (49). On the basis of this limited comparison, it would seem that, in addition to its strong dependence on temperature, [H] also presents a high day-to-day variability which is not accounted for

in the assumed model.

The ionosphere is a very variable layer. Its evolution along the day, although almost always following the same general pattern, may differ strongly from day to day with no obvious reason. One should then be very cautious in interpreting observations from one given day for the purpose of drawing conclusions of a general nature about the ionospheric morphology. For this reason, the isolated observations mentioned above are not in themselves an appropriate proof that the daytime ionization flux in the upper F-region is normally a sizeable fraction of the limiting flux. They do, however, show that an upward flux of protons through the diffusive barrier at or close to the maximum rate can actually take place.

If a limiting flux were allowed to exist indefinitely on both ends of a midlatitude (closed) protonospheric tube, the tube would eventually become so full that diffusive equilibrium would be established at the barrier and the flux would cease. On the other hand, if the tube is relatively empty, an upward flow of protons will be set up across the diffusive barrier in order to approach that steady-state condition; as long as the tube is empty enough, the upward flux will approach the limiting value. In other words, the physical precondition that is necessary to set up a limiting flow across the diffusive barrier is the existence of low pressures in the protonosphere. It is therefore convenient to consider what normally happens in a closed protonospheric tube, in order to establish the viability of the limiting flux condition as the normal state of the diffusive barrier.

Park [1970] has studied whistler measurements of the electron content in magnetospheric tubes of force during an 8-day quiet period

following a magnetic storm that severely depleted the protonosphere. He found that the magnetospheric tube content maintained a steady daily gain during that period, even after the content had exceeded its monthly median value. The daily gain of about 5×10^{12} cm^{-2} day^{-1} was due to a transfer of ionization from the ionosphere to the protonosphere of at least 2×10^{8} cm^{-2} sec^{-1} during daytime, which was only partially returned to the ionosphere during the night. Park's measurements are consistent with calculations of the downward flux needed to maintain the nighttime ionosphere [Geisler and Bowhill, 1965]. Moreover, they strongly indicate that, as long as quiet conditions prevail, the magnetosphere will take in more ionization every day. What allows this situation to exist for most days is that occasionally the protonosphere will be rapidly depleted due to a magnetic disturbance.

The fact that the magnetospheric content is constantly kept below its daytime steady-state value means that an upward flow across the diffusive barrier will exist during the day. The steady daily gain in protonospheric tube content during an 8-day period reported by Park suggests a large departure of the protonosphere from equilibrium with the daytime ionosphere; consequently, the flux across the diffusive barrier is expected to approach the limiting value.

For the polar regions, the prevalence of the upper low pressure conditions is due to the fact that the magnetic lines are open. The resulting plasma escape flux (polar wind) was calculated by Banks [1969b] who found values in the range $2 - 7 \times 10^{8}$ cm^{-2} sec^{-1}, with the higher fluxes corresponding to lower temperatures. Unlike the daytime mid-latitude escape flux from the ionosphere to the protonosphere, the

polar wind contributes to the planetary loss of hydrogen.

It is our feeling that Park's findings, together with Bank's polar wind calculations and the ionic composition measurements by Hoffmann and others comprise a body of experimental and theoretical evidence sufficiently thorough to establish that:

a) daytime upward fluxes in the range $4 \times 10^7 - 4 \times 10^8$ cm^{-2} sec^{-1} from the ionosphere to the protonosphere can and do occur at midlatitudes;

b) at least sometimes the ionosphere is transferring ionization to the protonosphere at the limiting rate, and this may indeed be the state of the ionosphere more often than not;

c) at least from a theoretical point of view, the limiting fluxes should either decrease or stay approximately constant with increasing solar activity, depending on how strong the solar-cycle variation of $[H]_{500}$ really is.

With regard to this last item, the presently availabe direct experimental evidence is unclear, because most measurements were made during years of low solar activity. More recently, measurements of the vertical flux of O^+ ions in the upper F-region were made through incoherent scatter radar techniques for a midlatitude location [Evans, 1971]. Under steady-state, the flux of O^+ ions that enters the charge-exchange region is practically equal to the proton flux that emerges from it and crosses the diffusive barrier. Therefore, it is possible to infer the magnitude of the upgoing proton flux from these measurements. Accordingly, Evans estimated daytime upward proton fluxes in the range 3×10^7 -

10^8 cm^{-2} sec^{-1} for 24 March 1970 at Millstone Hill. Table 2 summarizes the limiting flux values that are given by or can be inferred from the investigations mentioned above. The scarcity of information about these fluxes is obvious; however, the available results are not inconsistent with a general decrease of the limiting flux with solar activity. This decrease is theoretically predicted for the assumed neutral hydrogen concentrations and model atmosphere (Figure 11).

TABLE 2

References	Method	Date	S	Limiting Fluxes	
				Measured	Theoretical
Park, 1970	Whistler Measurements	16-24 June 65	75-80	$2\text{-}4 \times 10^8$	1.9×10^8
Hoffmann, 1969	Ionic Composition	15 Aug. 66	91	5×10^7	2.1×10^8
Brinton, 1969	Ionic Composition	2 Mar. 66	78	$>1.5 \times 10^8$	1.9×10^8
Evans, 1971	Incoherent Scatter	24 Mar. 70	171	$3 \times 10^7\text{-}10^8$	7×10^7

Table 2 - Experimental determinations of Ψ_D are compared with values of this flux predicted theoretically for the solar-cycle epoch of the measurement and model ionosphere and neutral atmosphere.

D. Influence of an Outgoing Topside Flux on the F2-Layer Concentrations: The Upper Sink

In the previous section, it was shown that the presently available information about the daytime ionization fluxes from the ionosphere to the protonosphere is consistent with the assumption that these fluxes attain their limiting values. Moreover, it puts these values roughly in the range $4 \times 10^7 - 4 \times 10^8$ cm^{-2} sec^{-1}. In the remaining part of this work, such assumption will be adopted as a working hypothesis. As a result, it will be shown that a theory can be developed to explain many important features of the daytime F2-layer, particularly its variations along the year and along the solar cycle.

Under steady-state conditions, the proton flux that crosses the diffusive barrier has to be fed by an upward flow in the same amount of O^+ ions from the F2-layer. Therefore, in order to study the impact of the limiting outflow condition on the ionosphere, one must start by studying the influence of an upward topside flux of O^+ ions on the F-region.

In the topside of the F-layer, both the production and loss terms are quite small when compared with the diffusion term in the continuity equation, so that the plasma concentration profile will approach the one given in Equation (14) (see Chapter 2).

$$n = n_1 e^{-h/2H} + n_2 e^{-h/H_n}$$

where H_n is scale height of atomic oxygen.

From Equation (9), one can derive the flux Ψ corresponding to such distribution of ionization:

$$\Psi = D \left[- \frac{n}{2H} - \frac{\partial n}{\partial h} \right] =$$

$$= D\, n_2 \left[\frac{1}{H_n} - \frac{1}{2H} \right] e^{-h/H_n}$$

$$= n_2 \frac{-g}{\nu(0^+, 0)} (r - 1)\, e^{-h/H_n} = n_2\, e^{-h/H_n}\, v_L$$

where v_L is the limiting velocity of 0^+ ions when diffusing upwards through the neutral oxygen atoms of the thermosphere. Since v_L increases with height with scale height H_n, the flux as expressed above is height-independent (as it should be, for continuity reasons) and depends only on the intensity n_2 of the full-flow mode.

The limiting velocity v_L is determined at each point by the collision frequency $\nu(0^+,0)$, which basically depends on the concentration of neutral oxygen atoms. The flux Ψ during the day has to match the upgoing flux of protons Ψ_D through the diffusive barrier. Therefore, the magnitude n_2 of the full flow mode is completely specified by the rate of ionization transfer from the ionosphere to the protonosphere. As for the diffusive equilibrium mode, its magnitude n_1 can be estimated if the transition between the diffusion and photochemical equilibrium regions is initially assumed to occur abruptly at a given height. If the reference height $h = 0$ is positioned at this transition height and if n_o is the concentration given there by photochemical equilibrium, then:

$$n_1 = n_o - n_2 = n_o - \frac{\Psi}{v_{Lb}}$$

where v_{Lb} is the limiting velocity at the base of the diffusion region. The distribution specified by Ψ and n_o in the diffusion region will then be:

$$n = n_o e^{-h/2H} - \frac{\Psi}{v_{Lb}} [e^{-h/2H} - e^{-h/H_n}]$$

At any fixed point, the concentration n will vary linearly with the flux Ψ. For $\Psi = 0$, the distribution will be given by $n = n_o \exp(-\frac{h}{2H})$. Since n_1 cannot be negative (a negative n_1 would produce negative concentration values at high altitudes), there is a maximum possible value for n_2 and therefore for the flux of O^+ ions:

$$n_2 \leq n_o$$

$$\Psi \leq n_o v_{Lb} = \Psi_L$$

The flux Ψ_L represents a limiting upward flux of O^+ ions and should be distinguished from the limiting proton flux Ψ_D across the diffusive barrier. In fact, since the proton flux must be matched by an equal upgoing flux of O^+ ions, we must have $\Psi_D \leq \Psi_L$ if the flow of protons is to be given by Ψ_D.

For $\Psi = \Psi_L$, the distribution will be given by $n = n_o \exp(-\frac{h}{H_n})$. At any fixed height h, the variation of $[O^+]$ with the upgoing flux Ψ will be given by a linear decrease as shown on Figure 13.

The accuracy of this approach is naturally limited by the fact that the transition between photochemical equilibrium and diffusion control is actually a gradual transition. Banks et al. [1969a] have

calculated $[O^+]$ profiles resulting from different values of Ψ/Ψ_L, for distributed losses and production. Figure 14, which was partly reproduced from their paper, shows that the O^+ concentrations both around the peak and in the topside can be severely reduced by an upward flux, provided that this flux be a significant fraction of the limiting O^+ flux Ψ_L.

We have seen on the previous section that the limiting daytime flux of protons Ψ_D across the diffusive barrier is generally in the range $4 \times 10^7 - 4 \times 10^8$ cm^{-2} sec^{-1}. In order to determine whether these fluxes have any sizeable effect on the ionospheric plasma distribution and abundance, we must now estimate how they compare with the maximum O^+ flux Ψ_L.

The theoretical determination of Ψ_L has been made by Geisler and Bowhill [1965]. In order to visualize the meaning and magnitude of this flux, it is convenient to consider initially a hypothetical layer in which the effect of recombination is lumped into a single point in which the concentration is set equal to zero as a lower boundary condition. This point, which is taken as the reference height h_o, roughly corresponds to the point where the effect of losses is strong enough to hinder the ionic diffusion, and therefore would be properly located around 200 km. Since at this height the optical depths for the solar EUV are still much less than unity, it is reasonable to assume a production function exponentially decreasing with height above the lower boundary. The steady-state continuity-momentum transport equation for this layer will then be, for an isothermal case and in the absence of neutral winds or electromagnetic drifts:

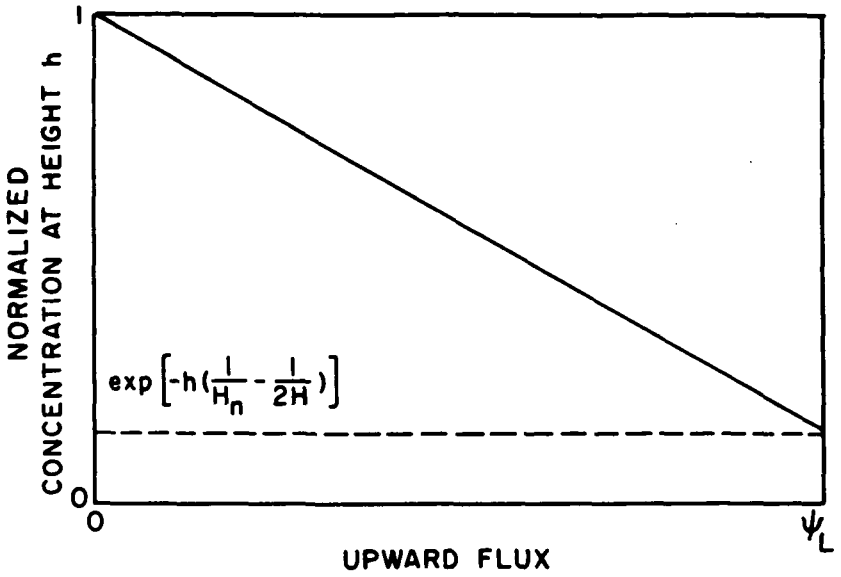


Fig. 13. LOCAL DEPLETION BY AN UPWARD FLUX. The variation of $[O^+]$ with Ψ is represented for the hypothetical case of an abrupt transition between photochemical and diffusion-controlled regimes.

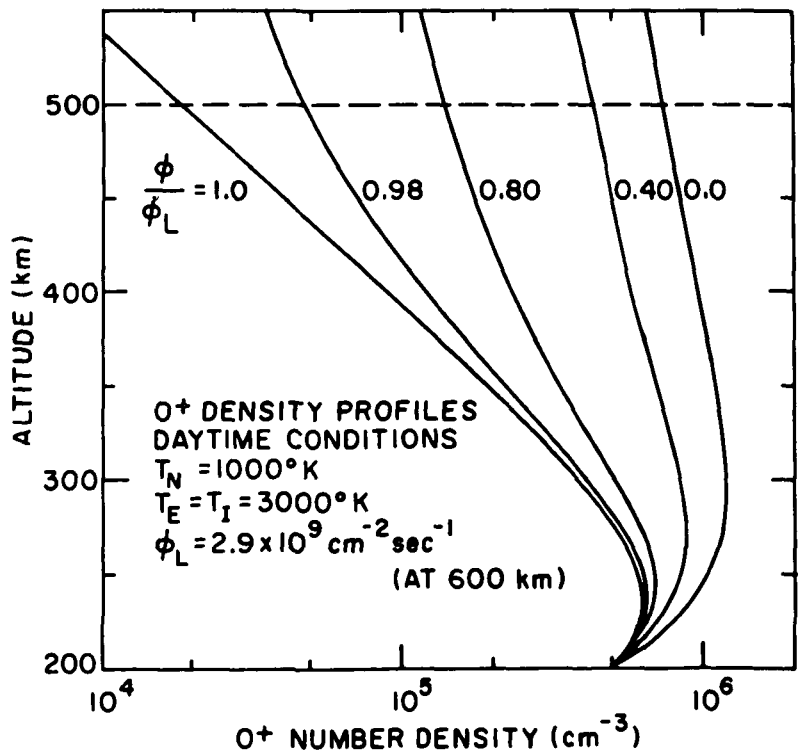


Fig. 14. IONOSPHERIC DEPLETION. This figure was reproduced from a paper by Banks [1969]. Notice how the concentration at, say, 500 km, varies almost linearly with Ψ, as in the more crude model depicted in the previous figure.

$$q_o e^{-z} + \frac{D_o}{H_n^2} e^z \left\{ \frac{\partial^2 [O^+]}{\partial z^2} + \left(1 + \frac{1}{r} \right) \frac{\partial}{\partial z} [O^+] + \frac{1}{r} [O^+] \right\} = 0 \qquad (50)$$

where:

$$r = \frac{T_e + T_i}{T_n}$$

$$z = \frac{h - h_o}{H_n}$$

The constants q_o and D_o are respectively the photoionization rate of oxygen atoms at the base and the diffusion constant at the base. The general solution of this equation is:

$$[O^+] = C_1 e^{-z/r} + C_2 e^{-z} - \frac{r}{2r - 1} \frac{q_o H_n^2}{D_o} e^{-2z}$$

The O^+ flux $\Psi(z)$ corresponding to this distribution can be derived from Equation (9) as:

$$\Psi = - \frac{D_o e^z}{r H_n} \left\{ [O^+] + r \frac{\partial}{\partial z} [O^+] \right\} =$$

$$= \frac{r - 1}{r} \frac{D_o C_2}{H_n} - q_o H_n e^{-z}$$

For increasing values of z, $\Psi(z)$ will tend to the following value:

$$\Psi_\infty = \frac{r - 1}{r} \frac{D_o C_2}{H_n} \qquad (51)$$

If an upper flux Ψ_∞ is now specified as an upper boundary condition, the constants C_1 and C_2 can be determined as:

$$C_2 = \frac{r}{r-1} \frac{H_n}{D_o} \Psi_\infty$$

$$C_1 = \frac{r}{2r-1} \frac{q_o H_n^2}{D_o} - C_2 \qquad \text{(because } [O^+] = 0 \text{ at } z = 0\text{)}$$

The ionic concentration profile resulting from the imposition of a flux Ψ_∞ at the topside will then be:

$$[O^+] = \frac{H_n}{D_o} \left\{ \left[\frac{r}{2r-1} q_o H_n - \frac{r}{r-1} \Psi_\infty \right] e^{-z/r} + \frac{r}{r-1} \Psi_\infty e^{-z} - \frac{r}{2r-1} q_o H_n e^{-2z} \right\} \tag{52}$$

Again, as in the case of diffusion without production, $[O^+]$ will vary linearly with Ψ_∞ for any fixed z. The coefficient of $\exp(-\frac{z}{r})$ cannot be negative, since this term will dominate the distribution at high altitudes. This condition reflects the existence of the maximum O^+ flux Ψ_L, which is given in this case by:

$$\Psi_L = \frac{r-1}{2r-1} q_o H_n \tag{53}$$

Since $q_o H_n$ is equal to the integrated production above h_o, it follows that the maximum upward flux that can be supported by this kind of layer is between one third and one half of the total number of ions produced per unit time above the lower sink (for $r \geq 2$). Since the layer itself is lossless, this also means that at least one half to two

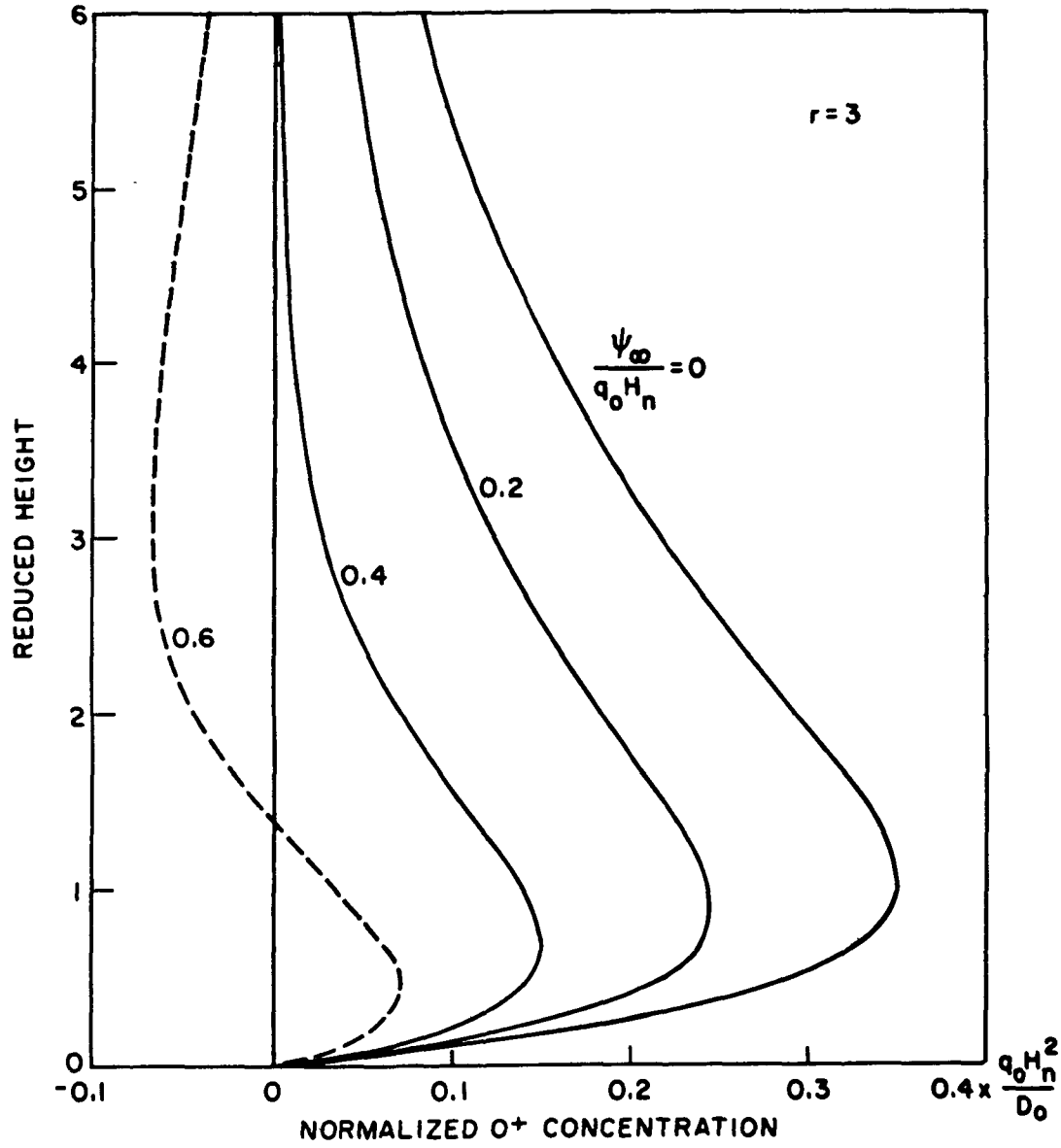


Fig. 15. CONCENTRATION PROFILES IN A LOSSLESS LAYER. The $[O^+]$ profiles given for r = 3 by Eq. (52) are plotted for Ψ_∞ = 0, 0.2 $q_o H_n$, 0.4 $q_o H_n$, and 0.6 $q_o H_n$. Clearly, the last solution is not realizable. In this case (r = 3), the maximum allowable topside flux is 40% of the total production above the sink.

thirds of the total ionic production, depending on the ratio between plasma and neutral temperatures, must go down to the lower sink and recombine.

The situation described above is roughly approached by the real F2-layer, since the recombination coefficient, varying in space with the scale height of the N_2 or O_2 molecules, has a much stronger height gradient than either the atomic oxygen photoionization rate or the diffusion coefficient, which vary with the scale height of the O atoms. The height of the sink will naturally depend on the amount of molecular neutral species present. The higher the $[N_2]/[O]$ and/or the $[O_2]/[O]$ ratio, the higher will the sink be with respect to the O^+ production profile, thereby leaving less ionization able to flow upwards.

The approach described above is useful in that it illustrates some qualitative aspects of the problem. It shows that the maximum available O^+ flux is a significant fraction of the total production of O^+ ions above the loss region. It can be seen in Figure 2 that this is only a small part of the total electron production. Most of the electron production is associated with the production of N_2^+ and O_2^+ ions, which contribute very little to the total ionospheric electron content. Of the remaining production of O^+ ions, most of the ions are produced below 200 km, i.e., in a region of heavy recombination where they are rapidly lost without having a chance to reach the F2-layer.

The O^+ ions produced above the recombination region are the main source for maintenance of the daytime F-2 layer, although they correspond to only a small part of the total electron production, which is of the order of 4×10^{10} cm^{-2} sec^{-1} for medium solar activity. Accordingly, an upward flux comparable in magnitude with the production of

these ions will be able to deplete the F2-layer, especially in the topside, even though it may be small when compared with the total electron production. For this reason the fact that the proton fluxes across the diffusive barrier are in the range $4 \times 10^7 - 4 \times 10^8$ cm^{-2} sec^{-1}, therefore being very small when compared with the total electron production, does not mean that they are unable to cause appreciable depletion in the concentrations of the F2-layer.

In order to assess the ability of these fluxes to deplete the F2-layer, a quantitative estimate must be made of the maximum O^+ flux Ψ_L. For this purpose, calculations based on a lossless layer are not suitable, because of the difficulty in specifying the right position for the lower sink and the sensitivity of the results to this parameter. A term representing distributed losses must then be added to the continuity equation. This is done by adding a term in the form $-\beta_o \exp(-az)[O^+]$ to the left-hand side of Equation (50), where a is 1.75 or 2 according to whether recombination is dominated by the N_2 or the O_2 molecules (uncertainties in the corresponding reaction rates make it impossible to know, at present, which of these two species dominates the O^+ recombination processes in the height range of interest). The appropriate lower boundary condition to this equation is that $[O^+]$ is required to approach zero when z tends to $-\infty$.

With the addition of the recombination term, Equation (50) will become an inhomogeneous second-order linear differential equation in which one of the coefficients is an exponential function of z. The fact that the coefficients are not all constant will considerably complicate the solution. Geisler and Bowhill [1965] have solved this equation for Ψ_∞ as a parameter, and obtained for the concentration profile

an equation similar to Equation (52), but with the coefficients being given by complicated functions of q_o, β_o, D_o and H_n, and the exponentials being replaced by infinite series of exponential functions of different linear functions of z. As in the previous cases, however, the variation of $[O^+]$ with Ψ_∞ for any fixed z is linear. Moreover, a maximum possible value for Ψ_∞ that will yield positive concentration values everywhere is also apparent. For r=2, and assuming that β follows the scale height of the O_2 molecules, their solution gives:

$$\Psi_L = 0.468\ q \left(\frac{2H_n D}{\beta}\right)^{1/3}$$

The expression above is height-independent, so that the parameters involved may refer to any particular point inside the F2-layer. For $r > 2$ (plasma temperature higher than T_n), the limiting flux will be higher. Returning to the case of the lossless layer, one can see from Equation (53) that for r = 3, the flux will increase by 20%, and for $r = \infty$ by 50% with respect to its value for r = 2. It seems reasonable to assume that the variation of Ψ_L with r will be of the same order for a layer with distributed losses, so that for r in the vicinity of 3 or 4:

$$\Psi_L \cong 0.6\ q \left(\frac{2H_n D}{\beta}\right)^{1/3} \qquad (54)$$

The main limitation on the accuracy of this determination derives from the omission in the continuity equation of a term expressing the effect of neutral winds and/or external electric fields. During the day, a wind from the equator to the poles blows in the thermosphere. Such a wind tends to drive the ionization down along the magnetic lines to the

recombination region, thereby lowering the amount of ionization that can flow upwards without recombining. For this reason, the value of Ψ_L given in Equation (54) should be regarded as an overestimate, since it was derived for an assumed absence of neutral winds.

Writing Equation (54) in terms of neutral temperatures and concentrations:

$$\Psi_L = 1.17 \times 10^{12} P \left\{ \frac{[O]^2}{[O_2]} \right\}^{1/3} \left(\frac{T_n}{1000} \right)^{1/3} \left(\frac{T}{1000} \right)^{1/3} \qquad (55)$$

where:

$$D = 1.29 \times 10^{16} \frac{T}{[O]}$$

$$\beta = 2 \times 10^{-11} [O_2]$$

$$P = \frac{q}{[O]} = \text{rate of photoionization of oxygen atoms}$$

Since the scale height of the oxygen atoms is twice that of the O_2 molecules, the ratio $[O]^2/[O_2]$ is height-independent. Moreover, assuming that the neutral concentrations at the turbopause (around the 120 km level) are constant throughout the year and the solar cycle, this ratio may be expressed by a constant, which for the CIRA model atmosphere is 7.7×10^{10} cm^{-3}. The main variation of Ψ_L is then due to the photoionization rate P, which is proportional to the incident solar EUV flux. For $T_n = 1000°$ k and $T = 1800°$ k:

$$\Psi_L = 6 \times 10^{15} P \qquad (56)$$

Recent satellite measurements of the solar EUV flux made over an extended period have provided accurate values for the solar radiation in this band when S is about 150 [Hall and Hinteregger, 1970]. For the EUV flux values given by those measurements and the ionization cross-sections used by Hinteregger et al. [1965], the resulting value for P is 2.5×10^{-7} sec^{-1}. In addition, Hall et al. [1969] reported a survey of several rocket measurements of the solar EUV flux performed between 1961 and 1968 which indicates that the intensity of the solar EUV radiation increases by 50% when S varies from 75 to 150. On the basis of these results, we may then estimate P to be about 1.6×10^{-7} sec^{-1} at solar cycle minimum and 3×10^{-7} sec^{-1} for S = 200. The corresponding estimates for Ψ_L are 1.0×10^9 cm^{-2} sec^{-1} for solar cycle minimum and 1.8×10^9 cm^{-2} sec^{-1} for S = 200.

It is of interest to notice that, due to insufficient information about the solar EUV flux and its solar cycle variation, it has been customary in the past to scale this quantity proportionally to the solar radio flux at 10.7 cm., since available measurements of the solar EUV flux existed only for solar-cycle minimum. Such procedure seems questionable in the light of the recent studies mentioned above. For this reason, values of P used in the past were somewhat higher than those adopted in this study. Banks [1968b], for example, used $P = 5 \times 10^{-7}$ sec^{-1}, which would give $\Psi_L = 3 \times 10^9$ cm^{-2} sec^{-1}. The late experimental determinations of the solar EUV flux favor lower values for P and hence for Ψ_L, indicating that the F2 region is more sensitive to an upgoing flux of O^+ ions than was thought before.

The daytime O^+ fluxes which are compatible with the observations are in the range 4×10^7 - 4×10^8 cm^{-2} sec^{-1}, with the higher values

possibly associated with solar cycle minimum conditions. On the other hand, the limiting O^+ fluxes are below 1 - 2 x 10^9 cm^{-2} sec^{-1}, with the lower values occurring during solar cycle minimum. The resulting Ψ/Ψ_L ratio may then be expected to be in excess of 40% at solar cycle minimum. As seen on Figure 14, this will cause strong reduction in the $[O^+]$ values in the topside. This will cause depletion in the ionospheric electron content, and, to a smaller extent, in the peak concentration, and a decrease in the height of the layer. With increasing solar activity, the higher exospheric temperatures will severely reduce the neutral hydrogen concentration, thereby reducing the upgoing fluxes. On the other hand, Ψ_L will increase due to higher production rates, so that Ψ/Ψ_L will decrease with solar activity. This will tend to eliminate the depletion of ionospheric concentrations due to the limiting outflow condition during solar cycle maximum, especially during parts of the year in which the exospheric temperatures are higher.

We are justified to state at this point that under certain conditions, and especially for lower exospheric temperatures (high Ψ) and smaller production rates (small Ψ_L), the limiting upward flow of protons through the diffusive barrier will effectively deplete the daytime F2-region at midlatitudes. Under these conditions, the protonosphere will actually act as an upper sink to the ionosphere, draining its ionization in the form of protons.

In order to estimate the effectiveness of the upper sink in controlling the ionospheric electron content, it is interesting to use the concentration profiles derived for the lossless layer. Calling n_T the columnar content of O^+ ions, one has from Equations (52) and (53):

$$n_T = H_n \int_0^\infty [O^+]\, dz = \frac{q_o H_n^3}{D_o} \frac{r}{2r-1} \left\{ (r - \frac{1}{2}) - (r - 1) \frac{\Psi_\infty}{\Psi_L} \right\} \qquad (57)$$

Therefore, the ionospheric content will vary linearly with the outgoing flux Ψ_∞. If n_{To} is the content resulting from a zero upward flux and n_{TL} from $\Psi_\infty = \Psi_L$, then:

$$\frac{n_{To}}{n_{TL}} = 2r - 1 \qquad (58)$$

For r = 3, the content may then vary by a factor of 5 depending on the value of the upgoing flux. As seen on Figure 15, the corresponding effect on the peak concentration is much more modest, and the variation in the height of the peak is a fraction of the neutral scale height H_n.

A final comment should be added to this section, regarding the reduction in Ψ_L caused by daytime poleward neutral winds blowing in the thermosphere. Although it is difficult to evaluate this effect properly, one can have an idea of the epoch of the solar cycle in which it might become important, by looking at the long-term variations in the intensity of the winds. As of now, very little experimental information is available on this subject, but some theoretical studies have been made. Cho and Yeh [1970] have derived wind velocities for different times of the year and degrees of solar activity by solving simultaneously the equations of motion for the neutral atmosphere and the ionospheric plasma. They have found that during solar cycle minimum the daytime wind velocities are much higher than for median or maximum solar activity, due to reduced ion drag. Daytime meridional wind velocities between 100 and 200m

sec^{-1} were derived for solar cycle minimum. For midlatitudes, winds of this magnitude will cause effective downward vertical drifts faster than 30 m/sec, which are certain to have a strong impact on the F layer. For this reason, the solar cycle minimum value of 10^9 cm^{-2} sec^{-1} for Ψ_L may be greatly overestimated, while the solar cycle maximum result is probably closer to reality. A lower boundary for Ψ_L for low solar activity, however, is provided by Park's suggestion of an upward proton flux of $2 - 4 \times 10^8$ cm^{-2} sec^{-1}, since this flux must be supplied by an equal amount of O^+ ions, which cannot exceed Ψ_L. Therefore, we can reasonably place Ψ_L in the range $4 \times 10^8 - 10^9$ cm^{-2} sec^{-1} for solar cycle minimum.

E. The Self-Consistent Limiting Proton Flux

In section A of this chapter, it was shown that the limiting proton flux depends, among other parameters, on the amount of O^+ ions existing on the topside of the F layer. An expression was derived for which Ψ_D is proportional to $[O^+]_{500}$ raised to the power $s = \frac{15\ r}{16(r+1)}$. This dependence of Ψ_D on $[O^+]_{500}$ was also derived by Geisler [1967], for $r = 2$. Using exospheric hydrogen concentrations of the order of 10^4 cm^{-3}, he estimated the limiting proton flux to be about 1.5×10^7 cm^{-2} sec^{-1}, a flux that would obviously have little effect on the F2-region concentrations. For the model hydrogen concentrations adopted in this work, however, we have derived limiting proton fluxes that are actually able to deplete the F2-layer concentrations, especially in the topside. This depletion will greatly reduce the parameter $[O^+]_{500}$, which is one of the controlling factors in determining Ψ_D. Therefore, if one uses in Eq. (47) a value of $[O^+]_{500}$ which is compatible with a zero flux condition,

the resulting value of Ψ_D will in general be inconsistent with the input conditions on which it is based.

In order to obtain a self-consistent value for Ψ_D, one must introduce in Eq. (47) the dependence of $[O^+]_{500}$ on Ψ_D, assuming the limiting outflow condition (i.e., assuming that the flux at the topside F region is given by Ψ_D). From Figure 14, it is clear that, for values of Ψ_D up to 98% of the limiting O^+ flux Ψ_L, a good approximation is:

$$[O^+]_{500} = n_o \left[1 - \frac{\Psi_D}{\Psi_L}\right]$$

where n_o is the O^+ concentration at 500 km that would be produced by a zero flux condition at the top of the F layer.

Substituting in Eq. (47a), one has, for r = 4:

$$\frac{\Psi_D}{\Psi_L} = X\,[H]_{500}\left(1 - \frac{\Psi_D}{\Psi_L}\right)^{3/4} \qquad (59)$$

where:

$$X \equiv 5.6 \times 10^{-5} \frac{n_o^{3/4}}{\Psi_L [O]_{500}^{1/4}}\; T_n^2\, T_i^{-1/2} \sin^2 I$$

The ratio Ψ_D/Ψ_L is very informative, because it shows the degree to which the ionosphere is being depleted due to the limiting outflow. For this reason, this ratio will be called the depletion factor. Figure 16 shows the variation of this quantity with $X[H]_{500}$, as given by Equation (59). When $X[H]_{500} > 1$, Ψ_D/Ψ_L will approach unity, in which case the ionosphere will be in a photodepletive regime. On the other hand, when $X[H]_{500} < 0.1$, the ionosphere will be almost unaffected by the outgoing flux of protons, since the latter will comprise only a small fraction of the limiting O^+ flux Ψ_L.

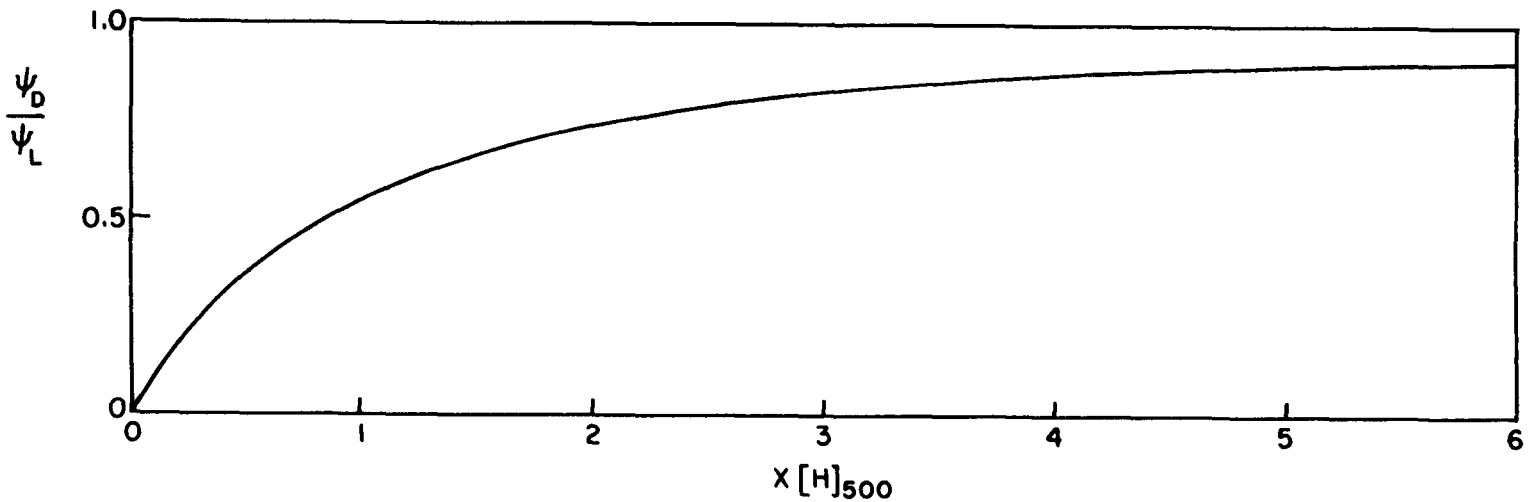


Fig. 16. SELF-CONSISTENT LIMITING FLUX. Notice that, for large enough values of $[H]_{500}$, the limiting proton flux Ψ_D will actually be limited by the maximum available supply Ψ_L of O^+ ions.

Unfortunately, the knowledge of both X and $[H]_{500}$, and particularly of the latter, are subject to considerable uncertainties. For this reason, it is not possible to foretell with confidence whether the daytime ionosphere is in a photodepletive regime or not, especially because the range of possible values for $X[H]_{500}$ encompasses unity. Table 2 shows the values of $X[H]_{500}$ that result from the adoption of typical estimates for the quantities involved, and for the model hydrogen concentrations given in Equation (49). The parameter X does not seem to have a strong solar-cycle variation, varying by less than a factor of two from solar cycle minimum to maximum. Most of the variation in the state of depletion of the ionosphere throughout the solar cycle, if it exists, will then be due to changes in $[H]_{500}$, which decreases by more than an order of magnitude from minimum to maximum of the solar cycle.

Table 3

	S = 75	S = 175	Units
n_o	2×10^5	1.0×10^6	cm^{-3}
$[O]_{500}$	1.75×10^7	8.73×10^7	cm^{-3}
Ψ_L	5×10^8	2×10^9	$cm^{-2}\ sec^{-1}$
T_n	1061	1576	°K
T_i	2000	3000	°K
$\sin^2 I$	.8	.8	
X	3.4×10^{-7}	3.2×10^{-7}	cm^3
$[H]_{500}$	2×10^6	10^5	cm^{-3}
$X[H]_{500}$	.68	3.2×10^{-2}	

If neutral hydrogen is present in the thermosphere in the amounts predicted by Eq. (47), it is clear that changes will occur in the state of depletion of the ionosphere during a solar cycle. A decrease in $X[H]_{500}$ from 0.68 to 0.03 is given in Table 3 from minimum to maximum of the solar cycle, showing that the depletion effect tends to be removed with growing solar activity. This helps to explain why the solar-cycle variation in ionospheric electron content is much stronger than the corresponding 50% variation in solar EUV flux which can be inferred from measurements of this quantity [Hall et al., 1969].

When $X[H]_{500}$ falls from 0.68 to 0.03, the ratio Ψ_D/Ψ_L falls from 0.44 to 0.03, i.e., it is reduced by about 16 times. Since the corresponding estimated increase in Ψ_L is a factor of 4, a decrease by a factor of 4 is predicted for Ψ_D from solar cycle minimum to maximum. Since Ψ_D is the flux of protons through the diffusive barrier, it constitutes the only source of ionization for the midlatitude protonospheric tube (disregarding the effect of cross-L diffusion, which is thought to be unimportant). The mechanisms by which the protonospheric plasma is transferred back to the nighttime ionosphere or to outer space are not sufficiently well known to permit assessment of their variation along the solar cycle; however, it seems reasonable to assume that the tubular losses are roughly proportional to the content of the tube. If this is so, a reduction in Ψ_D by a factor of 4 would mean that the mean protonospheric tubular content should decrease by the same factor from minimum to maximum solar activity. This is in disagreement with the results of whistler observations, which suggest that no appreciable solar-cycle variation exists in the tubular protonospheric content [Smith, 1961];

it does explain, however, the surprising disparity between the solar-cycle variations in the ionospheric and the protonosphere.

The discrepancy between this theory and the observed lack of solar-cycle variation in the protonosphere can be removed if either higher values or a smaller solar-cycle excursion is adopted for $[H]_{500}$. For higher values, it can be seen on Figure 16 that the assumed 20:1 variation in $[H]_{500}$ will correspond to smaller variations in Ψ_D/Ψ_L, and therefore in Ψ_D. If the assumed values of $[H]_{500}$ are multiplied by 8 for example, $X[H]_{500}$ will vary from 5.4 to 0.25 in a solar cycle and Ψ_D/Ψ_L will correspondingly vary from 0.91 to 0.22, requiring very little variation in Ψ_D. Such large values of $[H]_{500}$, however, are not expected from the existing information about this quantity.

Up to this point, we have considered solar-cycle variations in $[H]_{500}$ that result from corresponding variations in neutral temperature through the strong dependence of the rate of thermal escape on T_n. The neutral temperature, however, is also known to present semi-annual variations, with maxima in April and October [Paetzold et al., 1960]. The semi-annual variation in T_n has now been studied for more than a decade, and its amplitude was found to be roughly proportional to S, the solar radio flux at 10.7 cm [Jacchia et al., 1969]. The annual maximum-to-minimum difference in temperature due to this effort is about 100°K for solar cycle minimum, and may exceed 200°K for a year of high solar activity. For the model hydrogen concentrations specified in Eq. (47), such temperature variations will account for semi-annual fluctuations in $[H]_{500}$ with a maximum-to-minimum ratio of 1.75:1 at solar cycle minimum and 3:1 at solar cycle maximum. This will happen because the larger neutral temperatures occurring in the thermosphere around April and October will increase

the speed of thermal escape of the hydrogen atoms, thereby reducing their abundance.

The variation in the ionospheric depletion factor (and therefore in the state of depletion of the ionosphere) due to the semi-annual variation in $[H]_{500}$ will depend on the mean annual value of $X[H]_{500}$. If this value is well below unity, Ψ_D/Ψ_L will vary in the same proportion as [H], but its absolute value throughout the year will be small; in this case, the ionosphere will remain almost undepleted during the entire year, and no sizeable semi-annual variation in electron content would occur as a result of this effect. On the other hand, if $X[H]_{500}$ is well above unity, the depletion factor becomes rather insensitive to the amount of neutral hydrogen present in the thermosphere (see Figure 16); in this case, the ionosphere would remain essentially a photodepletive layer throughout the year, regardless of how much hydrogen is present, since the flux is being limited by the maximum O^+ flow Ψ_L. It is only when $X[H]_{500}$ is of the order of unity that a 3:1 variation in $[H]_{500}$ will produce a sizeable ionospheric effect, due to changes in the depletion factor. During April and October, when $[H]_{500}$ is lower, the depletion factor is also lower and therefore the electron content should be higher.

For the values of $X[H]_{500}$ given in Table 3, one should then expect semi-annual variations in electron content during solar cycle minimum, but not during solar cycle maximum. In actuality, however, just the opposite occurs: as will be shown in the next chapter, the ionospheric electron content does present a semi-annual variation with maxima in April and October, but this effect tends to disappear during

the years of low solar activity. The uncertainties in X and $[H]_{500}$, however, are such that it is possible to fit the observed behavior with this theory by choosing an ad hoc model for [H] (and its variation with T_n) within the range of possibilities consistent with present knowledge of this quantity. Consider, for instance, that the hydrogen concentrations are 3 times larger than given in Eq. (47); which seems possible since Reber's measurements of [H], which fit into the values given by that equation for low solar activity, are claimed to be correct with a factor of 3 [Reber et al., 1967]. The parameter $X[H]_{500}$ will then vary from 2 to 0.1 from solar cycle minimum to maximum. One might then consider $X[H]_{500}$ having 6 month-period fluctuations between 1.5 and 2.5 at low solar activity year and between 0.33 and 1 in a year of medium-to-high solar activity. The corresponding fluctuations in the ionospheric depletion factor are shown on Figure 17. It can be seen that in this case the weaker semi-annual variation in the depletion factor is predicted for solar cycle minimum, in agreement with the observations. With increasing solar activity, the ionosphere becomes more sensitive to variations in [H] and larger semi-annual excursions in the depletion factor occur. When $X[H]_{500}$ varies between 1 and 0.33, for instance, the depletion factor will vary between 0.54 and 0.26; this should have a noticeable effect on the electron content. It is difficult to tell the precise intensity of this effect in the real ionosphere. One may estimate its importance by considering the lossless layer with a lower sink, in which case the content will vary by a factor of 5 (for r = 3) when the depletion factor varies between 0 and 1.

For very high solar activity, even for the larger [H] values suggested above, $X[H]_{500}$ will fall to one-tenth of unity if the temperature

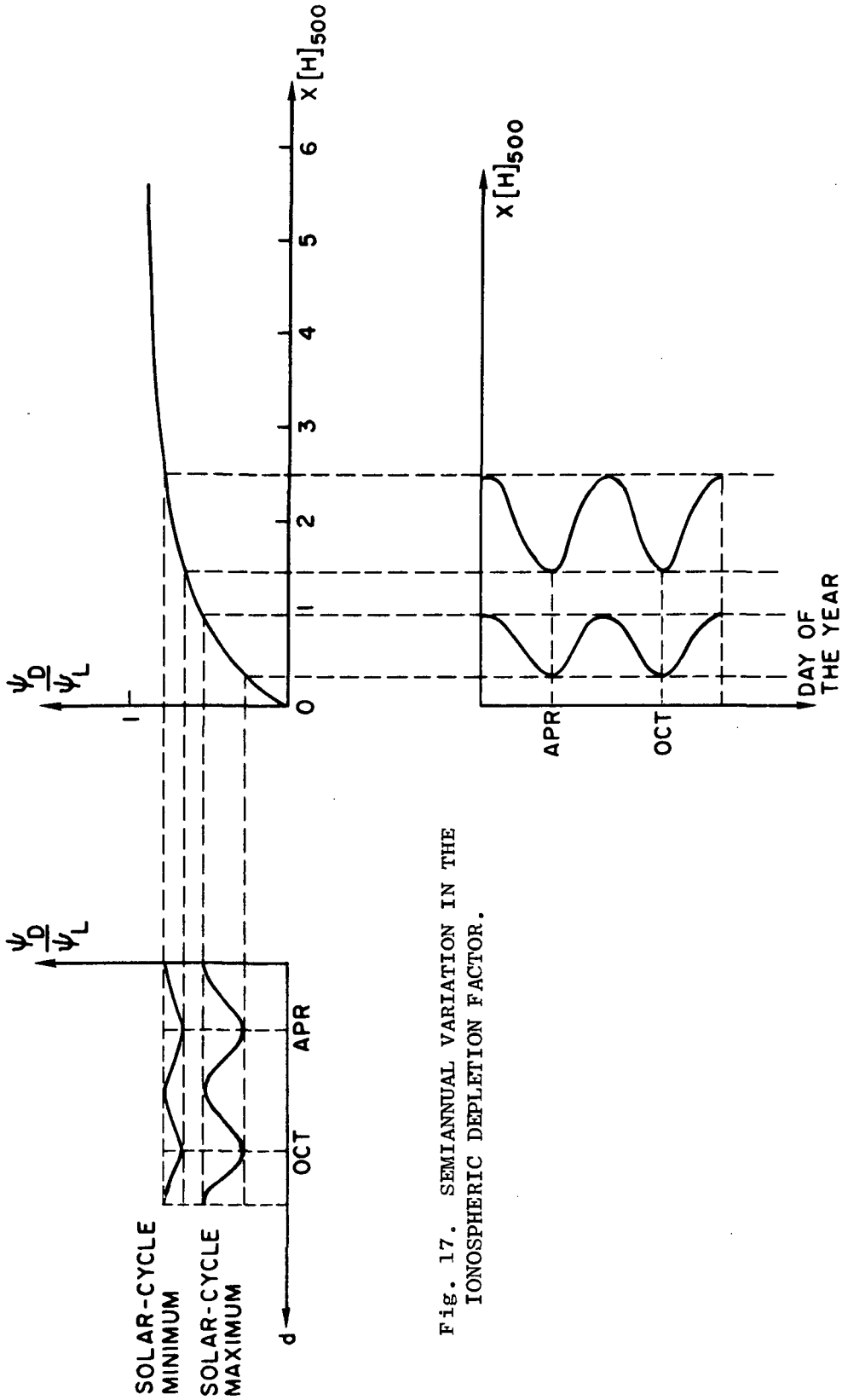


Fig. 17. SEMIANNUAL VARIATION IN THE IONOSPHERIC DEPLETION FACTOR.

variation given by Eq. (47) is adopted. The ionospheric semi-annual variations should then taper off, due to the prevalence of a low depletion factor throughout the year. This is not consistent with the observations, which show that the semi-annual variation in electron content persists through the peak of the solar cycle. This discrepancy can be removed if one considers a somewhat slower temperature variation in $[H]_{500}$ than that given by Eq. (47), or a smaller solar-cycle excursion in neutral temperature, for example. The resulting smaller solar-cycle excursion in $[H]_{500}$ is also indicated by the lack of solar cycle variation in the protonospheric tubular content.

Page Intentionally Left Blank

Chapter 5

LONG-TERM VARIATIONS IN IONOSPHERIC ELECTRON CONTENT

The present chapter is concerned with the study and interpretation of a large amount of ionospheric data collected at Stanford, Calif., from October 64 to December 69 on an almost continuous basis, thereby covering roughly one half of a solar cycle. The data consist of measurements of the Faraday rotation between a geostationary satellite and the Stanford station. The Faraday rotation Ω can be expressed as [Browne et al., 1956]:

$$\Omega = \frac{K}{f^2} < B \cos \theta \sec \chi > \int n_e \, dh$$

where K is 2.36×10^4 in MKS units, f is the frequency of the wave in Hz, θ is the angle between the geomagnetic field and the direction of propagation and χ is the zenithal angle of the direction of propagation. The mean $<B \cos \theta \sec \chi>$ is weighted by the function $n_e(h)$ along the path of propagation, and is equal to the value of B cos θ sec χ at an altitude of approximately 350 km in that path [Yeh and Gonzalez, 1960].

The use of this relation between the vertical columnar electron content and the Faraday rotation of a linearly polarized wave traveling through the ionosphere allows the evaluation of the vertical columnar ionospheric electron content from continuous monitoring at the ground of the polarization angle of a linearly polarized signal continuously transmitted by a geostationary satellite [Garriott et al., 1965].

The technique mentioned above has been used to measure the ionospheric electron content at Stanford for many years. From October 64

to July 66, this was done with the Syncom-3 satellite, and from January 67 to December 69 with the ATS-1 satellite. In recent years, other geostationary satellites have also been used at the Stanford receiving station, but in the present work we will confine ourselves to discussion of the Syncom-3 and ATS-1 data. The only long interruption in the time coverage of these data is the second semester of 1966.

The electron content data represent an integrated quantity, and as such they convey little direct information about the detailed structure of the ionosphere, especially in what regards the vertical distribution of ionization (some information about horizontal gradients might conceivably be obtained from observations of a single satellite from two different sites or through simultaneous use of two geostationary satellites). The great advantage in the use of these data in ionospheric research, however, resides in their very good time resolution and in the low operations costs involved in their continuous acquisition during long time intervals. For these reasons, measurements of electron content with geostationary satellites lend themselves extraordinarily well for: (a) study of rapid ionospheric disturbances that would be missed in any intermittent monitoring of the ionosphere; (b) observation of the long-term variations in the ionosphere, such as annual or solar-cycle effects.

The most important variation in the ionospheric electron content is the diurnal variation: Every day at sunrise the electron content starts increasing rapidly. The increase will generally slow down throughout the morning until a maximum is reached, which normally occurs at some time between 1100 LT and 1600 LT. After sunset, the electron content decays rather rapidly. An example of this variation is given on Figure 18. Many irregular features are superimposed on the daily regular pattern, since the ionosphere is subject to a

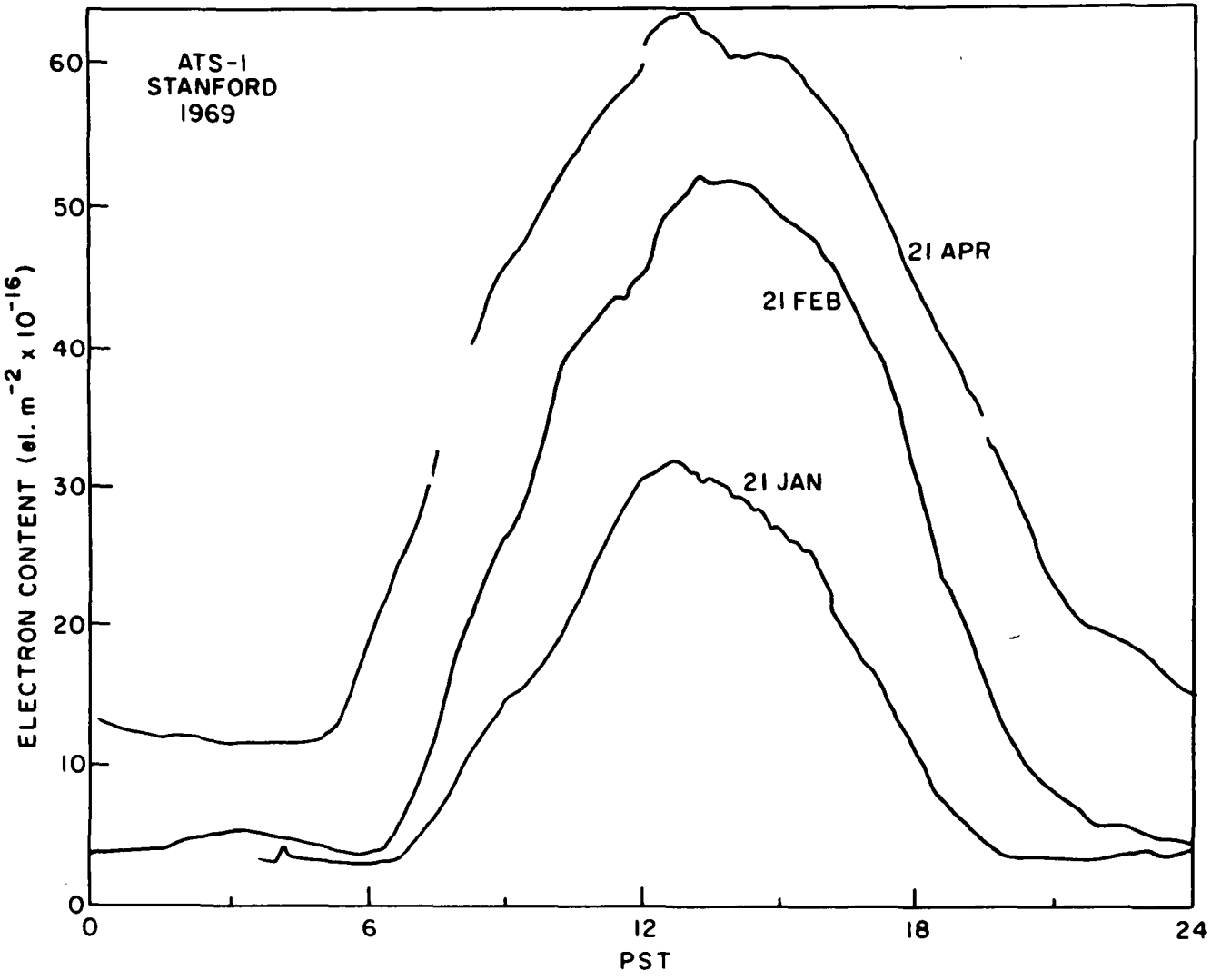


Fig. 18. TYPICAL DIURNAL CURVES OF ELECTRON CONTENT. Three curves are shown for a year of high solar activity; the increase in electron content from January to April reflects the presence of the semiannual variation.

number of perturbations. Even in the presence of these hardly predictable disturbances, a baseline diurnal electron content variation is clearly discernible, which essentially depends on the time of the year and the general level of solar activity. Recently, electron content data obtained with geostationary satellites have become sufficiently abundant to permit the synthesis of models of the diurnal variation in electron content and its dependence on epoch of the year and solar activity. By applying a best-fitting procedure to all presently availabe electron content data from Stanford, Waldman and da Rosa [1971] have developed one such model. The resulting curves of modeled electron content versus time of day, which represent the expected behavior of that quantity throughout the day at Stanford, are illustrated on Figure 19 for selected dates and solar activity levels.

The present work is mainly concerned with long-term variations in the daytime values of electron content at Stanford, which is taken to be representative of the midlatitude zone. These variations can be seen, for instance, in Figure 20, which shows the daily electron content values at 1300 LT for Stanford during an entire year. Even though a semiannual variation in this quantity is clearly visible, it is nevertheless considerably masked by the large day-to-day changes in the ionosphere. Many of such day-to-day changes can be associated with disturbances affecting the geomagnetic field; in general, however, despite the fact that they rival in size with the more regular yearly ionospheric variations, they are one of the least understood and least discussed components of the temporal changes in the ionosphere.

Since the day-to-day variations are basically random, they can be partly filtered out by subjecting the data to a running mean procedure.

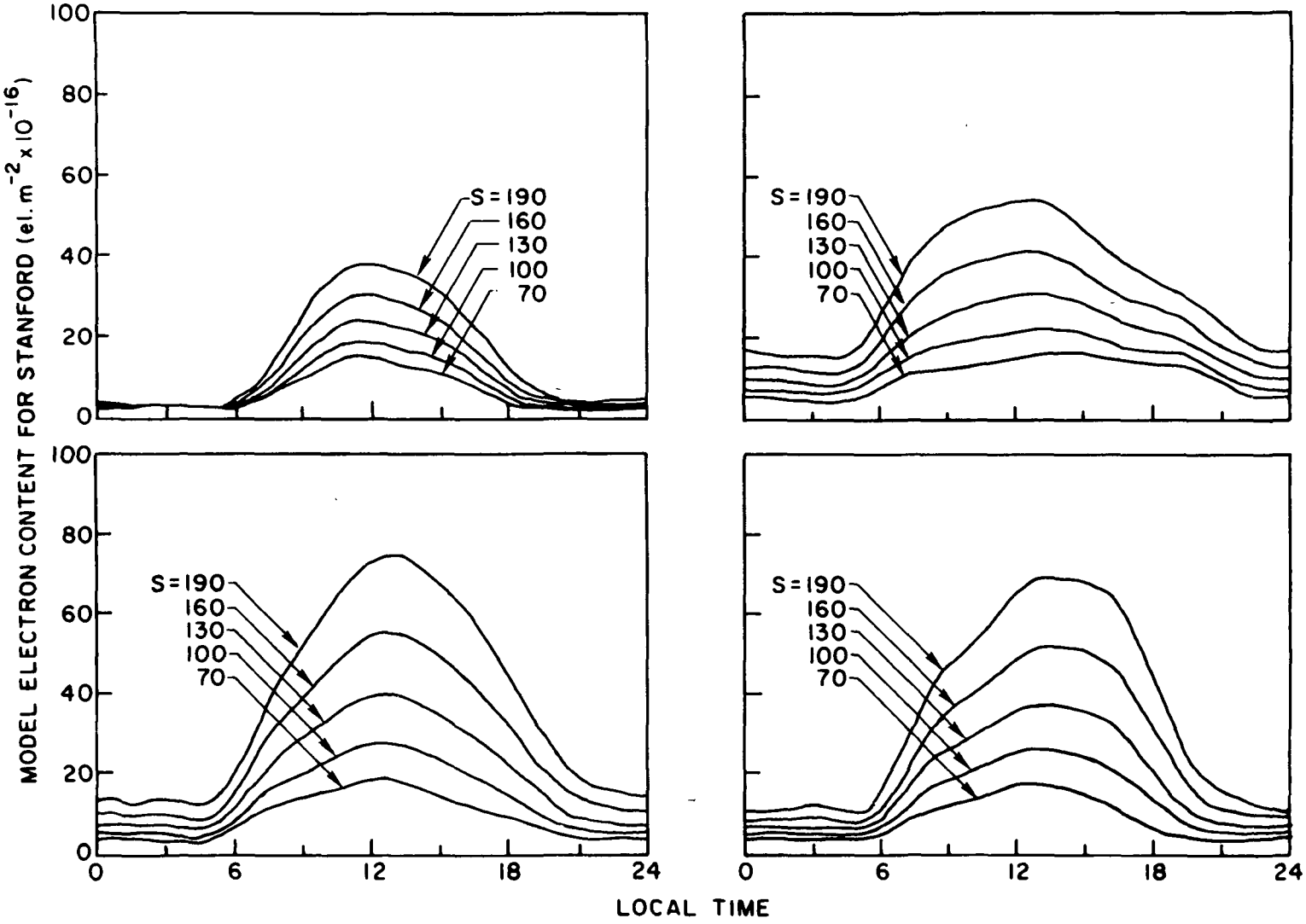


Fig. 19.. AN EMPIRICAL MODEL OF ELECTRON CONTENT. This figure, which is reproduced from Waldman and da Rosa [1971], illustrates the dependence of electron content on time of the day, epoch of the year, and solar activity, contained in a mathematical model of that quantity which was synthesized from all presently available electron content measurements from Stanford.

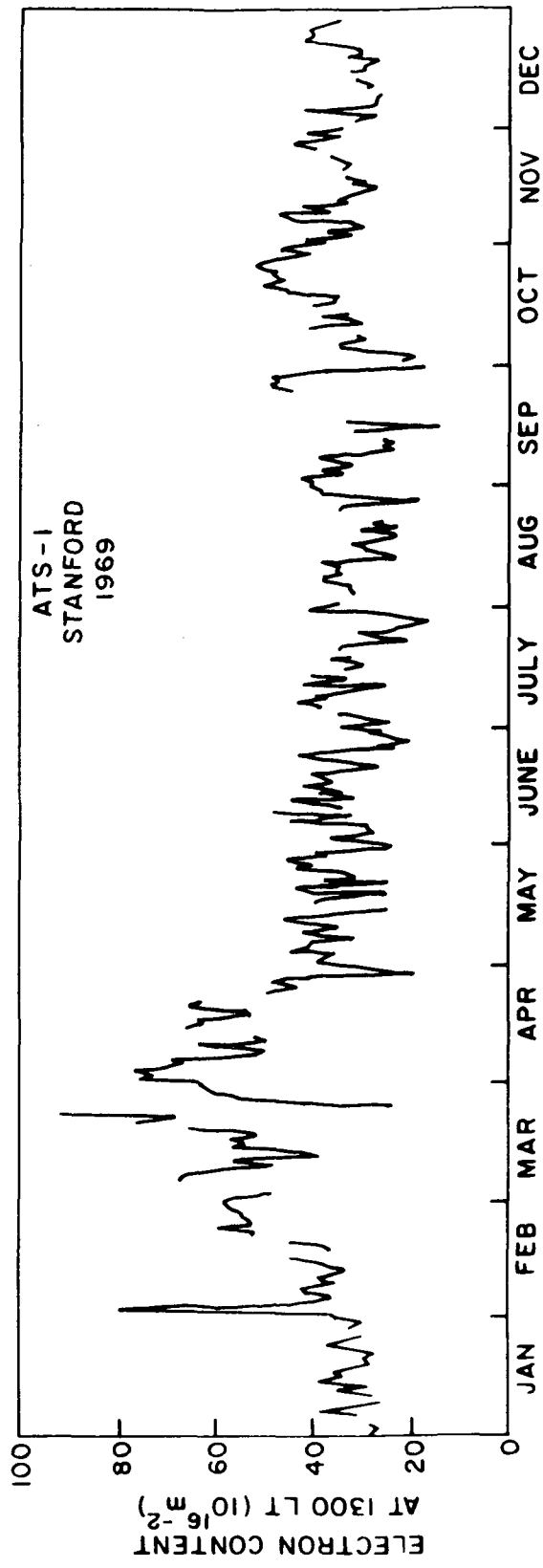


Fig. 20. DAILY VALUES OF ELECTRON CONTENT. Notice the large day-to-day variability of electron content, which constitutes the main difficulty in its prediction.

This has the advantage of rendering more evident the slow variations (periods of 6 months or more) that are the object of the present study; naturally, the averaging interval has to be much shorter than 6 months, so as to essentially preserve the variations to be studied. A one-month averaging interval, for example, would be quite appropriate for such purpose. However, taking monthly averages of the electron content values would conceal the important 27-day variations that result from solar rotation. In order to preserve the information contained in the latter variations, we have chosen to use mean electron content values taken over 7-day intervals. Although this interval is not long enough to effectively average out the day-to-day fluctuations, it does attenuate them considerably while preserving all regular ionospheric variations over longer intervals.

Figures 21 and 22 show the variation of the 7-day mean electron content at 1300 LT at Stanford from 1965 to 1969. In the following sections, this variation is discussed in the light of the theory developed on the previous chapter, and several additional intervening mechanisms are discussed. Considerations involving long-term variations in the nighttime ionosphere and short-term variations in the daytime ionosphere are introduced as needed, but the discussion is centered around long-term variations in the daytime ionosphere at midlatitudes.

A. The Solar-Cycle Variation in Daytime Electron Content

During the period covered by the data presented above, solar emissions in the EUV and x-ray bands were increasing from minimum to maximum of an 11-year cycle of solar activity. Therefore, these data supply information about the solar cycle variation in electron content.

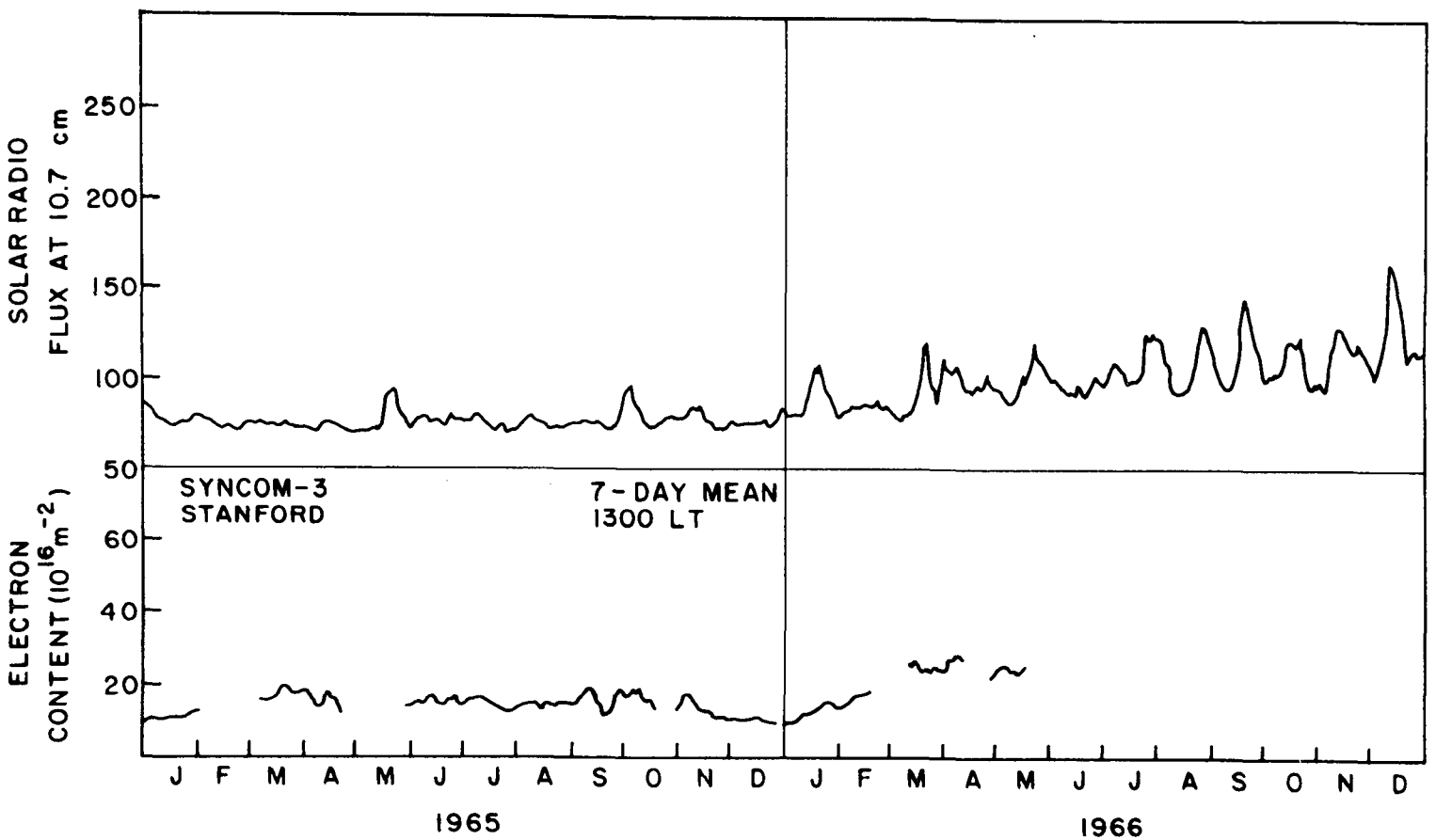


Fig. 21. DAYTIME ELECTRON CONTENT AT LOW SOLAR ACTIVITY. Also shown is the 10.7 cm solar radio flux.

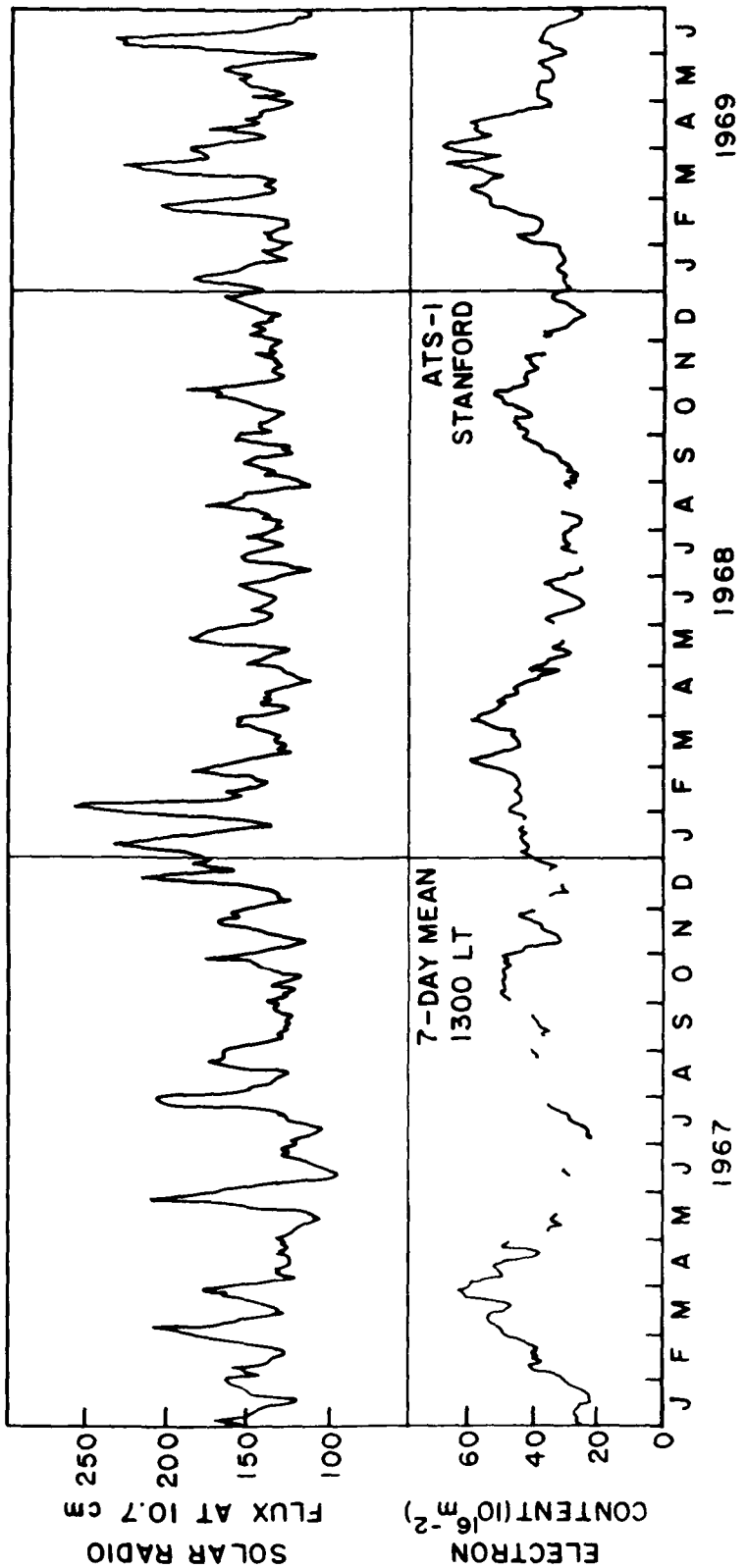


Fig. 22. DAYTIME ELECTRON CONTENT AT HIGH SOLAR ACTIVITY. Also shown is the 10.7 cm solar radio flux.

From 1965 to 1969, the mean annual electron content at 1300 LT grew from about 16 x 10^{16} m^{-2} to 40 x 10^{16} m^{-2}, therefore increasing by 150%. Meanwhile, the mean value of the 10.7 cm solar radio flux S roughly increased from 75 to 150.

The most immediate reason for the daytime electron content variation with solar activity is its dependence on the rate of production of electrons. An increase in solar activity normally brings about an increase in the incident solar EUV flux, which will cause proportional enhancements in the rate of photoionization and resulting plasma concentrations. Due to lack of information concerning the solar EUV flux, it had been customary in the past to scale that quantity proportionally to the solar radio flux. Such procedure would give an increase of 100% in the solar EUV flux from 1965 to 1969, accounting for 100% of the observed 150% variation of the daytime electron content during that period. Recently, however, it became clear that the solar EUV flux increases by only 50% when S varies from 75 to 150 [Hall et al., 1969]. Therefore, only 50% of the observed 150% increase in daytime electron content can be explained by the variation in the incident solar EUV flux.

The increase in the ionization production rates is only one of the effects of increasing solar activity. In addition, the neutral temperature in the thermosphere will increase due to more intense heating by the solar EUV and by corpuscular radiation. The expansion will affect the values of loss and production rates, as well as the diffusion coefficient, thereby modifying the resulting plasma distribution. This problem was studied by Garriott and Rishbeth [1963], who concluded that the total electron content is proportional to $T_n^{1/2}$. According to the CIRA model temperature, the daytime neutral temperature in the

thermosphere increases by about 40% when S goes from 75 to 150. An additional enhancment of something less than 20% is then expected in the daytime electron content from 1965 to 1969, due to thermal expansion in the neutral atmosphere.

The combined effects of increase in incident ionizing radiation and thermal expansion of the atmosphere will then account for an increase of the total daytime electron content by a factor of 1.5 x 1.2 = 1.8 between minimum and maximum of the present solar cycle. Since the observed increase was by a factor of 2.5, an additional gain of 40% has to be explained by another mechanism. Present knowledge of recombination rates and neutral winds in the thermosphere presently leaves a large room for speculation at this point. A strong negative neutral temperature dependence of γ_{O_2} and γ_{N_2} would certainly bridge the gap, but evidence of temperature dependence in these parameters favors a positive one. On the other hand, the additional increase in electron content can be attributed at least partly to solar-cycle variations in the velocity of neutral winds that blow toward the poles during the day, driving the ionization down along the geomagnetic lines and therefore intensifying the recombination processes. Since the daytime wind is controlled primarily by ion drag [Geisler, 1966], one can expect the winds to be faster during the period of low solar activity as some calculations have indicated [Cho and Yeh, 1970]. This would selectively depress the electron content at solar cycle minimum, accounting for some of the variation in electron content. The importance of this effect, and even its existence, is quite uncertain due to a number of facts. First, the winds are driven by atmospheric pressure gradients that are themselves expected to increase

with solar activity, partly or totally counter balancing the increased frictional resistance presented by the ions. Since these gradients are not well known, the results of neutral wind calculations are affected by much uncertainty. Besides, the solar-cycle variation of the height h_{max} of the peak electron concentration does not suggest any significant change in the importance of ionospheric neutral wind effects throughout the solar-cycle [Mayr and Mahajan, 1971]. Finally, if changes in neutral wind velocity play a major role in the solar-cycle variation of electron content, we should expect some geomagnetic dependence in this variation, with a maximum solar-cycle electron content excursion for $I = 45°$, where the winds are most efficient in dragging the ions down; as far as we know, this dependence has not been established so far.

Another possibility is that changes occur in the relative amount of atomic and molecular species at the turbopause during the course of a solar cycle. Any tendency of the atomic oxygen concentration (which control the photoionization rates in the F region) to dominate over $[N_2]$ and $[O_2]$ (which determine the recombination rates) at higher levels of solar activity would entail a corresponding increase in electron content. A recent survey of neutral composition measurements made with rockets during different phases of the solar cycle, however, does not show any solar-cycle variation in the ratio $[O]/[N_2]$ either.

Other mechanisms could be added to those mentioned above, but their role in the ionospheric solar-cycle variations is likewise difficult to establish firmly or rule out completely. In the previous chapter, a theory was developed that predicts a decrease with solar activity of the ionospheric depletion factor $r_d \equiv \Psi/\Psi_L$. If the higher values of r_d existing during solar cycle minimum are comparable with unity, its

reduction will certainly cause an additional enhancement in electron content. This will happen if the neutral hydrogen concentrations at solar cycle minimum are of the order of 10^6 cm^{-3}, such as given in mass spectrometric measurements [Reber et al., 1967]. Naturally, uncertainties in this parameter make it impossible to determine whether the reduction in r_d is by itself sufficient to account for the additional 40% increase in total electron content; such increase, however, can be obtained by even moderate variations in the depletion factor, which speaks for the viability of this hypothesis.

In order to estimate the variation in r_d required to produce the observed electron content variation, one must first establish how the latter depends on r_d. For a lossless F-layer on top of an infinite sink, it was shown on the previous chapter that, for $T_e + T_i = 3T_n$:

$$n_T = n_{To} \left(1 - \frac{4}{5} r_d\right)$$

so that a variation in r_d from unity to zero will multiply the F-layer content by 5. Bank's calculations of F-layer concentration profiles, which were made for distributed losses, roughly reproduce this result. Since the observed unexplained increase of 40% is in the total electron content, a somewhat higher enhancement is required for the F-region content, but certainly less than a factor of 5. We conclude that r_d does not vary from one to zero in a solar cycle, but rather over a narrower range.

For neutral hydrogen concentrations 3 times larger than those given in Equation (47) with the CIRA model neutral temperatures, we have derived $X\ [H]_{500} = 2$, and consequently $r_d = 0.7$, for solar cycle minimum.

The protonospheric tubular content does not show any appreciable solar-cycle variation, which suggests that the upward daytime flux Ψ from the ionosphere is likewise unaffected by solar activity. Assuming that the maximum possible O^+ flux Ψ_L increases by a factor of 2.5, for instance, r_d would go down to 0.28 for S = 150, corresponding to an increase of 80% in the F-region columnar content. This would require $[H]_{500} = 8 \times 10^5$ cm^{-3} for S = 150, which is also about 3 times larger than the values of this parameter given by Equation (47) for the CIRA neutral temperatures. That such values of hydrogen concentration are actually reached in the real atmosphere seems plausible, but not certain by any means. They are consistent with certain mass spectrometer measurements of [H] [Reber et al., 1967], and would be yielded by Equation (47) for neutral temperatures about 200°K lower than those given by the CIRA model atmosphere for around noon.

The reduction of r_d with increasing solar activity due to thermal depletion in the hydrogen concentrations can explain the disparity between the solar-cycle variations in the ionosphere and in the protonosphere. It can also explain the large observed solar-cycle excursion in ionospheric electron content, but only if the neutral hydrogen concentrations are assumed to be on the high side of the presently accepted wide range of possibilities. As an explanation of the above mentioned excursion, changes in the neutral hydrogen concentration do not seem more promising than, say, in the neutral wind velocity. Their main interest, however, resides in their potential for explaining other features of the temporal variations in the ionosphere, such as the semi-annual variation in electron content.

B. Semiannual and Annual Variations in Daytime Electron Content

Even before the presence of geostationary satellites in orbit allowed for a more systematic observation of long-term ionospheric variations, studies of low altitude satellite data had shown that the dependence of daytime electron content on solar activity was stronger during the equinoxes than during summer and winter [Bhonsle et al., 1965]. This result clearly indicated the presence of a semiannual variation with maxima around the equinoxes, at least during the years of high solar activity, which is confirmed in the present research.

To a certain extent, this behavior of the midday electron content contrasts with that of the peak electron concentration, N_m, for which the yearly variation is more strongly influenced by a seasonal component. While at solar-cycle minimum the daytime values of N_m are greater in summer than in winter, just the opposite occurs when solar activity becomes larger. This fact, which constitutes the seasonal anomaly, has been one of the most extensively discussed items of ionospheric morphology so far. As far as the electron content is concerned, it has been shown that the seasonal anomaly associated with it is smaller than the N_m seasonal anomaly [Taylor, 1966], as confirmed by the data presented in this chapter. This indicates that the seasonal anomaly is associated with seasonal changes in the shape of the ionospheric electron concentration profile.

The yearly variations of ionospheric electron content at Stanford are generally dominated by a semiannual component, with maxima in April

and October. This is in agreement with previous studies of the long-term behavior of midlatitude ionospheric electron content [Yeh and Flaherty, 1966]. The data discussed in the present work, however, permit a more systematic and precise assessment of this effect. During the period of low solar activity, the semiannual effect is small and tends to be partially masked by the seasonal variation, which is most pronounced in this phase of the solar cycle. The strength of the semiannual variation can be roughly characterized by the ratio, r_{SA}, between the mean daytime content at the equinoxes and the solstices. At solar cycle minimum, the data indicate $r_{SA} = 1.2$. On the other hand, during 1968 and 1969, r_{SA} was about 1.65. Clearly, the semiannual variation is more intense at solar cycle maximum than at minimum.

On the previous chapter, a mechanism was proposed that links the semiannual variation in ionospheric electron content to that which occurs in the neutral temperature at the thermosphere. The latter variation has now been known and observed for more than a decade, since it was first detected by Paetzold and Zschorner [1960]. It is characterized by temperature maxima in April and October-November, and minima in January and July. The amplitude of this variation is proportional to the mean annual 10.7 cm solar flux [Jacchia et al., 1969], and for S = 100 the total excursion in temperature due to this effect reaches 100°K. During 1968 and 1969, when S was about 150, the maximum-to-minimum difference in temperature due to the semiannual effect was then 150°K. During 1965 on the other hand, the temperature excursion due to this effect was only 70°K. These temperature variations will cause the atmosphere to be more expanded in April and October, thereby enhancing the electron content

in these months. Since the semiannual temperature excursion is less than 10% of the temperature itself, thermal expansion would account for only a 5% difference between extremes of the semiannual variation in electron content, because of its dependence on $T_n^{1/2}$ [Garriott and Rishbeth, 1963]. However, the observed differences are much higher than 5%, and therefore other processes, such as the one discussed on the previous chapter, must be considered.

According to Equation (47), an increment of 150°K in the neutral temperature will reduce the atomic hydrogen concentration to less than one half of its original value. Therefore large oscillations in [H] can be expected to occur at solar cycle maximum. Depending on the mean annual value of $X[H]_{500}$ during the day, these oscillations may cause considerable semiannual variations in the ionospheric depletion factor (see Figure 16), and therefore in its electron content. Since [H] and therefore the depletion factor are minimum in April and October, the electron content will present maxima at these months, as observed.

When the mean annual value of $X[H]_{500}$ is somewhat larger than unity, changes in [H] will not be as effective in altering the depletion factor, because the flow of ionization from the ionosphere is actually being limited by the maximum O^+ flux Ψ_L. From the considerations of the previous section, we may consider this situation to exist at solar cycle minimum. For this reason, and also because of the smaller semiannual temperature oscillation occurring at that time, the corresponding changes in r_d will be small; from Equation (47), a variation of 70°K in T_n will make [H] oscillate in the ratio 3:2. If the mean annual value of $X[H]_{500}$, as suggested on the previous section, is around 2 for solar-cycle minimum, we consider this quantity to vary, say, from 2.25

at the solstices to 1.5 at the equinoxes. The depletion factor would correspondingly vary from 0.76 to 0.66, as shown on Figure 16; the resulting ratio between the electron content in the equinoxes and the solstices would be:

$$\frac{1 - 0.8 \times 0.66}{1 - 0.8 \times 0.76} = 1.2$$

which checks very well with the value of r_{SA} observed during 1965 at Stanford.

As the solar cycle proceeds away from its minimum, the semiannual oscillations in temperature will become larger, increasing the ratio between [H] at the solstices and at the equinoxes. Besides, these changes in [H] will become more effective in altering the depletion factor, since the mean annual value of $X[H]_{500}$ will then fall below unity. Consequently, the semiannual variation in electron content should become more intense, as confirmed by observation. Eventually, as neutral hydrogen concentrations are depleted by increasingly faster thermal escape, the ionospheric depletion factor will vary proportionally to $X[H]_{500}$; by then, however, r_d will be so small that its variations will have little effect on the ionospheric electron content. Therefore, the semiannual variation in electron content due to such effect would eventually subside if solar activity grew to large enough values. Such behavior is not inconsistent with the data: in 1967, when S was around 140, the maximum-to-minimum ratio r_{SA} approached 2, while for 1968 and 1969 it was about 1.65, indicating that moderate solar activity causes a stronger semiannual variation in electron content than a solar-cycle maximum condition. However, it is clear that this semiannual variation persisted through the peak of

the solar cycle, suggesting that either the depletion of hydrogen concentration at the years of maximum solar activity is not enough to effectively uncouple the daytime ionosphere from the protonosphere throughout the year, or that the ionospheric semiannual variations are partly due to some other mechanism.

On the basis of the observed behavior of daytime electron content throughout the first half of the present solar cycle, it was suggested on the previous section that the mean annual value of r_d was around 0.3 for S = 150, which was the condition generally prevailing in 1968 and 1969. For [H] varying by a factor of 2.5 between equinoxes and solstices around such quiescent point, one would have r_d varying between 0.2 and 0.4 from Figure 16, which would result in $r_{SA} = 1.24$. Since the observations yield $r_{SA} = 1.65$, the ionospheric semiannual variation at solar cycle maximum can only be explained in its entirety by this mechanism if the semiannual variation of [H] under this condition is higher than would be predicted by Equation (47). Considerations involving the ionospheric response to short-term variations in solar activity (which occur mainly with a period of around 27 days) tend to give some support to this assumption, as explained later.

It is also possible that this discrepancy be explained by the interaction between the ionosphere and neutral winds: The higher daytime ionic concentrations existing during the equinoxes would slow down the winds through increased ion drag. This would lower the efficiency of the losses, thereby making electron content even higher. This possibility, however, must be regarded with caution, because the effect of neutral winds on an ionosphere under the limiting outflow condition is still unclear at present (see Chapter 4).

We must also consider alternative mechanisms for explaining the

semiannual variations in electron content. It has been suggested that this effect might be explained by considering only semiannual variations in the composition of the neutral air at around 120 km [Mayr and Mahajan, 1971]. This suggestion was based on a collection of composition measurements made with rocket-borne mass spectrometers, which indicated a possible semi-annual variation in the $[O]/[O_2]$ ratio of 120 km, with maxima in April and October. Such variation, if as large as can be inferred from the rocket data (which are affected by considerable spreads), and if due to changes in [O] only, were shown to be more than enough to account for the observed semiannual variation in the peak electron concentration N_m [Mayr and Mahajan, 1971]. However, an investigation of the $[O]/[N_2]$ ratio in the F_1 layer and its yearly changes did not show any significant semiannual component in those variations [Cox and Evans, 1970]. On the basis of these results, it would seem that the variations of $[O]/[O_2]$ mentioned above are caused by changes in $[O_2]$ rather than in [O], which would make them much less effective in controlling the electron concentrations in the daytime ionosphere. In fact, the effect of such changes in the ionosphere would also depend on the relative affinity of O_2 and N_2 for charge-exchange with O^+, which is still poorly known at present. In addition, it must be taken into account that composition changes in the neutral atmosphere that alter the balance between loss and diffusion of O^+ ions should affect the nighttime ionosphere too. Therefore, if they are the dominant factor responsible for the semiannual variation in daytime electron content, a comparable variation should also be expected in the nighttime ionosphere. Our data, however, tend to show that the ionospheric semiannual variation is basically a daytime phenomenon. The disappearance of semiannual variation in nighttime electron content is illustrated on Figure 23; while in 1965 no trace of such variation is

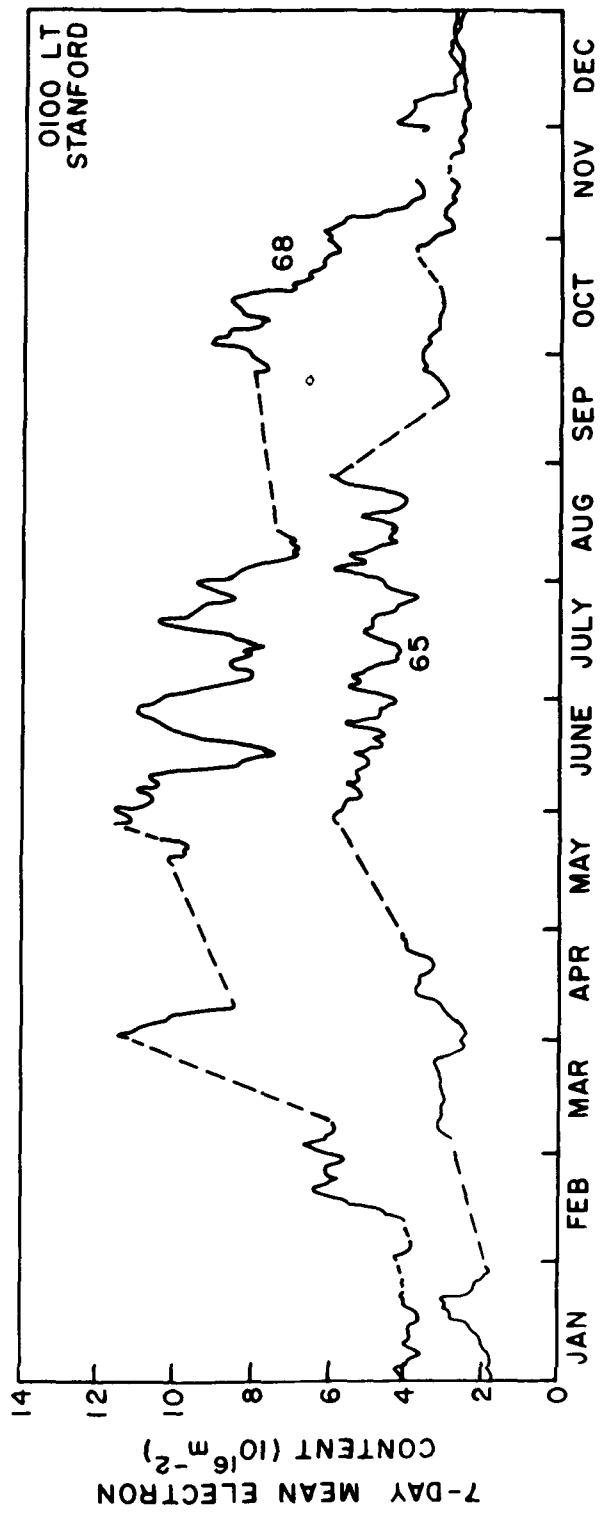


Fig. 23. *ANNUAL VARIATION OF NIGHTTIME ELECTRON CONTENT.* Two years of different solar activity levels are compared.

evident, only secondary peaks in nighttime electron content during 1968 can be noticed in April and October. Clearly, the semiannual effect on the nighttime electron content is much weaker than on the daytime ionosphere. Therefore, if the semiannual variation in daytime electron content is due mainly to changes in neutral composition, such semiannual changes in the neutral atmosphere must occur only during the day and not at night. However, this does not seem likely to happen, because the time constants involved in neutral composition modification are much longer than a day [c.f., Mange, 1955].

Other mechanisms for the explanation of the ionospheric semiannual variations do not seem promising at present. For example, neutral wind velocities during the equinoxes should assume an intermediate value between the summer and winter velocities, so that they can hardly be thought to produce a maximum of electron content at the equinoxes. Only two mechanisms, changes in the ionosphere-protonosphere coupling and in the composition of the neutral atmosphere, remain as plausible alternatives, although the former seems uncertain due to poor knowledge of the neutral hydrogen concentration and its yearly variations, while a critical examination of the latter will reveal some qualitative incongruities. We will tentatively resolve between these two mechanisms by considering the different kinds of semiannual variations that they would cause in different ionospheric parameters, and checking these predicted features against information that has been gathered from ionospheric data.

One of the main characteristics of ionospheric modification due to changes in the upward flux at the upper F-layer is that the electron content is much more strongly affected than the peak electron concentration N_m. This is quite compatible with the ionospheric observations. The semiannual variation in N_m at midlatitudes is quite weak, becoming

stronger at low latitudes with a maximum-to-minimum ratio of 1.6 [Mayr and Mahajan, 1971]. However, the same effect is much stronger in the ionospheric electron content, at least under midlatitude daytime conditions. This would tend to rule out neutral composition changes as the main source of daytime midlatitude ionospheric semiannual variations, because such changes would probably affect n_T and N_m in about the same proportion.

The larger effect on n_T than on N_m caused by semiannual oscillations in the upgoing daytime flux should reflect on semiannual variations in the midlatitude equivalent slab thickness, with maxima in April and October. Indeed, a recently developed analytic model of that quantity revealed a substantial semiannual variation in the slab thickness of the daytime ionosphere only [Klobuchar and Allen, 1970]. This indicates that the smaller semiannual effect present in the nighttime ionosphere affects n_T and N_m in the same rate, thus being possibly associated with composition changes in the neutral atmosphere.

It is also of interest to examine the variations in the height h_{max} of the F-layer peak that would be predicted from each hypothesis. As seen on Figure 14, the state of depletion of the ionosphere has a rather strong influence on the height of the peak. If on April and October the daytime ionosphere is less depleted than on the rest of the year, the layer should be higher. Indeed, Becker [1966] has found strong h_{max} semiannual variations with maxima in April and October at another midlatitude location. The behavior of h_{max} that would result from semiannual changes in the $[O]/[O_2]$ ratio depends on whether such changes are caused by variations in $[O]$ or $[O_2]$. If they are caused by

April and October maxima in [0], then they are consistent with the observed semiannual maxima in h_{max}, since an increase in [0] would lift the layer by slowing down diffusion of 0^+ ions. The level of the peak is located by $[0_2][0]$ = constant, so that even a 3:1 variation in [0] could not caused more than a 20 km shift in h_{max}. In contrast, Becker's observations revealed more than a 50 km semiannual excursion in h_{max} for high solar activity, and about 20 km at solar-cycle minimum. On the other hand, if the semiannual variation in $[0]/[0_2]$ is given by $[0_2]$ minima in April and October, there would be h_{max} minima at those times of the year, which is quite contrary to the observations.

The comparative morphology of the semiannual variations in the electron content, the peak electron concentration and the height of the peak does not seem to fit the hypothesis that they are caused mainly by neutral composition changes. It does seem to be consistent, however, with the idea of semiannual changes in the depletion state of the ionosphere. The latter assumption, however, is only applicable to the midlatitude ionosphere, which is connected to the protonosphere through magnetic lines. Nevertheless, a semiannual variation in the daytime peak electron concentration has been found on the geomagnetic equator, although much weaker than the corresponding variation at midlatitudes [Yonezawa, 1971].

Summarizing, the evidence seems to indicate that the semiannual variations in the daytime electron content at midlatitudes is mainly due to changes in the state of depletion of the ionosphere, which are caused by varying concentrations of neutral hydrogen in the thermosphere. Other mechanisms, such as neutral composition changes at the turbopause,

might also have a less important contribution, and would be mainly responsible for the smaller semiannual effects occurring in the nighttime and equatorial ionosphere.

C. Implications on the Protonosphere of Changes in the State of Depletion of the Ionosphere

The semiannual variation in the daytime ionospheric depletion factor suggested on this chapter, if it really exists, would have implications other than the observed semiannual variation in the daytime ionosphere. Since it implies that the daytime flux from the ionosphere to the protonosphere has minima on April and October, it would also result in protonospheric tubular content minima, unless tubular losses are then also correspondingly smaller. Actually, the yearly variation in protonospheric tubular content is dominated by an annual component, with the maximum occurring in December; semiannual variations in this quantity, if they exist, do not appear to be significant. The protonospheric yearly variations can only be reconciled with the proposed semiannual variations in the daytime ionospheric depletion factor if a similar variation in tubular losses is enough to offset the reduced daytime ionization input into the protonosphere that would occur around April and October. It is difficult to determine whether this happens or not, because the relative importance of different mechanisms through which ionization is removed from the protonosphere is still to be established. It is known, however, that the regular nighttime flux of ionization into the ionosphere, which is largely responsible for maintenance of the nighttime ionosphere, can only account for a fraction of the ionization gained by the protonosphere during the day [Park, 1970]. Other mechanisms, mostly associated with

stormy events, must be responsible for a substantial part of the protonospheric tubular losses. One of such events that have been associated with transfer of ionization from the protonosphere to the ionosphere are anomalous mighttime increases in electron content which are occasionally observed [da Rosa and Smith, 1967]. A study of these increases at Hawaii covering four years of data revealed that they occur almost always during the solstices, with very few occurrences during the equinoxes [Young et al., 1970]. Therefore, if the anomalous nighttime increases are responsible for a considerable part of the over all removal of ionization from the protonosphere, they might counterbalance the semiannual variation in daytime transfer of ionization from the ionosphere to the protonosphere, thereby accounting for the relative absence of semiannual variations in the protonosphere.

D. Short-term Variations in Daytime Electron Content

In addition to its seasonal variation, the daytime ionospheric electron content also presents variations with periods shorter than one season. The study of these short-term variations is an entirely different branch in ionospheric morphology, and we do not propose to examine this matter in detail in the present work. However, it seems pertinent to inquire at this point about the implications of the suggested semiannual changes in the daytime ionospheric depletion on such short-term variations. In this section, we will attempt to disclose some of such implications, as well as to check them briefly against the observations.

The emission of solar radiation does not occur uniformly over the Sun's surface. Instead, it may be somewhat concentrated on certain active areas that may persist for several months, particularly during the

years of high solar activity. Consequently, as the Sun rotates with a period of about 27 days, solar activity as seen from the earth will sometimes seem to vary with that period. This variation in the incident solar radiation is responsible for the appearance of 27-day variations in the earth's ionosphere, which are an important component of the ionospheric short-term variations. The thermospheric temperature during one of these 27-day periods will also vary accordingly, but its sensitivity to this short-term component of solar activity is much smaller than to the long-term one [Jacchia, 1963]. As seen on section A of this chapter, the increase in neutral temperature throughout the solar cycle, both through the mechanisms of thermal expansion and increased hydrogen escape, is responsible for a large portion of the corresponding increase in daytime electron content. Therefore, since the neutral temperature responds much more weakly to the 27-day oscillations in solar activity than to its 11-year cycle, we should also expect the electron content dependence on the 27-day component of S to be much weaker than on the long-term component of S. Indeed, a preliminary investigation of this subject has revealed that, both at Hawaii and Stanford, the long-term electron content sensitivity to solar radio flux is about twice as large as the short-term sensitivity [da Rosa et al, 1971]. It is interesting to notice that this difference in ionospheric response to S cannot be attributed to any difference between the short-term and long-term correlation between S and the solar EUV flux, because continuous satellite measurements of the latter have revealed that its variations over 27-day periods are similar in magnitude and spectral character to those observed over 11 years [Hall and Hinteregger, 1970].

The difference between the states of daytime ionospheric depletion prevailing in the equinoxes and solstices should also affect the ionospheric response to short-term changes in solar activity. During the winter, for instance, the amount of neutral hydrogen present should be maximum, rendering the ionosphere relatively insensitive to changes in T_n, due to its continuously depleted state. During the equinoxes, the reduced concentrations of neutral hydrogen would make it possible for it to control the state of ionospheric depletion, making the ionosphere more sensitive to changes in T_n. This is illustrated, for instance, on Figure 24, which covers the 1967-68 winter. It is interesting to see that, while the 27-day oscillations in S are decreasing in intensity from January to March of 1968, the ionospheric response to those oscillations is actually increasing in magnitude.

The remaining part of the regular short-term ionospheric variations is the important day-to-day variability illustrated on Figure 18. Satellite measurements of the solar EUV flux over an extended period have revealed a rather smooth time variation in that quantity, so that the day-to-day ionospheric variability is probably associated with some other factor. One possibility is that the neutral hydrogen concentration, and hence the ionospheric depletion factor, are also very variable.

Finally, we should consider the diurnal variation itself. As mentioned previously, there is a large day-night asymmetry in the concentration of neutral hydrogen. Consequently, we may expect that [H] varies during the day, with a minimum around noon. This would result in the ionosphere being more depleted during the morning and afternoon than around midday. This might be the explanation for the considerable disparity existing between the typical curve of measured electron content

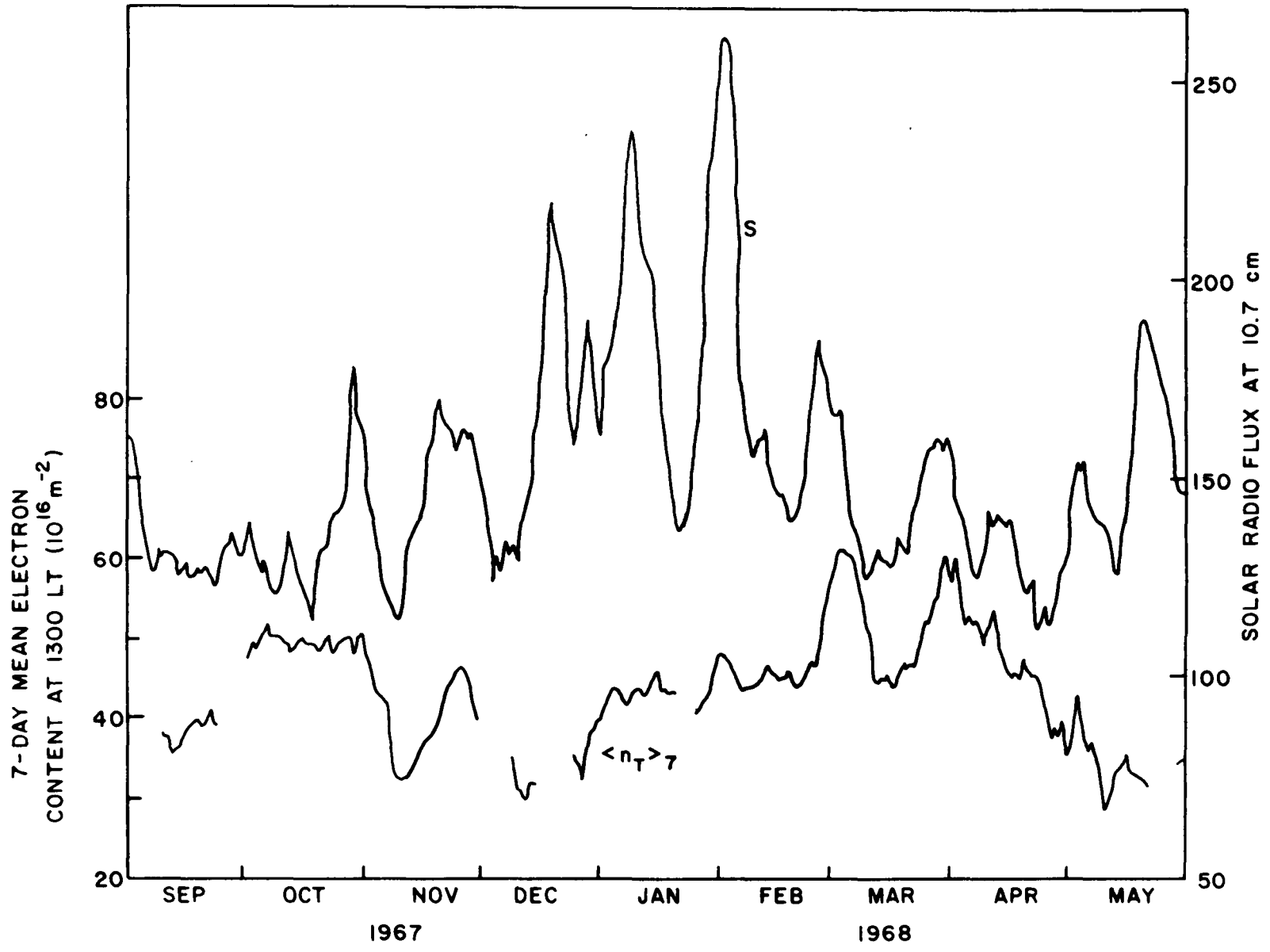


Fig. 24. SHORT-TERM DEPENDENCE OF DAYTIME ELECTRON CONTENT ON SOLAR ACTIVITY. The time variation of the seven-day mean electron content at 1300 LT is compared with that of solar activity, with regard to short-term effects.

versus time of day and the diurnal curves of electron content generated on the computer using a "zero flux" condition [e.g., Stubbe, 1970].

Chapter 6

CONCLUSIONS

The main purpose of this work was to re-examine the coupling between the ionosphere and the protonosphere in the light of new evidence which shows that the flux of ionization between those two layers is much higher than theory had previously predicted. In particular, it was shown that, because of ionospheric depletion caused by such large fluxes during the day, previous theoretical determinations of the limiting daytime flux of ionization from the F-region to the protonosphere are not adequate. A self-consistent calculation of the limiting flux was then made in order to yield a more general expression for that quantity.

Unlike the previously calculated theoretical limiting flux, the self-consistent limiting flux is not strictly proportional to the concentration of neutral hydrogen in the thermosphere. Instead, it departs from such proportionality when $[H]_{500}$ reaches about 10^6 cm^{-3}, saturating for higher values of $[H]_{500}$ at a level given by the limiting upward flux of oxygen ions from the F-layer. This merely reflects the fact that, even in the presence of an arbitrarily large quantity of hydrogen atoms in the thermosphere, the outgoing flux of protons cannot exceed the maximum flux of oxygen ions that can be supported by the F-layer, because the creation of each escaping proton requires the neutralization of one oxygen ion.

From the available experimental evidence concerning the ionic composition profiles in the upper F-region, the behavior of protonospheric tubular content observed with whistlers, and the F-region fluxes measured with radar, the conclusion was drawn that the daytime flux of protons

from the ionosphere to the protonosphere usually approaches its limiting value. The implications of this condition on the behavior of the main ionospheric parameters cannot be fully appreciated at present, because of current lack of knowledge on the thermospheric neutral hydrogen concentration and its variations. It can be said, however, that such daytime fluxes can have a significant effect on the profile of $[O^+]$, which dominates the bulk of the ionosphere. The daytime ionosphere is depleted, with a corresponding reduction in the columnar electron content, and, to a smaller extent, in the peak electron concentration.

The extent of ionospheric depletion is strongly dependent on the concentration of neutral hydrogen, provided that saturation has not been approached. On the other hand, it was found that the exchange of protons between the ionosphere and the protonosphere, being normally set at the limiting rate during the day, may have a significant effect on the global balance of neutral hydrogen, tending to reduce the planetary loss of this element. For this reason, a fully self-contained solution to the problem of ionosphere-protonosphere coupling would have to include the interaction between the limiting flux and the concentration of neutral hydrogen. Such an approach, however, seems hardly possible at present, since the problem of the global circulation and escape of hydrogen, even disregarding the storage of hydrogen atoms in the protonosphere in the form of protons and the loss of hydrogen through the polar wind, is very difficult.

A more promising line of inquiry at present is the study of ionospheric variations resulting from assumed corresponding variations in the neutral hydrogen concentration under the condition of limiting daytime flux. Very little experimental information exists about such variations

but one can infer them theoretically from the long-term variations in T_n through the well known temperature dependence of the hydrogen escape rate.

A study of the influence of the ionosphere-protonosphere coupling on the long-term ionospheric variations was made on this basis, revealing that:

a) the semiannual variation in ionospheric electron content can be due to corresponding variations in the neutral hydrogen concentration caused by the well-known semiannual variation in thermospheric neutral temperature;

b) the solar-cycle variation in daytime ionospheric electron content can only partly be explained in terms of changes in the solar ionizing radiation and in the neutral concentration profiles, and this gap can be closed when one considers the decreased coupling between ionosphere and protonosphere at high levels of solar activity;

c) the disparity between electron content solar-cycle variations in the ionosphere and in the protonosphere can be understood in terms of the smaller coupling with increasing solar activity.

Although this study was mainly confined to long-term variations, it is quite possible that changes in the daytime limiting flux have an important role in short-term ionospheric variations too. For example, the strong day-to-day variations in daytime ionospheric electron content, which still remain largely unexplained, might be associated with changes in the coupling flux. In addition, the pronounced minimum in the neutral hydrogen concentration around noon may have a role in the diurnal ionospheric

variation, reenforcing the electron content peak at midday. This is particularly interesting when one considers that the typical diurnal curve of ionospheric electron content is quite different from what the numerical simulations have yielded under the "zero topside flux" condition. The existence of these effects, however, remains very uncertain until more information is gathered about the fluxes between the ionosphere and the protonosphere.

The present work also underlines the need for more information concerning the behavior of neutral hydrogen in the thermosphere. Although it is a minor component of the neutral air in the heights of interest, and although its importance in the production and loss of electrons is insignificant, the role of neutral hydrogen in regulating the coupling between the F-region and the protonosphere makes it an important element for ionospheric studies.

REFERENCES

Page Intentionally Left Blank

Banks, P., Collision frequencies and energy transfer: ions, Planetary Space Sci., 14, 1105, 1966.

Banks, P., Collision frequencies and energy transfer: electrons, Planetary Space Sci., 14, 1085, 1966.

Banks, P., Ion temperature in the upper atmosphere, J. Geophys. Res., 72, 3365, 1967.

Banks, P. M., and T. E. Holzer, Features of plasma transport in the upper atmosphere, J. Geophys. Res., 74, 6304, 1969.

Banks, P. M., and T. E. Holzer, High-latitude plasma transport: the polar wind, J. Geophys. Res., 74, 6317, 1969.

Becker, W., The seasonal anomaly of the F region at midlatitudes and its interpretation, Electron Density Profiles in the Ionosphere and Exosphere, North-Holland Publishing Co., Amsterdam, 218, 1966.

Bhonsle, R. V., A. V. da Rosa, and O. K. Garriott, Measurements of the total electron content and equivalent slab thickness of the midlatitude ionosphere, Radio Sci. J. Res., 69D, 929, 1965.

Brinton, H. C., M. W. Pharo, III, H. G. Mayr, and H. A. Taylor, Jr., Implications for ionospheric chemistry and dynamics of a direct measurement of ion composition in the F2 region, J. Geophys. Res., 74, 2941, 1969.

Browne, I. C., J. V. Evans, J. K. Hargreaves, and W. A. S. Murray, Radio echoes from the moon, Proc. Phys. Soc. London, 70, 901, 1956.

Cho, H. R., and K. C. Yeh, Neutral winds and the behavior of the ionospheric F2 region, Radio Science, 5, 881, 1970.

CIRA 1965: Cospar International Reference Atmosphere: 1965, North-Holland Publishing Co., Amsterdam, 1965.

Cox, L. P., and J. V. Evans, Seasonal variation of the O/N_2 ratio in the F1 region, J. Geophys. Res., 75, 6271, 1970.

Crank, J., and P. Nicolson, A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type, Proc. Camb. Phil. Soc., 43, 50, 1947.

Dalgarno, A., M. B. McElroy, and R. J. Moffett, Electron temperatures in the ionosphere, Planetary Space Sci., 11, 463, 1963.

da Rosa, A. V., Thermal behavior of the ionosphere and observations of the exosphere and the ionosphere by means of distant earth satellites, Tech. Report No. 2, Radioscience Laboratory, Stanford Electronics Laboratories, Stanford University, Stanford, Calif., 1965.

da Rosa, A. V., and F. L. Smith, III, Behavior of the nighttime ionosphere, J. Geophys. Res., 72, 1829, 1967.

da Rosa, A. V., H. Waldman, J. Bendito, and O. K. Garriott, Influence of the solar activity on the ionospheric electron content, presented at the 1971 Spring meeting of the U. S. National Committee of URSI, Washington, D.C., April 1971.

Evans, J. V., R. F. Julian, and W. A. Reid, Incoherent scatter measurements of F region density, temperatures, and vertical velocity at Millstone Hill, Technical Report 477, Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts, 1970.

Evans, J. V., Observations of F region vertical velocities at Millstone Hill: II variation of velocity with altitude, Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts, 1971.

Fehsenfeld, F. C., A. L. Schmeltekopf, and E. E. Ferguson, Some measured rates for oxygen and nitrogen ion-molecule reactions of atmospheric importance, including $O^+ + N_2 \rightarrow NO^+ + N$, Planetary and Space Sci., 13, 219, 1965.

Fox, L., Numerical solution of ordinary and partial differential equations, Addison-Wesley Publishing Co., Inc., 236-246, 1962.

Garriott, O. K., and H. Rishbeth, Effects of temperature changes on the electron density profile in the F2 layer, Planetary Space Sci., 11, 587, 1963.

Garriott, O. K., F. L. Smith, III, and P. C. Yuen, Observations of ionospheric electron content using a geostationary satellite, Planetary Space Sci., 13, 829, 1965.

Geisler, J. E., and S. A. Bowhill, An investigation of ionospheric-protonosphere coupling, Aeronomy Report No. 5, Aeronomy Laboratory, University of Illinois, Urbana, Illinois, 1965.

Geisler, J. E., Atmospheric winds in the middle latitude F region, J. Atmos. Terr. Phys., 28, 703, 1966.

Geisler, J. E., On the limiting daytime flux of ionization into the protonosphere, J. Geophys. Res., 72, 81, 1967.

Hall, L. A., J. E. Higgins, C. W. Chagnon, and H. E. Hinteregger, Solar-cycle variation of extreme ultraviolet radiation, J. Geophys. Res., 74, 4181, 1969.

Hall, L. A., and H. E. Hinteregger, Solar radiation in the extreme ultraviolet and its variation with solar rotation, J. Geophys. Res., 75, 6959, 1970.

Hanson, W. B., and I. B. Ortenburger, The coupling between the protonosphere and the normal F region, J. Geophys. Res., 66, 1425, 1961.

Hanson, W. B., and T. N. L. Patterson, Diurnal variation of the hydrogen concentration in the exosphere, Planetary Space Sci., 11, 1035, 1963.

Herman, J. R., and S. Chandra, The influence of varying solar flux on ionospheric temperatures and densities: a theoretical study, Planetary Space Sci., 17, 815, 1969.

Hinteregger, H. E., L. A. Hall, and G. Schmidtke, Solar XUV radiation and neutral particle distribution in July 1963 thermosphere, Space Research V, North-Holland Publishing Co., Amsterdam, 1175, 1965.

Hoffman, J. H., C. Y. Johnson, J. C. Holmes, and J. M. Young, Daytime midlatitude ion composition measurements, J. Geophys. Res., 74, 6281, 1969.

Holzer, T. E., and P. M. Banks, Accidentally resonant charge exchange and ion momentum transfer, Planetary Space Sci., 17, 1074, 1969.

Jacchia, L. G., Electromagnetic and corpuscular heating of the upper atmosphere, Space Research III, North-Holland Publishing Co., Amsterdam, 3, 1963.

Jacchia, L. G., J. W. Slowey and I. G. Campbell, A study of the semiannual density variation in the upper atmosphere from 1958 to 1966, based on satellite drag analysis, Planetary Space Sci., 17, 49, 1969.

Johnson, C. Y., Ion and neutral composition of the ionosphere, presented at the IQSY/COSPAR Joint Symposium, Imperial College of Science and Technology, London, 1967.

Klobuchar, J. A., and R. S. Allen, A first-order prediction model of total-electron-content group path delay for a midlatitude ionosphere, Air Force Surveys in Geophysics No. 222, AFCRL-70-0403, Air Force Cambridge Research Laboratories, Bedford, Mass., 1970.

Liwshitz, M., and S. F. Singer, Thermal escape of neutral hydrogen and its distribution in the earth's thermosphere, Planetary Space Sci., 14, 541, 1966.

Mange, P., Diffusion processes in the thermosphere, Ann. Geophys., 11, 153, 1955.

Mayr, H. G., and K. K. Mahajan, Seasonal variation in the F2 region, J. Geophys. Res., 76, 1017, 1971.

Meier, R. R., and D. K. Prinz, Absorption of the solar Lyman alpha line by geocoronal atomic hydrogen, J. Geophys. Res., 75, 6969, 1970.

Mitra, A. P., A review of D-region processes in non-polar latitudes, J. Atmos. Terr. Phys., 30, 1065, 1968.

Nagy, A. F., E. G. Fontheim, R. S. Stolarski, and A. E. Beutler, Ionospheric electron temperature calculations including protonospheric and conjugate effects, J. Geophys. Res., 74, 4667, 1969.

Nisbet, J. S., Photoelectron escape from the ionosphere, J. Atmos. Terr. Phys., 30, 1257, 1968.

Paetzold, H. K., and H. Zschorner, Bearings of Sputnik 3 and the variable acceleration of satellites, Space Research I, North-Holland Publishing Co., Amsterdam, 24, 1960.

Park, C. G., Whistler observations of the interchange of ionization between the ionosphere and the protonosphere, J. Geophys. Res., 75, 4249, 1970.

Ratcliffe, J. A., and K. Weekes, The ionosphere, in Physics of the Upper Atmosphere, edited by J. A. Ratcliffe, Academic Press, New York and London, 377, 1960.

Reber, C. A., J. E. Cooley, and D. N. Harpold, Upper atmosphere hydrogen and helium measurements from the Explorer 32 satellite, Space Research VIII, North-Holland Publishing Co., Amsterdam, 993, 1968.

Richtmyer, R. D., Difference methods for initial value problems, Interscience, New York and London, 96-109, 1957.

Shea, M. F., R. D. Sharp, and M. B. McElroy, Measurements and interpretation of low-energy photoelectrons, J. Geophys. Res., 73, 4199, 1968.

Smith, R. L., Properties of the outer ionosphere deduced from nose whistlers, J. Geophys. Res., 66, 3709, 1961.

Spitzer, L., Jr., Physics of fully ionized gases, Interscience, 1962.

Stubbe, P., Temperature dependence of the rate constants for the reactions $O^+ + O_2 \rightarrow O_2^+ + O$ and $O^+ + N_2 \rightarrow NO^+ + N$, Planet. Space Sci., 17, 1221, 1969.

Stubbe, P., On the possibility of evaluating the hydrogen concentration in the upper atmosphere by comparing the bottomside and topside electron contents, Proceedings of the Symposium on the Future Application of Satellite Beacon Experiments, Max-Planck-Institut fur Aeronomie, Lindau/Harz, W. Germany, 1970.

Stubbe, P., Simultaneous solution of the time dependent coupled continuity equations, and equations of motion for a system consisting of a neutral gas, an electron gas, and a four component ion gas, J. Atmos. Terr. Phys., 32, 865, 1970.

Taylor, G. N., The seasonal anomaly in electron content, Electron Density Profiles in Ionosphere and Exosphere, North-Holland Publishing Co., Amsterdam, 543, 1966.

Waldman, H., and A. V. da Rosa, Prognostication of ionospheric electron content, Technical Report in press, Radioscience Laboratory, Stanford University, Stanford, California, 1971.

Wand, R., Ionospheric heating results obtained at Arecibo, Aeronomy Report No. 19, Thomson Scatter Studies of the Ionosphere--An Informal Conference Record, University of Illinois, Urbana, Illinois, 1967.

Wand, R. H., Electron-to-ion temperature ratio from radar Thomson scatter observations, J. Geophys. Res., 75, 829, 1970.

Whitten, R. C., and I. G. Poppoff, Ion kinetics in the lower ionosphere, J. Atmos. Sci., 21, 117, 1964.

Yeh, K. C., and V. H. Gonzales, Note on the geometry of the earth magnetic field useful to Faraday effect experiments, J. Geophys. Res., 65, 3209, 1960.

Yeh, K. C., and B. J. Flaherty, Ionospheric electron content at temperate latitudes during the declining phase of the sunspot cycle, J. Geophys. Res., 71, 4557, 1966.

Yonezawa, T., The solar-activity and latitudinal characteristics of the seasonal, non-seasonal and semi-annual variations in the peak electron densities of the F2-layer at noon and at midnight in middle and low latitudes, submitted to J. Atmos. Terr. Phys., 1971.

Young, D. M. L., P. C. Yuen, and T. H. Roelofs, Anomalous nighttime increases in total electron content, Planetary Space Sci., 18, 1163, 1970.