INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
34100 TRIESTE (ITALY) - P.O.B. 586 - MIRAMARE - STRADA COSTIERA 11 - TELEPHONE 2240-1
CABLE: CENTRATOM - TELEX 460392 - 1

H4.SMR.203 - 20

" SPRING COLLEGE ON GEOMAGNETISM AND AERONOMY "

(2 - 27 March 1987)

" Ionospheric effects on earth-space propagation "

presented by :

J.A. KLOBUCHAR
Ionospheric Physics Division
Air Force Geophysics Laboratory
Air Force Base
Hanscom, MA 01731
U.S.A.

These are preliminary lecture notes, intended for distribution to participants only.

AFGL-TR-84-0004
ENVIRONMENTAL RESEARCH PAPERS, NO. 866

Ionospheric Effects on Earth-Space Propagation

JOHN A. KLOBUCHAR

27 December 1983

IONOSPHERIC PHYSICS DIVISION PROJECT 4643
AIR FORCE GEOPHYSICS LABORATORY
HANSCOM AFB, MA 01731

This report has been reviewed by the ESD Public Affairs Office (PA) and is releasable to the National Technical Information Services (NTIS).

"This technical report has been reviewed and is approved for publication"

FOR THE COMMANDER

for H. Whitney
HERBERT C. CARLSON
Branch Chief

ROBERT A. SKRIVANEK
Division Director

Qualified requestors may obtain additional copies from the Defense Technical Information Center. All others should apply to the National Technical Information Service.

If your address has changed, or if you wish to be removed from the mailing list, or if the addressee is no longer employed by your organization, please notify AFGL/DAA, Hanscom AFB, MA 01731. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document requires that it be returned.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: AFGL-TR-84-0004	2. GOVT ACCESSION NO. / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): IONOSPHERIC EFFECTS ON EARTH-SPACE PROPAGATION	5. TYPE OF REPORT & PERIOD COVERED: Scientific
	6. PERFORMING ORG. REPORT NUMBER: ERP, No. 866
7. AUTHOR(s): John A. Klobuchar	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Air Force Geophysics Laboratory (LIS), Hanscom AFB, Massachusetts 01731	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 62101F, 46430901
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Geophysics Laboratory (LIS), Hanscom AFB, Massachusetts 01731	12. REPORT DATE: 27 December 1983
	13. NUMBER OF PAGES: 31
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
Ionospheric time delay
Ionospheric effects on satellite systems
Ionospheric range errors

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The ionosphere produces several effects upon radio waves that must travel through it. For radio waves such as those that propagate to and from satellites used for communication or navigation purposes, and for ranging to space objects from satellite detection and tracking radars, the effects of the ionosphere may be important depending upon the system operating frequency and the state of the ionosphere at the time. In this report the various effects that the ionosphere has upon radio waves which must traverse it are discussed and the main characteristics of the ionosphere, which are important

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. (Contd.)

to systems design engineers and others who must consider the effects of the ionosphere upon space systems, are presented.

Contents

Illustrations

Ionospheric Effects on Earth-Space Propagation

1. INTRODUCTION

Radio waves undergo several effects when they pass through the earth's ionosphere. One of the most important effects that the ionosphere has upon radio waves that traverse it is a retardation, or group delay, on the modulation or information carried on the radio wave that is due to its encounter with the free, thermal electrons in the earth's ionosphere. Other effects the ionosphere has on radio waves include: (1) RF carrier phase advance; (2) Doppler shift of the RF carrier of the radio wave; (3) Faraday rotation of the plane of polarization of linearly polarized waves; (4) angular refraction or bending of the radio wave path as it travels through the ionosphere; (5) distortion of the waveform of transmitted pulses, and (6) amplitude and phase scintillation. With the exception of scintillation effects, which are not described in this report, all the other effects listed here are proportional, at least to first order, to the total number of electrons encountered by the wave on its passage through the ionosphere or to their time rate of change. In fact, phase scintillation also is merely the short term time rate of change of total electron content (TEC) after the longer term variations have been removed.

In this report, a short description is given of each ionospheric TEC effect upon radio waves, along with a representative value of the magnitude of each of these

(Received for publication 21 December 1983)

effects under normal ionospheric conditions. This is followed by a discussion of the important characteristics of average ionospheric TEC behavior and the temporal and spatial variability of TEC.

1.1 Group Path Delay

The additional time delay, over the free space transit time, of a signal transmitted from above the ionosphere to a user on, or near, the earth's surface is given by Millman:[†]

$$\Delta t = \frac{40.3}{cf^2} * \text{TEC} \qquad \text{(seconds)} \tag{1}$$

where the TEC is the total number of electrons along the path from the transmitter to the receiver, c is the velocity of light in m/sec, and f is the system operating frequency in Hertz. The TEC is generally expressed as the number of electrons in a unit cross section column of one square meter area along this path.

A plot of time delay vs system operating frequency for TEC values from 10^{16} to 10^{19} el/m^3 is given in Figure 1. These two values represent the extremes of observed TEC in the earth's ionosphere. Note that, at a system operating frequency of 1 GHz, for example, a TEC of 10^{18}, a value frequently exceeded in many parts of the world, would produce a time delay of 134 nanoseconds or 132 ft (40.2 m) of range error. At a system operating frequency of 100 MHz this same TEC value would produce a range error of over 13,000 ft or 4 km! Obviously, the TEC parameter is of potentially great importance to precision satellite ranging systems.

1.1.1 TWO-FREQUENCY IONOSPHERIC TIME DELAY CORRECTIONS

If the navigation or ranging system bandwidth is large enough so that two, fairly widely spaced bands can be used for ranging, the ionospheric time delay error can be reduced to an acceptable level automatically, and can be made transparant to the system user. Because the ionospheric delay is a function of frequency we can write:

$$\Delta t_1 = \frac{k}{cf_1^2} * \text{TEC}, \quad \Delta t_2 = \frac{k}{cf_2^2} * \text{TEC}$$

where Δt_1 is the ionospheric error on frequency f_1 and Δt_2 is the ionospheric error on the frequency f_2. If the normal system operational frequency is f_1 and we choose f_2 at a lower frequency for ionospheric correction purposes, we obtain

[†]See Reference 7, page 29.)

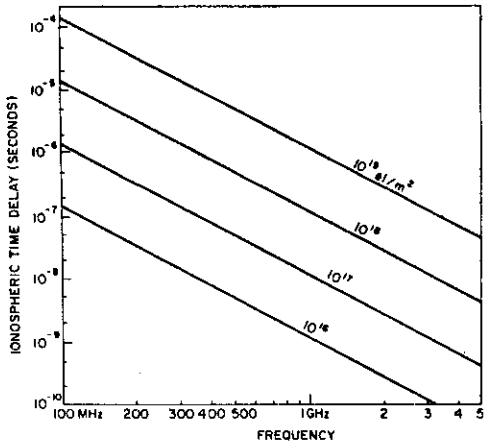


Figure 1. Time Delay vs Frequency for Various Values of TEC

$$\delta(\Delta t) = \frac{k}{c} * \text{TEC}\,(1/f_2^2 - 1/f_1^2)$$

$$= \Delta t_1 (f_1^2 - f_2^2)/f_2^2$$

or

$$\Delta t_1 = f_2^2/(f_1^2 - f_2^2) * \delta(\Delta t)\,. \tag{2}$$

The value $\delta(\Delta t)$ is obtained from the difference of the simultaneous measurements of the total range, including ionospheric time delay, at the two frequencies, f_1 and f_2, since the geometric distance is, of course, the same at all frequencies. The quantity $f_2^2/(f_1^2 - f_2^2)$ is called the ionospheric scaling factor. For small ratios of f_1/f_2 this factor is much larger than unity and the required precision of the differential measurement may be unreasonably large. A plot of this quantity, normalized by f_1, is given in Figure 2. In this derivation the contribution of receiver noise to the differential measurement accuracy has not been considered.

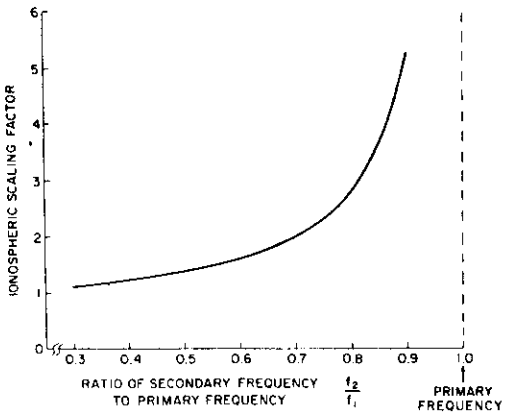


Figure 2. Ionospheric Scaling Factor vs Ratio of Secondary (lower) to Primary (higher) Frequency

1.1.2 AN EXAMPLE OF A TWO-FREQUENCY IONOSPHERIC TIME DELAY SYSTEM

The Department of Defense is currently testing an advanced navigation system, called the NAVSTAR-Global Positioning System (GPS),[1,2] which uses coherently derived, identical modulation on two carrier frequencies, called L1 and L2, to measure the ionospheric group path delay directly and thereby correct for ionospheric time delay. The ratio of frequencies used in the GPS system is exactly 154/120, with the higher frequency (L1) at 1575 MHz. The two carrier frequencies transmitted by the GPS system are the 154th and 120th harmonics of 10.23 MHz. This 10 MHz frequency is bi-phase modulated on both carriers with a psuedo random code resulting in a $[(\sin x)/x]^2$ shaped spectrum of width 20 MHz to the first nulls. A user with knowledge of the transmitted code collapses the received spectrum to equivalent carriers with 10 MHz modulation. The 10 MHz modulation is transmitted with a known phase difference on the two carriers, and the received modulation phase difference is a direct measure of the ionospheric group path delay.

1. Demaro, R.P. (1981) The all-purpose satellite, IEEE Spectrum, 18:35-40.
2. Milliken, R.J., and Zaller, C.J. (1978) Principle of operation of NAVSTAR and system characteristics, Navigation: Journal of the Institute of Navigation, 25(No. 2):95-106.

For the GPS carrier and modulation frequencies the ionospheric group path delay at frequency L1, as obtained from Eq. (2) is:

$$\Delta t_1 = -1.5457\ \delta(\Delta t)$$

where $\delta(\Delta t)$ is the difference between the ionospheric time delay measured at the two frequencies. This difference in range is directly related to absolute ionospheric time delay as, of course, the satellite is at the same range at both frequencies. The only frequency dependent parameter in range measurements is the ionospheric time delay effect, assuming the satellite transmitted modulation phase at L1 and L2 is known and the receiving system frequency dispersive characteristics can be independently measured and corrected for.

For a typical daytime high solar activity TEC value of 10^{18} el/m^2 column the $\delta(\Delta t)$ measured by a GPS receiver would be 54 nanoseconds or 53.1 ft (18.2 m) of ionospheric error. For a direct measure of absolute TEC from the modulation phase delay at L2 minus L1 we have:

$$\text{TEC} = \frac{1.852}{\cancel{1.645}} * 10^{16} * \delta(\Delta t)$$

where $\delta(\Delta t)$ is measured in nanoseconds.

Since, at 10.23 MHz, one complete cycle of modulation phase of 360° is 97.75 nanoseconds, we obtain: TEC = 0.7745×10^{16} el/m^2 per degree of 10.23 MHz modulation phase difference, or TEC = 278.8×10^{16} el/m^2 per cycle of modulation phase difference. Thus, the cycle ambiguity in absolute values of TEC is trivial to resolve using the GPS system as a means of determining ionospheric time delay.

Absolute ionospheric time delay measurements can be made with an accuracy better than 1 to 2 nanoseconds, depending upon the received signal-to-noise ratio on both frequencies. For the power levels transmitted by the GPS satellites, an omni-directional receiving antenna, and a receiver with a modulation tracking bandwidth of approximately 15 Hz, the differential modulation phase has been measured to within approximately ± 2 nanoseconds. The contribution of receiver noise for the two-frequency ionospheric time delay corrections on the GPS system has been considered by Cretcher.[3]

3. Cretcher, C.K. (1975) Ionospheric effects in NAVSTAR-GPS, in Effects of the Ionosphere on Space Systems and Communications, J.M. Goodman, Ed., Naval Research Laboratory, Washington, D.C. US Government Printing Office, Stock No. 008-051-00064-0.

1.2 RF Carrier Phase Advance

In addition to group path delay, or modulation time delay, over the free space delay, the phase of the carrier of radio frequency transmissions is changed by the ionosphere. The RF phase is advanced with respect to its phase in the absence of an ionosphere. This effect is extremely important in determining space object velocities by means of range rate measurements. The amount of phase increase or phase path decrease can be expressed as:

$$\Delta\phi = \frac{1.34 \times 10^{-7}}{f} \text{ TEC} \quad \text{(cycles)} \tag{3}$$

where f is the system operating frequency in Hertz, and TEC is in el/m^2 column. In practice, the amount of this phase advance cannot readily be measured on a single frequency and two, coherently derived, frequencies are required for this measurement.

1.2.1 DIFFERENTIAL CARRIER PHASE

In addition to the dual frequency identical modulation transmitted from the GPS satellites for ionospheric group path correction, these satellites also transmit two, coherently-derived carrier frequencies for ionospheric differential carrier phase measurements. For the pair of frequencies used by GPS, approximately 1.2 and 1.6 GHz, the differential phase shift, referenced to the lower frequency, is:

$$\Delta\phi \quad \frac{1.34 \times 10^{-7}}{f_2} \times \frac{(m^2 - 1)}{m^2} \times \text{TEC} \quad \text{(cycles)} \tag{4}$$

where $m = f_1/f_2 = 1.2833$. For the GPS frequencies $\Delta\phi = 4.31 \times 10^{-17} *$ TEC or $2.32 * 10^{16}$ el/m^2 per complete 2π cycle of differential carrier phase between L1 and L2, measured at L2. The differential carrier phase [Eq. (4)] is related to the differential modulation phase, [Eq. (2)] simply by the ratio of carrier to modulation frequencies. With a reasonable carrier signal-to-noise ratio, this differential carrier phase can be measured to within a few degrees, or less than approximately $0.04 * 10^{16}$ el/m^2. Since the TEC is generally much greater than 2.32×10^{16}, corresponding to 2π of differential carrier phase, there is a $2n\pi$ ambiguity in the differential phase measurement.

The differential carrier method of measuring TEC cannot, in practice, be used to measure absolute values of TEC by itself due to the large $2n\pi$ phase ambiguity in the measurement, but this is not important for navigation systems which require a correction only for range rate errors due to the ionosphere between two measurement times.

The US Navy Navigation Satellite System, NNSS.* determines position for stationary and slowly moving vehicles by measuring satellite transmitted RF carrier phase changes as a function of low-orbit satellite motion across the sky. This method of positioning requires only range rate information. The primary NNSS frequency is 400 MHz. A second RF carrier at 150 MHz is used only for ionospheric range rate corrections. While various techniques have been proposed for determining the absolute TEC from the differential carrier phase information received from the NNSS satellites, they all involve assumptions concerning some a-priori knowledge of the ionosphere, and they cannot be used in the general case.

As an ionospheric monitoring tool the combination of differential carrier phase and differential modulation phase provides an excellent means of determining ionospheric electron content along the ray path to the satellite. The absolute value of TEC can be determined by the group delay technique and relative TEC changes can be measured with great accuracy by the differential carrier phase technique.

1.2.2 SECOND DIFFERENCE OF CARRIER PHASE

The second difference in phase between an RF carrier and that of its upper and lower sidebands can be used to measure absolute values of TEC, as described by Burns and Fremouw.[4] If three, coherently derived frequencies, $f - f_m$, f, and $f + f_m$ are transmitted the second difference of phase is given by:

$$\Delta_2\phi = (\phi_u - \phi_c) - (\phi_c - \phi_L) = \phi_u + \phi_L - 2 \times \phi_c .$$

From Eq. (3),

$$\Delta\phi = \frac{1.34 \times 10^{-7}}{f} \times \text{TEC} \quad \text{(cycles)}$$

thus

$$\Delta_2\phi = \frac{2.68 \times 10^{-7} f_m^2}{f(f^2 - f_m^2)} \text{ TEC} \quad \text{(cycles)}$$

when $f^2 >> f_m^2$

$$\Delta_2\phi = \frac{2.68 \times 10^{-7} f_m^2}{f^3} \text{ TEC} \quad \text{(cycles)} . \tag{5}$$

---

*See References 30 and 31, Reference page 29.

4. Burns, A.A., and Fremouw, E.J. (1970) A real-time correction technique for transionospheric ranging error, IEEE Trans. on Antennas and Propag. AP-18(No. 6):

For a carrier frequency of 100 MHz a modulation frequency of 1.93 MHz would be required to give 2π of second differential phase for a TEC value of 10^{18} el/m^2. A value of $\Delta 2\phi$ of 2π for 10^{18} el/m^2 is a reasonable compromise between the requirement for minimizing chances of an ambiguity in absolute TEC and accuracy in measuring TEC relative changes. The second difference of carrier phase has been used with the DNA-002 satellite to make estimates of the absolute values of TEC.[5]

1.3 Doppler Shift

Since frequency is simply the time derivative of phase, an additional contribution to geometric Doppler shift results due to changing TEC. This additional frequency shift is generally small compared to the normal geometric Doppler shift, but can be computed by:

$$\Delta f = \frac{d\phi}{dt} = \frac{1.34 \times 10^{-7}}{f} \frac{d}{dt} \text{ TEC} \qquad \text{(Hertz)}$$

For high orbit satellites where the diurnal changes in TEC are greater than geometric ones, an upper limit to the rate of change of TEC is approximately $0.1\ 10^{16}$ el/m^2 per second. This value yields an additional frequency shift of less than one-tenth of a Hertz at 1.6 GHz which would not be significant compared with a typical required receiver loop bandwidth of at least a few Hertz. At 400 MHz a similar rate of change of TEC would produce a frequency shift of approximately 0.3 Hz, probably still not significant.

During times of severe phase scintillation, which can occur even at GHZ frequencies, the TEC may not change in a consistent, rapid manner to yield greater ionospheric Doppler shifts, but the phase of the incoming RF signal can have a large random fluctuation superimposed upon the changes associated with the normal rate of change in TEC. This large, random component may actually spread out the spectrum of the received signal sufficiently to cause the receiver to lose phase lock, as the receiver signal phase may have little energy remaining in the carrier, and instead may be spread over several Hertz, with little recognizable carrier remaining. A knowledge of phase scintillation rates is required to determine the spread of received signal phase.

5. Fremouw, E.J., Leadabrand, R.L., Livingston, R.C., Cousins, M.D., Rino, C.L., Fair, B.C., and Long, R.A. (1978) Early results from the DNA wideband satellite experiment-complex signal scintillation, Radio Science, 13(No. 1):167-187.

1.4 Faraday Polarization Rotation

When a linearly polarized radio wave traverses the ionosphere the wave undergoes rotation of the plane of polarization. At frequencies of approximately 100 MHz and higher the amount of this polarization rotation can be described by:

$$\Omega = \frac{2.36 \times 10^{-5}}{f^2} \int B \cos\theta \; N \, dl \qquad (6)$$

where the quantity inside the integral is the product of electron density times the longitudinal component of the earth's magnetic field, integrated along the radio wave path. Many ionospheric workers have used this effect, named for Michael Faraday who first observed polarization changes in an optical experiment, to make measurements of the TEC of the ionosphere. Since the longitudinal magnetic field intensity changes much slower with height than the electron density of the ionosphere, the equation can be rewritten as:

$$\Omega = \frac{K}{f^2} B_L * \text{TEC} \qquad (7)$$

where $B_L = \overline{B \cos\theta}$ is taken at a mean ionospheric height, usually near 400 km, $K = 2.36 \times 10^{-5}$ and TEC is $\int N \, dl$. Typical values of polarization rotation for northern mid-latitude stations viewing a geostationary satellite near their station meridian are given in Figure 3 as a function of system frequency and total electron content. In fact, the largest portion of TEC data available today from stations throughout the world have come from Faraday rotation measurements from geostationary satellite VHF signals of opportunity.

For satellite navigation and communication designers, however, the Faraday polarization rotation effect is a nuisance. If a linearly polarized wave is transmitted from a satellite to an observer, on, or near, the surface of the earth, the amount of polarization rotation may be nearly an odd integral multiple of 90 degrees, thereby giving no signal on the receiver's linearly polarized antenna, unless the user is careful to realign his antenna polarization for maximum received signal.

As shown in Figure 3, at 4 GHz, a commercial satellite transponder frequency band, the amount of Faraday rotation can be a tenth of a radian, well in excess of that required for dual, linear orthogonal channel separation.

The Faraday rotation problem is overcome by the use of circular polarization of the correct sense at both the satellite and at the user's receiver. Generally the mobile user finds it difficult to utilize circular polarization due to the continual vehicle directional changes; thus he settles for a received linear polarization. The 3 dB loss between transmitted circular polarization and receiver linear polarization is a necessary price to pay for user antenna maneuverability and simplicity.

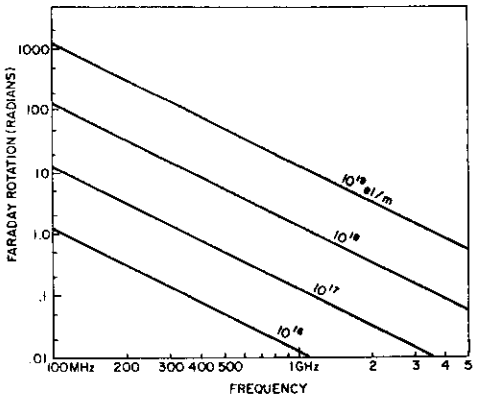


Figure 3. Faraday Polarization Rotation vs Frequency for Various Values of TEC

1.5 Angular Refraction

The refractive index of the earth's ionosphere is responsible for the bending of radio waves from a straight line geometric path between satellite and ground. This angular refraction or bending produces an apparent higher elevation angle than the geometric elevation. Millman and Reinsmith[6] have derived expressions relating the refraction to the resultant angular bending. Perhaps the easiest expressions to use, as given by Millman and Reinsmith[6] relate the ionospheric range error to angular refraction:

$$\Delta E = \frac{R + r_o \sin E_o \; (r_o \cos E_o)}{h_i (2\, r_o + h_i) + r_o^2 \sin^2 E_o} * \frac{\Delta R}{R} \qquad (8)$$

where E_o is the apparent elevation angle, R is the apparent range, ΔR is computed from $\Delta R = (40.3/f^2) * TEC$, r_o is the earth's radius, and h_i is the height of the centroid of the TEC distribution, generally between 300 and 400 km.

6. Millman, G. H., and Reinsmith, G. M. (1974) An Analysis of the Incoherent Scatter-Faraday Rotation Technique for Ionospheric Propagation Error Correction, Report No. R74EMH2, General Electric Company, Syracuse, New York.

For low elevation angles and satellites well above most of the ionization, $R >> r_c \sin E_o$, and the angular refraction can be expressed as:

$$\Delta E = \frac{\cos E_o}{2\, h_i} \Delta R . \qquad (9)$$

Typical values of elevation refraction error for a TEC of 10^{18} el/m^2 column are shown in Figure 4 for several frequencies. Note that, at the lowest frequency, 100 MHz, near the horizon the refraction is well over 1.5 degrees! The curves shown in Figure 4 have been constructed using the approximation derived by Millman and Reinsmith[6] for low elevation angles given in Eq. (9).

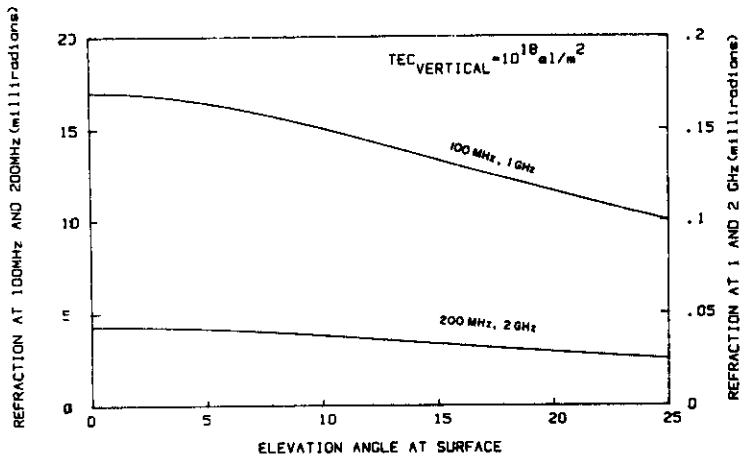


Figure 4. Refraction in Elevation Angle vs Elevation Angle for Indicated Frequencies and Values of TEC

Generally, the range error itself is the main ionospheric problem for advanced navigation systems, and elevation angle errors are insignificant. Satellite detection radar systems, on the other hand, do have the requirement to know accurate pointing elevation angles for their large aperture arrays, though generally the accurate tracking is done by using range rate information, and elevation angle is of secondary importance as long as the beamwidth of the antenna is large enough to see the target.

Errors in the azimuth of radio waves transmitted through the ionosphere can also occur; they depend upon azimuthal gradients in TEC which are generally small and which usually can be neglected in practical cases.

1.6 Distortion of Pulse Waveforms

Two characteristics of the ionosphere can produce distortion of pulses of RF energy propagated through it. The dispersion, or differential time delay due to the normal ionosphere, as derived by Millman[7] is proportional to $1/f^2$, and produces a difference in pulse arrival time across a bandwidth Δf of:

$$\Delta t = \frac{80.6 * 10^6}{c f^3} \Delta f \times TEC \qquad (10)$$

where c is the velocity of light in m/sec, f and Δf are expressed in Hertz, and TEC is in el/m^2 column. The dispersive term for pulse distortion is thus proportional to TEC. When the difference in group delay time across the bandwidth of the pulse is the same magnitude as the width of the pulse it will be significantly disturbed by the ionosphere. Millman and Bell[8] also derived mathematical relationships for ionospheric dispersive effects on an FM Gaussian shaped pulse.

In addition to pulse distortion by the dispersive effects due to the TEC of the normal background ionosphere, radio pulses are also modified by scattering from ionospheric irregularities. Yeh and Liu[9] have computed pulse mean arrival time and mean pulsewidth due to both dispersion and scattering.

2. IONOSPHERIC TOTAL ELECTRON CONTENT (TEC)

2.1 Average TEC Behavior

The ionospheric parameter responsible for the effects described in Section 1 of this report is the total number of free electrons, TEC, or its rate of change, along the path from a satellite to a ground station. The greatest contribution to TEC comes from the F2 region of the ionosphere. A typical daytime mid-latitude, high solar maximum electron density profile is illustrated in Figure 5. The curve on the left side of Figure 5 is the log of N_e plotted vs height as normally shown by

7. Millman, G.H. (1965) in Modern Radar, Ed. by R.S. Berkowitz, John Wiley & Sons, New York.
8. Millman, G.H., and Bell, C.D. (1971) Ionospheric dispersion of an FM electromagnetic pulse, IEEE Trans. Antennas Propag. AP-19(No. 1).
9. Yeh, K.C., and Liu, C.H. (1979) Ionospheric effects on radio communication and ranging pulses. IEEE Trans. Antennas Propag. AP-27(No. 6).

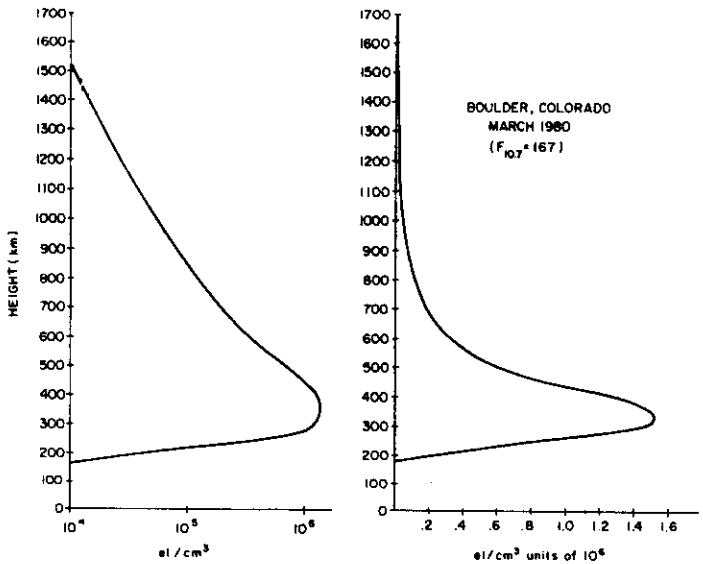


Figure 5. Typical Profile of Electron Density vs Height. In (a), log N_e is plotted; in (b) N_e is plotted on a linear scale

ionospheric workers. Since the TEC is represented by the area under the curve of linear plot of N_e vs height, the right hand plot of Figure 5 illustrates the actual linear plot. Note that most of the contribution to TEC occurs near the peak of the F2 region, which is fortunate, as ground-based ionosondes have been used since the 1930's to make continuous, routine measurements of the density at the peak of the F2 region. Ionosondes measure foF2, which is related to N_{max} by:

$$(foF2)^2 = 80.6\ N_{max}$$

where foF2 is in MHz, and N_{max} is in units of 10^6 el/cc.

In the 1950's and 1960's continuing to a more limited extent even today, upwards of 150 ionosondes were operated to provide improved prediction capability for long distance high frequency propagation by means of ionospheric refraction. Various models of foF2 were developed for this purpose, one of the more popular ones being

commonly known as ITS-78[10] after the report number which described the model. This model, among other things, characterized the 10-day average world-wide behavior of foF2 by Fourier temporal components and Legendre polynomial geographic coefficients ordered by magnetic, rather than geographic, latitude. The success of this experimental, data based, or empirical model, in representing the actual world-wide foF2 is due to the large amount of data available from ionosondes in many regions of the world. Other characteristics of this model are discussed in Dandekar.[11]

For the TEC parameter, data availability have been, and will likely continue to be, much more sparse. First, TEC measurements have been made generally by using measurements of Faraday polarization rotation using VHF signals of opportunity transmitted from geostationary satellite telemetry transmitters. A few lunar reflected Faraday rotation measurements in the late 1950's and early 1960's and the TEC obtained from a few low orbit satellites did not contribute significantly to our knowledge of world-wide TEC behavior, at least not for modelling average ionospheric conditions. Only since the early to mid-1960's have TEC values been obtained on a more-or-less regular basis. Even today fewer than one dozen stations regularly contribute TEC data, which can be used in TEC modelling purposes, to a world data center.

Fortunately, most of the contribution to TEC comes from near the F2 region density peak where models of foF2 are available. These foF2 models can be combined with some limited knowledge of topside ionospheric thickness obtained from topside sounders, along with topside in-situ density measurements, to produce a complete ionospheric height profile model. The most well known of these models is the one by Bent,[12] which uses ITS-78 coefficients for foF2, and topside exponentials for computing TEC. A representation of world-wide average behavior of TEC is illustrated in Figure 6 for 2000 hrs UT. To first order the TEC contours shown in Figure 6 move westward along magnetic, rather than geographic, latitude lines, at the earth's rotation rate. The Bent model was constructed using solar maximum data from the 1968-69 period and had to be adjusted upward somewhat to account for the much higher 1979-1980 solar maximum than that of 1968-1969. This adjustment was necessary to adequately represent the actual TEC values from stations making observations in March 1980, which was near the maximum of the second highest solar cycle ever recorded in the over 200-year history of solar

10. Barghausen, A. L., Finner, J. W., Proctor, L. L., and Schultz, L. D. (1969) Predicting Longterm Operational Parameters of High-frequency Sky-wave Telecommunication Systems, ESSA Tech Report, ERL 110-ITS 78.

11. Dandekar, B. S. (1982) Ionospheric Modeling, AFGL-TR-82-0024, AD A115243.

12. Llewellyn, S. K., and Bent, R. B. (1973) Documentation and Description of the Bent Ionospheric Model, AFCRL-TR-73-0657, AD 772733.

cycle observations. The Bent model, appropriately adjusted for high solar cycle values, does however, represent fairly well the average behavior of TEC for many locations tested. Other world-wide ionospheric electron density profile models from which average TEC can be obtained includes one by Ching and Chiu,[13] and Chiu,[14] Kohnlein,[15] the 4-D model,[16] and the International Reference Ionosphere (IRI).[17] The characteristics of some of these models are described in Dandekar.[10]

Other empirical models of TEC have been developed directly from TEC data alone, though these have necessarily been limited in temporal and geographic extent by the available data base. These include models of TEC over Europe and the Mediterranean[18] for low and medium activity portions of the 11-year solar cycle, and a model of TEC over the Indian sub-continent for both solar minimum and for an average solar maximum.[19] Models of the slab thickness parameter, the ratio of TEC/N_{max} have been developed for specific regions such as the one for northern Europe by Kersley,[20] and one for the eastern USA by Klobuchar and Allen,[21] from which TEC can be obtained from a model of foF2. An algorithm designed for an approximate 50 percent correction to world-wide TEC, for use in an advanced navigation system, has been developed by Klobuchar.[22]

All of the models listed here, and the list is by no means complete, are empirical models which attempt to correct for average TEC behavior only. However, the variability from average TEC behavior can be large and may be important to some radio wave systems which must propagate signals through the ionosphere.

2.2 Temporal Variability of TEC

2.2.1 VARIABILITY FROM MONTHLY MEAN TEC VALUES

The ionosphere is a weakly ionized plasma and the resultant TEC is a function of many variables including solar ionizing radiation, neutral wind and electric field effects, neutral composition and temperature changes. A monthly overplot of curves of diurnal changes in TEC for a northern mid-latitude station for twelve months during a solar maximum period is shown in Figure 7. The standard deviation from monthly mean diurnal behavior is approximately 20-25 percent, during the daytime hours when the absolute TEC values are greatest. Figure 8 shows the standard deviation from monthly average TEC behavior for the mid-day hours for a number of stations during the solar maximum period 1968-1969. Again 20-25 percent is a good value for the standard deviation from the monthly average behavior. The standard deviation is somewhat higher during the nighttime hours, but the absolute TEC values are much lower during those periods.

(Due to the large number of references cited above, they will not be listed here. Refer to References, page 29.)

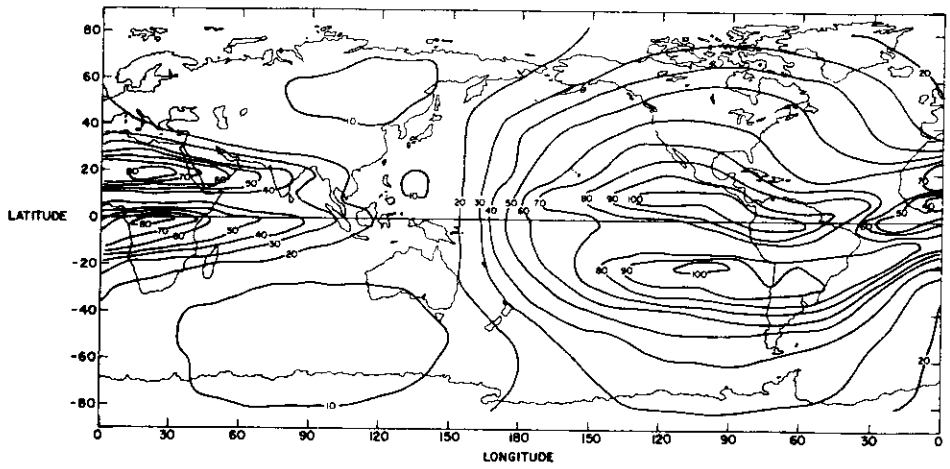


Figure 6. Contours of Vertical TEC, in Units of 10^{16} el/m^2 Column, for 2000 UT, March 1980

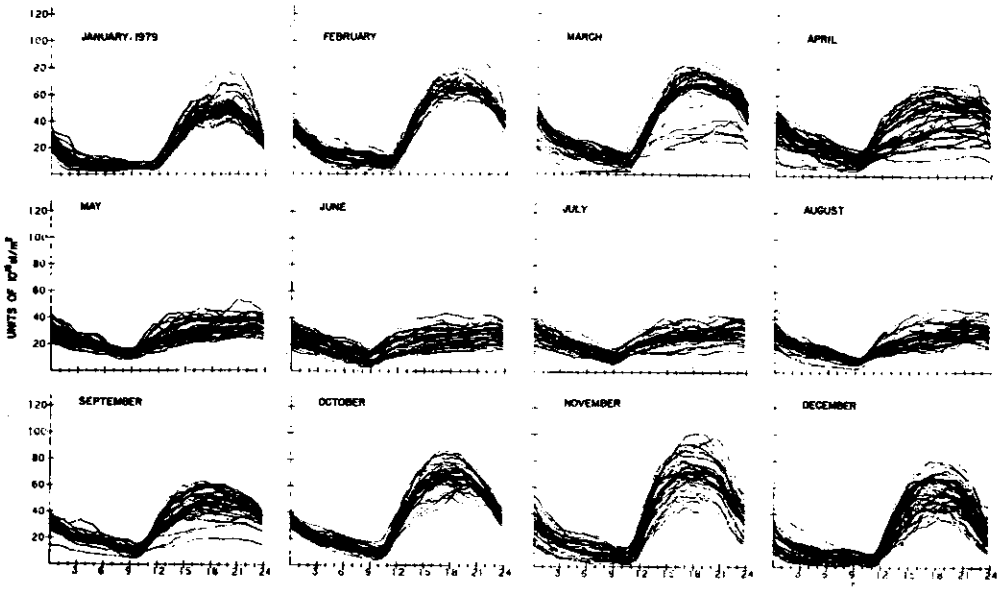


Figure 7. Monthly Overplots of TEC Diurnal Curves vs UT for Hamilton, Massachusetts for 1979

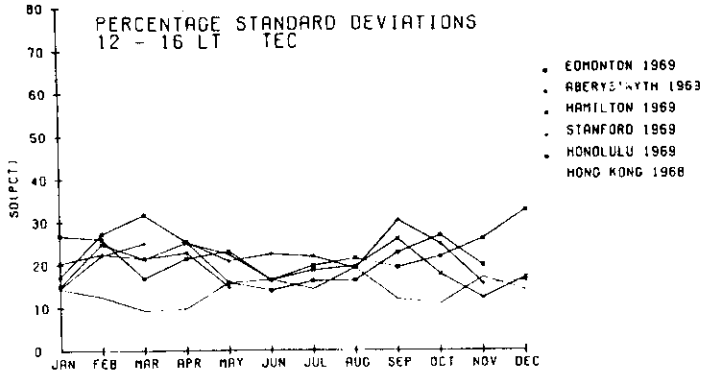


Figure 8. Percentage Standard Deviations for Daytime TEC From the Stations Indicated

If a satellite ranging system has error requirements such that it must correct for monthly average ionospheric time delay, but still can tolerate the approximate 20-25 percent variability of TEC from monthly average conditions, approximately 70-80 percent of the ionospheric effect on the system can be eliminated by the use of an average TEC model such as the one constructed by Bent.[12] If the system only requires an approximate 50 percent rms correction of the ionospheric time delay the algorithm developed by Klobuchar[22] can be used. On the other hand, if corrections for some portion of the remainder of the ionospheric time delay are required, after a state of the art TEC model, such as the Bent one, has been used to take out the monthly mean TEC, then the short term (a few hours) temporal variability as well as the geographic variability, of TEC must be considered.

2.2.2 SHORT TERM TEMPORAL VARIABILITY OF TEC

The correlation time of departures of TEC from monthly average curves has been studied by Donatelli and Allen.[23] They concluded, for the mid-latitude station they studied, that the useful prediction time was a function of local time, season, and long term sunspot activity. However, in most cases they found no significant improvement over the use of monthly mean predictions when they used actual data over three hours old. The longest useful prediction interval occurred, fortunately,

23. Donatelli, D.C., and Allen, R.S. (1981) Time cells for adaptive predictions of total electron content, Radio Sci. 16(No. 2):261-269.

during solar maximum daytime hours when absolute TEC values are highest. During solar minimum periods their useful prediction time interval was often as short as one hour.

The TEC prediction study done by Donatelli and Allen[23] was for data in which the predictions were for the same geographic location and direction in the sky as the measurements. If the prediction is for a different location the temporal correlation will be lower.

2.2.3 GEOGRAPHIC VARIABILITY OF TEC

The variability of TEC at the same local time, but as a function of distance has been studied by Klobuchar and Johanson.[24] They used TEC data from two sets of stations, one aligned approximately along an east-west direction, with the other set of stations aligned along an approximate north-south direction. Their results are shown in Figures 9a and 9b for the east-west and the north-south station alignments, respectively. No significant difference in correlation distance was found with season.

The percent improvement, P.I., in TEC from the average value, if a measured value is available from a location where the correlation coefficient between the two locations is r, is related to the correlation coefficient, r, by:

$$P.I. = 100 * [1 - (1 - r^2)^{0.5}]$$ (see Gautier and Zacharisen[25]).

Note that a correlation coefficient of 0.7 explains only 29 percent of the variance between the data at station pairs; hence a measurement at one station, with a correlation coefficient of 0.7 between data sets with a second station would result in an improvement at the second station over the average predicted value of only 29 percent.

2.3 TEC in the Near-Equatorial Region

All of the preceeding sections have concentrated on the behavior of TEC in the mid-latitude regions of the world, mainly because most of the available data are from that region. The near-equatorial region deserves special mention due to the fact that the highest TEC values in the world occur in this region, as shown in Figure 6. This region extends to approximately ± 20-25 degrees on either of the magnetic equator, with the highest TEC values not at the equator, but rather at the

24. Klobuchar, J.A., and Johanson, J.M. (1977) Correlation Distance for Mean Daytime Electron Content, AFGL-TR-77-0185, AD A048117.

25. Gautier, T.N., and Zacharisen, D.H. (1965) Use of space and time correlations in short term ionospheric predictions, IEEE Conf. Rec. 1st Annual IEEE Communications Convention, pp. 671-676.

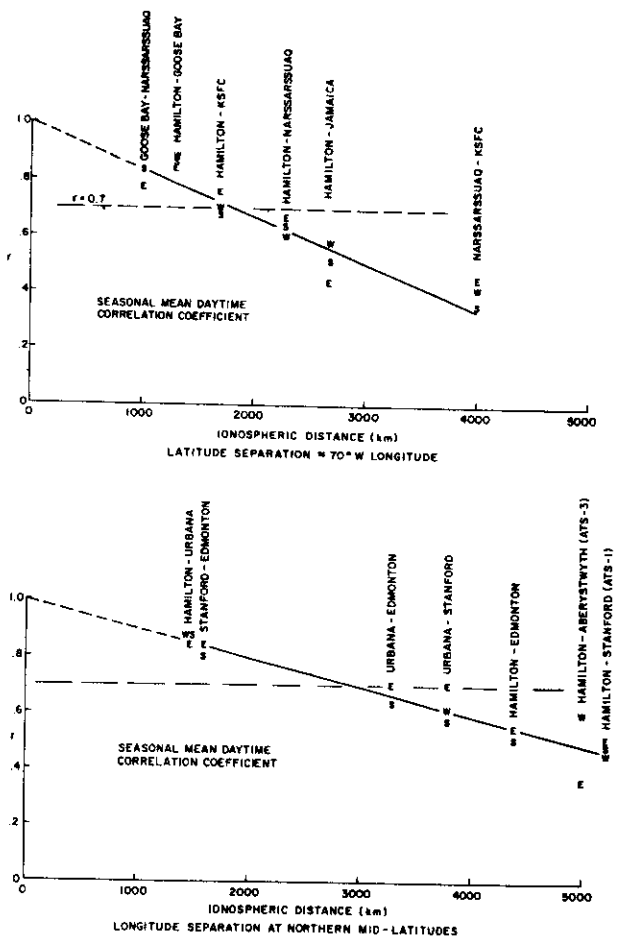


Figure 9. Correlation Coefficient vs Station Separation in Latitude (9a) and Longitude (9b)

so called "equatorial anomaly" regions located at approximately ± 15 degrees from the magnetic equator. The regions of highest TEC values at 2000 hrs UT are clearly seen in Figure 6 near 100 degrees west longitude.

Most of the day-to-day geographic variability of TEC in the equatorial anomaly region during solar minimum conditions can be explained by the variability of ecuatorial electrojet strength. Unfortunately, no such similar TEC data are available for solar maximum.

An example of the high temporal variability of TEC for solar maximum conditions for Ascension Island, a station located near the peak of the southern TEC equatorial crest region, is shown in Figure 10. Note the extremely large day-to-day TEC variability in the afternoon and evening hours in some months. Any satellite ranging system requiring ionospheric TEC corrections in the near-equatorial region should not use the mid-latitude standard deviation values of approximately 20-25 percent to represent the variability of the near-equatorial region.

2.4 TEC in the Auroral and Polar Cap Regions

Since most available TEC values have been measured from radio signals transmitted from geostationary satellites, which can be viewed only at low elevation angles from high latitudes, knowledge of the variability of TEC in the auroral and polar cap regions is sparse. In the American longitude sector, where the magnetic latitudes are lowest for a given geographic latitude, there is considerable TEC data from Goose Bay, Labrador. At Goose Bay the aurora passes southward even during moderately magnetically disturbed periods. The behavior of TEC during those periods can be highly irregular, especially during the nighttime hours when TEC values often exhibit rapid changes and occasionally even exceed the daytime maximum values briefly. While the occurrence of general auroral activity may be predictable, the specific large increases in TEC, likely due to auroral precipitation, are not individually predictable, but may be statistically characterized as a function of magnetic activity.

In the polar cap region a negligible amount of TEC data exists. The absolute TEC values are likely lower in this region than in the mid-latitudes, and the variability of the polar cap TEC is probably very high.

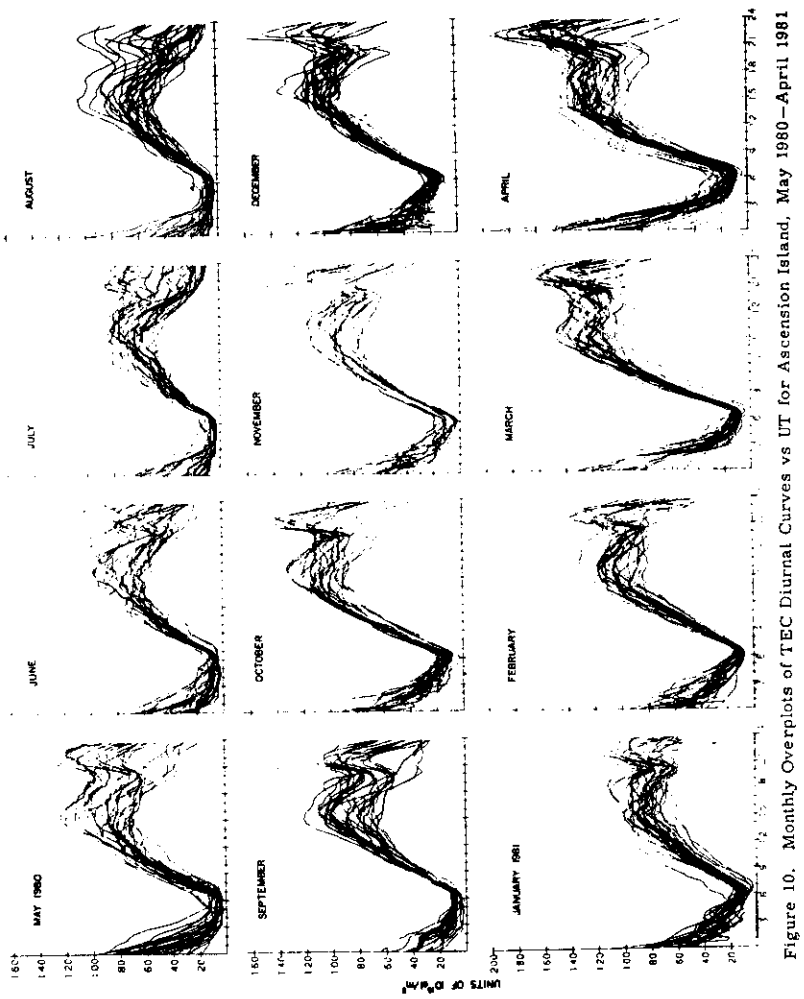


Figure 10. Monthly Overplots of TEC Diurnal Curves vs UT for Ascension Island, May 1980–April 1981

2.5 Protonospheric Electron Content

Most of the available TEC data has been taken by measurements of Faraday rotation of single frequency radio waves transmitted from geostationary satellites to ground observers. The electron content obtained from Faraday rotation observations, while made from radio waves transmitted from satellites at geostationary satellite height above the earth's surface, includes only the contribution of electrons up to heights of approximately 2000 km. This is because the integrated product of the longitudinal component of the earth's magnetic field times the electron density, above approximately 2000 km, is negligible. The only measurements of the additional contribution of electrons above the Faraday maximum height have been made using signals from an ionospheric beacon on the geostationary satellite ATS-6. Davies[26] has reviewed the overall results of the ATS-6 experiment. A summary of typical protonospheric electron content data is shown in Figure 11 taken from Klobuchar et al.[27] Note that the protonospheric values are fairly low in absolute value.

During the nighttime hours when the ionospheric TEC is low, the protonospheric contribution may become a fairly large percentage of the total number of electrons between a satellite at geostationary height and an observer on, or near the earth's surface. Unfortunately, no protonospheric electron content are available during solar maximum conditions.

2.6 Short Term Variations in TEC

The time rate of change of TEC, in addition to the normal diurnal variations, also has periodic variations due to perturbations of the ionospheric F region from various potential sources as geomagnetic substorms, meteorological sources such as weather fronts, shock waves from supersonic aircraft, volcanic explosions, rocket launches, and other miscellaneous sources. While these short term variations in TEC cover a large range of periods and amplitudes, common periods range from 20 to over 100 min with amplitudes of a few percent of the background TEC. A 10 percent ionospheric disturbance with respect to the background TEC is very rare, while a 1 percent TEC perturbation is common. Titheridge[28] and Yeh[29] have made studies of the statistics of travelling ionospheric disturbances (TID's), in TEC for mid-latitude regions.

A system which requires correction for the rate of change of TEC cannot rely on models of TEC to provide reliable information on short term rate of change of TEC information, and can use available TID information only in a statistical manner.

(Due to the large number of references cited above, they will not be listed here. See References, page 29.)

The only recourse for a system significantly affected by rate of change of TEC is to use a dual frequency measurement technique to measure directly the ionospheric contribution to range rate.

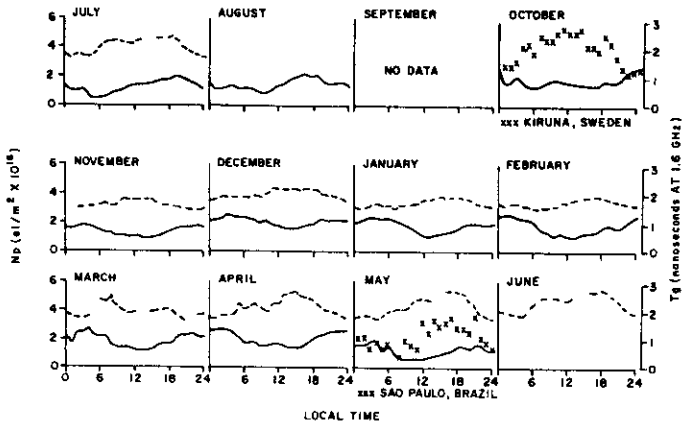


Figure 11. Monthly Average Plasmaspheric Electron Content vs Local Time for Aberystwyth, Wales (dashed line) and for Hamilton, Massachusetts (solid line). Also ploted are values for Kiruna, Sweden for October 1975 and from Sao Paulo, Brazil for May 1975

3. CONCLUSIONS

There are at least three categories of systems potentially affected by ionospheric time delay. For the first category of user the potential systems effects may be small, at least under any naturally occurring worst case ionospheric conditions. In the second category, a user may require a nominal correction for average ionospheric time delay, but is able to tolerate the 20-25 percent standard deviation from average conditions. He should expect at least a 50 percent correction for ionospheric time delay effects using a relatively simple time delay algorithm, and up to 70-80 percent for a state of the art, fairly complex model. These model corrections can be improved by the use of actual ionospheric measurements within a reasonable temporal and spatial frame. For the third category of user, ionospheric model corrections, even updated with near-real-time measurements, may not be

sufficient to correct for ionospheric time delay, and the system must then make its own ionospheric correction. Fortunately, the ionosphere is a dispersive medium and the use of identical modulation on two, widely-spaced frequencies will allow a direct measurement to be made of ionospheric range delay. Two coherently-derived carrier frequencies may be used to obtain accurate time rate of change information for TEC. Details of measuring ionospheric effects directly by a system's use of multiple frequencies are available in Burns and Fremouw.[4]

References

1. Demaro, R. P. (1981) The all-purpose satellite, IEEE Spectrum, 18:35-40.
2. Milliken, R. J., and Zaller, C. J. (1978) Principle of operation of NAVSTAR and system characteristics, Navigation: Journal of the Institute of Navigation, 25(No. 2):95-106.
3. Cretcher, C. K. (1975) Ionospheric effects in NAVSTAR-GPS, in Effects of the Ionosphere on Space Systems and Communications, J. M. Goodman, Ed., Naval Research Laboratory, Washington, D.C. US Government Printing Office, Stock No. 008-051-00064-0.
4. Burns, A. A., and Fremouw, E. J. (1970) A real-time correction technique for transionospheric ranging error, IEEE Trans. on Antennas and Propag. AP-18(No. 6):
5. Fremouw, E. J., Leadabrand, R. L., Livingston, R. C., Cousins, M. D., Rino, C. L., Fair, B. C., and Long, R. A. (1978) Early Results from the DNA wideband satellite experiment-complex signal scintillation, Radio Science, 13(No. 1):167-187.
6. Millman, G. H., and Reinsmith, G. M. (1974) An Analysis of the Incoherent Scatter-Faraday Rotation Technique for Ionospheric Propagation Error Correction, Report No. R74EMH2, General Electric Company, Syracuse, New York.
7. Millman, G. H. (1965) in Modern Radar, Ed. by R. S. Berkowitz, John Wiley & Sons, New York.
8. Millman, G. H., and Bell, C. D. (1971) Ionospheric dispersion of an FM electromagnetic pulse, IEEE Trans. Antennas Propag. AP-19(No. 1):
9. Yeh, K. C., and Liu, C. H. (1979) Ionospheric effects on radio communication and ranging pulses, IEEE Trans. Antennas Propag. AP-27(No. 6):
10. Barghausen, A. L., Finner, J. W., Proctor, L. L., and Schultz, L. D. (1969) Predicting Longterm Operational Parameters of High-frequency Sky-wave Telecommunication Systems, ESSA Tech. Report, ERL 110-ITS 78.
11. Dandekar, B. S. (1982) Ionospheric Modeling, AFGL-TR-82-0024, AD A115243.

12. Llewellyn, S.K., and Bent, R.B. (1973) Documentation and Description of the Bent Ionospheric Model, AFCRL-TR-73-0657, AD 772733.

13. Ching, B.K., and Chiu, Y.T. (1973) A phenomenological model of global ionospheric electron density in the E, F1, and F2 regions, J. Atmos. Terr. Phys. 35:1615.

14. Chiu, Y.T. (1975) An improved phenomenological model of ionospheric density, J. Atmos. Terr. Phys. 37:341.

15. Kohnlein, W. (1978) Electron density models of the ionosphere, Rev. of Geophys. and Space Phys. 16:341.

16. von Flotow, C.S. (1978) Ionospheric forecasting at Air Force Global Weather Central, in Effects of the Ionosphere on Space and Terrestrial Systems, J.M. Goodman, Ed., US Government Printing Office, Stock No. 008-051-00069-1.

17. Rawer, K. (1981) in International Reference Ionosphere-IRI79, Editors, Lincoln, J.V., and Conkright, R.O., World Data Center A for STP, NOAA Boulder, Colorado.

18. Klobuchar, J.A. (1973) Numerical Models of Total Electron Content over Europe and the Mediterranean and Multi-Station Scintillation Comparisons, AGARDOGRAPH: AG-166A.

19. Klobuchar, J.A., Iyer, K.N., Vats, H.O., and Rastogi, R.G. (1977) A numerical model of equatorial and low latitude total electron content for use by satellite tracking systems for ionospheric corrections, Indian Journal of Radio and Space Physics, 6:

20. Kersley, L. (1980) An Empirical Model of Ionospheric Slab Thickness, AGARD-CPP-284.

21. Klobuchar, J.A. and Allen, R.S. (1970) A First-Order Prediction Model of Total-Electron-Content Group Path Delay for a Midlatitude Ionosphere, AFCRL-70-0403, AD 711365.

22. Klobuchar, J.A. (1975) A First-Order, Worldwide, Ionospheric Time Delay Algorithm, AFCRL-TR-75-0502, AD A018862.

23. Donatelli, D.C., and Allen, R.S. (1981) Time cells for adaptive predictions of total electron content, Radio Sci. 16(No. 2):261-269.

24. Klobuchar, J.A., and Johanson, J.M. (1977) Correlation Distance for Mean Daytime Electron Content, AFGL-TR-77-0185, AD A048117.

25. Gautier, T.N., and Zacharisen, D.H. (1965) Use of space and time correlations in short term ionospheric predictions, IEEE Conf. Rec. 1st Annual IEEE Communications Convention, pp. 671-676.

26. Davies, K. (1980) Recent progress in satellite radio beacon studies with particular emphasis on the ATS-6 radio beacon experiment, Space Sci. Reviews, 25:357-430.

27. Klobuchar, J.A., Buonsanto, M.J., Mendillo, M.J., and Johanson, J.M. (1978) The contribution of the plasmasphere to total time delay, in Effects of the Ionosphere on Space and Terrestrial Systems, J.M. Goodman, Ed. US Government Printing Office, Stock No. 008-051-00069-1.

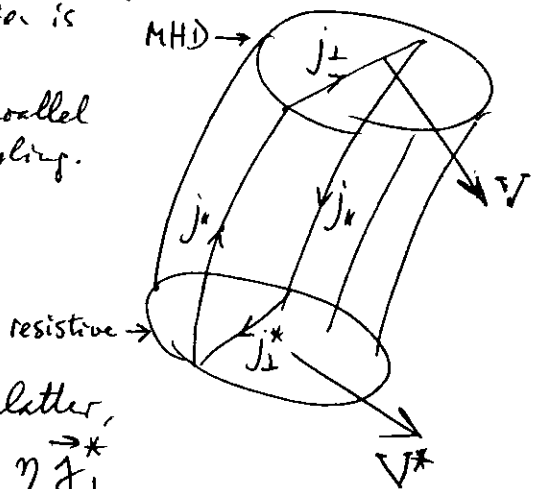
28. Titheridge, J.E. (1968) Periodic disturbances in the ionosphere, J. Geophys. Res. 73:243-252.

29. Yeh, K.C. (1972) Travelling ionospheric disturbances as a diagnostic tool for thermospheric dynamics, J. Geophys. Res. 77:709-719.

30. Black, H.D. (1980) The transit system, 1977; Performance, plans and potential, Philos. Trans. R. Soc. Lond. A294:217-236.

31. Kouba, J. (1983) A review of geodetic and geodynamic satellite Doppler positioning, Rev. Geophys. and Space Phys. V21(No. 1):27-40.

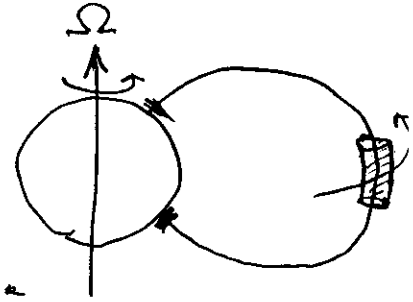
All the above is an example of the general case in which a collisionless plasma region is magnetically coupled to a resistive medium. The parallel currents $j_\parallel$ mediate this coupling. In the MHD region, we have $\vec{E} + \vec{V} \times \vec{B} = 0$, which determines E, which is then "mapped" onto the resistive region. In the latter, we have $\vec{E}^* + \vec{V}^* \times \vec{B}^* = \eta \vec{j}_\perp^*$



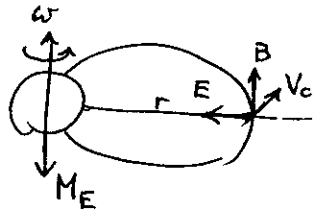
Where η is the resistivity (tensor). Note that $\vec{E}^* + \vec{V}^* \times \vec{B}^*$ is the electric field in the frame that moves with the resistive medium.

Only if $\vec{E}^* + \vec{V}^* \times \vec{B}^* \equiv 0$, will there be no $j_\perp^*$, hence no energy dissipation, and also no $j_\parallel$ (hence no $j_\perp$ in the MHD region). Therefore, the condition that there be no interaction between the two regions is that in both the frozen-in condition apply. This means that V^* must be such that there is no electric field in that frame of reference: both V^* and V must correspond to what we have defined as moving field lines! In other words, the resistive region must co-move (magnetic field line-wise!) with the MHD region, or vice versa.

This indeed happens with the plasmasphere, the extension of the low and middlatitude ionosphere into the magnetosphere. Suppose it initially is "at rest", while the Earth and the ionosphere rotate underneath. A dynamo action will set in (mediated by field-aligned currents) that will, as a result, exert a drag force on the plasma blob so as to induce it to co-rotate. In the final state, all the plasma on the closed field lines will corotate (co-move!) with the resistive plasma of the ionosphere (in this case it is the resistive medium which "wins" and the MHD medium which "yields"!). All this is assuming that no other electric fields are in operation (see later), except the corotational field.



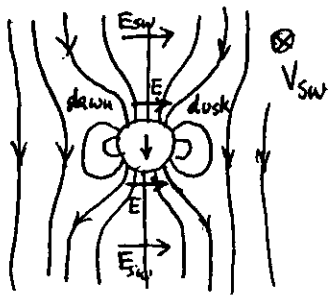
The latter is a field observed by an inertial observer. Assuming a pure dipole field rotating about its own axis with angular velocity ω, the corotational electric field can be obtained from $\vec{E} + \vec{V}_c \times \vec{B} = 0$. We can write for the corresponding potential on the equator



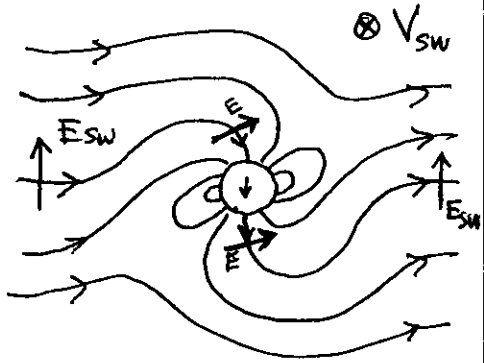
$$\Phi_{corot} = -\frac{\omega M_E}{r} \qquad (\text{for } \Phi_\infty = 0)$$

Returning to the Solar wind dynamo situation, note that all class II field lines (Fig 2 of paper) are connected to the polar ionosphere. These field lines are called open field lines, their intersection with the ionosphere is generally called the polar cap.

The solar wind electric field is thus mapped along the open field lines onto the polar cap ionosphere. Equipotentials in the polar cap are obtained by mapping down geometrically the equipotentials of the solar wind. They will depend very much on the particular field line interconnection topology. But there are some general common features. For simplicity, consider just a dipole in an external uniform B-field directed southward. The "solar wind" flows into the paper. Note in the figure how the solar wind electric field maps into a dawn-dusk electric field over the polar caps.



Now consider a "horizontal" external field (i.e., in the ecliptic). The SW electric field still maps onto the polar caps in a dawn-dusk direction!



This is why in general the direction of the polar cap electric field is dawn-dusk, although the details (size of polar cap, dawn-dusk position of polar cap, etc) depend strongly on $\vec{B}_{SW}$ and the interconnection topology. When $\vec{B}_{SW}$ is directed northward, there is no interconnection in principle, and the polar cap should shrink to zero. Fortunately, it doesn't.