IONOSPHERIC LIMITATIONS TO TIME TRANSFER BY SATELLITE

by

S. H. Knowles

Naval Research Laboratory
E.O. Hulburt Center for Space Research
Space Science Division, Ionospheric Effects Branch

ABSTRACT

The ionosphere can contribute appreciable group delay and phase change to radio signals traversing it; this can constitute a fundamental limitation to the accuracy of time and frequency measurements using satellites. Because of the dispersive nature of the ionosphere, the amount of delay is strongly frequency-dependent ($1/f^2$). At the 1.5 GHz frequency band used in the GPS systems the vertical-incidence ionospheric excess delay is typically 5 nanoseconds during the daytime (based on a total electron content of $10^{17}/m^2$), and 20% of that at night. Even at X-band, the total daytime ionospheric delay of about 1 nanosecond is enough to make compensation necessary for implementation of extremely precise time transfer schemes such as the coherent satellite link proposed by Knowles. Calibration using models is an unreliable procedure because of the variable nature of the ionosphere. It is possible in principle to provide a self-calibration by observing at two frequencies simultaneously. While this technique has on occasion proven successful in reducing ionospheric errors, a fuller understanding of the underlying phenomenon is necessary in order to understand the basic limitations it mandates in time transfer techniques.

The ionosphere is known to be highly variable, both in terms of large-scale changes and in terms of smaller-scale irregularities. Much work in recent years has focused on 'pathological' irregularities in the polar regions and equatorial region that amount to variations of 10% to several hundred percent in the total electron content. Recent

efforts to investigate ionospheric disturbances using radio astronomy interferomters have shown that the mid-latitude region, where most time transfer experiments take place, also has a prevalent irregularity distribution. This work has been undertaken by our NRL group using interferometers at Green Bank, W. VA, and Magdalena, NM, and also by a Dutch group using the Westerbork interferometer (3). These experiments measure the differential electron content between two or more antenna locations, and relate more closely to the ionospheric effect seen by a time transfer experiment than most other techniques. Easily measurable ionospheric irregularities have been seen in mid-latitude regions. Our group has observed a typical differential electron content of about $10^{15}/m^2$ over a baseline of 35 km. A day-night correlation with TEC values is observed; the night-time ionosphere is more irregular during periods of high geomagnetic activity. By combining results from our work with that of the Dutch group, who used a baseline that was considerably shorter, it is possible to obtain information about the size distribution of the irregularities. The irregularity amplitude is approximately proportional to the baseline over the range studied. Only limited information is available about another parameter of interest, the time scale distribution of these irregularites. Irregularities have been observed on all time scales from seconds to hours, as illustrated by data samples. The typical irregularity in the mid-latitude region is about 1% of the total electron content. While this is less than the extreme variations seen in polar regions, calibration of these anomalies must also take place to achieve full accuracy in time transfer and its associated use for position location.

INTRODUCTION

The ionosphere can contribute appreciable group delay to signals traversing it. This delay depends on both the frequency and the state of the ionosphere. The equation for the refractive index of the ionosphere can be expressed as:

$$n = 1 - K_1 \frac{n_e}{f^2} \pm K_2 \frac{n_e B \cos\theta}{f^3} + O(1/f^4), \qquad (1)$$

where n_e is the electron density

B is the earth's magnetic field

θ is the angle between the wave normal and the magnetic field, and

K_1, K_2 are numerical constants (Mathur et al., 1970).

If n_e is replaced by a typical total electron content along a vertical column (TEC), an estimate can be obtained of the additional group delay due to the ionosphere. Such a curve is plotted in figure 1 for a typical daytime TEC of $2 \times 10^{17}/m^2$. The curve is for a vertical path and is multiplied by the secant of the zenith distance of the line-of-sight. The dominant term from equation (1) has a shape inversely proportional to frequency squared, and a total magnitude of about 5 nanoseconds at the Global Positioning System (GPS) frequency. The typical night-time delay is about 25% of this. Klobuchar (1984) discusses the limitations in attempting to correct for the delay with a model due to the variations in TEC. Such a model can be expected to be accurate to about $\pm$ 25%.

A method often mentioned for correcting accurately for the ionospheric group delay is dual-frequency measurements. Because of the dispersive nature of the ionospheric delay, measurements at two frequencies can be used to determine what the delay constant is and correct for it. In principle this method can be very accurate, because it measures the TEC along the same line of sight that is to be used for the observation. Although the dual-frequency method has been used in a number of experiments, and is to be implemented for GPS, its accuracy has not been critically examined. Possible inaccuracies in this method include error multiplication caused by a finite frequency difference, and the effect of higher-order terms in equation (1). Figure 1 contains an estimate of the residual component if this correction is accurate to 1%. One residual component of concern is the $1/f^3$ component in equation (1). This component gives rise to two possible phase and group velocities in a wave propagating through the ionosphere: The difference in velocities depends on the angle of the propagation direction to the earth's magnetic field. This effect gives rise to Faraday polarization rotation, but it also causes a difference in group delay between the two modes. Figure 1 shows the "worst-case" magnitude of this effect. While the net time delay change caused is zero, a propagation bifurcation results that causes a dual arrival time measurement. This can be a serious limitation on lower-frequency timing measurements.

When considering the limitations provided by the ionosphere to time transfer measurements, it is important to consider the complete temporal and spatial spectrum of ionospheric irregularities. Klobuchar (1984) discusses the large-scale (global) component of TEC variations. Recent experimental information has been accumulating about medium-scale (100 m - 1000 km) irregularities in TEC that are important in contributing to time measurement inaccuracies. It has been deduced from in-situ satellite measurements that medium-scale irregularities of 1-10 % of TEC exist commonly at all times in

the arctic, and during nighttime in equatorial regions (Szuszczewicz et al, 1983). Although it has been commonly assumed that medium-scale irregularities in the "normal" mid-latitude ionosphere are insignificant, recent research using radio astronomy interferometers have shown that this is not the case.

In a radio interferometer, two or more large radio antennas are pointed at the same distant natural radio source. The outputs of these antennas are combined coherently, after correction for differential delay, to make a Michaelson interferometer. The output of this interferometer measures any source of differential phase delay. Since the ionosphere is the dominant dispersive effect, it is possible to measure the difference in total electron content by combining measurements at two frequencies in an appropriate manner. To this end, the Naval Observatory's Green Bank, West Virginia interferometer and the newer Very Large Array facility in New Mexico have been used to conduct preliminary investigations into the feasibility and usefulness of such a tool. The radio interferometer system at Green Bank, whose primary mission is to measure geodetic parameters, observes continuously a succession of radio sources at quasi-random points on the local sky. (Klepczynski et al., 1979). Simultaneous observations at 2695 MHz (11.1 cm. wavelength) and 8085 MHz (3.7 cm. wavelength) provide a measurement of the differential electron content between the ionospheric paths of any of the possible antenna pairs. The continuous availabililty of data from this array makes its use attractive for ionospheric monitoring. At the time the present data was recorded, the array consisted of 3 antennas of 85-foot diameter located along a 2.5 kilometer track that is oriented at an azimuth of 62.05 degrees, and one remote station with a 45-foot diameter antenna located about 35 km away. An additional remote station has since been added. The Very Large Array, described in detail by Thompson et al (1980), is an array of 27 antennas located near Socorro, NM, and designed with maximum flexibility for a variety of astrophysical experiments. Its 27 antennas allow significantly more complete spatial sampling of ionospheric phase differences than the Green Bank array. The maximum baseline available is approximately the same as the Green Bank array.

The differential phase path between two antennas due to the ionosphere is:

$$\Delta L = \delta L_{2I} - \delta L_{1I} = K_1 \frac{\int^2 n_e dl - \int^1 n_e dl}{f^2} = K_1 \frac{(D.E.C.)}{f^2} \qquad (2)$$

so that the differential phase delay is proportional to the difference in total electron content along the two parallel paths. In the context of this article, D.E.C. will be used to refer to the experimentally determined difference in ionospheric path length as between the locations of different antennas. It may be due to a number of different mechanisms; e.g., ionospheric gradients (Komesaroff, 1960), irregularities, diurnal effects on TEC, etc. This quantity ($\int^2 n_e\, dl - \int^1 n_e\, dl$) will henceforth be called differential electron content = D.E.C. Converting from measurement units of delay to units of phase, a power of frequency is cancelled, so that:

$$\Delta\phi_I = \frac{K\,(D.E.C.)}{f} \tag{3}$$

where $K = \frac{2\pi K_1}{C}$

and $\Delta\phi_I$ is the measured phase change due to the ionosphere.

Another important parameter to consider is the effective area intercepted by each antenna primary beam as it transits the ionosphere, as well as the overlap between beams. The differential phase effects on an interferometer from the ionosphere are described by a special case of scintillation-scattering-diffraction theory; in particular one may expect thin-screen diffraction theory to serve as a basis. This type of theory is complex and has been discussed by a number of authors (Crane, 1977; Rufenach, 1975), although relatively few (Whale and Gardiner, 1966) discuss the expected phase perturbations. While it is beyond the scope of this article to develop a complete theory for this phenomenon, the basic remark can be made that if the two antenna beams do not overlap the phase fluctuations will be fully developed, while for beams that largely overlap the phase fluctuations will be largely cancelled. With the 27 antennas of the VLA, it is possible to do a more complete reconstruction of the ionosphere. If the antenna beams do not intersect while passing through the ionosphere, 27 independent points in the ionosphere, located along a Y-configuration at distances from each other of up to 35 km, will be sampled by the VLA.

The differential electron content measured by an interferometer pair is a sum of the true difference in the free electron density between the paths over each antenna, and a second order difference in between the paths over each antenna, and a second order difference in the path length caused by the curvature of the earth. The formula for the curvature effect may be shown to be in phase units

$$\Delta\phi = 4.85 \times 10^{-11} \frac{(TEC)}{f} \cdot \left(\frac{1}{\sin\Theta_2} - \frac{1}{\sin\Theta_1}\right) \tag{4}$$

where Θ_1, Θ_2 are the elevation of the source at each antenna, f is the frequency in MHz, and $\Delta\phi$ is the differential phase in units of degrees. The difference $(1/\sin\Theta_2 - 1/\sin\Theta_1)$ can be either positive or negative, depending on the source geometry, but is always less than d/r where d is the distance between antennas, and r is the earth's radius. Since this is a refraction effect due to the sphericity of the earth, it does not depend on the height of the electron layers doing the refracting. The total measured differential electron content is thus the sum of the differential electron content due to variations in total electron content between the two locations and that due to the sphericity of the earth.

For the baselines used in our data sample, both effects are of comparable magnitude and must be considered. To measure the irregularity component, the

other must be subtracted. In order to do this, some estimate must be made of the TEC. In analyzing the Green Bank data TEC values obtained from satellite Faraday rotation data by the Air Force Global Weather Control (1981) were used for this purpose. Values obtained at Sagamore Hill, MA were used as the closest approximation available to the Green Bank, WV site. A time correction was made in the TEC data to compensate for the difference in longitude. In order to simplify the date analyses only phases from two baselines, one of 35 km and one of 1.5 km, were considered. Data taken with the U.S.N.O.'s Greenbank interferometer during November 1983 together with smaller samples from 1980-81, was analyzed to isolate ionospheric effects. The periods include a wide range of geomagnetic activity.

Figure 2 shows the estimates of variance of estimates of K for the long and short Green bank baseline for the month of November 1983. Data has been grouped into periods of morning twilight, day, evening twilight, and night to emphasize diurnal effects. Occasional data points have been bridged. Both the higher absolute value of the variance on the long baseline and its increased variability are evident. Figure 3 shows a weighted average of the variance for the four time periods for the entire month. The day variance for the long baseline is clearly greater during the daytime, while this effect is not present for the short baseline. An obvious interpretation of this is that D.E.C.'s of $5 \times 10^{14}/m^2$ are consistently present during the daytime over a 35 km baseline, while irregularities over a 1.5 km baseline are consistently less than $1.5 \times 10^{14}/m^2$. The phase of the effect agrees with that of TEC measurements, which is normally four times higher during daylight hours. Although the constant portion of the long baseline variance may be due to the ionosphere, it is more reasonable to ascribe this to equipment effects.

Although the variance correlates well with the diurnal TEC effect, it would be expected to correlate more directly with an index that would predict irregularities in the ionosphere. In Figure 4 the daytime long baseline variance is plotted on the same graph as the geomagnetic index A_p. A general correlation is evident. No noticeable correlation of the daytime variance with the daytime maximum TEC was evident. This is an indication that irregularities are being measured, rather than some differential of the normal daily gradient.

Another method of displaying the results of the experiment is to follow the tracks of one or more sources and compare the measured differential ionospheric effect with that predicted from independent total electron content data. This is done in Figures 5 and 6 for the radio source 0355+508, which was observed at night, and the radio source 1749+701, which was observed during the day. Also plotted on these figures is the differential path length ionospheric effect predicted from the total electron content data. The data from the nighttime source is seen to be generally smoother, although variable from night to night. That from the daytime source always, and that from the nighttime source sometimes, has fluctuations of $1\text{-}2 \times 10^{15}$ in differential electron content. The data is too sparsely sampled to enable a detailed Fourier analysis of the time scale of the fluctuations, but it appears that a typical time scale is of the order of 1-2 hours. For the baseline of 35 km, the observed 1-2 hour time scale means that the ionospheric irregularities

observed are typically either changing with that time scale or moving with a velocity of at least 20 km/hr (7m/sec) with respect to the earth's surface. This estimated speed is much lower than currently accepted values for the ionospheric superrotation rate of 50-150 m/sec.

A sample detection of an ionospheric irregularity using the VLA is shown in Figure 7. A method similar to that described above has been used to separate phase excursions at frequencies of 1.4 GHz and 5 GHz into dispersive, or ionospheric, and non-dispersive, or tropospheric terms. The data shown includes spacings of 17 km., 9 km., 3 km., and 0.4 km., during a period of 24 hours beginning at 08 hours local time on 21 January 1979. The figure shows observed phase in units of degrees at 1.4 GHz (100 degrees = 2.5×10^{15} D.E.C.). During this period, the source 1311+678, which is circumpolar, was observed until about 14.7 hours local time; then source 0552+398 was observed until the first source was again visible at about 0.7 hours. Source-change times are indicated by the long, vertical bars extending through the figure.

The deep protrusion in the 17 km. and 9 km. ionospheric outputs may be interpreted as a gradient in TEC which lasted for a duration of about 7 hours, with a mean epoch of about 16 hours local time. Its apparent intensity is magnified by the fact that the observations at that time were low in the sky. After correcting for a secant effect multiplication of about 3, an irregularity of 6×10^{14} in D.E.C. is indicated; this is consistent with the Green Bank results.

The 9 km spacing is generally similar to the 17 km spacing, but is decreased in amplitude by about 50%, as would be expected for an ionospheric gradient. Thus, the total D.E.C. may be larger depending on the size of the irregularity. A consistency check is provided by the tropospheric output, which shows no significant change during this period.

DISCUSSION

The radio astronomy technique, can measure ionospheric structure on a variety of scales, and provides a measure of an effect integrated through the ionosphere, rather than the marginal electron content measurement made by either bottomside or topside (Reinisch and Xuegin, 1982) sounders. During the observational periods described here, irregularities of the order of from 2×10^{14} to 2.5×10^{15} in differential electron content were observed. This amounts to about 0.5 to 5 % of the measured TEC, and indicates the frequent presence of significant irregular structure in the ionosphere on this scale.

A Dutch group (Spoelstra and Kelder, 1984) have recently used the Westerbork interferometer to measure ionospheric disturbances. Although their longest baseline available was only 2.7 km., they were able to easily measure ionospheric irregularities. They found the typical size to be consistently greater than their maximum baseline, which is in agreement with the present size estimates. Their conclusions about the greater prevalence of disturbances during the daytime are in general accord with ours, although the Dutch group did not provide numerical estimates of differential electron content.

Both the Green Bank results and the sample of data from the VLA show clear evidence of horizontal ionospheric gradients in TEC persisting for periods of several hours, while the Dutch group observed periods of less than fifty minutes. The mid-latitude radio interferometer observations clearly are not related to the ionospheric wind, which is known to have velocities of 50-150 kilometers per second (Regan and Rodriquez, 1981; Rino and Livingston, 1982). They are most easily accounted for by the phenomenon of traveling ionospheric disturbances, or TID's. This class of disturbances, as pointed out by Evans et al (1983), includes fluctuations with a wide variety of periods, which, in spite of many observations with various techniques, are not well-defined.

From the limited observations undertaken by our group and the Dutch group, it is possible to make a rough estimate of the approximate amplitude vs. distance relationship of these medium-scale irregularities. This is shown in figure 8. A similar graph that gives typical time scales is not yet possible due to inadequate data.

SUMMARY

Ionospheric compensation is necessary for the most precise time transfer and frequency measurements, with a group delay accuracy better than 10 nanoseconds. A priori modeling is not accurate to better than 25%. The dual-frequency compensation method holds promise, but has not been rigorously experimentally tested. Irregularities in the ionosphere must be included in the compensation process.

REFERENCES

1. Air Force Global Weather Central, Omaha, Neb., suppliers of data base for T.E.C. Measurements.

2. Aarons, J.H., E. Whitney and R.S. Allen, "Global Morphology of Ionospheric Scintillation", Proc. IEEE 59, 159-172 (1971).

3. Coffey, H.E. (ed.), "Geomagnetic and Solar Data", J.G.R. 89, 2411 (1984).

4. Hamaker, J.P., "Atmospheric Delay Fluctuations with Scale Sizes Greater than One Kilometer, Observed with a Radio Interferometer Array", Radio Science 13, p 873, 1978.

5. Klepczynski, W.J., G.H. Kaplan, D.D. McCarthy, F.J. Josties, R.L. Branham, K.J. Johnston and J.S. Spencer, "Progress Report of the USNO/NRL Green Bank Interferometer Program", NASA Conference Publication 2115, Radio Interferometry Techniques for Geodesy", Proceedings of a conference held at M.I.T., Cambridge, MA., June 19-21, 1979, p. 63.

6. Klobuchar, J.A., "A Review of Ionospheric Effects in Earth-Space Propagation", submitted to Proceedings of the Sixteenth PTTI Meeting.

7. Knowles, S.H., W.B. Waltman, F.J. Kelly and S. Odenwald, "Observations of the Radio Noise Background in the Frequency Range 150-180 kHz", in press for Radio Science (1985).

8. Mathur, N.C., M.D. Grossi and M.R. Pearlman, "Atmospheric Effects in Very Long Baseline Interferometry", Radio Sci. 5, 1253 (1970).

9. Regan, R.D. and P. Rodriguez, "An Overview of the External Magnetic Field with Regard to Magnetic Surveys", Geophys. Surveys 4, 256 (1981).

10. Reinisch, B.W. and H. Xueqin, "Automatic Calculation of Electron Density Profiles from Digital Ionograms 1. Automatic O and X Trace Identification for Topside Ionograms", Radio Sci. 17, 421 (1982).

11. Rino, C.L. and R.C. Livingston, "On the Analysis and Interpretation of Space-Receiver Measurements of Transionospheric Radio Waves", Radio Sci. 17, 845 (1982).

12. Spoelstra, T.A.T., and H. Kelder, "Effects Produced by the Ionosphere on Radio Interferometry", Radio Sci., 19, 779, 1984.

13. Szuszczewicz, E.D., R.T. Tsunoda, R. Narcisi, and J.C. Holmes, "Coincident Radar and Rocket Observations of Equatorial Spread-F", Geophys. Res L. 7, 537 (1980).

14. Szuszczewicz, E.D., M. Singh, P. Rodriquez, and S. Mango, "Ionospheric Irregularities and their Potential Impact on Synthetic Aperture Radars", Radio Science 18, p. 765 (1983).

15. Thompson, A.R., B.G. Clark, C. Wade and P.J. Napier, "The Very Large Array", Ap. J. Supp. 44, 151 (1980).

16. Whale, H.A. and C.W. Gardiner, "The Effect of a Specular Component on the Correlation Between the Signals Received on Spaced Antennas", Radio Sci. 1, 557, 1966.

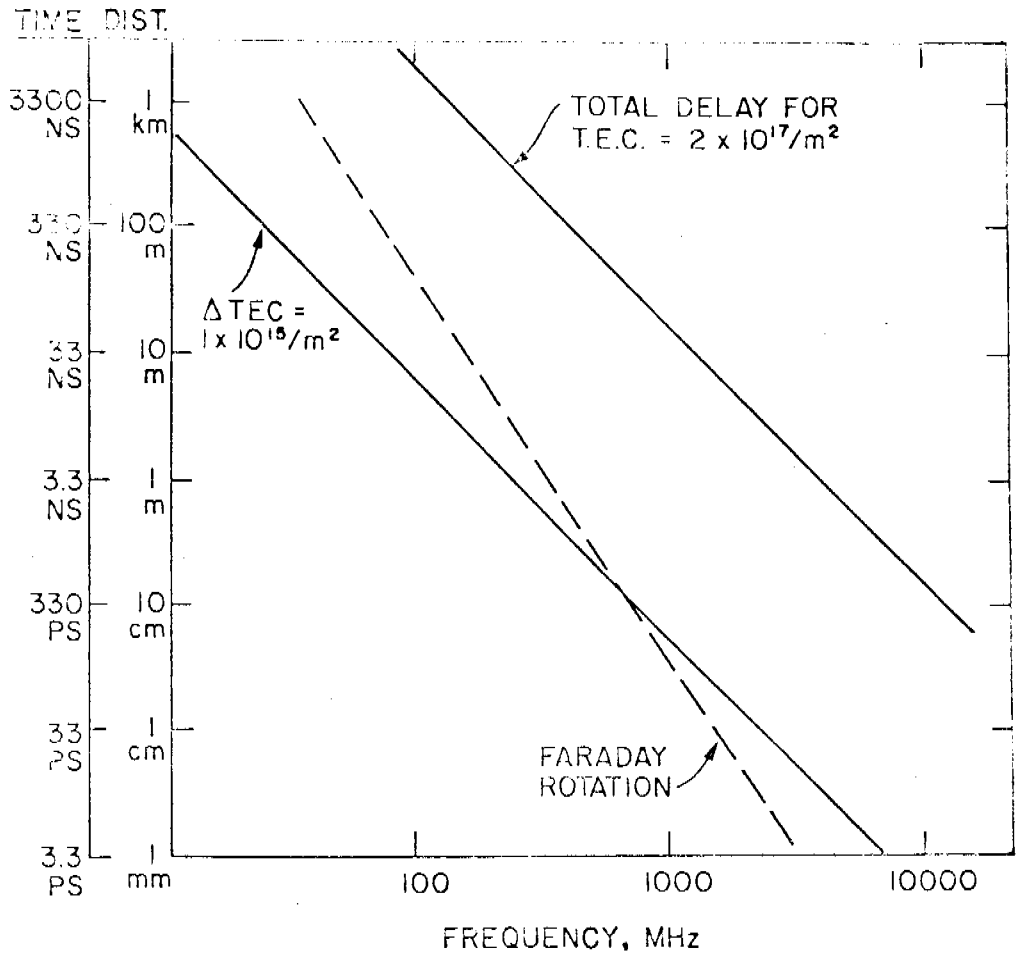

Fig. 1 Ionospheric effects on group delay.

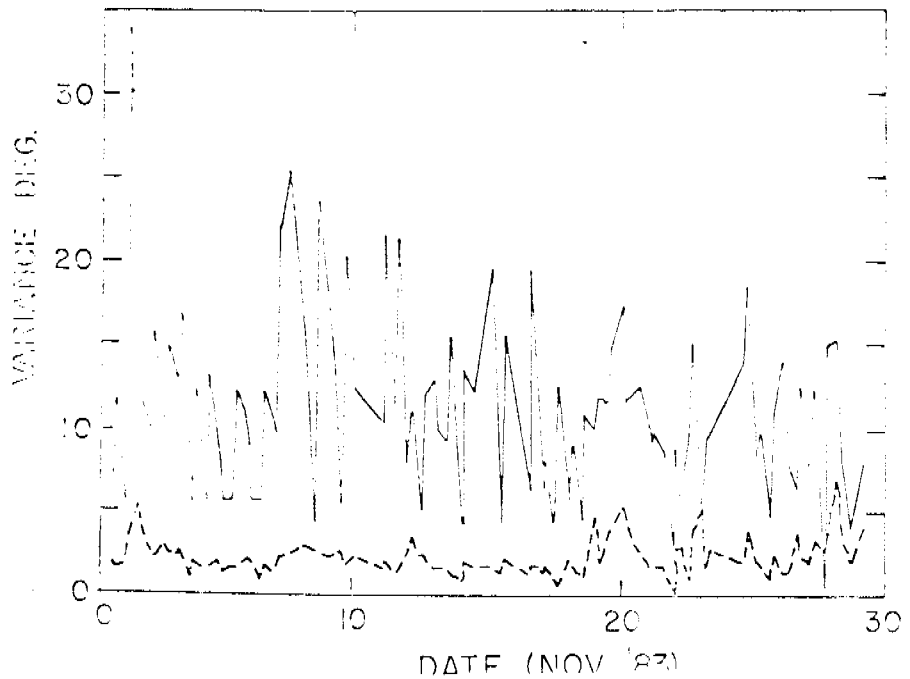

Fig. 2. Scatter in ionospheric phase measurments for Green Bank interferometer for the month of November, 1983. Dotted line marks short baseline; solid line marks long baseline. Data is grouped into morning twilight, day, evening twilight and night.

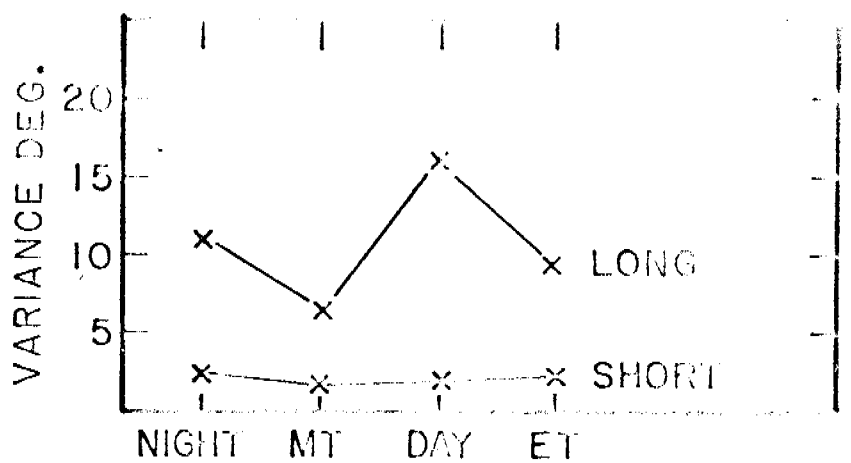

Fig. 3 All data of Figure 2 grouped to show behavior of variance as a function of time of day.

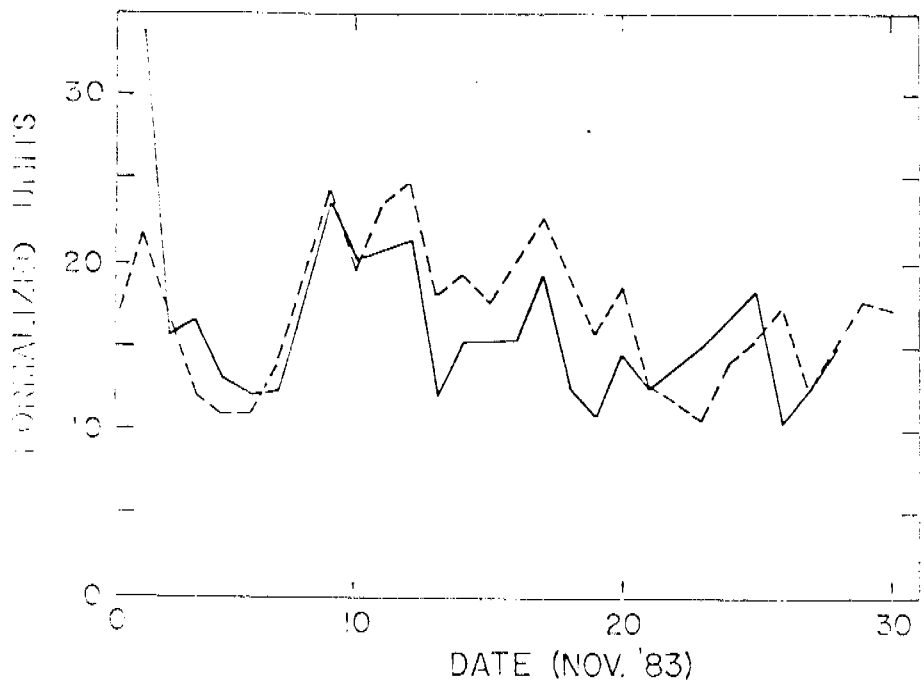

Fig. 4 Comparison of Green Bank long baseline variance for daytime with geomagnetic index A_p. Normalized units used on ordinate.

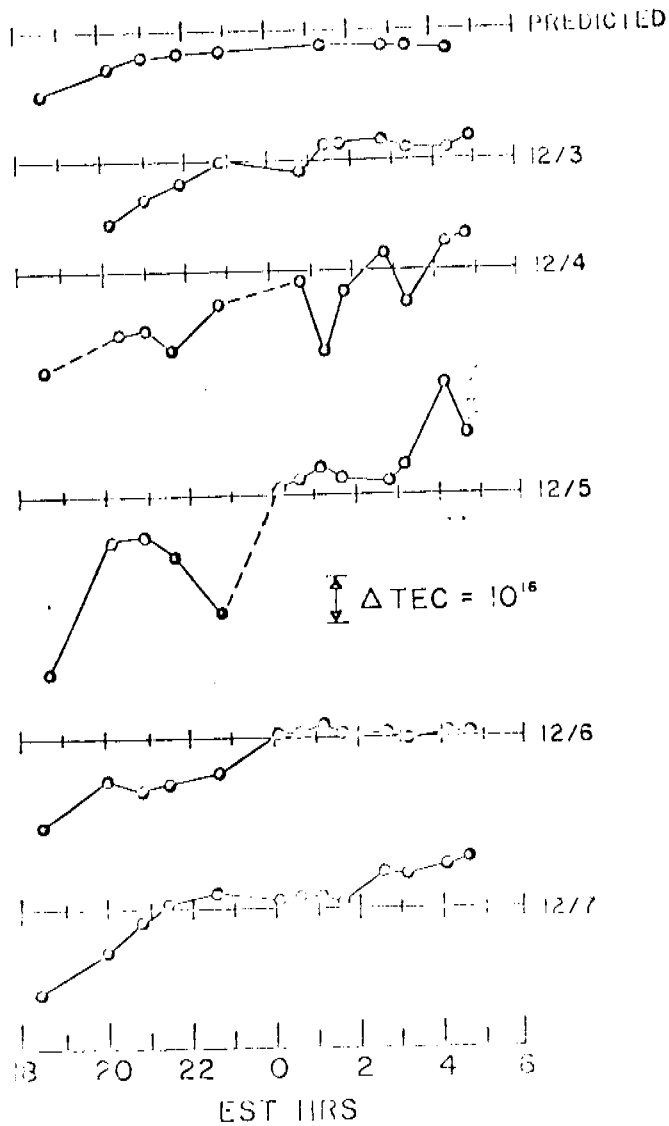

Fig. 5 Differential electron content measurement as a function of time for 0355+308, a nighttime source. Observations taken during early December, 1981.

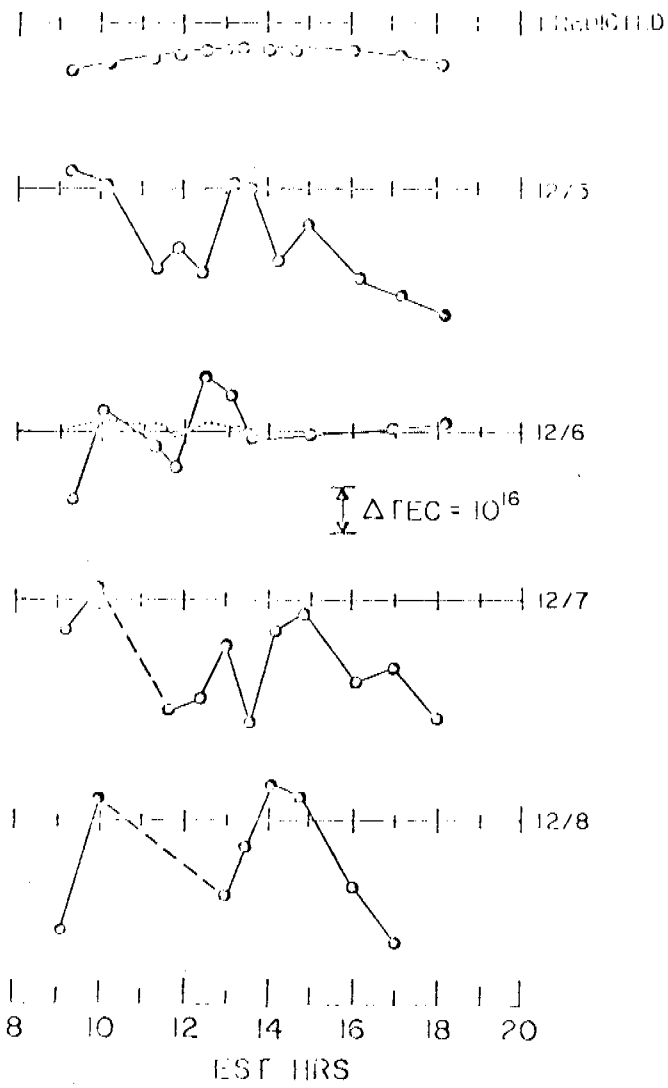

Fig. 6 Differential electron content measurement as a function of time for 1749+701, a daytime source. Period of observations same as Fig. 5.

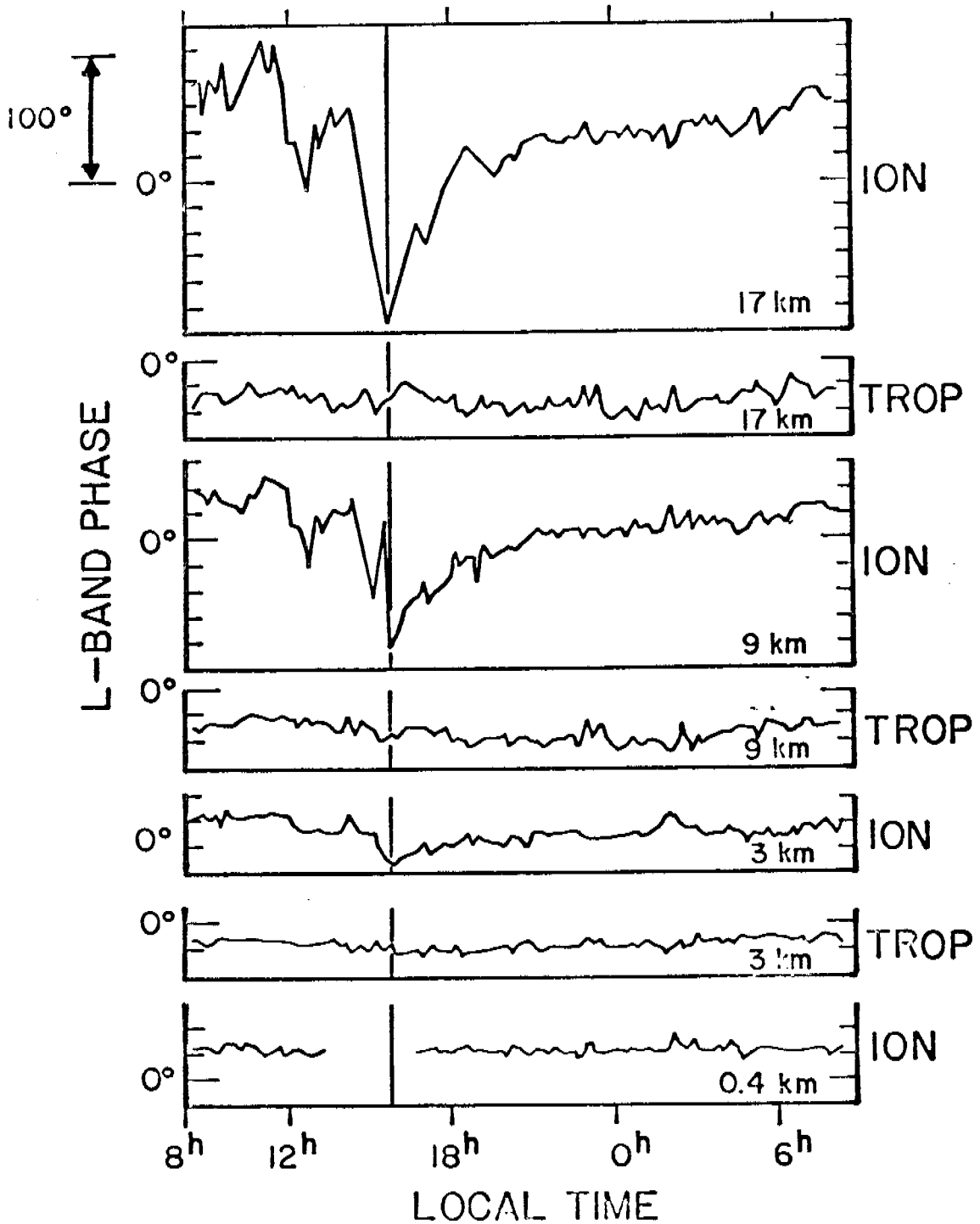

Fig. 7 Sample ionospheric measurements made using a series of baselines with the Very Large Array, illustrating the detection of an ionospheric disturbance. The graph shows both non-dispersive, or tropospheric and dispersive or ionospheric differential phases. Long vertical bars mark source changes.

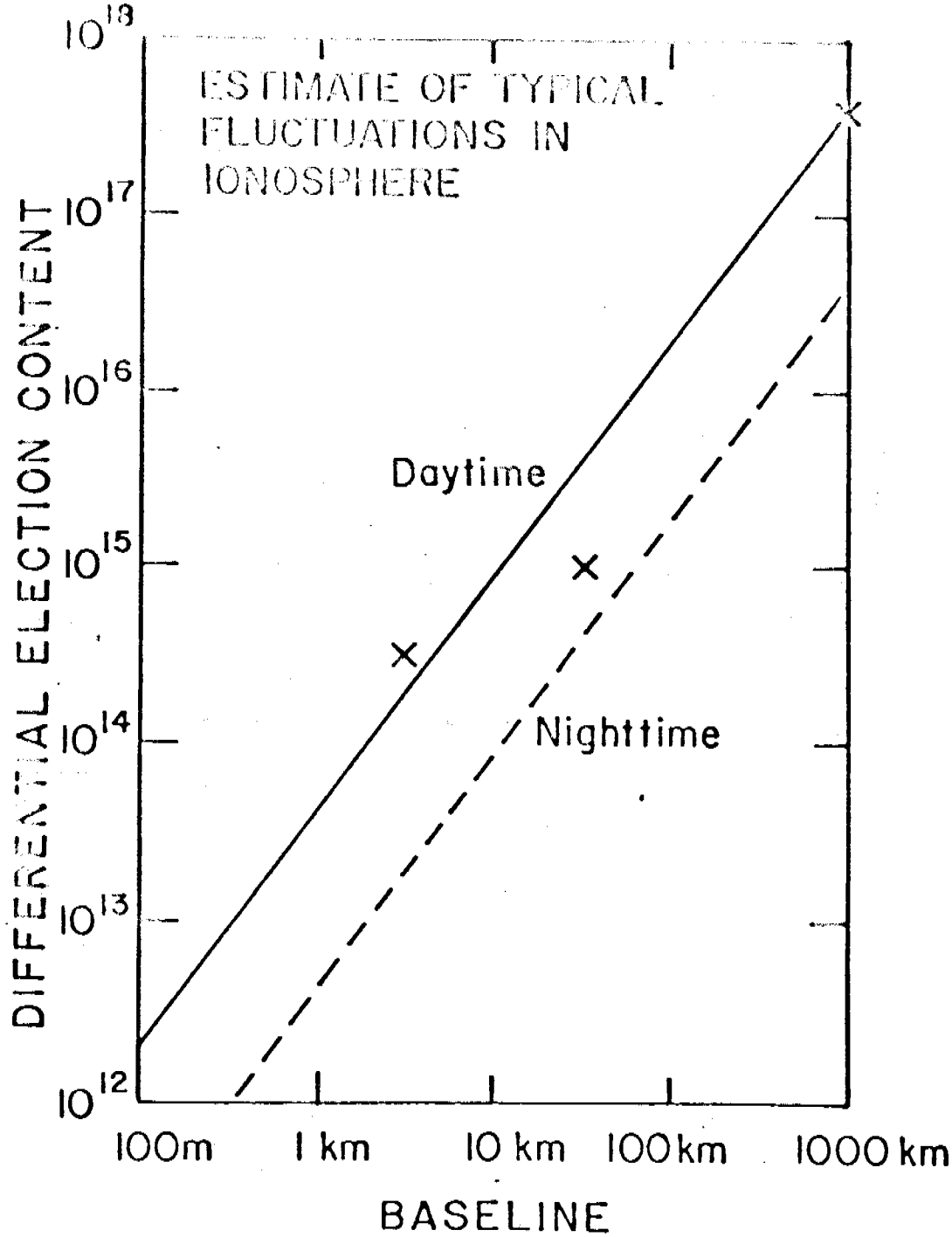

Fig. 8 Amplitude vs. size parameters of ionospheric irregularities.

QUESTIONS AND ANSWERS

LARRY D'ADDARIO, NATIONAL RADIO ASTRONOMY OBSERVATORY: This information about the variations between stations on relatively short baselines would also suggest that, for a single station, there would be a lot of variation in direction if you are looking at the sky. Thus, knowledge of the zenith total electron content would not be a good predictor of the path delay in other directions where you are usually looking. Is that right?

MR. KNOWLES: I would agree with that statement.

MR. KLOBUCHAR: If I could interject a comment here, that's certainly true, but you have to look at what detail is true. If Steve would show the last view graph.... (the view graph was shown here). You will notice that the two real points that you said were actually observed, were down around ten to the fifteenth. The point way up at the top here at around a thousand kilometers above ten to the seventeenth is actually a difference in the diurnal curve between two points. The total time delay might be on the order of ten to the seventeenth, and you are seeing effects here on the order of ten to the fifteenth, one or two percent. Ten to the sixteenth is half a nanosecond at L-band, so we are talking about much less than a nanosecond, differential. If that's a problem for your time transfer, then you are in trouble, because you always get fluctuations of this order. For the VLBI folks, it is a serious problem, and Steve's work is very important in this regard, because this size of irregularity is one that really hasn't been measured before. It is very important to continue that kind of work, because there are irregularities at all levels, and no one up until this work has done anything in that distance range that I am aware of. The Dutch work and this work are the only ones.