AGARD-CP-209

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARD CONFERENCE PROCEEDINGS No. 209

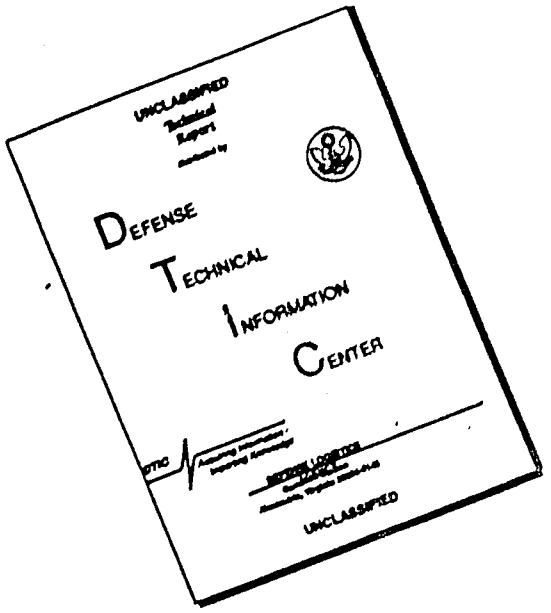
on

Propagation Limitations of Navigation and Positioning Systems

NORTH ATLANTIC TREATY ORGANIZATION

DISTRIBUTION AND AVAILABILITY
ON BACK COVER

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

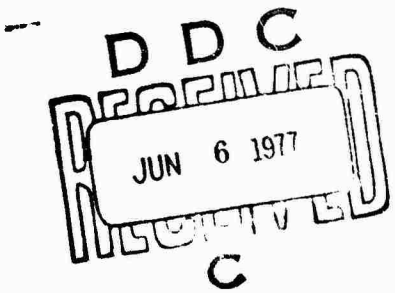
NORTH ATLANTIC TREATY ORGANIZATION

ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT

(ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD)

AGARD Conference Proceedings No.209

PROPAGATION LIMITATIONS OF NAVIGATION AND POSITIONING SYSTEMS



Papers presented at the Electromagnetic Wave Propagation Panel Specialists' Meeting
held in Istanbul, 20–22 October 1976.

THE MISSION OF AGARD

The mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

- Exchanging of scientific and technical information;
- Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;
- Improving the co-operation among member nations in aerospace research and development;
- Providing scientific and technical advice and assistance to the North Atlantic Military Committee in the field of aerospace research and development;
- Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field;
- Providing assistance to member nations for the purpose of increasing their scientific and technical potential;
- Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Program and the Aerospace Applications Studies Program. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

Published February 1977

ISBN 92-835-0189-6

Printed by Technical Editing and Reproduction Ltd
Harford House, 7-9 Charlotte St, London, W1P 1HD

ELECTROMAGNETIC WAVE PROPAGATION PANEL

CHAIRMAN: Mr P.Halley
Ingenieur en Chef au CNET, France

DEPUTY CHAIRMAN: Dr H.J.Albrecht
FGAN, Germany

PROGRAM COMMITTEE MEMBERS

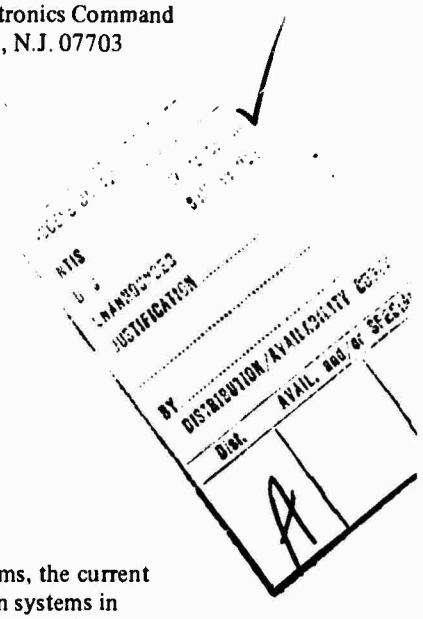
Dr Jules Aarons (Co-Chairman)
Senior Scientist
Air Force Geophysics Laboratory
Hanscom AFB, MA 01731
USA

Dr Haim Soicher (Co-Chairman)
AMSEL NL-RH-4
US Army Electronics Command
Ft. Monmouth, N.J. 07703
USA

Dr J.Belrose
Canada

Dr B.Burgess
UK

Mr P.Halley
France



PANEL EXECUTIVE

Cdr. D.G.Carruthers, US Navy

THEME

In order to assess the requirements for navigation and positioning systems, the current status of these systems, and the limitation the propagation medium places on systems in being or planned, the AGARD EPP held a Specialists' Meeting in Istanbul, October 20–22 1976, on this subject. The concept of the meeting was to outline requirements and progress in systems and to plan programs for future studies to reduce any propagation limitations on navigation systems.

Accordingly, the meeting reviewed various subjects including civilian requirements for sea, air and ground navigation systems both short and long range, and propagation study needs of Loran C, Omega, NAVSTAR, and Aerostat. In addition, requirements for other systems were outlined and possible propagation problems discussed. Proposed future research and development programs was the subject of a round table discussion. The aim was to mitigate the propagation problems that systems will and have encountered.

CONTENTS

IONOSPHERIC EFFECTS IN NAVSTAR GPS

Bradford W. Parkinson
Colonel, United States Air Force
Program Director
NAVSTAR Global Positioning System
Space And Missile Systems Organization
El Segundo, California

Edward M. Lassiter
The Aerospace Corporation
El Segundo, California

C. K. Cretcher
The Aerospace Corporation
El Segundo, California

SUMMARY

This paper reviews the impact of ionospheric effects upon the user system performance in the positioning process. System performance of ionospheric delay corrections in terms of ionospheric modeling and in terms of dual frequency receiver calibration are discussed. The effects of signal degradation by phase and amplitude scintillation are reviewed.

1. INTRODUCTION

The NAVSTAR Global Positioning System (GPS) is a satellite-based navigation system that will provide extremely accurate three-dimensional position fixes and timing information to properly equipped users anywhere on or near the earth. The system will be available continuously regardless of weather conditions, and will find extensive utilization in impro weapons delivery accuracies, range instrumentation, etc. Furthermore, it will provide an ultimate sa in the number and cost of navigation and position-fixing systems currently employed or projected.

NAVSTAR GPS is a joint-Service program managed by the Air Force with participation of the Navy, Army, Marines, and Defense Mapping Agency. The system concept evolved from USAF and Navy studies initiated in the mid-1960's. Current program plans call for the deployment of six satellites in 1977 to permit demonstration and evaluation tests over the continental United States (CONUS). The system is expected to be expanded through deployment of additional satellites into an operational 24-satellite system.

Approval to proceed with the Phase I (concept validation) portion of the program was granted in December 1973. The next major decision is scheduled for March 1978. The GPS could attain a worldwide two-dimensional capability in the early 1980's with full three dimensional navigation by the mid-1980's.

2. SYSTEM DESCRIPTION

The NAVSTAR GPS consists of three major segments: space, control, and user, as shown in Figure 1. The concept of operations can be described as follows. Each satellite carries an atomic clock with stabilities of the order of 1 part in 10^{13}. This clock is used to generate timing for the dual frequency pseudo-random noise (PRN) spread spectrum L-band navigation signals which the satellites radiate continually. These navigation signals contain information regarding the satellite ephemerides and clock behavior. Geographically dispersed monitor sets receive these signals and send the received information to the Master Control Station (MCS) located in the CONUS. The MCS processes the data and calculates predicted satellite positions, velocities, and clock drifts to be uploaded and stored in the satellite memories for later broadcast to the users. The MCS insures that the satellite clocks are synchronized within a few nanoseconds. Users perform navigation using the signals radiated by the satellites. If the users had precision clocks (synchronized to the satellites), they could make passive range measurements to the satellites and determine their position as the intersection of three spheres, each centered at the satellite. To avoid requiring all users to be equipped with precision clocks, enough satellites will be deployed so that all users will have at least four satellites available. The simultaneous reception of four navigation signals allows three independent range difference equations to be formed which may be used to calculate the intersection of three hyperboloids of revolution uniquely defining the user position.

Space Segment. In its operational configuration, the GPS satellite constellation will consist of 24 satellites in circular, 10,900 nmi orbits with an inclination of 63 degrees (See Fig. 2). They will be deployed in three planes, each plane containing eight satellites. This satellite constellation insures that at least six satellites are always in view from any point on the earth and that on the average, nine satellites are in view, thus insuring satellite coverage for three-dimensional positioning and navigation on a worldwide basis.

The satellite is shown in Figure 3. The design life is five years with enough expendables to last seven years. The on-orbit weight is 950 pounds and the end of life power available is 400 watts. The power is supplied by five square meter solar arrays that continuously track the sun. Three nickel-cadium batteries are available for eclipse operations. The satellite is three-axis stabilized by means of four skewed reaction wheels. A hydrazine propulsion system is on-board for station keeping maneuvers, and to provide a momentum dump capability for the attitude control system. A 12-element shaped beam helix array provides an earth coverage antenna pattern for the L-band navigation signals.

Each satellite transmits two spread-spectrum PRN navigation signals, one signal at 1575 MHz and a second signal at 1227 MHz. The signals are coherently generated and can be used to determine and correct for effects of the ionospheric signal delay. Both navigation signals consist of a sequence of

binary digits (PRN sequences) bi-phase modulated onto the carriers at a rate of 10.23 Mbps. The ephemeris and satellite clock data are modulated onto the PRN sequence at a rate of 50 bps. Thus, the basic navigation data is "spread" from a bandwidth of approximately 100 Hz to a bandwidth of approximately 20 MHz by the PRN sequence. The PRN sequences are unique to each satellite and mutually orthogonal permitting the use of a common carrier frequency for all satellites in the constellation. An acquisition signal is quadraphase modulated with the navigation signal at 1575 MHz to also provide a rapid acquisition capability for all users.

Control Segment. The control segment consists of a Master Control Station (MCS), widely separated Monitor Sets (MS), and an Upload Station (ULS). Redundant master control, and upload stations are planned for the operational system. The widely spaced monitor sets, located on U.S. controlled territory, will passively track the satellites accumulating ranging data from the navigation signals. These data will be transmitted, along with meterological and status information, to the MCS in the CONUS. At the MCS the ranging data will be corrected for transmission delays, e.g., ionospheric and tropospheric delays, relativistic effects and processed by a filter algorithm to provide best estimates of satellite position, velocity, acceleration e.g., due to solar pressure variations, and satellite clock drift relative to system time. Additionally, MS clock drifts relative to system time, polar wander parameters and tropospheric correction residuals are estimated. Subsequent post-event data processing will be done to generate progressively refined information defining: a) the gravitational field influencing the satellite motion, b) MS locations, and c) other observable system influences. The data, thus derived, will be used to generate future navigation messages to be loaded into the satellite memories via the ULS, also located in the CONUS, at least once a day.

User Segment. The user set consists of an antenna, receiver, data processor with software, and control/display unit. Some configurations will be integrated with auxiliary sensors such as inertial measuring units. The receiver measures pseudo-range and pseudo-range-rate using the navigation signal from each of four satellites. The processor converts these data to three-dimensional position, velocity, and system time. The position solution is developed in earth-centered coordinates, which are subsequently converted and presented on the display unit as either geographic coordinates, UTM grid coordinates, or any other coordinate system desired by the user.

Projections of total user equipment needs for the U.S. Department of Defense exceed 25,000. Although the system is designed to meet military requirements, it will also have potential for civil and international use. In order to minimize user equipment costs, the space and control segments are designed to emphasize low user costs. During the system definition studies, six classes of user requirements were identified and defined by the military services to meet their operational performance requirements. The various classes of user requirements are summarized in Table 1.

3. NAVIGATION ALGEBRA

The technique for achieving the desired system accuracies calls for making transit time measurements of RF signals encoded as pseudo-random noise modulation on an L-band carrier. A precision timing system carefully synchronized between satellite transmitters must be used. Users with precision clocks could then passively range to the satellites and determine their own position as the calculated intersection of at least three spheres centered at the transmitters. To avoid requiring all users to be equipped with precision clocks, enough satellites will be deployed so that all users have available at least four satellites simultaneously. Then synchronized time is not required because the simultaneous reception of four signals produces three independent range differences which may be used to calculate the intersection of three hyperboloids of revolution.

The range differencing solution is primarily conceptual, since it is just one possible method of determining user position and clock bias (a total of four unknowns) from the four range measurements. The basic equation relating ranges and transit times is termed a pseudo-range equation which in an error-free system is given by (see Figure 4):

$$r_i = \left[(\underline{s}_i - \underline{x}) \cdot (\underline{s}_i - \underline{x})\right]^{\frac{1}{2}} + b \qquad (3\text{-}1)$$

where

r_i = the time delay multiplied by the speed of light measured from satellite i

$\underline{x}$ = vector from earth center to user position

$\underline{s}_i$ = vector from earth center to satellite i

b = user clock phase bias (multiplied by the speed of light).

The variable r_i has been termed a pseudo-range since it does not conform to Euclidean distance because of the additional term b. Given four independent pseudo-ranges, one can solve for the user position and time. When more than four pseudo-ranges are available and/or noise is introduced into the measurements, the determination of $\underline{x}$ becomes a process of optimal estimation.

An iterative solution for user position can be defined as follows where matrix notation will be employed for convenience; matrices will be denoted by capital letters:

$$X \triangleq (x_1, x_2, x_3)^T = \text{user coordinates in a reference coordinate frame}$$

and

$$S_i \triangleq (s_{i1}, s_{i2}, s_{i3})^T = \text{satellite coordinates in the reference coordinate frame}$$

The four simultaneous pseudo-range measurements are then given by

$$m_i = \left[(S_i - X)^T (S_i - X)\right]^{\frac{1}{2}} + b - b_i + P_i \quad (i = 1,2,3,4) \tag{3-2}$$

where p_i is the noise and b_i the satellite clock phase bias in measurement i.

Let the direction cosine of the i^{th} measurement vector along the j^{th} coordinate be defined by

$$\gamma_{ij} \triangleq (S_{ij} - X_j) / \left[(S_i - X)^T (S_i - X)\right]^{\frac{1}{2}} \tag{3-3}$$

Then, defining the geometric matrices

$$\underset{(4X4)}{G} = \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \\ \Gamma_3 \\ \Gamma_4 \end{bmatrix} \qquad \underset{(4X16)}{B} = \begin{bmatrix} \Gamma_1 & 0 & 0 & 0 \\ 0 & \Gamma_2 & 0 & 0 \\ 0 & 0 & \Gamma_3 & 0 \\ 0 & 0 & 0 & \Gamma_4 \end{bmatrix}$$

where

$$\Gamma_i = (\gamma_{i1}, \gamma_{i2}, \gamma_{i3}, 1)$$
$$0 = (0,0,0,0)$$

the user state and the measurement matrices are

$$\underset{(1X4)}{Y} \triangleq \left[x_1, x_2, x_3, b\right]^T$$

$$\underset{(1X4)}{M} = (m_1, m_2, m_3, m_4)^T$$

the augmented satellite state matrix is

$$\underset{(1X16)}{S} = (S_1^T, b_1, S_2^T, b_2, S_3^T, b_3, S_4^T, b_4)^T$$

and the noise matrix is

$$\underset{(1X4)}{P} = (p_1, p_2, p_3, p_4)^T$$

thus permitting the following matrix equation to be written in place of Eq. (3-2):

$$GY = BS - M + P \tag{3-4}$$

The minimum variance estimate of Y, assuming no statistical a priori information, is then:

$$\hat{Y} = G^{-1} (B\hat{S} - M) \tag{3-5}$$

where $\hat{S}$ is an estimate of S.

Since G and B are functions of system geometry and therefore $\hat{Y}$, an iterative solution for $\hat{Y}$ is indicated. However, sufficiently accurate initial estimates of $\hat{Y}$ will generally be available and only a few iterations, at most, will be required.

The covariance matrix for the error δY in the estimate of Y from Eq. (3-5) is

$$\text{cov } \delta Y = (G^T R^{-1} G)^{-1} = \left(G^T \left[B(\text{cov } \delta S)B^T + \text{cov } P\right]^{-1} G\right)^{-1} \tag{3-6}$$

where Eq. (3-[illegible]) defines the matrix R and $\delta S = S - \hat{S}$.

GDOP (geometric dilution of precision) is calculated by setting the total range measurement error covariance matrix R (consisting of satellite position error and measurement error) equal to the identity matrix, resulting in

$$\text{cov } \delta Y = (G^T G)^{-1} \tag{3-7}$$

The GDOP factor is then defined as the square root of the projection of cov δY onto the user coordinate of interest.

If variances σ_r^2 and correlations ρ are assumed for R, the result is

$$\text{cov } \delta X = \sigma_r^2 (1 - \rho)(G^T G)^{-1} \tag{3-8}$$

upper left hand 3X3

$$\text{cov } \sigma b = \sigma_r^2 \rho + (1 - \rho)(G^T G)^{-1}_{4X4} \tag{3-9}$$

Thus, the corresponding standard deviation in the user coordinate is obtained by multiplying the GDOP factor by $\sigma_r(1 - \rho)^{\frac{1}{2}}$. For the clock bias this GDOP multiplier is only valid for uncorrelated errors ($\rho = 0$). If, in fact, quantities σ_r and ρ can be found which produce user coordinated variances closely approximating the variances produced by error source modeling, then the quantity $\sigma_r(1 - \rho)^{\frac{1}{2}}$ is termed the user equivalent ranging error (UERE). Under these conditions, system/user geometry is characterized by GDOP factors and pseudo-ranging errors by UERE.

Figure 5 shows the cumulative probability distribution function for the GDOP factor along the vertical axis and the GDOP factor for radial error in the horizontal plane for GPS. Since in fact satellite position error projections and atmospheric refraction modeling errors are geometry dependent, the GDOP/UERE concept is only an approximation. GDOP considerations tend to indicate ideal satellite locations consisting of one directly overhead and three equally spaced near the horizon. Since tropospheric and ionospheric become larger near the horizon, this is not truly optimal for minimizing user error. However, atmospheric errors are only a portion of the total pseudo-ranging error and GDOP is still a reasonable figure of merit, particularly on a statistical basis.

The specific major error sources and their contributions in terms of UERE are shown in Table 2 for both single and dual frequency receiver users. It is currently anticipated that all users demanding high accuracies will use dual frequency calibration. A less sophisticated model will be available for users not requiring precise ionospheric delay calibration.

Total system errors have been calculated for worldwide population averaged over 24 hours. The expected single measurement horizontal and vertical distributions are shown in Figure 6.

For the system accuracy results shown in Figure 6, the satellite position errors and signal propagation errors are modeled in terms of variances and correlations to provide the required covariance matrix R. The following sections discuss the errors contributed by the ionosphere.

4. IONOSPHERIC SIGNAL DELAY AND DEGRADATION

A transionospheric RF signal experiences an excess delay over the free space propagation time between a satellite and a subionospheric receiver. The excess time delay is proportional to the number of free electrons encountered along the signal path. The integration number of electrons encountered per unit area is commonly referred to as the total electron content (TEC). For the GPS system, TEC is measured in units of nanosecond of signal delay at 1.6 GHz. For frequencies above 200 MHz, the delay is essentially inversely proportional to the square of the frequency. This frequency dependence is the basis for using the dual-frequency measurements to calibrate the ionosphere delay.

In addition to signal delay by TEC, there are several types of ionospheric distrubances which can cause signal delay and degradation. Ionospheric disturbances can be caused by geomagnetic storms and by solar activity. During a severe storm, TEC increases. Although the magnitude of TEC increase may not be predictable by a single-frequency user, such storms are infrequent and those associated with geomagnetic disturbances are often predictable hours to days in advance. Traveling ionospheric disturbances are produced by auroral activity and nuclear explosions. Their short duration and infrequent occurrence limit their role in GPS considerations. Scintillations are high-frequency, noise-like amplitude and phase fluctuations of the received signal which have been defracted by irregularities in the electron density distribution. Affecting systems with carrier frequencies as high as 6 GHz, signal scintillation is by far the most serious problem in transionospheric propagation.

The effects of TEC time delay and of signal scintillation are discussed in subsequent sections.

5. TOTAL ELECTRON CONTENT MODELS

To provide TEC predictions for single frequency users in the field, it is necessary to relate (model) TEC to currently measured and predicted parameters, time of day and season, geographic location and satellite-user geometry.

Between 1969 and 1972, a series of studies were undertaken to demonstrate the feasibility of modeling the ionospheric delay to accuracies commensurate with GPS requirements using generally available ionospheric parameter inputs.

Five different studies were undertaken, either at USAF request or under USAF contract. The evaluations of the models developed in those studies were made using Faraday rotation data taken over mid-latitude North America and Hawaii. An HF signal propagating in the earth's magnetic field splits into two modes of slightly different frequency with different phase speeds. This causes the plane of polarization to rotate about the wave path. TEC is essentially proportional to the Faraday rotation angle. Two errors are inherent in this technique; perhaps the highest 10 percent of the electrons in the upper ionosphere do not contribute measurably to the rotation angle, and the constant of proportionality is not known precisely.

In all of the studies, TEC was converted to the vertical TEC above the ground station by an obliquity factor, which is a function of elevation angle and mean ionospheric height. The latter was assumed constant, and does not induce significant errors for the domain considered.

A preliminary TEC model was developed by J. Klobuchar and R. Allen at the Air Force Cambridge Research Laboratories (Ref. 2). The model was constructed from Faraday rotation data, taken at the Sagamore Hill Radio Observatory near the solar cycle maximum, and analytically predicted TEC as a function of measured or predicted values of f_0F_2, which is the lowest radio frequency that can normally penetrate the dominant F_2 layer of the ionosphere. The model, applicable only near the time and location of the data gathering, served to demonstrate the feasibility of analytic ionospheric delay modeling with a relatively simple analytic function depending only on f_0F_2. When this model, which amounted to a smoothing of the original data, was used to predict TEC, overall rms errors of 1 to 4 nsec were encountered.

The goal of the next four studies, performed between 1970 and 1972, was to determine the predictability of the total vertical electron content of the ionosphere in the mid-latitude Western Hemisphere by evaluating derived models against NASA Faraday rotation data. These data were available for a number of stations and time periods, including maximum and minimum of a solar cycle.

Stanford University used the Faraday rotation data to synthesize a TEC model over reference stations where large quantities of data exist and to determine geographical extrapolation functions for other stations (Ref. 3). A commonly predicted solar activity index is required as input. Values predicted by this process were compared to converted Faraday rotation measurements, with overall residuals at the various evaluation stations lying between 2.7 and 6.8 nsec.

The University of Illinois ionospheric model used a weighted sum of near-real-time measurements and measurement residuals at other observation locations to predict the value at the evaluation stations (Ref. 4). When measurement residuals are used, the predicted value and weighting coefficients require prior accumulation of data over the geographic region of interest to determine estimates of residual statistics and running monthly means. Both approaches produced rms residuals of 1.2 to 13 nsec.

The Applied Physics Laboratory (APL) at Johns Hopkins Univeristy constructed a predictive model based on assumed electron density altitude profiles constructed from independent analytical and empirical studies and requiring commonly available predicted ionospheric parameters as input (Ref. 5). The resulting evaluation rms residuals range from 1.5 to 12 nsec.

The Bent ionospheric model was developed for NASA and is based on a large accumulation of worldwide bottomside and topside ionosphere soundings taken over the 1960's (Ref. 6). An ionospheric electron density altitude profile was established from these data that requires two commonly available ionospheric parameters. (This model construction and that of APL are completely independent of the evaluation data). The DBA Systems evaluation of this model (by Dr. Bent) produced overall rms residuals between 2.6 and 8.8 nsec, depending on the evaluation station.

The accuracies resulting from these models (without real-time updates) are summarized in Figure 7. These studies served to validate the concept of ionospheric modeling in the North America and Hawaii domain.

The difficulties with the modeling approach for calibrating the ionospheric delay are:

a. A large amount of data must be transferred to the satellite and to the user to provide a sufficiently accurate model.

b. The model cannot accommodate ionospheric disturbances without near-real-time update data, an operation inconsistent with current system operations.

c. Model errors will also depend strongly on time of day, geographic location and satellite evaluation angle.

For users requiring less than full system accuracy and using less sophisticated equipment, a simplified model approach will be satisfactory and has been selected for Phase I evaluation. The transmitted data should be sufficient to describe a model which eliminates 50 to 60 percent of the ionospheric delay. More recently, global ionospheric models, developed for geophysical research have features which would overcome some of the above difficulties (Ref. 7, 8, and 9). It seems rather clear that ionospheric models, no matter how complex, are incapable of predicting the smaller scale structures of the ionosphere. While ionospheric models may be required for the single frequency user, the dual frequency calibration technique, discussed in the next section, will provide superior performance for the user requiring the ultimate in accuracy.

6. DUAL FREQUENCY RECEIVERS

From the dispersion relation of a high frequency electromagnetic wave in plasma, it is sufficiently accurate for L-band signals in GPS applications to characterize the frequency dependence of the ionospheric delay in terms of the inverse squared frequency. Thus, two L-band frequencies can be used to calibrate the ionospheric delay in GPS. These two frequencies must be transmitted coherently and must be sufficiently separated in frequency. The equations determining the calibration accuracy are derived below.

The pseudo-range equation provides

$$m_1 = p + d + (K/f_1^2) + n_1$$

where

m_1 = measured pseudo-range in channel i

p = true pseudo-range (including user clock bias)

d = pseudo-range uncertainties independent of frequency and receiver noise

K = parameter proportional to total electron content along with measurement path

f_1 = signal frequency in channel i

n_1 = receiver noise in channel i

Ignoring d and assuming measurements are made in two channels, permits K to be eliminated algebraically to obtain

$$p = (m_1 - n_1) + F\left[(m_1 - n_1) - (m_2 - n_2)\right] \tag{6-1}$$

where

$$F = f_2^2/(f_1^2 - f_2^2)$$

The second term in Eq. (6-1) comes from the solution for K, whereas the first term would occur even if K were known precisely and a single channel measurement is used.

Since the user does not know n_1 (by definition), the estimate of p will be

$$\hat{p} = m_1 + F(m_1 - m_2) \tag{6-2}$$

The error δp in the user estimate is then

$$\delta p = \hat{p} - p = n_1 + F(n_1 - n_2) \tag{6-3}$$

Note that n_1 and n_2 will be correlated.

If n_1 and n_2 have equal standard deviations of σ_n, and a correlation factor ρ, the standard deviation in pseudo-range error is

$$\sigma_p = \left\{\sigma_n^2 + \left(\frac{\sqrt{2} f_1 f_2}{f_2^2 - f_1^2}\right)^2 \sigma_n^2 (1 - \rho)\right\}^{1/2} \tag{6-4}$$

The first term in brackets represents the contribution due to single channel measurement error. The second term arises from the ionospheric calibration process alone. Using the current values for the two transmitted frequencies in GPS L_1 = 1.575 GHz and L_2 = 1.227 GHz, the relative contributions of these two sources are indicated in Figure 8. The sensitivity of the delay calibration to the selection of the secondary frequency L_2 is also shown. For the selected frequencies this figure shows that the total receiver measurement error is approximately 2.6 times the single channel measurement error for ρ = 0.

Although the ionospheric delay is not exactly proportional to the inverse of the frequency squared, higher order terms will be a negligible source of error. More detailed discussions of the dual frequency receiver ionospheric delay calibration are given in Refs. 10 and 11.

7. SUMMARY

In conclusion, we know that the ionosphere produces position errors at L-band frequencies which can be significant for high precision users (as much as 20 feet). The use of ionospheric modeling techniques can reduce the effects of these errors by 50 - 75 percent. One method of determining in real-time the degradation and delay due to ionospheric effects is the use of a dual frequency receiver. Two frequency calibration effectively reduces the ionospheric problem to a small noise number. An additional benefit is that GPS will allow large scale mapping and modeling of the ionosphere.

REFERENCES

1. Cretcher, C. K., "NAVSTAR and Global Positioning System: System Description, Program Plan, and User Performance" May 1974

2. Klobuchar, J. A. and Allen, R. S., "A First Order Prediction Model of Total Electronic Content for a Mid-Latitude Ionosphere" AFCRL-70-0403, Air Force Cambridge Research Laboratories, Bedford, Mass., July 1970.

3. Waldman, H. and De Rose, A. Y., "Prognostication of Ionosphere Electron Content" SAMSO TR-71-82, DDC AD 731095 & 6, Stanford University, Palo Alto, California, August 1971.

4. Rao, N. N., Youakim, M. Y., and Yeh, K. C., "Feasibility Study of Correcting for the Excess Time Delay of Transionospheric Navigational Ranging Signals" SAMSO TR-71-163, DDC AD-729797, University of Illinois, Urbana, Illinois, July 1971.

5. Pisacane, V. L., Feen M. M., and Sturmanis, M., "Prediction Techniques for the Effect of the Ionosphere on Pseudo-Ranging from Synchronous Altitude Satellites" SAMSO TR-72-22, DDC AD 749486, Applied Physics Laboratory, Silver Springs, Maryland, August 1972.

6. Bent, R. B., Llewellyn, S. K., and Wallock, M. K., "Description and Evaluation of the Bent Ionospheric Model" SAMSO TR-72-239 DDC AD-753081-6, DBA Systems, Inc., Melbourne, Florida, October 1972.

7. Nisbet, J. S., Radio Science, 6, 437, 1971.

8. Chiu, Y. T., J. Atm. and Terr. Phys., 37, 1563, 1975.

9. Ching, B. K., and Chiu, Y. T., J. Atm. and Terr. Phys., 35, 1615, 1973.

10. "Global Positioning System Final Report" SAMSO TR-74-182, DDC AD-921524L, General Dynamics, San Diego, California, February 1974.

11. "Global Positioning System Final Report" SAMSO TR-74-183, DDC AD-921525L, Philco Ford Corporation, Palo Alto, California, February 1974.

TABLE 1

User Requirement Classes and Applications

CLASS	ACCURACY	USER DYNAMICS	IMMUNITY TO ELECTROMAGNETIC JAMMING	TYPICAL USER
A	HIGH	MEDIUM	HIGH	PHOTO RECONNAISSANCE
B	HIGH	HIGH	MEDIUM	HELICOPTER
C	(LOW COST)	MEDIUM	(IMMUNITY TO NATURAL EMI)	GENERAL ENROUTE NAVIGATION
D	HIGH	LOW	HIGH	LAND VEHICLE
E	MEDIUM	LOW	HIGH	GROUND TROOP
F	HIGH	LOW	MEDIUM	SHIP

TABLE 2

Error Budget

SOURCE	CONTRIBUTION - m (1σ)	
	SINGLE FREQUENCY	DUAL FREQUENCY
EPHEMERIS	1.5	1.5
SATELLITE CLOCK AND ELECTRONICS	0.9	0.9
TROPOSPHERE MODEL	1.5	1.5
IONOSPHERE MODEL	4.9	-
RECEIVER NOISE	1.5	2.4*
MULTIPATH	1.2	1.8*
RSS	5.7	3.8

*Increase due to noise propagation from second frequency measurement

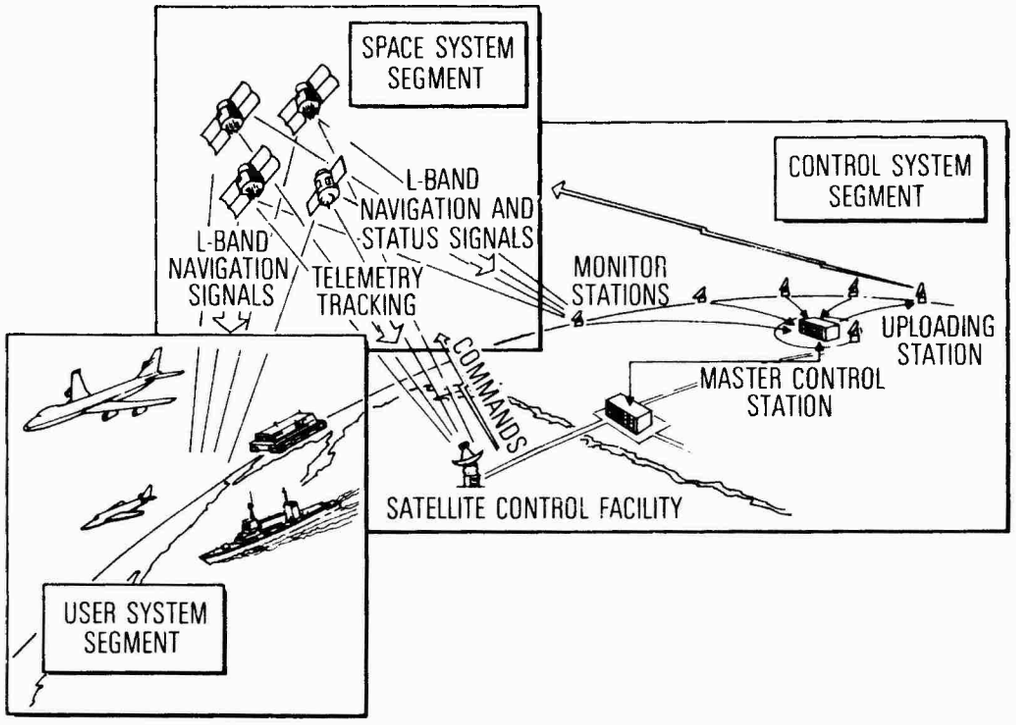
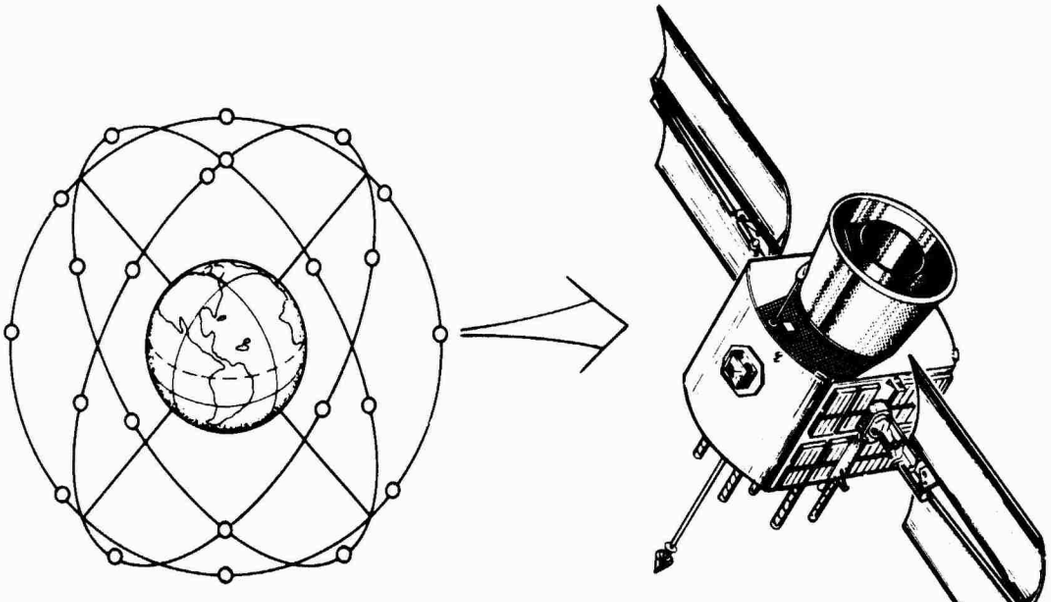


Figure 1



ORBITAL CONFIGURATION

- 12-hr PERIOD
- CIRCULAR ORBITS
- 63° INCLINATION
- EIGHT SATELLITES PER PLANE

SPACECRAFT CONFIGURATION

- 950 lb (typical)
- 450 Watts (typical)
- NAV SIGNALS
  - 1200 MHz
  - 1600 MHz
- 10^{-13} CLOCK STABILITY
- 5-YEAR DESIGN LIFE

Fig.2 GPS space segment

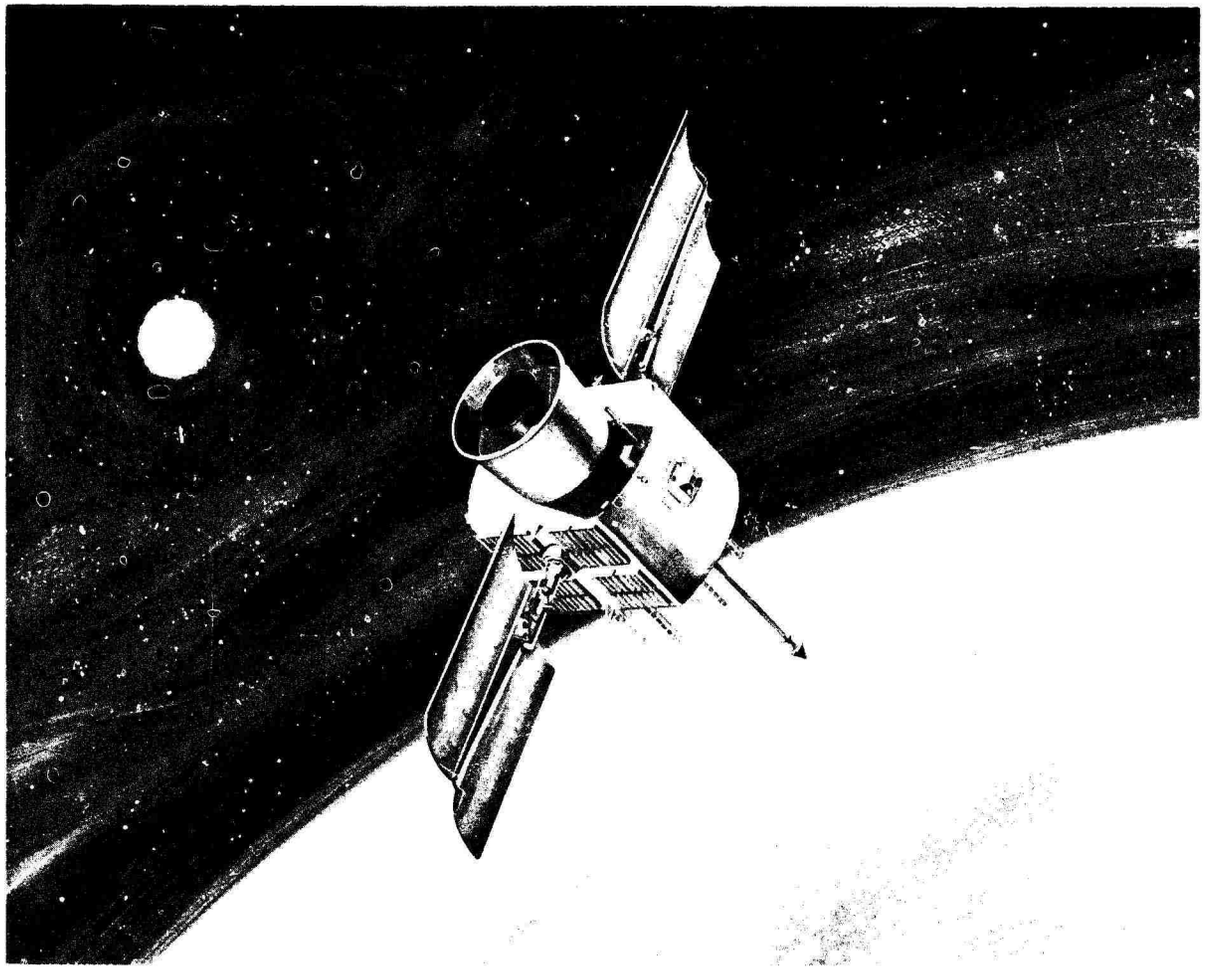


Figure 3

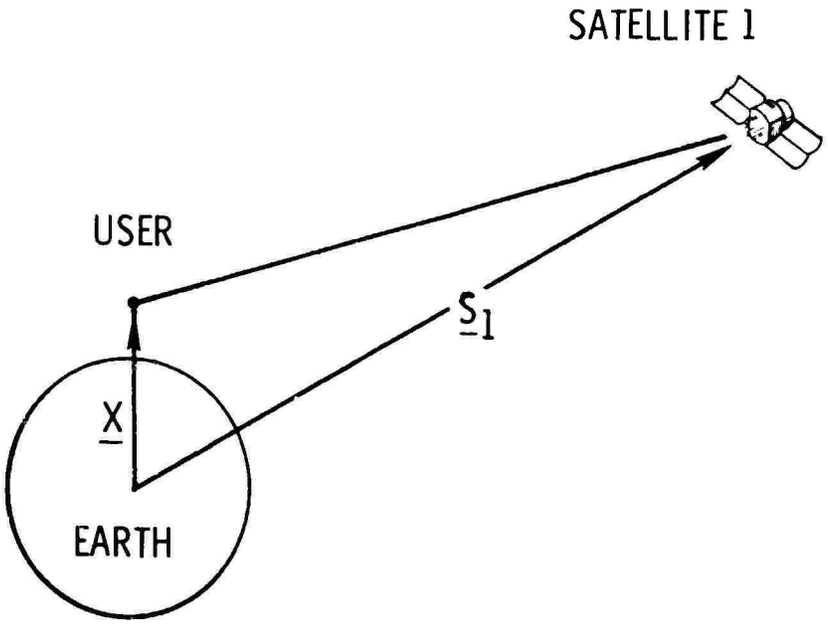


Figure 4

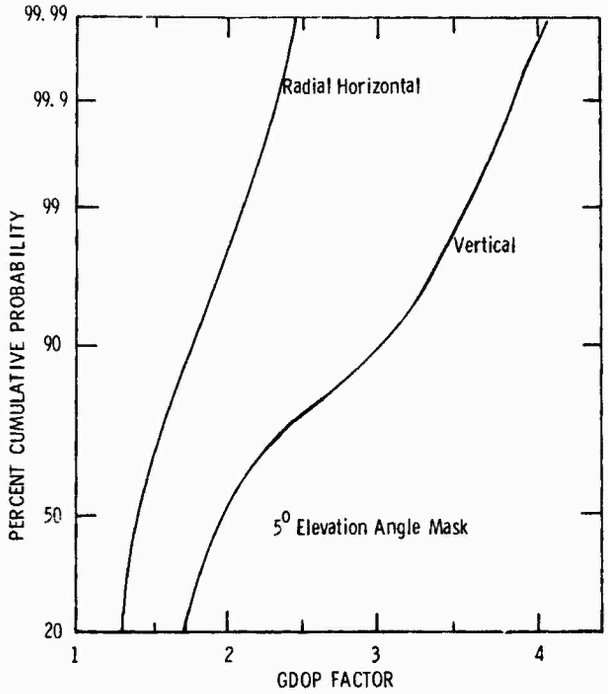


Figure 5

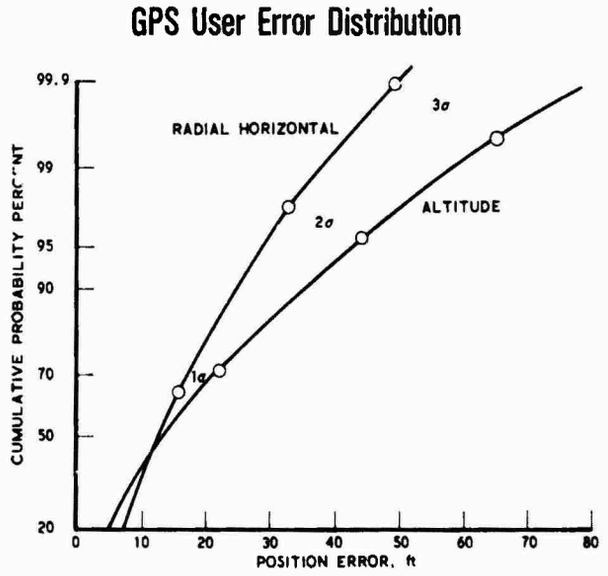


Figure 6

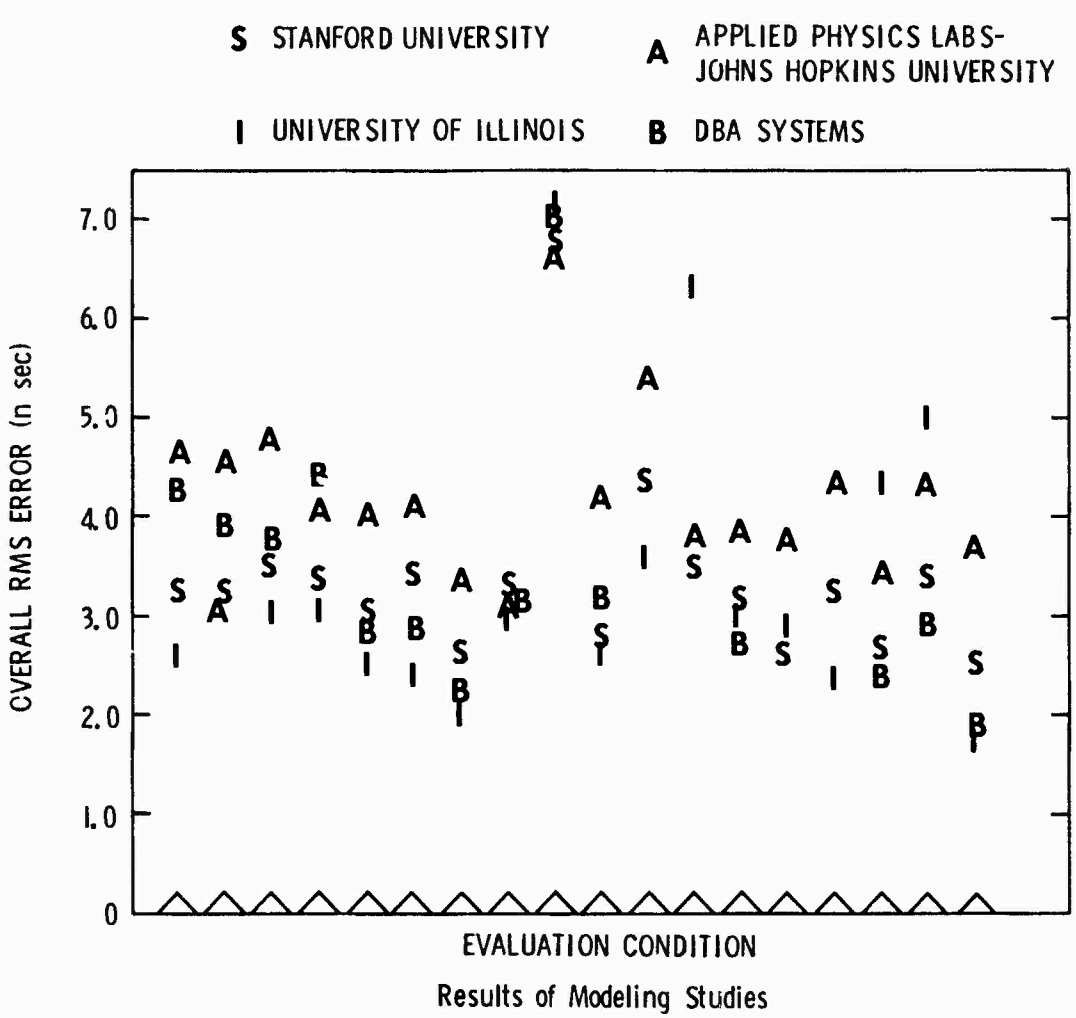


Fig.7 Results of modeling studies

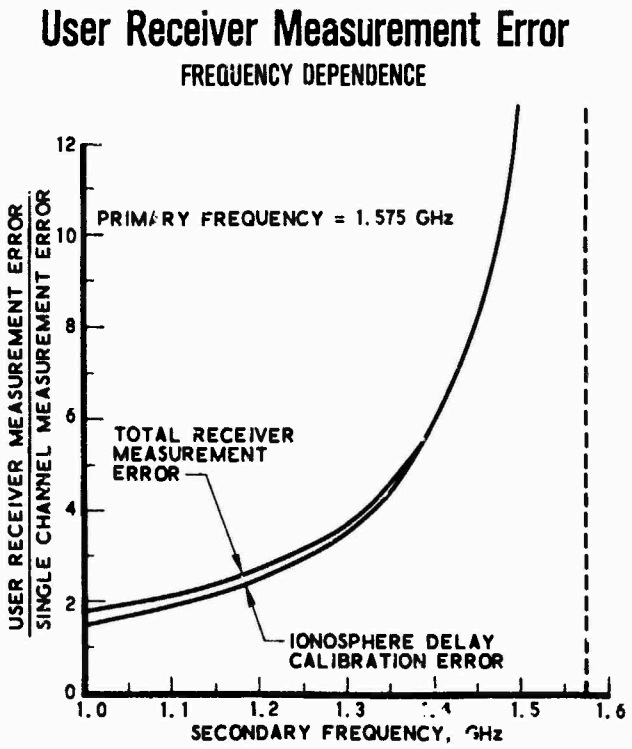


Fig.8 Frequency dependence of receiver measurement error

PLASMASPHERIC SIGNAL TIME-DELAY EFFECTS IN SATELLITE NAVIGATION SYSTEMS

H. Soicher
Communications/Automatic Data Processing Laboratory
U.S. Army Electronics Command, Fort Monmouth, NJ 07703
United States of America

SUMMARY

A satellite navigation concept requires determination of time-delays that satellite signals experience when traversing the distance between satellite and observer. A pulse propagating this distance is slowed down by an amount directly proportional to the total number of free electrons along its path. Ranging accuracy requirements mandate compensation for this additional signal delay.

The total electron content (TEC) for high-orbit satellites includes the ionospheric as well as plasmaspheric electron content.

The ATS-6 Radio Beacon Experiment provides the opportunity to determine: the ionospheric content (up to heights of ∿1500 km) by means of the Faraday rotation technique; the total content (from observer to satellite) by means of the dispersive-group-delay technique; and the plasmaspheric content (the difference between the two).

Since delay-time modeling efforts to date have been based on TEC obtained by the Faraday rotation technique, the plasmaspheric delay is not compensated for and thus comprises a source of error. In an effort to determine the magnitude of this error, observations of plasmaspheric content were performed at a midlatitude station, an auroral station and a station in the vicinity of the geomagnetic equatorial anomaly. Diurnal, day-to-day, and latitudinal variations of the contents were observed. The ratio of the plasmaspheric-to-ionospheric contents, which determines the corrections that must be applied to the present modeling scheme, varied substantially from day to night.

Introduction

A global satellite-navigation concept (Parkinson, 1974) requires that the user must measure the time-delay experienced by satellite-emitted signals in traversing the distance between satellite and user. Simultaneous measurements of the propagation time of signals originating from four different satellites permit the user to determine his position and clock bias if satellite ephemerides and signal propagation velocity are known.

A pulsed signal propagating through the ionosphere is slowed down somewhat from its free space velocity, thus resulting in an apparent range that is larger than the equivalent free-space range. The difference between the apparent range and the true range, or between the true velocity and the free space velocity, is the quantity needed to improve range accuracy. Either of these quantities are directly proportional to the total electron content (TEC) along the path of the propagating signal. Thus, if the total electron content is known, or is measured, a perfect correction to ranging can be performed. The total electron content may be measured in real time, provided the user has dual-frequency capabilities (Burns & Fremouw, 1970). However, a substantial reduction in the cost of user equipment could be realized if the navigation system utilized only one frequency. Then time delays due to TEC would have to be determined through empirical modeling techniques (Bent, et al., 1975; Klobuchar, 1975) based on existing and future global electron content data. The appropriate delay corrections would be transmitted to the user via the navigation signal.

The Radio Beacon Experiment (Davies, et al., 1972; Soicher, 1975) aboard the geostationary Applied Technology Satellite ATS-6 permits the determination of TEC by two independent methods. The first--the Faraday polarization rotation technique--has been widely used in obtaining TEC data. Faraday rotation is a terrestrial-magnetic-field dependent phenomenon. Since its magnitude is heavily weighted near the earth, it is considered to provide integrated electron content values for altitudes below ∿1500 km. The second--the dispersive-group-delay technique, in which the phase of the modulation envelope between a carrier and its sideband is compared at two frequencies--is independent of the terrestrial-magnetic-field and, hence, yields the integrated electron content between the observer and the satellite-signal source. The Faraday content, N_F, and the dispersive-group-delay content, N_T, therefore, yield the total electron content up to ∿1500 km and geostationary altitudes, respectively. The difference between N_T and N_F yields the content above ∿1500 km, which is referred to as the plasmaspheric content, N_P.

To date, delay-time modeling efforts have been based on TEC obtained by the Faraday rotation technique. Hence, the delay due to the plasmaspheric electron content has not been accounted for. Since the navigation satellite system under consideration will utilize satellites at ∿20,000 km, the signals will be slowed down by free electrons in both the ionosphere and the plasmasphere. Thus the corrections for TEC supplied by a model based on the Faraday technique will not adequately compensate for the total signal time delay. This paper aims to present the magnitude of the ionospheric and plasmaspheric time delay values at midlatitude stations, an auroral station, and at a station in the vicinity of the geomagnetic anomaly during the quiet phase of the solar cycle.

THE DATA

The ATS-6 Satellite was launched into a geostationary orbit in May 1974. During Phase I of its operation, when the satellite was located at 94°W, the U.S. Army Electronics Command operated observation stations at Fort Monmouth, NJ (40.18°N, 74.06°W); Richmond, Florida (25.60°N, 80.40°W); and Sao Paulo, Brazil (23.50°S, 46.50°W). In May 1975, the satellite was moved to 35°E for Phase II of its operation. During this phase, the observation stations were moved to Haifa, Israel (32.87°N, 35.09°E) and Kiruna, Sweden (61.84°N, 20.41°E). In August 1976, the satellite was moved to its final parking position at ∿134°W, with some of the observation stations being reactivated and others being established.

The Faraday observations were performed with the 140 MHz satellite beacon emissions. The dispersive-group-delay observations were performed with the 140 MHz and 360 MHz beacons modulated with sideband displacement of $\Delta f = +1$ MHz.

The daily variations in vertical TEC measured by the Faraday technique and by the dispersive-group-delay technique at Kiruna, Sweden for Oct. 75 are shown in Figs. 1 and 2, respectively. The equivalent signal-delay-times normalized to a frequency of 1.6 GHz, which is in the satellite navigation band, are also indicated in these figures. The monthly mean of N_F varied from a predawn minimum of ∿1.3 TEC units (1 TEC unit $\equiv 10^{16}$ e/m^2) to a maximum of ∿9.8 TEC units at ∿1100 UT. The monthly mean of N_T varied from a predawn minimum of 2.9 TEC units to a maximum of ∿11.5 TEC units, which also occurred at 1100 UT. The monthly standard deviation of N_F varied from a predawn minimum of ∿0.2 TEC units to a maximum of ∿2 TEC units at ∿1215 UT. The monthly standard deviation for N_T varied from ∿0.6 TEC units to 1.9 TEC units at the same times.

The daily variations in the vertical plasmaspheric content, N_p, at Kiruna for October 1975 are shown in Fig. 3. In general, the diurnal behavior of N_p appears similar to that of N_F and N_T, i.e., the content increased during the day and decreased at night. The monthly mean of N_p was at a minimum prior to midnight and at a maximum at ∿1300 UT, (between ∿0.9 TEC units and ∿2.2 TEC units). This was different from the behavior of N_p at a midlatitude station (Soicher, 1976). The monthly standard deviation of N_p varied from ∿0.4 TEC units to ∿0.8 TEC units.

The daily variations of the plasmaspheric-to-ionospheric content ratio (N_p/N_F) at Kiruna for Oct. 75 are shown in Fig. 4. The monthly average for this ratio varied from ∿19% (just prior to noon) to ∿120% (just prior to sunrise). The ratios dropped abruptly at sunrise, reached a minimum in the morning hours, from whence it increased gradually until it increased abruptly at 0130 UT. This abrupt predawn increase in the ratio is accounted for by the decrease in N_F during that period, while N_T remained unchanged, thus increasing N_p. Since N_F is sensitive to changes in density distributions whereas N_T is not, the changes in the ratio indicate that during the predawn period, upward electron fluxes were generated by temperature increases. The electrons streamed into higher altitudes where they were no longer counted by the Faraday rotation method, thus decreasing N_F. N_T, however, was unaffected by this redistribution, since its electrons were still counted by the dispersive-group-delay method. Since some of the N_F electrons transferred into the plasmasphere, an increase in N_p was also noticed. The increase in N_p and the simultaneous decrease in N_F caused the abrupt increase in their ratio just prior to dawn. The abrupt decrease in the ratio after dawn was due to ionospheric electron production buildup. Since production and loss processes are insignificant at plasmaspheric heights and N_p is varied by diffusive fluxes to and from the ionosphere, the effects of sunrise on the ionosphere are much larger than on the plasmasphere, resulting in the sharp decrease of the ratio during that period.

The daily variations of the plasmaspheric-to-total content ratio (N_p/N_T) for Kiruna, Sweden for October 1975 are shown in Fig. 5. The behavior of this ratio is similar to that of N_p/N_F shown in Fig. 4. The mean monthly average varied from ∿16% just prior to noon to ∿52% just prior to sunrise.

In order to determine the time-delay correction at an arbitrary geographic location, time of day, and season, the modeling base must include data from a global network of stations operating continuously. The following comparative data is presented for three world ionospheric regions: the high latitude, the midlatitude and the South Atlantic anomalous regions.

Figures 6 and 7 depict a comparative mean monthly variation of TEC measured by the Faraday technique at Fort Monmouth, Richmond, and Sao Paulo for February 1975 and May 1975, respectively. The solar zenith angle dependence of TEC is shown in the relative values of TEC recorded at the two northern hemispheric midlatitude stations. The comparatively higher TEC values recorded at Sao Paulo, even if seasonal effects are ignored (compare Figs, 6 and 7), were due to the passage of the ray path from satellite to observer through the Equatorial Anomaly region. At equatorial regions, ionospheric plasma which move upward from an imposed E-W electric field diffuses downward again, but obliquely, along the geomagnetic lines of force. Instead of returning to the source, the plasma arrives at two places: one to the south and one to the north of the equator. The electron content is then depleted at the magnetic equator, but is enhanced at the anomalous regions. The relatively higher TEC values for February as compared to May, confirm topside sounder results (Eccles & King, 1969), that the anomaly content crests are larger in the summer hemisphere. In addition, the ray path was in the vicinity of the South Atlantic Anomaly region where the magnetic field is abnormally weak and the loss of radiation belt particles is rapid. TEC enhancements are attributed to ionization produced by the precipitating energetic trapped electrons that mirror low in the anomaly (Zmuda, 1966; DeMendonca, 1965).

The comparative variation of TEC at a midlatitude station and at a station near the auroral zone is shown in Fig. 8 for Oct. 1975. The consistently higher TEC values at Haifa were due to solar zenith angle influence on TEC. The passage of the ray path from Kiruna to the satellite through the density trough (Muldrew, 1975; Liszka, 1966) was also responsible for the observed depressed TEC values.

Finally, the comparative variation of the plasmaspheric electron content, N_p, at Fort Monmouth and Sao Paulo during May 1975 is shown in Fig. 9. In general, the mean monthly values of N_p at both stations lay between ∿1 and ∿2 TEC units. The mean diurnal variation at Sao Paulo indicates that N_p had higher values during the day than during the night. The times of N_p enhancement were close to those at which the equatorial anomaly is formed and maintained (Eccles & King, 1969), while the times of N_p depression corresponded to disappearance of the anomaly. It is therefore possible, that during the formation of the anomaly and during its maintenance, the plasmaspheric and the ionospheric electron densities are increased.

CONCLUSIONS

Data from the Radio Beacon Experiment of the ATS-6 has yielded information on the magnitude and variations of ionospheric and plasmaspheric contents at midlatitude, auroral, and equatorial regions. The group-path-delay of a navigation signal caused by free electrons in the plasmasphere cannot be neglected when compared to the delay due to the ionosphere. The group-path-delay prediction models currently available are based on Faraday data, which do not adequately compensate for the total delay. During the night at an auroral station they may be off in the mean by ∿120% while during the day by ∿20%. At a midlatitude station, they may be off by more than 100% at night and by ∿13% during the day. For a station in the vicinity of the equatorial anomaly, they may be off by more than 100% at night and by less than 10% during the day. Fortunately, the highest ratio of plasmaspheric to ionospheric delay times occur at night when the total time delay (the quantity of interest) is small. The relatively small magnitude of the diurnal, day-to-day, and seasonal variation of N_p suggests that present ionospheric models may be corrected by adding a constant offset (say an average monthly value of N_p) to the ionospheric values.

The data reported here was taken during the quiet phase of the solar cycle. During such a phase, the total group delay is generally small, and modeling schemes yield corrections within the accuracy requirements of the proposed navigation systems. It remains to be seen, however, if the observed absolute values of the plasmaspheric content, or the observed ratio of plasmaspheric-to-total delays, will be larger during the maximum phase of the solar cycle. If such ratios are maintained during the maximum of the cycle, neglecting the plasmaspheric content will cause errors which exceed the accuracy requirements of the navigation system.

ACKNOWLEDGEMENTS

The author wishes to thank the following individuals and their institutions under whose guidance the data reported in this paper was collected: Dr. S. Westerlund, Kiruna Geophysical Institute, Kiruna, Sweden; Dr. Z. Houminer and Mr. A. Shuval, Radio Observatory, Haifa, Israel; Prof. P. Kaufmann, Mackenzie University, Sao Paulo, Brazil; Mr. D. Monger, Naval Observatory, Richmond, Florida; and Mr. F. J. Gorman, Jr., US Army Electronics Command, Fort Monmouth, NJ.

REFERENCES

BENT, R. B., LLEWELLYN, S. K., NESTERCZUK, G., and SCHMID, P. E., 20-22 Jan. 1975, "The Development of a Highly Successful World-Wide Empirical Ionospheric Model and its Use in Certain Aspects of Space Communications and World-Wide Total Electron Content Investigations," Paper 1-4 in Proc. 1975 Symposium on the Effects of the Ionosphere on Space Systems and Communications (sponsored by U. S. Naval Research Laboratory, Washington, D. C.).

BURNS, A. A. and FREMOUW, E. J., November 1970, "A Real-Time Correction Technique for Transionospheric Ranging Error," IEEE Transactions on Antennas & Propagation, vol. AP-18, pp. 785-790.

DAVIES, K., FRITZ, R. B., and GRUBB, R. N., March/April 1972, "The ATS-F/G Radio Beacon Experiments," J. Environmental Sciences, vol. 15, pp. 31-35.

DE MENDONCA, F., 1965, "Ionospheric Electron Content Measurements in Regions of Low-Magnetic Dip and Through the Brazilian Magnetic Anomaly," Space Research V, pp. 687-701.

ECCLES, D. and KING, J. W., June 1969, "A Review of Topside Sounder Studies of the Equatorial Ionosphere," Proc. IEEE, vol. 57, no. 6, pp. 1012-1018.

KLOBUCHAR, J. A., 25 Sept. 1975, "A First Order, World-Wide Ionospheric Time-Delay Algorithm," AFCRL-TR-0502.

LISZKA, L., Oct. 1966, "Latitudinal and Diurnal Variations of the Ionospheric Electron Content Near the Auroral Zone in Winter," Radio Science, vol. 1 (New Series), no. 10, pp. 1135-1137.

MULDREW, D. B., June 1, 1965, "F-Layer Ionization Troughs Deduced from Alouette Data," J. Geophys. Res., vol. 70, no. 11, pp. 2635-2650.

PARKINSON, B. W., 3-5 Dec. 1974, "NAVSTAR: Global Positioning System - An Evolutionary Research and Development Program," Proc. Sixth Annual Precise Time & Time Interval (PTTI) Planning Meeting, U. S. Naval Research Lab., Washington, D. C., Report No. X-814-75-117.

SOICHER, H., Jan. 24, 1975, "The ATS-6 Radio Beacon Experiment," Nature, vol. 253, pp. 252-254.

SOICHER, H., 1-4 June 1976, "Diurnal, Day-to-Day, and Seasonal Variability of N_F, N_T, and N_P at Fort Monmouth, N. J.," Proc. COSPAR Symposium on "The Geophysical Use of Satellite Beacon Observations," Boston University, Boston, Massachusetts.

ZMUDA, A. J., April 1, 1966, "Ionization Enhancement from Van Allen Electrons in the South Atlantic Magnetic Anomaly," J. Geophys. Res., vol. 71, no. 7, pp. 1911-1917.

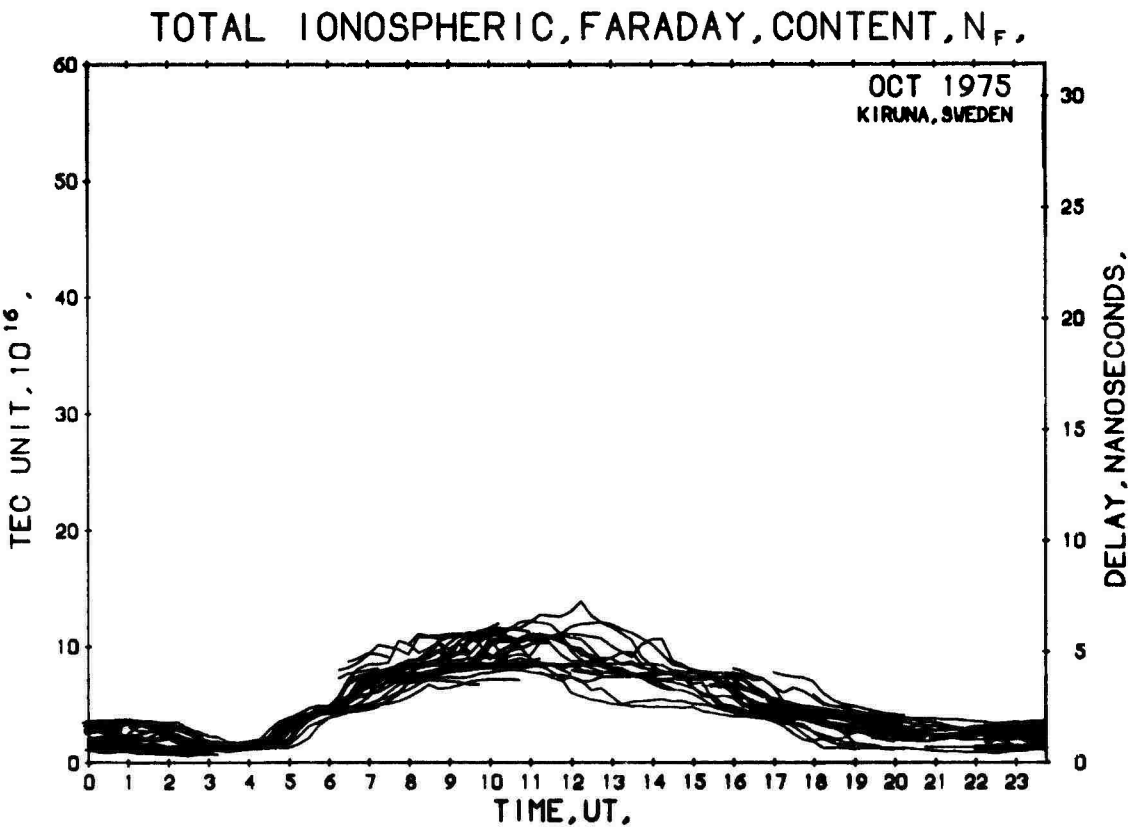


Fig.1 Total ionospheric (Faraday) electron content (N_F) at 15-minute intervals at Kiruna, Sweden, October 1975 (left ordinate: 1 TEC unit $\equiv 10^{16}$ e/m^2; right ordinate: signal delay normalized to 1.6 GHz)

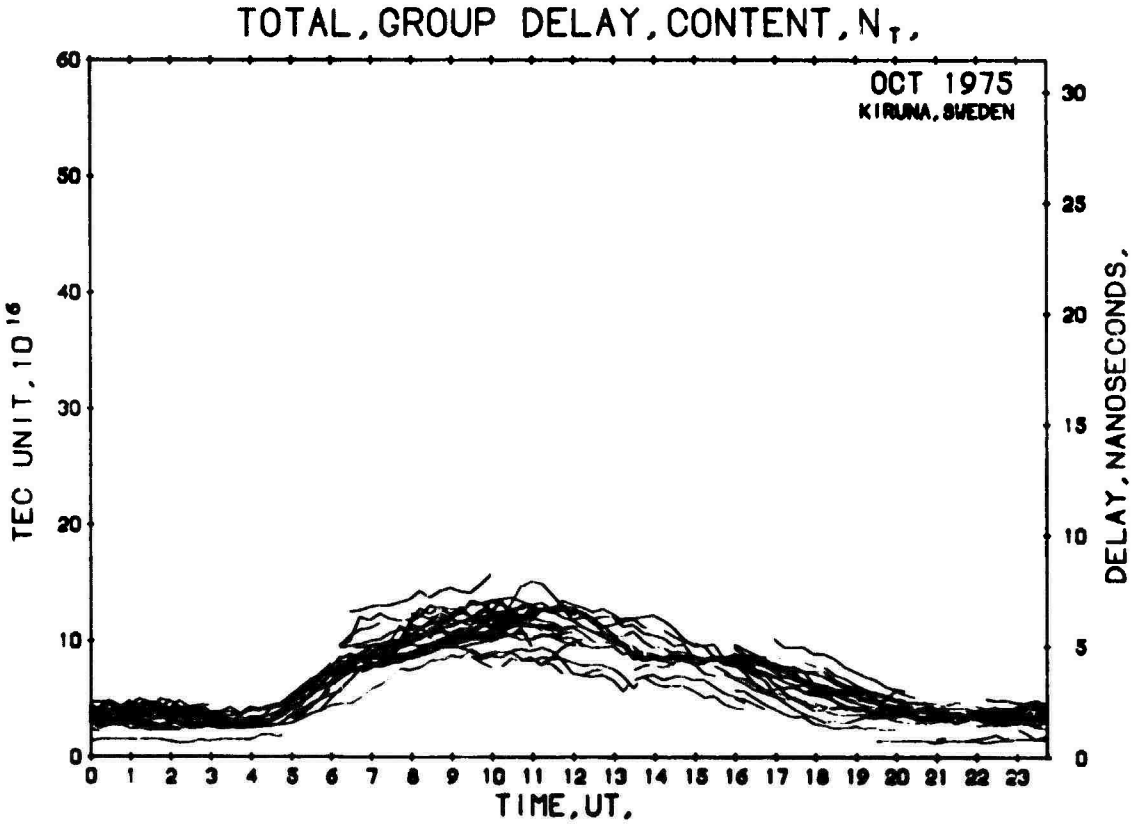


Fig.2 Total (group delay) electron content (N_T) at 15-minute intervals at Kiruna, Sweden, October 1975 (left ordinate: 1 TEC unit $\equiv 10^{16}$ e/m^2; right ordinate: signal delay normalized to 1.6 GHz)

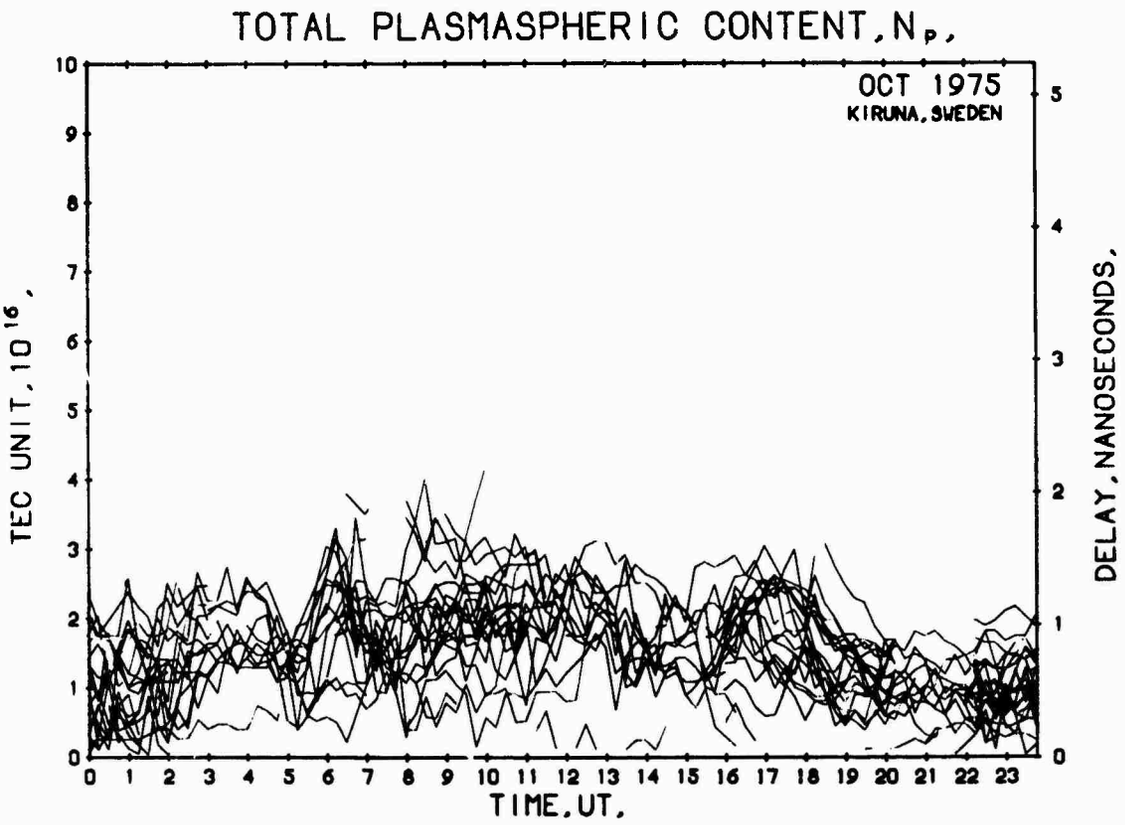


Fig.3 Total plasmaspheric content (N_p) at 15-minute intervals, Kiruna, Sweden, October 1975 (left ordinate: 1 TEC unit $\equiv 10^{16}$ e/m^2; right ordinate: signal delay normalized to 1.6 GHz)

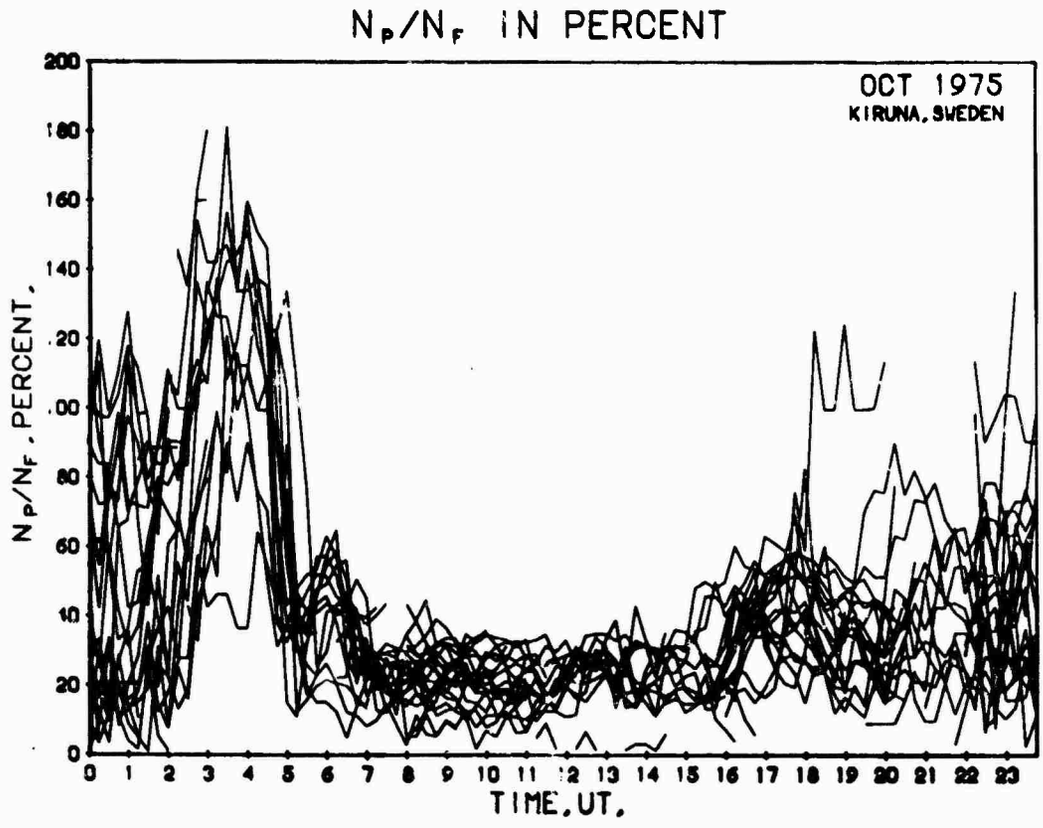


Fig.4 Ratio of plasmaspheric-to-ionospheric contents (N_p/N_F) at Kiruna, Sweden, October 1975

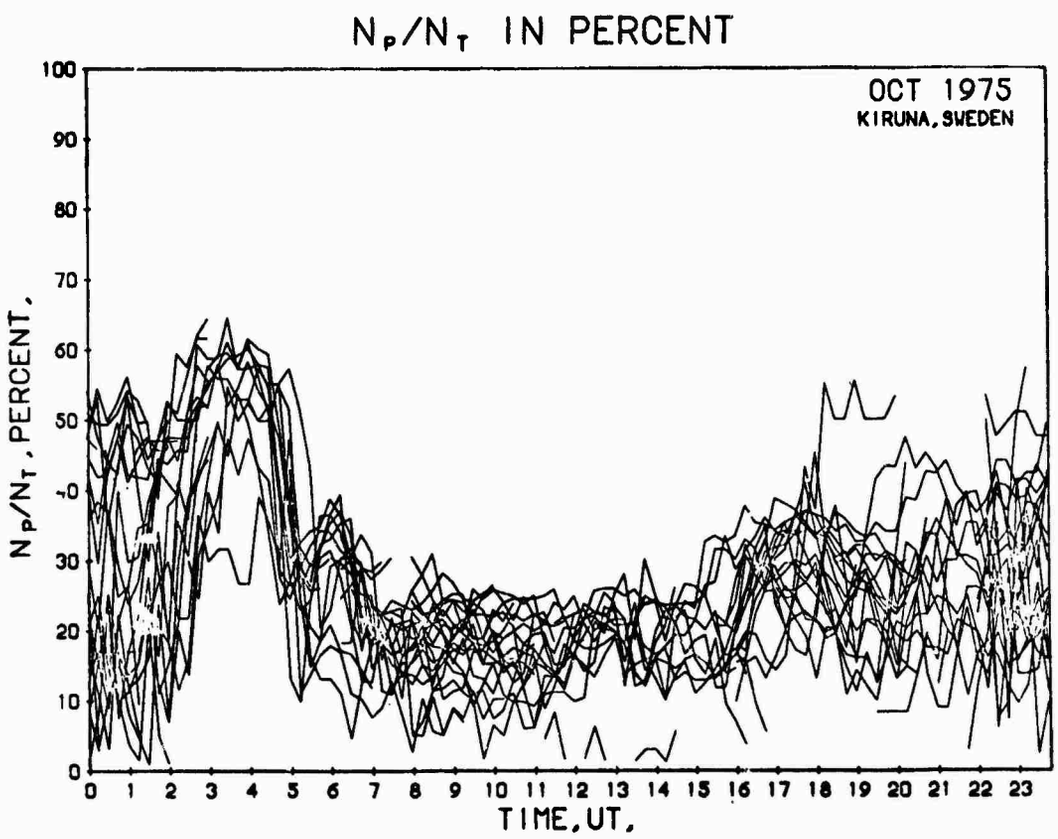


Fig.5 Ratio of plasmaspheric-to-total contents (N_P/N_T) at 15-minute intervals, Kiruna, Sweden, October 1975

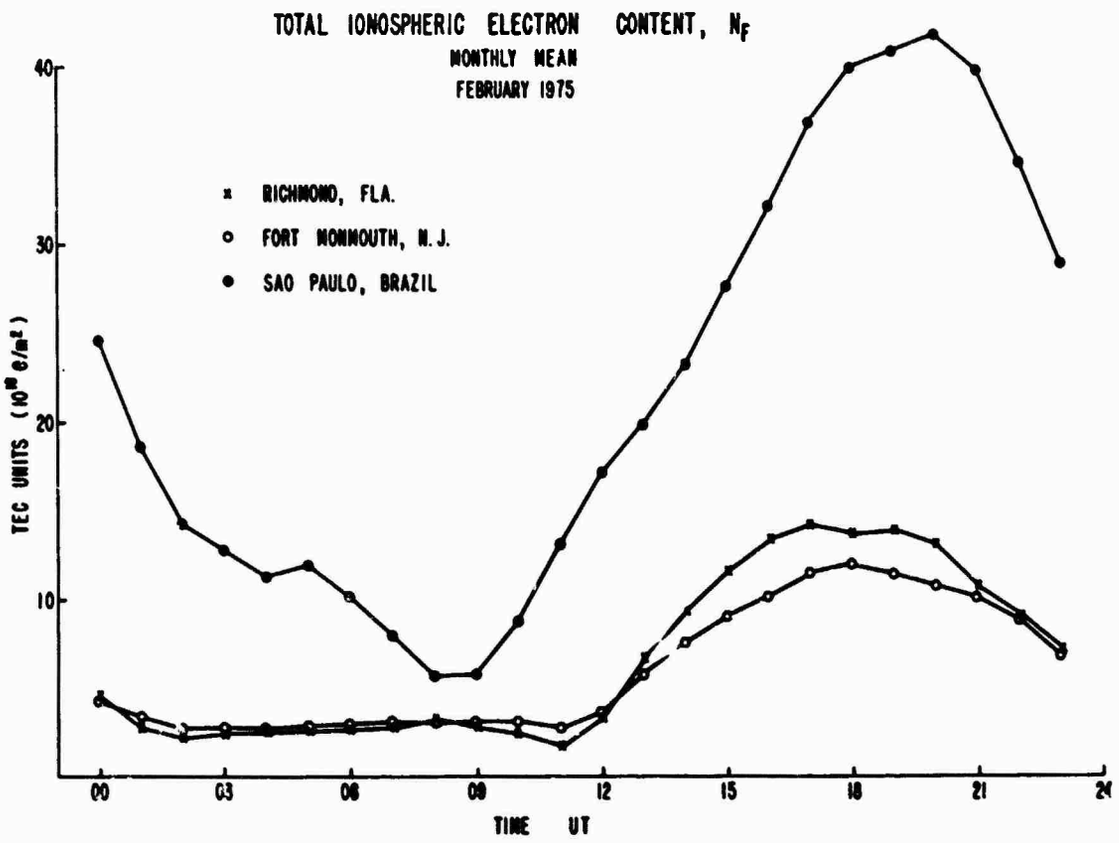


Fig.6 Mean monthly variation of N_F at Sao Paulo, Brazil (LT = UT – 3); Fort Monmouth, N.J. (LT = UT – 5); and Richmond, Florida (LT = UT – 5), during February 1975

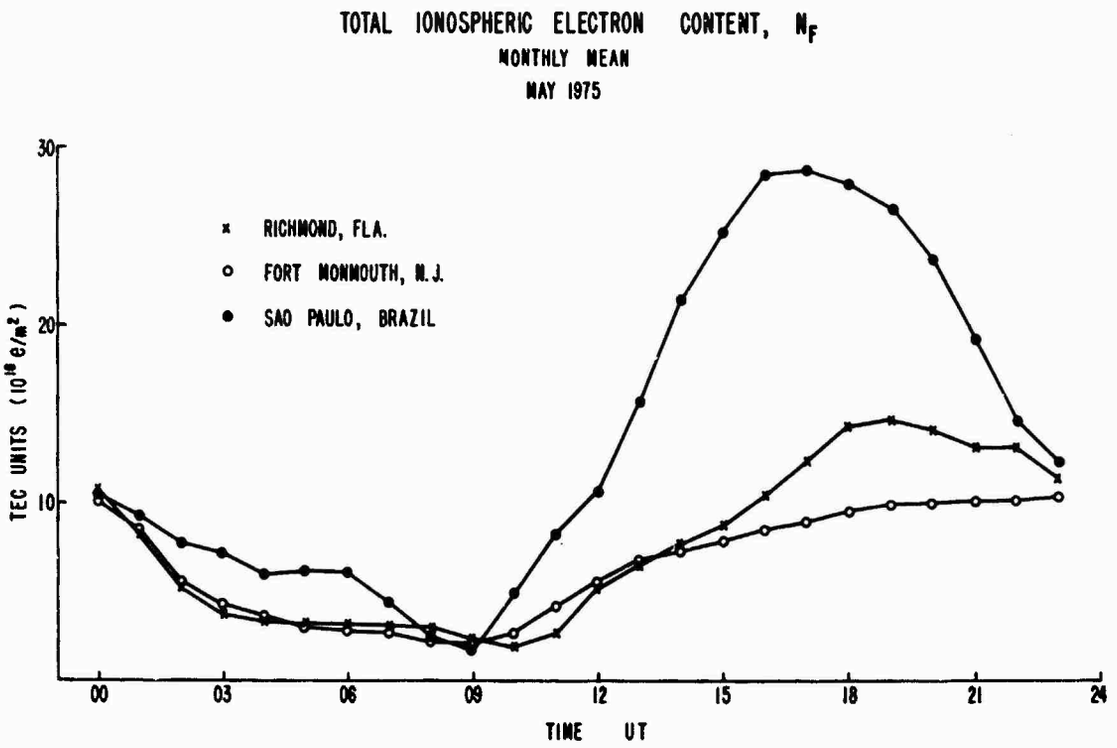


Fig.7 Mean monthly variation of N_F at Sao Paulo, Brazil (LT = UT – 3); Fort Monmouth, N.J. (LT = UT –5); and Richmond, Florida (LT = UT – 5), during May 1975

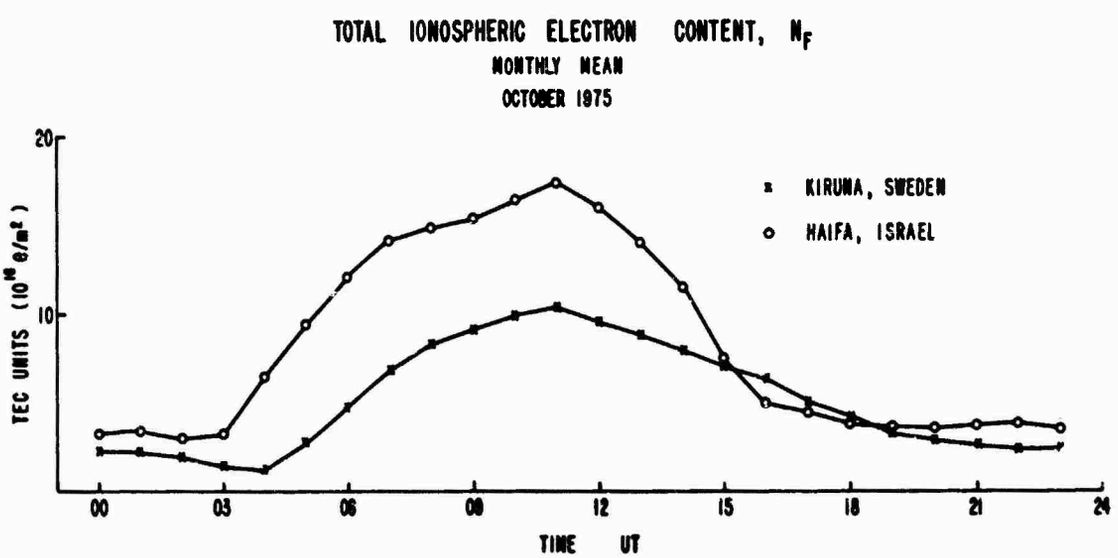


Fig.8 Mean monthly variation of N_F at Kiruna Sweden (LT = UT + 1); and Haifa, Israel (LT = UT + 2), October 1975

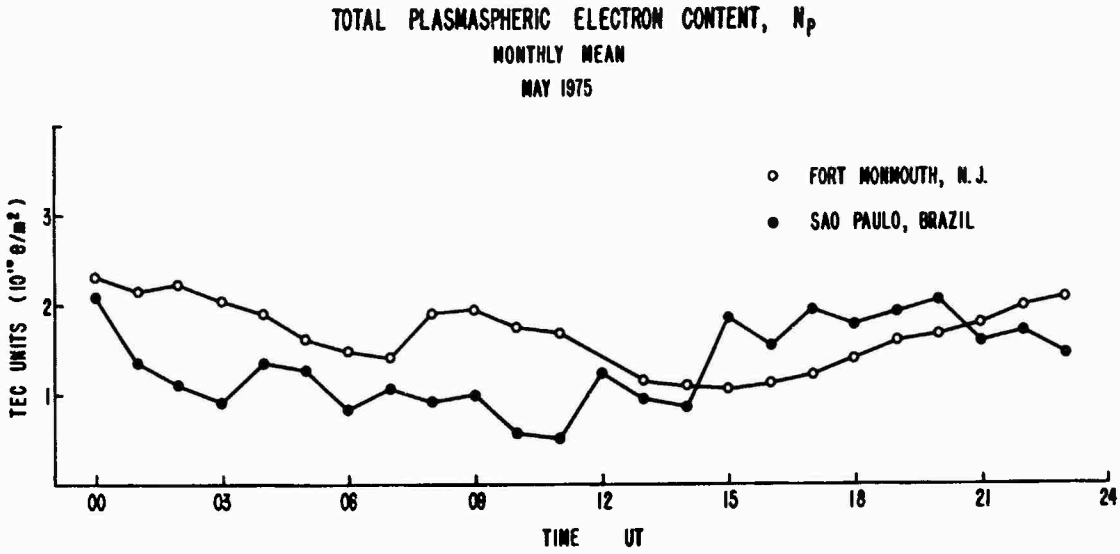


Fig.9 Mean monthly variation of N_p at Fort Monmouth, N.J.; and Sao Paulo, Brazil, May 1975

DISCUSSION

T.R.Larsen, Norwegian Defence Research Establishment, Kjeller, Norway

Did you have any measurements of N_T during severe ionospheric disturbances? (incl PCA events if available)

Author's Reply

Yes. A magnetic storm has definite effect on the ionospheric and plasmaspheric electron contents. Following Sudden Commencement (SC) both contents generally increase substantially for a few hours. This is followed by a drastic decrease in the contents which last for a few days. The response of the contents to magnetic storms depends on geomagnetic location, phase of the diurnal variation, and intensity of storm.

IONOSPHERIC TIME DELAY CORRECTIONS FOR ADVANCED SATELLITE RANGING SYSTEMS

J. A. Klobuchar
Air Force Geophysics Laboratory
Air Force Systems Command
Hanscom AFB, MA 01731

SUMMARY

As the requirements for greater positioning accuracy for satellite ranging systems have increased the time delay produced by the ionosphere has become a more important consideration over most of the world. For some advanced ranging systems advantage can be taken of the fact that the ionosphere is a dispersive medium and the ionospheric contribution to range error can be nearly eliminated by the use of two relatively widely spaced, system operating frequencies. For other systems only a narrow eperational band of frequencies can be utilized, and yet the ionospheric range error is, at times, the greatest source of system range error. For these systems numeric model representations of the ionospheric range error are generally used. The complexity and the need for real time update of these numeric models are a function of the user's requirements and his available computer resources. In this paper some general comments are made concerning ionospheric time delay models and a specifi algorithm designed for the single frequency user of the NAVSTAR-Global Positioning System (GPS) is described in detail.

1. INTRODUCTION

As the requirements for greater positioning accuracy for satellite ranging systems over most of the world have increased, the time delay produced by the ionosphere has become a more important consideration. For some advanced ranging systems advantage can be taken of the fact that the ionosphere is a dispersive medium and the ionospheric contribution to range error can be nearly eliminated by the use of two, relatively widely spaced, system operating frequencies. For other systems only a narrow operational band of frequencies can be utilized and yet the ionospheric range error is, at times, the greatest source of system range error. For these systems numeric model representations of the ionospheric range error are generally used. The complexity and the need for real time update of these numeric models are a function of the user's requirements. In this paper some general comments are made concerning ionospheric time delay models and a specific algorithm designed for the single frequency user of the NAVSTAR/Global Positioning System (GPS) is described in detail.

2. IONOSPHERIC MODEL REPRESENTATION

2.1 Model Choice

For systems which cannot utilize the dispersive characteristics of the ionosphere by operating with two, widely spaced frequencies to independently measure and subtract the ionospheric time delay, some type of numeric model representation of ionospheric time delay is generally utilized. For systems operating over a restricted geographic region, such as a satellite detection radar, which can typically view approximately 15 degrees of earth centered angle around the radar station, the numeric model can be a sub-model of a world-wide representation. That is, it can utilize fewer coefficients than a world-wide model and the coefficients can be better adapted to the ionospheric gradients known to exist over a particular station. Also, model updates can be made by near-real-time measurements of an ionospheric F2 region parameter at some point near the radar. If only a single-value measurement is available the numeric model is generally unchanged in shape, but modified only in absolute value, by the percentage ratio between the actual measurement and the model at the particular ionospheric location of the measurement. If several measured values are available the model gradients, as well as the absolute values, can be changed. Other systems may not require a state of the art ionospheric time delay model, or may not be able to incorporate the full model into the system because of software and system operation constraints.

2.2 NAVSTAR/GPS Ionospheric Model

The Global Positioning System, GPS, will offer a precise positioning service to within a few feet rms, on a global basis, to United States Department of Defense users. The time delay effects of the ionosphere, which can be up to 300 feet in satellite range for this system, are reduced to the same magnitude as the other sources of system error by the use of two L-band frequencies from which the ionospheric contribution to range is measured. Many system users may have only single frequency navigation capability, however, and, for this class of user, a simple, minimum sized algorithm of world-wide ionospheric time delay has been developed. This algorithm which requires only eight coefficients, corrects for approximately 50 percent of the ionospheric range delay on a world wide rms basis. A state of the art ionospheric numeric representation was not made available for single frequency GPS users due to the large number of coefficients and relatively complex computations required to be performed by the user. Even a state of the art ionospheric time

delay model would not satisfy the system operational requirements at times; thus, users who require the greatest accuracy at all times must use the two frequency capability of the GPS. Those users who do not utilize the GPS two frequency capability will still have the availability of a minimum sized algorithm which will correct for more than one-half of the rms ionospheric range error.

3. CHARACTERISTICS OF IONOSPHERIC TIME DELAY

The ionospheric time delay is directly proportional to the total electron content along the path between the satellite and the user. Total Electron Content (TEC) is a function of many variables; among them are local time, season, geographic location, and state of solar and magnetic activity. The TEC is highest in most regions of the world within a few hours of local noon and is greatest at regions located approximately plus and minus 20 degrees away from the geomagnetic equator. This equatorial anomaly region, as it is called, is the region of the world where single-frequency GPS users will encounter the greatest errors. Statistics of daily values of TEC from individual stations ranging from auroral to equatorial locations show that the rms difference from the monthly mean values during the midday hours is approximately 20 to 30 percent. Thus, an accurate specification of the monthly mean values alone will represent an elimination of 70 to 80 percent (rms) of the ionospheric error. The problem, then, is how to best represent the monthly mean behavior of ionospheric time delay, particularly during those times of the day and in those regions of the world when the time delay is greatest, using the minimum number of coefficients and computation time.

4. DESIGN OF THE ALGORITHM

The model algorithm was designed to meet several criteria. First, it was designed to fit best during those times of the day when the ionospheric delay is the largest, that is, during the afternoon period. Second, it was designed to have versatility in the coefficient array so that tradeoffs can be made in the number of coefficients used to represent the latitude dependence and the coefficients can be weighted for the latitudes of greatest interest, such as the Continental United States (CONUS). Alternatively, the coefficients can be weighted for best fit in those geographic locations where the time delays are greatest, namely the regions located approximately plus and minus 20 degrees either side of the geomagnetic equator. Finally, the model was simplified by making approximations to geometry, wherever possible, to make its computer running time shorter and storage-space requirements smaller.

The available experimental values of TEC from which time-delay values are obtained are not sufficient, by themselves, to form the basis of a world wide time-delay model. However, there are sufficient TEC data to yield important information on the statistical behavior of time delay for locations representative of several regions of the world. The available data also show, quite well, the diurnal behavior of TEC for various seasons and solar-flux values. These data have been used to determine the form of the diurnal algorithm and they have been supplemented with the Bent model (Llewellyn and Bent, 1973) representation for those regions where actual data were not available, particularly for the shape of the latitude gradient.

5. MODEL ALGORITHM

The average monthly diurnal behavior of time delay at any location, as a function of time of day, has been represented by a simple positive cosine wave dependence, with an additional constant offset which we shall call the DC term. In Figure 1 two examples of actual ionospheric time-delay (T_g) data are shown with a best-fit cosine-type curve representation of each set of actual data included. While the cosine curve is, of course, not a perfect fit to the actual data, it can be made to fit reasonably well during the maximum of the day when the absolute T_g values are greatest. The cosine function used is:

$$T_g = DC + A \cos \frac{(t - \phi)2\pi}{P}$$

where DC, A, ϕ, and P are the four parameters required to obtain a complete diurnal representation of T_g at any location. Only the positive half of the period of the cosine curve is used in this simple representation; thus, when the absolute value of the quantity $(t - \phi)2\pi/P$ becomes greater than 90 degrees, T_g is represented only by the DC term.

The simple positive cosine representation of the actual T_g is designed to fit best near the hours of maximum T_g. The fit near sunrise and well after sunset is expected to be poor; however, during these times the absolute value of T_g is much smaller than near the peak of the diurnal curve. The lower, later nighttime values are represented only by the constant DC term.

A check of the cosine diurnal fit, as compared with the actual monthly mean values, for a few months representative data from eight different TEC observing stations, revealed that the cosine algorithm was within approximately 10 percent of the actual 24-hr rms error.

As a further computation efficiency, since the cosine diurnal dependence is used only for the positive half of the waveform, the first two terms of the cosine expansion are used instead of the actual cosine function itself. Thus, the algorithm becomes:

$$T_g = DC + A\left[1 - \frac{x^2}{2} + \frac{x^4}{24}\right], \qquad (2)$$

where

$$x = \frac{(t - \phi)2\pi}{P},$$

and t is local time at the ionospheric point.

The error between the expansion and the actual cosine reaches a maximum of 0.02A at the points where it just intersects the DC value. Since the cosine fit is too low at these points, the series approximation to the cosine is actually slightly better than the cosine.

Other forms of diurnal algorithms, such as a Gaussian suggested by Antoniadis and daRosa (private communication) and a statistical F distribution have been considered. The 2-term expansion of the cosine was chosen as a reasonable compromise of accuracy versus algorithm difficulty.

6. GEOMETRY CALCULATION REQUIREMENTS

In determining the TEC, ionospheric time delay, from a user location to a satellite, several geometrical calculations must be made. The TEC must be found at the geographic point where the ray path intersects the mean ionospheric height, rather than at the user location. This point is taken here at a mean height of 350 kilometers. For a satellite at the lowest elevation considered here, 5 degrees, this point is approximately 14 degrees of earth angle away from the user location. The obliquity, or slant factor, also must be calculated for the mean ionospheric location. Finally, since the TEC is best ordered in geomagnetic, rather than geodetic coordinates, the assumption is made that the TEC is only a function of geomagnetic latitude and local time. Any two points on the earth having the same geomagnetic latitude are assumed to have the same TEC at the same local time. Thus, a conversion to geomagnetic latitude is required, and a calculation is made of local time at the geodetic longitude of the sub-ionospheric location. If the exact form of all the necessary geometry calculations is used, the computer time used would be greater than that required for the TEC algorithm itself. Therefore, in the derivation of these geometrical functions, simplifying assumptions were made in many cases to reduce the calculation complexity. A discussion of each of the steps in the geometry and the simplifying assumptions follows.

6.1 Earth Angle

The ionospheric time delay is computed at the point where the ray from the satellite intersects the ionosphere, rather than at the observer. For an observer located near the earth's surface looking at a satellite at 5 degrees elevation, the ionospheric intersection location is approximately 14 degrees of earth-centered angle from the observer. If the satellite is located along the observer's meridian at 5 degrees elevation angle, the 14 degree latitude difference can represent a significant gradient, particularly near the equatorial anomaly. If the 14 degree earth angle is all in longitude, at 40 degrees latitude, this represents a time difference of 1.2 hours, during which time the ionospheric time delay could have changed considerably.

The exact earth angle representation is:

$$A = 90 - el - Z, \qquad (3)$$

where

$$Z = \sin^{-1}(.948 \cos el),$$

for a mean ionospheric height of 350 km. A useful approximation is (R.S. Allen, private communication, 1975):

$$A = \frac{445}{el + 20} - 4. \qquad (4)$$

In Figure 2 both the actual and the approximate earth angle calculations versus satellite elevation angle are given. This approximation is less than 0.2 degrees in error for all elevations greater than 10 degrees. It is off by only 0.4 degrees and 0.3 degrees at elevations of 5 and 0 degrees, respectively.

6.2 Ionospheric Geodetic Latitude and Longitude

Given the station coordinates, elevation angle, and earth angle to the ionospheric point, its coordinates may be found by the following:

$$\phi_I = \sin^{-1} \{ \sin \phi_o \cos A + \cos \phi_o \sin A \cos az \} , \quad (5)$$

$$\lambda_I = \lambda_o + \sin^{-1} \{ \frac{\sin A \sin az}{\cos \phi_I} \} , \quad (6)$$

where ϕ_o is the station latitude, λ_o is the station longitude. If a flat earth approximation is used the following simplification can be made to the above two equations:

$$\phi_I = \phi_o + A \cos az , \quad (7)$$

$$\lambda_I = \lambda_o + \frac{A \sin az}{\cos \phi_I} . \quad (8)$$

In Figure 3 the geometry of the ionospheric location calculations is illustrated. At latitudes of 40 degrees and equatorward, the maximum error of these approximations is approximately 1 degree. At 60 degrees latitude, errors in determining ionospheric latitude reach a maximum of -3 degrees at 5 degrees elevation angle for a few azimuths. Ionospheric longitude errors reach only 2 degrees at the same location and direction. At high latitudes above 75 degrees an ionospheric latitude equal to the observer location is recommended, regardless of observer's elevation angle. Time delays at these latitudes are generally much lower and more highly variable than in the equatorial and mid-latitude regions. Thus, the time delay errors introduced by neglecting the conversion from user position to mean ionospheric position above 75 degrees will not significantly increase the overall worldwide rms errors of the model.

6.3 Conversion from Geodetic to Geomagnetic Latitude

Ionospheric time delay is, to an approximation, best ordered by geomagnetic, rather than geographic latitude. Thus, it is necessary to perform a transformation from geodetic to geomagnetic latitude. In this representation there is no true longitude dependence of ionospheric time delay, only that which comes about due to the differences between geomagnetic and geodetic latitudes as a function of longitude. The transformation from geodetic to geomagnetic latitude, assuming that the earth's magnetic field can be represented by an earth centered dipole, is given by (Davies, 1966):

$$\sin \Phi = \sin \phi \sin \phi_p + \cos \phi \cos \phi_p \cos (\lambda - \lambda_p) , \quad (9)$$

where

$$\phi_p = 78.3°N ,$$

$$\lambda_p = 291.0°E .$$

The approximation

$$\Phi = \phi + 11.6° \cos (\lambda - 291) \quad (10)$$

represents the exact form to within 1 degree at all geomagnetic latitudes equatorward of ±40 degrees geomagnetic latitude. At geomagnetic latitude up to ±65 degrees, it is within ±2 degrees. Figure 4 is a numeric representation of the error of the approximation to geomagnetic latitude. Over the entire CONUS region the error in the approximate form is less than 1 degree.

6.4 Finding Local Time

Given the universal time and the approximate longitude of the ionospheric point, λ_I, the local time at the ionospheric point is simply:

$$t = \frac{\lambda_I}{15} + UT . \text{ (hours)} \quad (11)$$

If t is greater than 24, make t = t - 24 hours.

6.5 The Obliquity Factor

The vertical time delay at the sub-ionospheric point must be multiplied by an obliquity factor defined as the secant of the zenith angle at the mean ionospheric height. An average ionospheric height of 350 km is assumed. The exact obliquity or slant factor is:

$$SF = \sec \{ \sin^{-1} [.948 \cos el] \} \quad . \qquad (12)$$

An approximation which is within 2 percent of the exact value for elevation angles from 5° to 90° is:

$$SF = 1 + 2 \left[\frac{96 - el}{90}\right]^3 \quad . \qquad (13)$$

The approximation saves 3 trigonometric functions at the expense of finding the cube of a number. Both the exact and the approximate slant factors are shown in Figure 5.

7. ALGORITHM LATITUDE COEFFICIENTS

The four coefficients used in the time delay algorithm are all functions of latitude, season, solar flux, and magnetic activity. A complete representation of all these variables in each of the coefficients would be complex indeed, if even possible. What has been done here is to address only the monthly average latitude dependence of each of these four coefficients for various seasons of a solar maximum year when the highest ionospheric time delays are expected. Possible methods of determining the daily variation of the algorithm coefficients, or of simply including a daily update factor, are discussed in a later section of this report.

7.1 The DC Term

The DC term in the model algorithm represents the nighttime behavior of the ionosphere. Recent evidence from combined Faraday and group-delay measurements from signals transmitted from the geostationary satellite, ATS-6, shows that the Faraday data, which is really representative only of TEC up to approximately 2000 km of vertical height, is considerably in error in the nighttime. The Bent model also was integrated in vertical height only up to 2000 km. The new group-delay data available thus far show a nearly constant additional TEC above the TEC normally measured by the Faraday effect. Thus, with this model representation it is a simple matter to increase the DC term to account for the additional electrons not measurable by the Faraday Effect.

The magnitude of the DC term versus latitude is shown in Figures 6 and 7. Since the DC term is generally much smaller than the amplitude term, A, it may be sufficient to represent its latitude dependence by a simple linear dependence with geomagnetic latitude either side of a reference latitude, or by a single constant for all latitudes. In the final algorithm for the GPS the DC term was represented by a constant 5 nanoseconds for all latitudes and times.

7.2 The Amplitude Term

The amplitude of the cosine fit is the most important term in the algorithm, and is the most highly variable with latitude, season, and solar flux. To closely represent the amplitude term as a function of geomagnetic latitude several coefficients are required. From Figures 6 and 7 the amplitude dependence with geomagnetic latitude can be seen. Note that the CONUS region extends approximately from 35°N to 55°N geomagnetic latitude. It is obvious from the curves in Figures 6 and 7 that no single, fixed shape will match the latitude dependence of the amplitude term. Therefore, it is necessary to use some form of polynomial representation for the latitude dependence of amplitude. Because of the numerous inflection points in the curve for some seasons a polynomial of degree 5 or 6 is necessary to obtain a close fit. However, if the equatorial minimum in the amplitude term, which is particularly obvious during the equinoxes, is smoothed over, a third-degree polynomial can do a fair job of representing the amplitude behavior versus geomagnetic latitude.

7.3 The Phase of the Diurnal Variation

Happily, the phase term in the ionospheric time-delay algorithm is not a highly variable function. A plot of the phase term for a northern midlatitude station, for all the monthly averages of a six-year period covering from solar maximum to well down towards the solar minimum, revealed that the phase of the diurnal maximum rarely changed more than plus and minus one hour from a mean of 14 hours local time. Results from the Bent model, in Figures 6 and 7, and from actual monthly mean TEC data from several stations, show that the phase lies between 12.5 and 15.5 for most of the latitude range covering the mid and equatorial latitudes. Therefore a constant value of 14 hours local time was taken for the phase of the diurnal variation. If a particular sub-model is required for a limited geographic region, perhaps some improvement can be made by making the phase a function of time of the year; but, in this attempt at a worldwide representation, a constant value for all latitudes and seasons is deemed adequate.

7.4 The Diurnal Period

The period of the diurnal cosine representation is, of course, not fixed at 24 hours. In fact it is nearly always significantly higher than 24 hours and is a function of geomagnetic latitude and season as shown in Figures 6 and 7. As is the case with the amplitude term, a third-degree polynomial with geomagnetic latitude was used as the function to represent the latitude dependence of the cosine period term. Again, the

equatorial anomaly region will not be perfectly represented by a third-degree polynomial; but, the fit should be adequate.

8. OPERATIONAL MODEL USE

The eight coefficients required for the ionospheric time delay algorithm will be transmitted down the satellite to user link, along with the necessary orbital elements of the satellite. Each coefficient will have approximately one percent digitization accuracy. The geometric calculations and necessary software to utilize the ionospheric time delay coefficients will be an integral portion of the complete user software package.

The coefficients are to be uploaded into each satellite once each day by the satellite master control facility. Initially, coefficients will represent ten day mean conditions for a relatively small range of solar flux conditions made up well in advance from existing model representations of ionospheric time delay. After the system becomes operational the capability will exist within the system to generate its own ionospheric coefficients by using measured values of time delay from two frequency receivers located at the various satellite monitoring stations. Because of the relatively small number of coefficients required to represent the ionospheric time delay it is also possible to load the entire initial coefficient set for all seasons and solar fluxes on a Read-Only-Memory specifically made for this purpose. A 256X8 chip could conveniently hold all the necessary coefficients.

9. ACCURACY OF THE ALGORITHM

The algorithm of ionospheric time delay was tested against actual values of ionospheric TEC available for 490 station months for a total of 18 stations covering a representative portion of the globe and both solar maximum and minimum conditions. Figure 8 shows the distribution of stations and times from which data were used for model tests. Notice that a significant portion of the data is from regions other than the mid-latitudes and that many parts of the world are represented in the available data.

Summary statistics of ionospheric time delay for each station for all available hours and seasons were computed in units of rms error in nanoseconds at 1.6 GHz. Also, the residual rms error for each station after correction by use of the time delay algorithm was computed. These summary results are shown in Figure 9 with the dashed line indicating the uncorrected ionospheric time delay and the solid line representing the residual time delay after correction. Two stations are shown in more than one column in Figure 9 because the actual ionospheric TEC data available from the station came from more than one satellite and are from different time periods.

The correction of the algorithm for the entire 490 station months gave an overall rms reduction of the ionospheric time delay of 62 percent. This is when the model was used against monthly mean TEC data. When the 62 percent correction is combined with the approximate 25 percent rms deviation of the actual data from the monthly mean the overall improvement is approximately 55 percent. Thus, the goal of a 50 percent algorithm correction has been met.

10. CONCLUSIONS

Military systems ranging accuracy requirements will become greater with improved satellite detection radars and with the implementation of the NAVSTAR/Global Positioning System. As these accuracy requirements become greater the ionosphere will become a greater potential system error source unless the time delay effects of the ionosphere are accounted for by means of appropriate models or real time measurements with a two-frequency operational system. Also, different ionospheric correction schemes will have to be developed for each type of system to maximize the efficiency of the correction process. Realistic tradeoffs must be made between ionospheric correction accuracy and computer resources devoted to implementing the ionospheric correction. Consideration must be given to real time ionospheric monitoring for those single frequency systems which require greater than 70 to 80 percent ionospheric correction. Corrections greater than approximately 90 percent rms will probably be beyond the capability of real time correction unless it is incorporated as part of the system by using a widely spaced two frequency scheme.

An example of an ionospheric correction model algorithm designed for a specific system, the NAVSTAR/Global Positioning System, when used in a single frequency mode, has been described in detail.

11. REFERENCES

Davies, K.,(1966), Ionospheric Radio Propagation, Dover Publications, New York.

Llewellyn, S.K. and R.G. Bent, (1973), Documentaion and Description of the Bent Ionospheric Model, AFCRL-TR-0657 and SAMSO-TR-73-252.

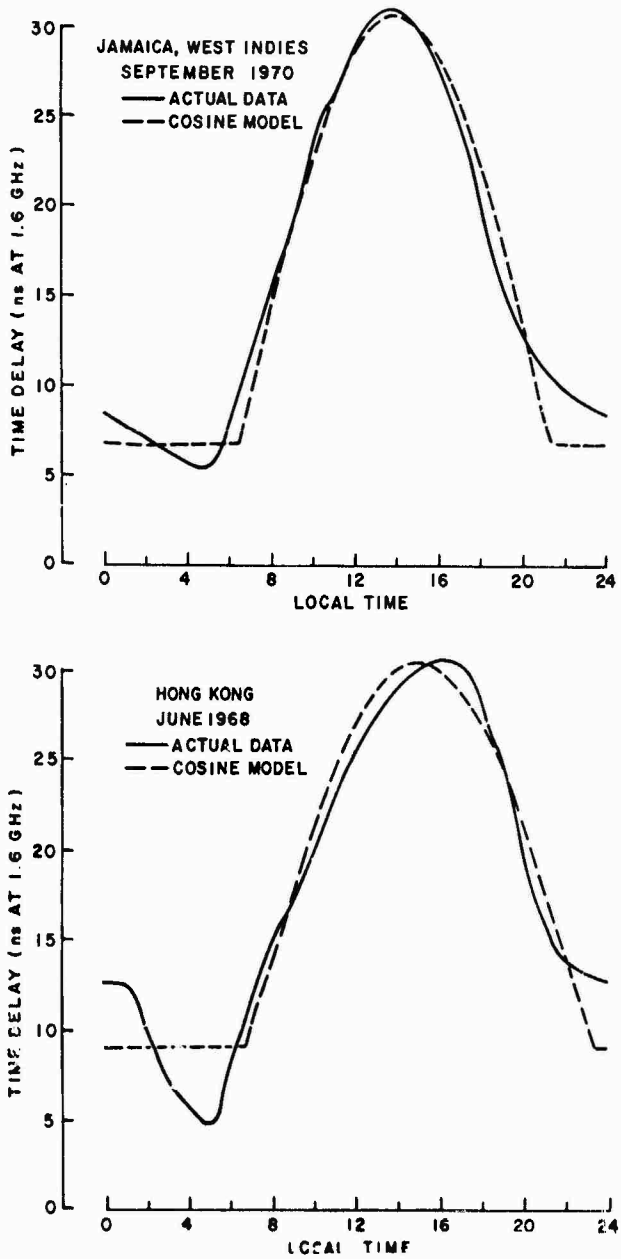


Figure 1. Two examples of actual monthly average time delay data along with cosine model fit to data.

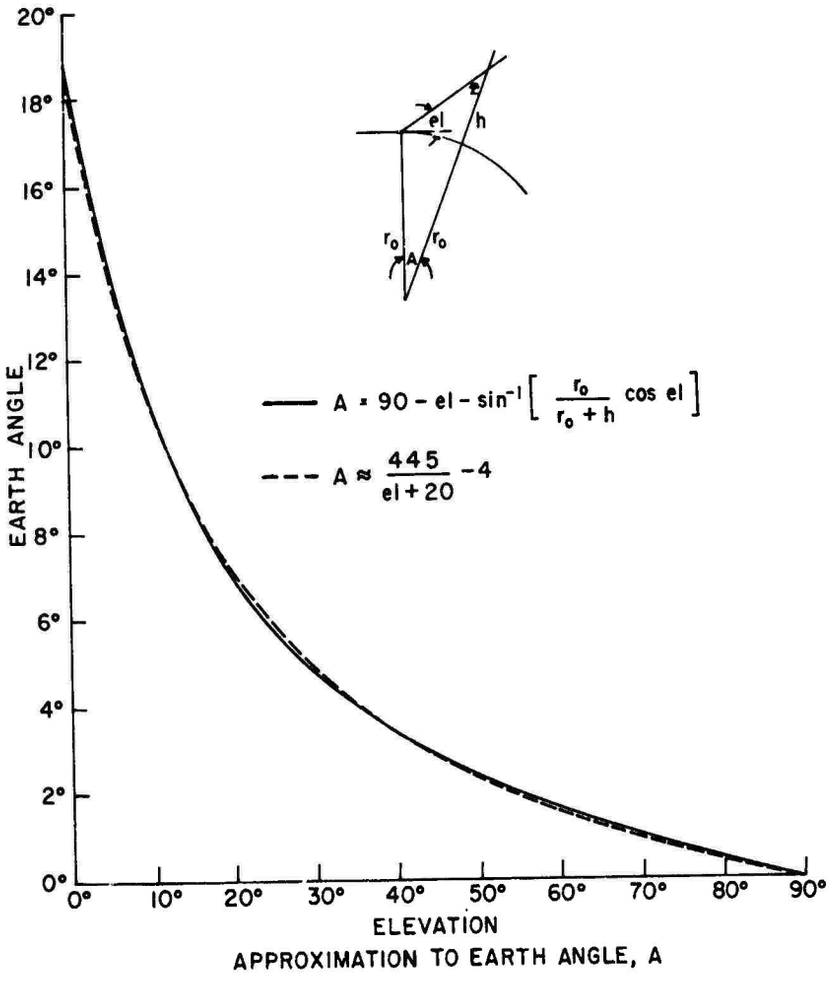
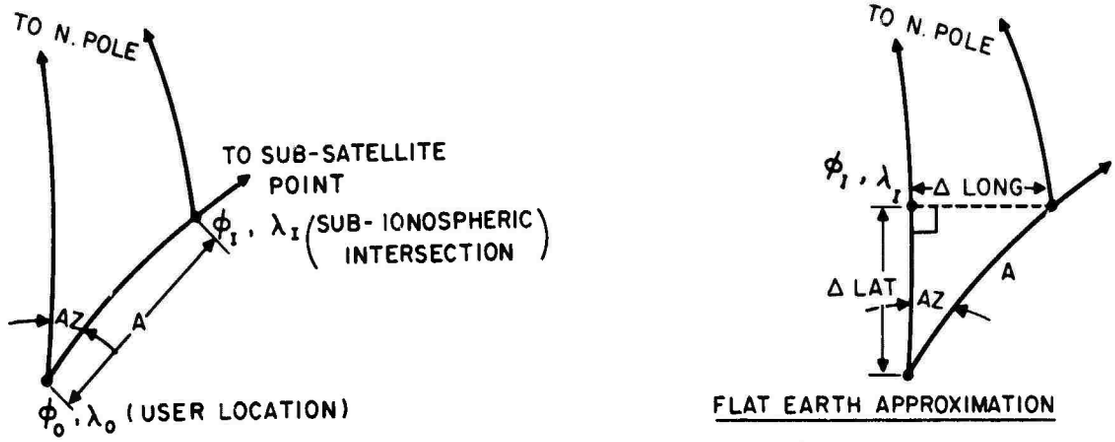


Figure 2. Approximation to Earth Angle, A.



APPROXIMATIONS TO FINDING SUB-IONOSPHERIC COORDINATES

Figure 3. Approximations to geometrical calculations for finding sub-ionospheric coordinates.

GEOCENTRIC DIPOLE MINUS APPROXIMATION
GEOM LAT = GEOG LAT + 11.6(COS(GEOG LONG - 291))
VALUES LESS THAN ONE DEGREE NOT SHOWN

GEOG LONG (EAST)

30	45	60	75	90	105	120	135	150	165	180	195	210	225	240	255	270	285	300	315	330	345	
-4	-3	-2	-1					-1	-2	-3	-4	-4	-4	-4	-3	-2			-2	-3	-4	75
-3	-2	-1							-2	-2	-3	-3	-3	-3	-2				-1	-2	-3	70
-2	-2	-1							-1	-2	-2	-3	-2	-2	-1					-1	-2	65
-2	-2	-1							-1	-2	-2	-2	-2	-2						-1	-2	60
-2	-1									-1	-2	-2	-2	-1							-1	55
-1	-1									-1	-1	-1	-1								-1	50
-1											-1	-1	-1									45
																						40
																						35
																						30
																						25
																						20
																						15
																						10
																						5
																						0
																						-5
																						-10
																						-15
																						-20
																						-25
																						-30
																						-35
																						-40
1	1									1	1	1										-45
1	1								1	1	1	1	1									-50
2	2	1							1	2	2	2	1									-55
2	2	2						1	2	2	2	2	2	1							1	-60
3	2	2	1					1	2	2	3	2	2	1							1	-65
3	3	2	2				1	2	3	3	3	3	2	1							2	-70
4	4	4	3	2			2	3	4	4	4	4	3	2	1					1	2	-75
30	45	60	75	90	105	120	135	150	165	180	195	210	225	240	255	270	285	300	315	330	345	

Figure 4. A plot of the error in the approximation of the geographic to geomagnetic coordinate conversion. Errors less than one degree are left blank.

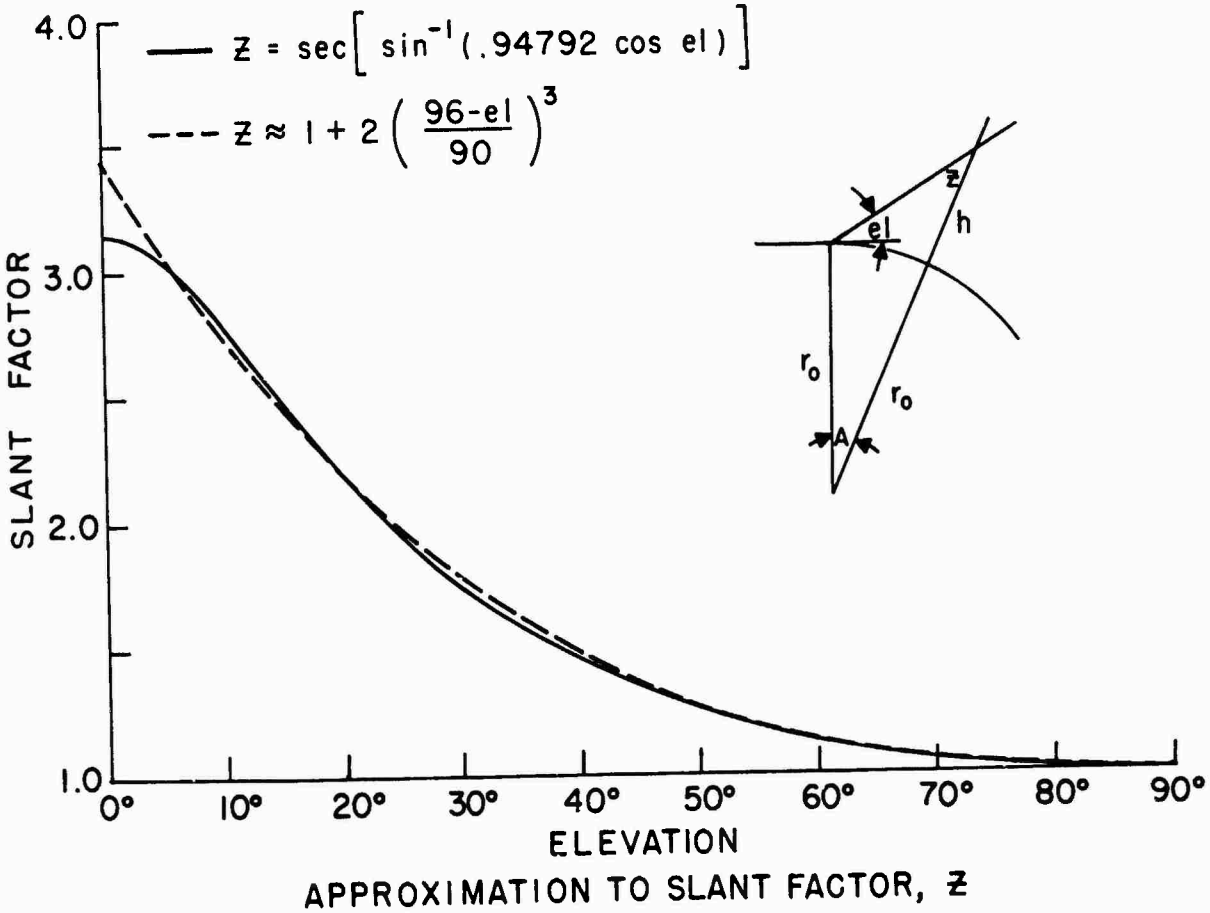


Figure 5. Approximation to slant factor, SF.

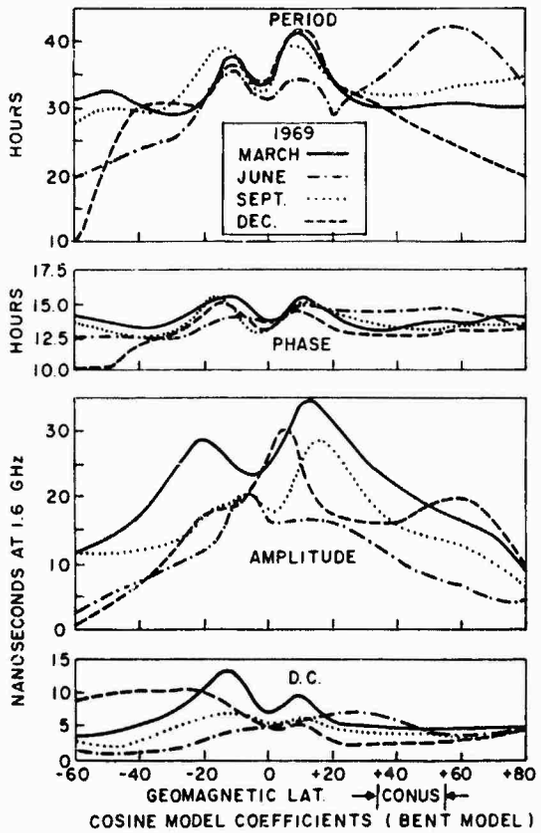


Figure 6. Geomagnetic latitudinal variation of the cosine model coefficients for March 1969, a solar maximum year.

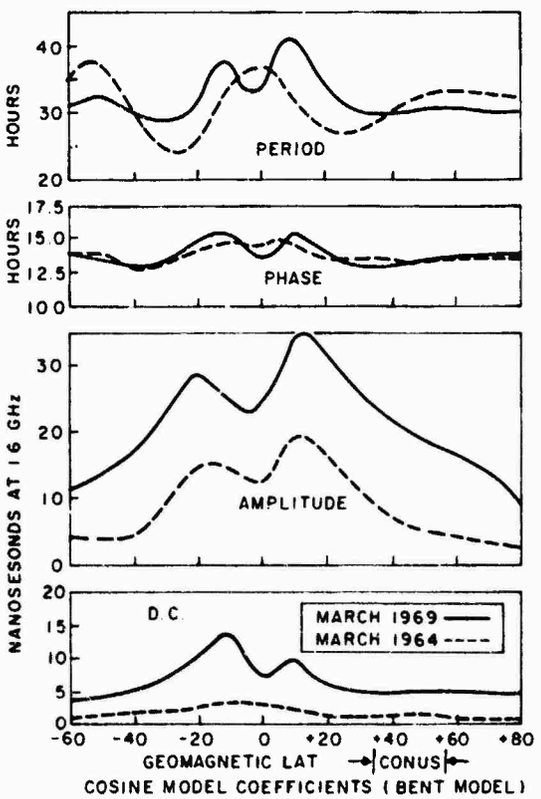


Figure 7. A comparison of the geomagnetic latitudinal variation of the cosine model coefficients for March of the solar minimum year 1964 with the solar maximum year of 1969.

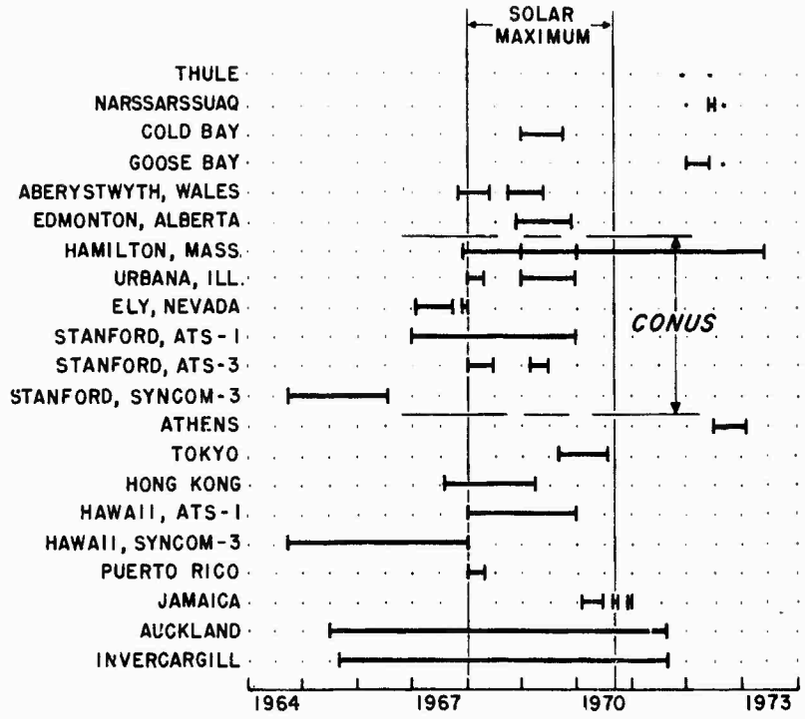


Figure 8. Distribution of times and locations when monthly average TEC data were available for algorithm testing.

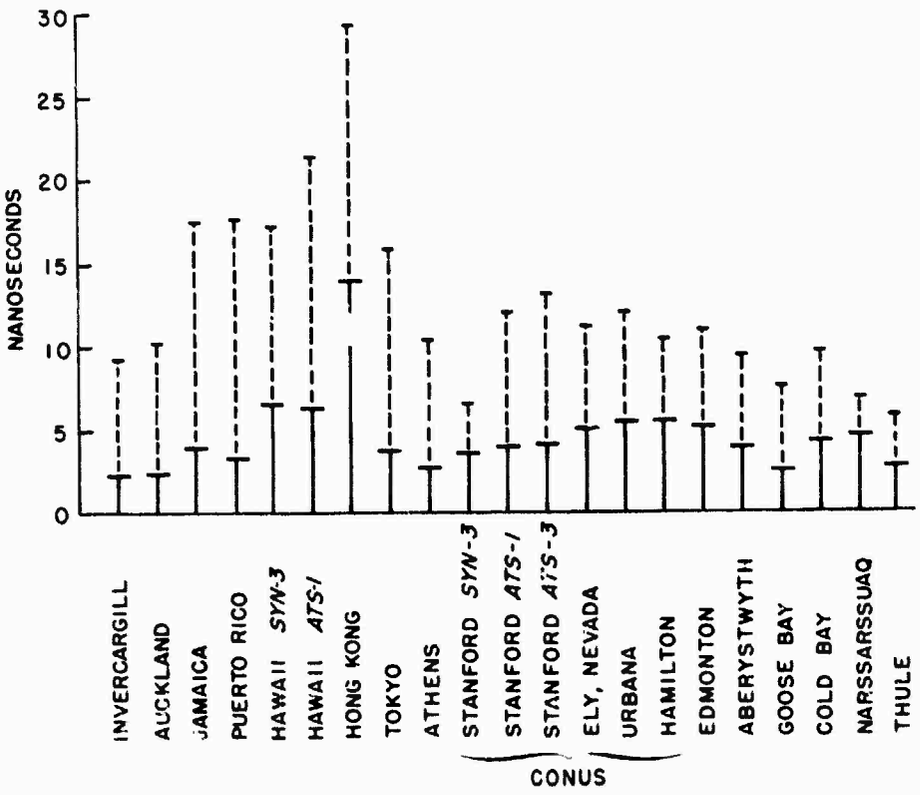


Figure 9. Results of ionospheric time delay algorithm by station. Dotted line indicates overall r.m.s. error for each station with no ionospheric correction. Solid line indicates remaining ionospheric error after application of algorithm.

DISCUSSION

G.H.Millman, General Electric Company, Syracuse, New York 13201, USA

You mentioned that range correction can be accomplished to an accuracy of 50%. Have you examined techniques to attain a higher accuracy?

Author's Reply

Yes. Certainly more complex techniques giving greater accuracy exist; but, in the tradeoff between number of coefficients versus accuracy we believe we are nearly optimum.

H.Soicher, US Army Electronics Command, Fort Monmouth, NJ 07703, USA

Time delay errors at Station B may be updated by continuous observation at Station A with measurement capabilities. Have any attempts been made to determine the geographic extent of separation of A and B for which a good correlation exists?

Author's Reply

This correlation distance is a function of season of the year; but, generally it is of the order of 1000 miles. This is for the average of several hours of values. Instantaneous values do not correlate well, even at much shorter distances as the scale size of traveling ionospheric disturbances is of the order of 100 kilometers wavelength and they limit the short term predictability.

G.H.Millman

The range connection required is dependent upon the particular system in question. For example, for a radar system which requires high range accuracy. it may be necessary to utilize a combination of correction techniques such as incoherent scatter or transit satellite emissions in conjunction with an average prediction model.

Author's Reply

True. One of the reasons we chose only an approximate 50% error reduction was that, due to the impossibility of updating except once per day, real time measurements do little good for this system. For systems which require more accurate models, they are available, but with large computer resource requirements, which give from 70 to 80% ionospheric correction. Measurements at the actual location of the required correction can improve upon this, providing they are not more than 1 or 2 hours old.

TRANSIT SATELLITE OBSERVATIONS OF IONOSPHERIC IRREGULARITIES

George H. Millman
General Electric Company
Syracuse, New York 13201, USA

and

Roy E. Anderson
General Electric Company
Schenectady, New York 12301, USA

SUMMARY

The characteristics of the phase fluctuations imposed by ionospheric irregularities on the 150- and 400-MHz coherent frequencies emanating from the polar-orbiting U.S. Navy Navigation Satellites (TRANSIT) are evaluated in terms of the phase coherence of ionospheric propagation. The observations were conducted at the General Electric Radio-Optical Observatory (geographic coordinates: 42.85°N, 74.07°W) during the three day period immediately following the severe magnetic disturbance of April 21-22, 1970.

In addition to the differential phase measurements, the two signal amplitudes were also recorded throughout the orbital passes.

From the second derivative with respect to time of the differential phase fluctuations, an estimate is made of the nonlinear phase shift introduced by the ionosphere.

For an assumed altitude of the irregularities, estimates are made of the overall spatial extent of the inhomogeneities producing the phase and amplitude fluctuations.

By correlating the Radio-Optical Observatory data with the differential phase and 150-MHz amplitude measurements recorded simultaneously at the Lincoln Laboratory - Millstone Hill facility (geographic coordinates: 42.62°N, 71.36°W) during the April 23-25, 1970 period, the two-dimensional spatial configuration of the irregularities responsible for the perturbations are deduced.

From the differential phase data, the characteristics of the equatorward edge of the midlatitude trough are evaluated.

A correlation is made of the equatorward edge of the trough with the equatorward edge of the phase and amplitude scintillation regions.

1. INTRODUCTION

The detection and tracking of satellites at great distances by ground based radars can be accomplished utilizing either extremely high-power transmitters or the coherent integration, i.e., predetction integration, of many pulse returns. To achieve coherent integration which is used to enhance the reflected signal-to-system noise ratio, phase coherence is required from pulse to pulse. The sources of phase instability are the radar equipment, the target configuration and the propagation medium. The phase instability of the ionosphere which is the only source discussed in this paper is brought about by irregularities of electron density.

Ionospheric irregularities can also influence the limits of resolution and accuracy of systems that depend on the transmissions of radio frequency signals between satellites and the earth's surface. The effects must be considered for systems, such as satellite-borne-synthetic aperture radars and navigation satellite systems, operating at frequencies up to approximately 3000 MHz.

In the case of a synthetic aperture radar, phase coherence is required to attain the desired resolution. The term, synthetic aperture, refers to the large effective antenna aperture achieved by a vehicle-borne side-looking radar that records data in a coherent mode as the vehicle moves. Coherent detection and processing of many sequential pulse transmissions greatly improves the azimuth resolution of the radar antenna over the resolution that would be acquired by the same physical structure if the data were not coherently detected and processed.

In this paper, the characteristics of the phase and amplitude fluctuations imposed by ionospheric irregularities on the 150- and 400-MHz coherent frequencies emitted from the polar-orbiting U.S. Navy Navigation satellites (TRANSIT) are evaluated in terms of the phase coherence of ionospheric propagation.

2. EXPERIMENTAL CONSIDERATIONS

The passive monitoring of the 150- and 400-MHz transmissions from the TRANSIT satellites was conducted at the General Electric Radio-Optical Observatory (geographic coordinates: 42.85°N latitude, 74.07°W longitude; corrected geomagnetic latitude: 56.41°N) during the three day period following the severe magnetic disturbance of April 21-22, 1970.

A total of 19 satellite passes were observed in the time interval between 1300 hours GMT, April 23, 1970, and 0600 hours GMT, April 25, 1970. The TRANSIT satellites are in a near circular, polar orbit at approximately 1025-1140 km altitude with periods ranging from 106.4 to 106.9 minutes (NASA, 1974).

In addition to the differential phase between the 150- and 400-MHz transmissions, the two signal amplitudes were recorded throughout the orbital passes.

Figure 1 is a portion of a typical analog chart recording of TRANSIT signals displaying a smooth monotonically-varying phase difference which is suggestive of quiet ionospheric conditions. The severe amplitude and phase perturbations, also shown in Figure 1, are indicative of the existence of ionospheric irregularities along the propagation path. An interesting feature of the data in the vicinity of 2217 GMT is the disclosure of 150-MHz amplitude fluctuations in the absence of phase fluctuations.

3. THEORY OF ANALYSIS

The phase shift, ϕ, imposed by the ionosphere on radio waves traversing the medium is given by

$$\phi = \frac{e^2}{m_e c f} N_s \tag{1}$$

where ϕ is in radians, e is the electron charge (4.8×10^{-10} esu), m_e is the electron mass (9.1×10^{-28} gm) c is the free space velocity in cm/sec, f is the transmission frequency in Hz, and N_s is the integrated electron density along the ray path, i.e., electron content, in electrons/cm^2.

In the derivation of this expression, the higher order frequency terms involving both the magnetic field parameters and second (and higher) order electron density terms have been neglected.

The phase shift function at a particular time, t, can be approximated by a Maclaurin's series as

$$\phi = \frac{K}{f} \left[N_s(0) + \frac{dN_s(0)}{dt} t + \frac{d^2N_s(0)}{dt^2} \frac{t^2}{2} + \cdots \right] \tag{2}$$

where $K = (e^2/m_e c) = 8.4396 \times 10^{-3}$.

The first term in this expression is the phase bias and is identical to Equation (1). The second term is the Doppler frequency shift while the third term defines the phase nonlinearity introduced by the ionosphere.

It is noted that the nonlinear phase perturbations in the ionosphere are the factors which can impose limitations on the effectiveness of coherent integration.

The differential phase between two harmonically-related coherent signals transmitted from a satellite and detected on the ground can be expressed by (Millman, 1974)

$$\Delta \phi = - \frac{K}{f_1} \left(\frac{b^2 - a^2}{a} \right) N_s \tag{3}$$

where a and b are constants related to the transmitted frequencies, f_1 and f_2, according to

$$a f_1 = b f_2 \tag{4}$$

The differential phase between the two frequencies is obtained from

$$\Delta \phi = a \phi_1 - b \phi_2 \tag{5}$$

where ϕ_1 and ϕ_2 are the total phase shifts encountered in the troposphere and the ionosphere by the radio waves traversing a space-earth propagation path.

The derivation of the phase difference relationship, Equation (3), is based on the supposition that the two transmitted signals travel along the same ray path. In other words, the ionospheric refraction effects are assumed to be negligible. This assumption is justified, for the most part, for frequencies in the VHF and UHF range except perhaps for propagation at the very low elevation angles (Millman, 1967).

The absolute phase of a satellite transmission received on the ground cannot be measured directly. Thus, the relative electron content along the paths from the satellite to a ground terminal can only be inferred from the differential phase measurements.

According to Equations (2) and (3), the magnitude of the nonlinear phase shift, ϕ_{nl}, can be defined in terms of the function

$$\phi_{nl} = \frac{t^2 f_1}{2f} \left(\frac{a}{b^2 - a^2} \right) \frac{d^2(\Delta \phi)}{dt^2} \tag{6}$$

For the TRANSIT satellite-differential phase data recorded in this experiment, the constants in Equation (6) are given by a = 1, b = 3/8 and f_1 = 150 MHz.

It follows that, for a coherent integration time, t, of 1.0 sec and $(d^2(\Delta \phi)/dt^2) = 1$ radian/sec^2, the nonlinear phase shift at 450 MHz evaluates to 11.1°.

4. DISCUSSION

From the second derivative with respect to time of the differential phase between the 150- and 400-MHz signals, an estimate can be made of the nonlinear phase shift introduced by the ionosphere on radar transmissions traversing the medium.

The differential phase data that were analyzed are presented in Figures 2 and 3. The latter is an example of the severe phase distortion encountered during a major portion of the orbit. Figure 4 is a plot in expanded form of 3 minutes of the data shown in Figure 3.

The phase fluctuations about the least square fit which is also indicated in Figures 2 through 4 were used in the calculation of the second derivative with respect to time.

Figures 5 and 6 are histograms of the second time derivative of the differential phase corresponding to the data presented in Figures 2 and 4, respectively. As a first approximation, the data appear to be normally distributed about zero having a maximum on the order of 12 rad/sec^2 and 25 rad/sec^2, respectively.

When replotting the histogram data in the form of a cumulative distribution function, shown in Figure 7, it becomes apparent that the distributions are quite different.

The statistical parameters of the two distributions are presented in Tables 1 and 2. For the April 24th phase perturbation record, the median and upper decile time-varying differential phase shift evaluate to 1.05 and 4.50 rad/sec^2, respectively, with the corresponding electron content variation becoming 2.172×10^{10} and 9.307×10^{10} electrons/cm^2/sec^2. According to Equation (6), for a 1 sec time integration, this implies a nonlinear phase shift at 450 MHz of 11.7° and 50.0°, respectively.

TABLE 1

STATISTICAL CHARACTERISTIC OF THE TIME-VARYING DIFFERENTIAL PHASE SHIFT OF TRANSIT SIGNALS RECORDED ON APRIL 24, 1970, 2233-2249 HOURS GMT

Statistical Parameter	$d^2(\Delta\phi)/dt^2$ (rad/sec^2)	d^2N_s/dt^2 (electrons/cm^2/sec^2)	Phase Shift at 450 MHz* (deg)
Upper Decile	4.50	9.307×10^{10}	50.0
Upper Quartile	2.60	5.377×10^{10}	28.9
Median	1.05	2.172×10^{10}	11.7
Lower Quartile	0.32	6.618×10^{9}	3.6
Lower Decile	0.10	2.068×10^{9}	1.1

* 1-second time integration

TABLE 2

STATISTICAL CHARACTERISTICS OF THE TIME-VARYING DIFFERENTIAL PHASE SHIFT OF TRANSIT SIGNALS RECORDED ON APRIL 25, 1970, 0543-0600 HOURS GMT

Statistical Parameter	$d^2(\Delta\phi)/dt^2$ (rad/sec^2)	d^2N_s/dt^2 (electrons/cm^2/sec^2)	Phase Shift at 450 MHz* (deg)
Upper Decile	7.33	1.516×10^{11}	81.5
Upper Quartile	4.74	9.803×10^{10}	52.7
Median	2.65	5.481×10^{10}	29.4
Lower Quartile	0.74	1.530×10^{10}	8.2
Lower Decile	0.18	3.723×10^{9}	2.0

* 1-second time integration

With respect to the April 25th data, the medium and upper decile values of $d^2(\Delta\phi)/dt^2$ are 2.65 and 7.33 rad/sec^2, respectively, which convert to an electron content - time variation of 5.481×10^{10} and 1.516×10^{11} electrons/cm^2/sec^2. Thus, the nonlinear phase shift at 450 MHz evaluates to 29.4° and 81.5°, respectively.

It is of interest to note that, in terms of local mean time at the subionospheric geographic coordinates of the 350-km penetration point of the ray, the nonlinear phase shift appears to be greatest near local midnight than during the daylight hours. Further analysis will have to be performed to determine whether this is a relevant characteristic of the ionospheric irregularities causing phase fluctuations.

Since both the 150- and 400-MHz amplitudes and the differential phase were recorded continuously as the satellites traversed a horizon-to-horizon trajectory, it was possible to determine whether any spatial correlation existed between the ionospheric irregularities producing the amplitude and phase fluctuations.

In all 19 satellite passes that were observed, both amplitude and phase fluctuations were always detected. The over-all extent of the phase and amplitude scintillation regions are plotted in Figures 8 and 9, respectively, in terms of the corrected geomagnetic latitude and local mean time. The time and latitude coordinates apply to the sub-ionospheric point of the ray penetration at the 350-km altitude level.

The corrected geomagnetic latitudes used in this analysis are those derived by Gustafsson (1970) and are a modification of the coordinate system developed by Hakura (1965).

According to the experimental observations of radio star noise by Hewish (1952) fluctuations in intensity are always accompanied by the fluctuations in the phase difference between the antennas of an interferometer. However, it is now generally known that the converse does not apply.

It is seen that, for the most part, the over-all spatial extent of the 150-MHz amplitude fluctuations correlated with those of the phase fluctuations. However, there were many occasions in which no correlations existed; that is, the amplitude fluctuations were not accompanied by fluctuations in phase.

A possible explanation for the discrepancy in the data presented in Figures 8 and 9 is that the presence of both the phase and amplitude irregularities was estimated visually from analog chart recordings similar to the one depicted in Figure 1. In the case of the differential phase presentation, the 6π radians phase difference (measured at 150 MHz) extended over 42 mm of the chart paper. It is estimated that phase variations appearing within 0.5 mm can be readily observed. This corresponds to a phase measurement resolution of approximately 13°. Thus, the lack of correlation between the phase and amplitude fluctuations is applicable for phase variations less than about 13°.

Figures 8 and 9 reveal that, in general, both the phase and VHF amplitude fluctuations occurred north of the Observatory which is situated at a corrected geomagnetic latitude (CGL) of 56.41°N. It should be mentioned that, during the data collection period, the geomagnetic activity index, k_p, remained in the 3 to 4- interval.

For an assumed irregularity altitude of 350 km, the equatorward edges of the phase irregularities during the daytime were located at corrected geomagnetic latitudes ranging from approximately 57°N to 70°N. In the case of the VHF amplitude irregularities, the equatorward edges were between 58°N and 67°N CGL.

During the nighttime hours, however, for which there is a sparsity of data, the lower boundary extended from 52°N to 55°N CGL for the phase irregularities and from 52°N to 58°N CGL for the VHF amplitude irregularities.

These results which are applicable to moderate magnetic activity conditions, compare favorably with Aarons' (1973) model of F-layer high-latitude irregularities as deduced from amplitude fluctuations of radio signals from satellites and radio stars.

The dimensions of the ionospheric irregularities in the direction of the satellite orbit can be estimated when their height distribution is known.

Table 3 lists the statistical characteristics of the apparent over-all extent of the scintillation regions based on the assumption that the irregularities were located at an altitude of 350 km. As indicated in Figures 8 and 9, the total number of phase and amplitude irregularities available for this analysis was 28 and 21, respectively.

TABLE 3

STATISTICAL CHARACTERISTICS OF THE APPARENT OVER-ALL EXTENT OF IONOSPHERIC PHASE AND AMPLITUDE SCINTILLATION REGIONS

Statistical Parameter	Irregularity Extent (km)	
	Phase Scintillation	Amplitude Scintillation
Upper Decile	1737	1749
Upper Quartile	1270	1670
Median	717	1475
Lower Quartile	397	1126
Lower Decile	179	300
Mean	909	1340

It is seen that the median and upper decile values of the phase irregularity dimensions are on the order of 717 and 1737 km, respectively, while, for the amplitude scintillation case, the dimensions increase to 1475 and 1749 km. These results give evidence that the phase producing fluctuation regions are more patchy and smaller than those causing amplitude scintillations.

By correlating the Radio-Optical Observatory data with the differential phase and 150-MHz amplitude measurements recorded simultaneously at the Lincoln Laboratory - Millstone Hill Field Station (geographic coordinates: 42.62°N, 71.36°W) during the April 23-25 period, it was possible to deduce the spatial configuration of the inhomogeneities responsible for the perturbations. The ground distance between the two facilities is approximately 211 km.

The calculations, summarized in Table 4, were based on the assumption that the irregularities are field-aligned and located at 350-km altitude. The orientation of the magnetic field, i.e., declination angle, at that altitude was evaluated by assuming a spherical harmonic model for the earth's magnetic field. The set of 48 spherical harmonic coefficients derived by Jensen and Cain (1962) was employed in the model.

TABLE 4

APPARENT OVER-ALL EXTENT OF IONOSPHERIC IRREGULARITIES PROJECTED ALONG AND ORTHOGONAL TO MAGNETIC FIELD LINES

Date of Observation	Time (GMT)	Irregularity Extent (km)			
		Phase Scintillation		Amplitude Scintillation	
		Along Field Line	Orthogonal to Field Line	Along Field Line	Orthogonal to Field Line
23 April 1970	1346 - 1403	1927	181	2384	191
24 April 1970	1439 - 1455	1962	130	1582	126
24 April 1970	1646 - 1703	1782	142	1890	144
24 April 1970	1835 - 1849	1666	148	1639	142
24 April 1970	2204 - 2221	1685	123	1698	132
25 April 1970	0543 - 0600	2381	261	2557	293

It is seen that the apparent over-all dimensions of the irregularities in the direction of the magnetic field lines, associated with the phase fluctuations, varied between 1666 and 2381 km while the transverse dimensions were on the order of 123 to 261 km.

For the amplitude scintillation case, the over-all dimensions ranged between 1582 and 2557 km along the magnetic north-south direction and between 126 and 293 km orthogonal to the field lines.

In addition to providing a means for the investigation of the characteristics of ionospheric inhomogeneities, the differential-phase technique also furnishes data on the midlatitude trough which is a region of low electron density located in the F-region north of the Observatory.

The effect of the trough on the TRANSIT radio signals is clearly evident in Figures 2 through 4. Instead of smooth, monotonically varying functions, the differential phase could either undergo a rapid change in magnitude or remain somewhat constant as the satellite traverses the trough region.

The trough is considered to be a nighttime phenomenon. However, the daytime trough has been observed (Feinblum and Horan, 1973). The presence of the daytime trough is shown in Figure 10 which is a plot of the equatorward edge of the trough as a function of the local mean time at the 350-km altitude subionospheric point. Of a total of 16 daytime and 3 nighttime satellite passes, the trough was not detected in 5 of the daytime and in 1 of the nighttime observations.

According to Figure 10, for the April 23rd data, the CGL coordinates of the southern boundary of the trough appear to be a function of time of day. At about local noon, the trough is located in the high latitudes, i.e., approximately 72°N CGL, and descends to the south, i.e., 60°N CGL, at 1800 hours local time.

It is of interest to note that, on the following day, April 24th, the daytime trough was not observed. In addition, the southern boundary of the trough did not vary with time and was located at approximately 65°-67°N CGL.

An analysis of Alouette I data has shown that the low electron density trough is found to follow magnetic latitude contour lines (Thomas and Andrews, 1968). According to Rycroft and Thomas (1970), for the summer nights between 2100 and 0700 hours local time, the average invariant magnetic latitude of the center of the trough, Λ, is a function of the geomagnetic planetary index, k_p, and can be represented by the relationship

$$\Lambda = 61.0° - (1.8° \pm 0.4°)\, k_p \qquad (7)$$

The location of the equatorward wall of the trough is described by Gassmann (1973) in terms of the expression

$$\cos^2 \Lambda = \frac{1}{6(1 - 0.1Q)} \tag{8}$$

where Q is the polar magnetic activity index. Assuming that $Q = k_p = 3$, it follows that $\Lambda = 60.8°$.

Since the invariant magnetic latitude and the corrected geomagnetic latitude differ only slightly, Gassmann's model is thus in agreement with the experimental data acquired on the afternoon of April 23rd.

Figure 11 is a plot of the CGL of the equatorward edge of the trough as a function of the CGL of the equatorward edge of the scintillation region, i.e., scintillation boundary. It is seen that the scintillation boundary was always south of the trough wall.

For the phase producing irregularities, the latitude coordinates of the scintillation boundary and the trough wall are not correlated. Similar results were reported by Evans (1973).

However, in the case of the amplitude producing irregularities, when neglecting the data in the vicinity of 53°N CGL which corresponds to the April 25th measurements depicted in Figure 3, a linear relationship is found to exist between the latitudes of the scintillation boundary and the trough equatorward edge. The possible existence of a correlation between these two parameters was hypothesized by Aarons and Allen (1971).

5. CONCLUSIONS

A statistical analysis of a nighttime phase perturbation record revealed a median time varying phase shift of 2.65 rad/sec^2 with the corresponding electron content time variation evaluating to 5.481×10^{10} electrons/cm^2/sec^2. For a one second time integration, this implies a nonlinear phase shift at 450 MHz of 29.4°. A median nonlinear phase shift at 450 MHz of a daytime phase perturbation record was computed to be 11.7°.

It was found that, in general, the spatial extent of the 150-MHz amplitude fluctuations correlated with those of the phase fluctuations. However, there were several occasions in which the amplitude fluctuations were not accompanied by fluctuations in phase. The lack of correlation is possibly attributed to the resolution of the analog chart recordings which only permitted resolving phase variations greater than about 13°. The phase and amplitude fluctuations, for the most part, occurred north of the Observatory.

During the daytime, the equatorward edges of the phase producing irregularities were concentrated between 57°N and 70°N corrected geomagnetic latitude. During the nighttime, the lower boundary ranged between 52°N and 55°N corrected geomagnetic latitude. The calculations were based on assumed irregularity altitude of 350 km.

The median value of the over-all dimensions of the phase irregularities was found to be 717 km as compared to 1475 km for the VHF amplitude irregularities.

Utilizing TRANSIT satellite data recorded simultaneously at Millstone Hill, the two-dimensional configuration of the irregularities were deduced. The apparent over-all extent of the irregularities in the direction of the magnetic field lines was about an order of magnitude greater than that in the transverse direction.

It was determined that the equatorward edge of the scintillation regions was always south of the equatorward edge of the trough and that the corrected geomagnetic latitude of the amplitude scintillation equatorward edge correlated with the latitude of the trough equatorward edge.

6. ACKNOWLEDGMENTS

The assistance of R.W. Garrett (formerly of the General Electric Company, Schenectady, New York) in the performance of the experiment and W.J. Nealon, III (formerly of the General Electric Company, Syracuse, New York) in the reduction and analysis of the data is greatly appreciated.

7. REFERENCES

Aarons, J., "A Descriptive Model of F-layer High-Latitude Irregularities as Shown by Scintillation Observations," Journal of Geophysical Research, Vol. 78, pp 7441-7450, November 1, 1973.

Aarons, J. and R.S. Allen, "Scintillation Boundary during Quiet and Disturbed Magnetic Conditions," Journal of Geophysical Research, Vol. 76, pp 170-177, January 1, 1971.

Evans, J.V., "Millstone Hill Radar Propagation Study: Scientific Results - Part II," Lincoln Laboratory, Report No. ESD-TR-73-260, November 1973.

Feinblum, D.A. and R.J. Horan, "HILION - A Model of the High-Latitude Ionospheric F2 Layer and Statistics of Regular Ionospheric Effects at Ft. Churchill, 1968," Bell Laboratories, October 1973.

Gassmann, G.J., "Analog Model 1972 of the Arctic Ionosphere," Air Force Cambridge Research Laboratories, Air Force Surveys in Geophysics No. 259, AFCRL-TR-73-0151, March 1973.

Gustafsson, G., "A Revised Corrected Geomagnetic Coordinate System," Arkiv för Geofysik, Vol. 5, pp 595-617, 1970.

Hakura, Y., "Tables of Maps of Geomagnetic Coordinates Corrected by the Higher Order Spherical Harmonic Terms," Report Ionosphere Space Research Japan, Vol. 19, pp 121-157, 1965.

Hewish, A., "The Diffraction of Galactic Radio Waves as a Method of Investigating the Irregular Structures of the Ionosphere," Proceedings of the Royal Society, Vol. 124, pp 494-514, August-October 1952.

Jensen, D.C. and J.C. Cain, "An Interim Geomagnetic Field," Journal of Geophysical Research, Vol. 67, pp 3568-3596, August 1962.

Millman, G.H., "An Analysis of Ionospheric Electron Content Measurements Utilizing Satellite-Emitted Signals," General Electric Technical Information Series Report No. 74EMH24, December 1974.

Millman, G.H., "A Survey of Tropospheric, Ionospheric and Extraterrestrial Effects on Radio Propagation Between the Earth and Space Vehicles," in Propagation Factors in Space Communications, AGARD Conference Proceedings No. 3, pp 3-55, Technivision, 1967.

National Aeronautics and Space Administration, "Satellite Situation Report," Goddard Space Flight Center, Vol. 14, No. 1, January 1974.

Rycroft, M.J. and J.O. Thomas, "The Magnetospheric Plasmapause and the Electron Density Trough at the Alouette I Orbit," Planetary Space Science, Vol. 18, pp 65-80, January 1970.

Thomas J.O. and M.K. Andrews, "Transpolar Exospheric Plasma 1, Plasmasphere Termination," Journal of Geophysical Research, Vol. 73, pp 7407-7417, December 1, 1968.

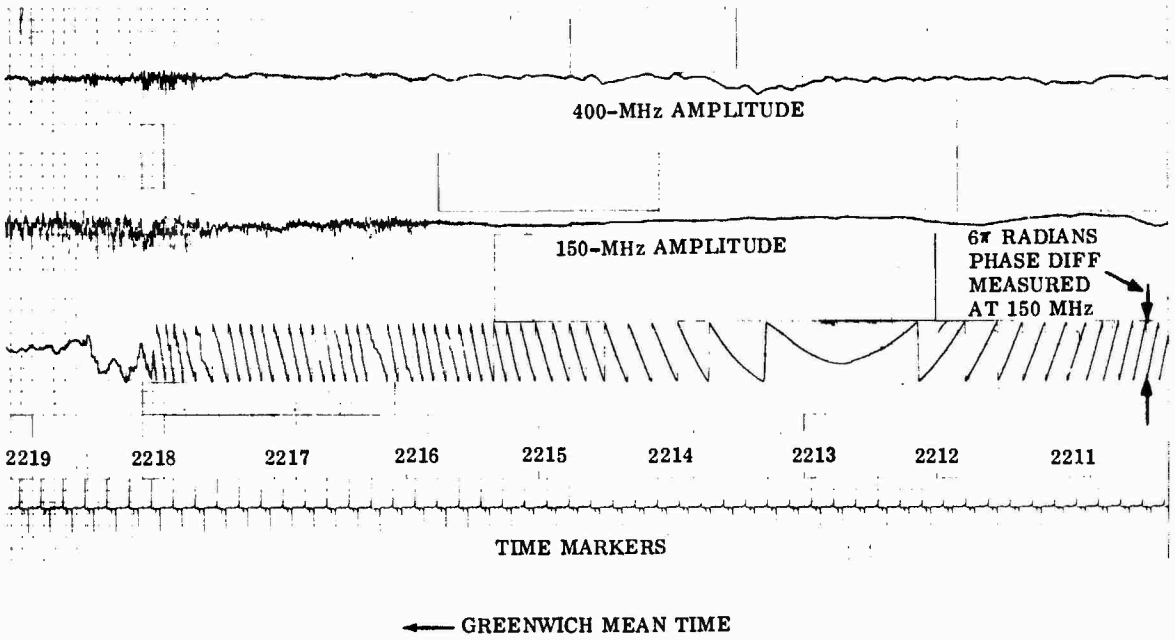


Figure 1. Portion of an Analogue Chart Recording of the Signals From TRANSIT Satellite, Catalogue No. 2965, April 24, 1970, 2212-2219 Hours GMT, Displaying Ionospheric Amplitude and Phase Fluctuation Effects

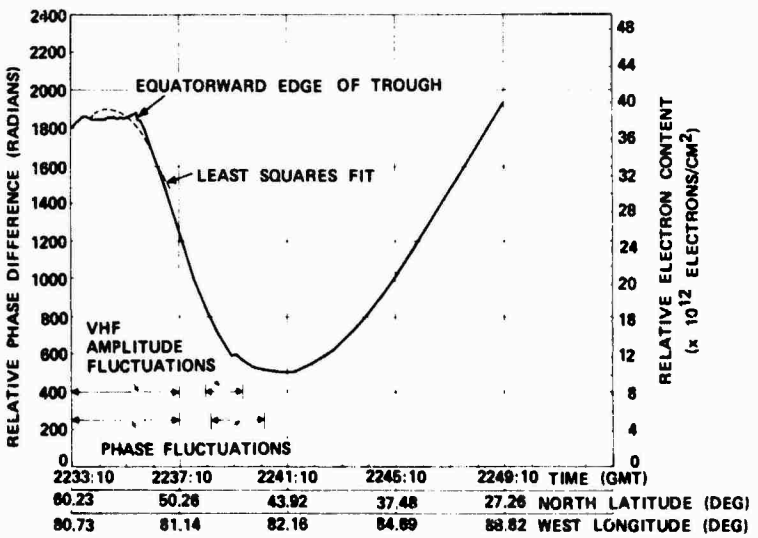


Figure 2. Differential Phase Measurements of the Transmissions From TRANSIT Satellite, Recorded on April 24, 1970, 2233-2249 Hours GMT

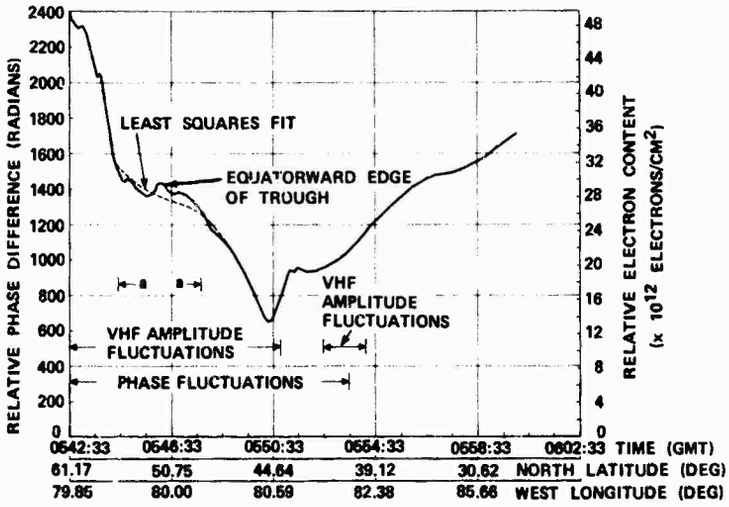


Figure 3. Differential Phase Measurements of the Transmissions From TRANSIT Satellite, Recorded on April 25, 1970, 0542-0600 Hours GMT

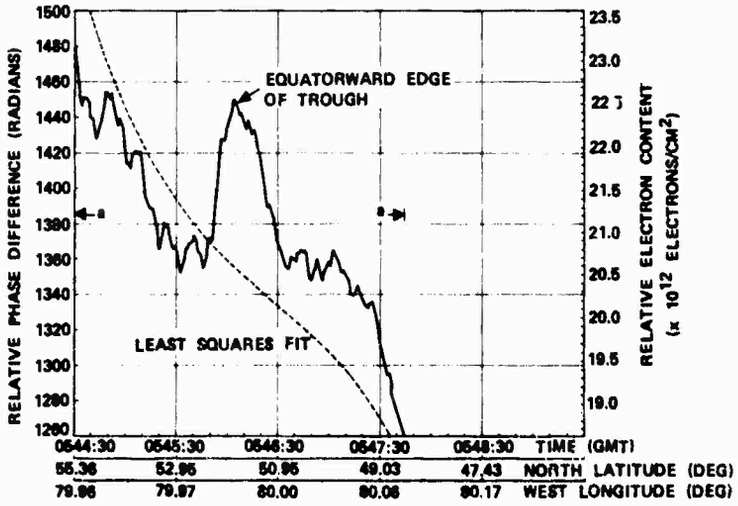


Figure 4. Expanded Portion of Differential Phase Measurement Data From TRANSIT Satellite, Recorded on April 25, 1970, 0544-0547 Hours GMT

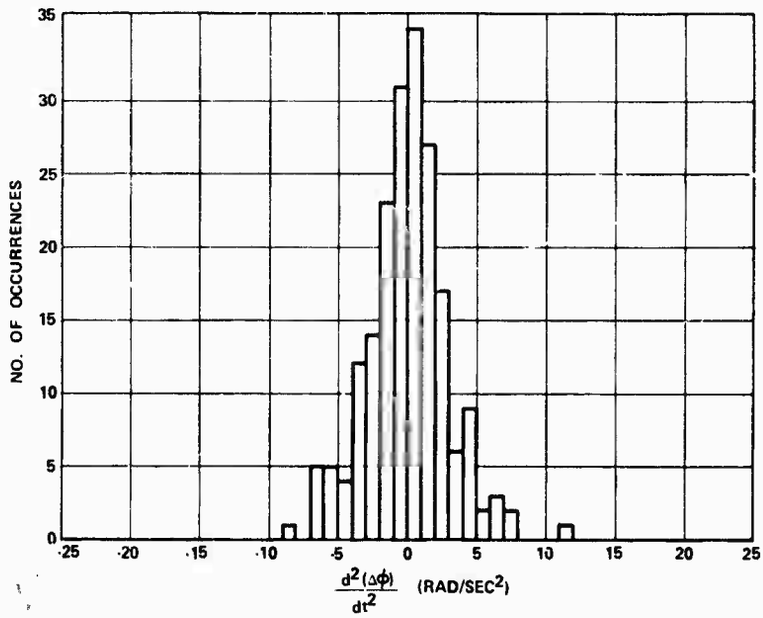


Figure 5. Histogram of the 2nd Derivative with Respect to Time of the Phase Fluctuations, TRANSIT Satellite, Recorded on April 24, 1970, 2233-2236 Hours GMT

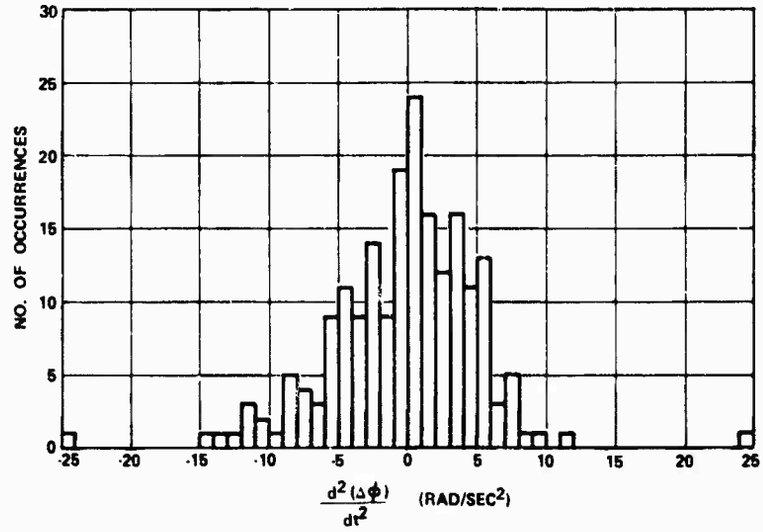


Figure 6. Histogram of the 2nd Derivative with Respect to Time of of the Phase Fluctuations, TRANSIT Satellite, Recorded on April 25, 1970, 0544-0547 Hours GMT

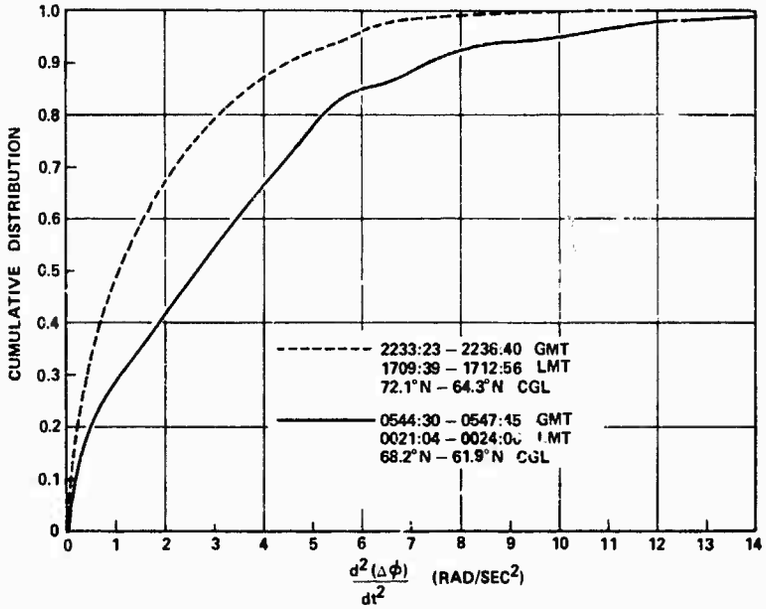


Figure 7. Cumulative Distribution Function of the 2nd Derivative with Respect to Time of the Differential Phase Fluctuations

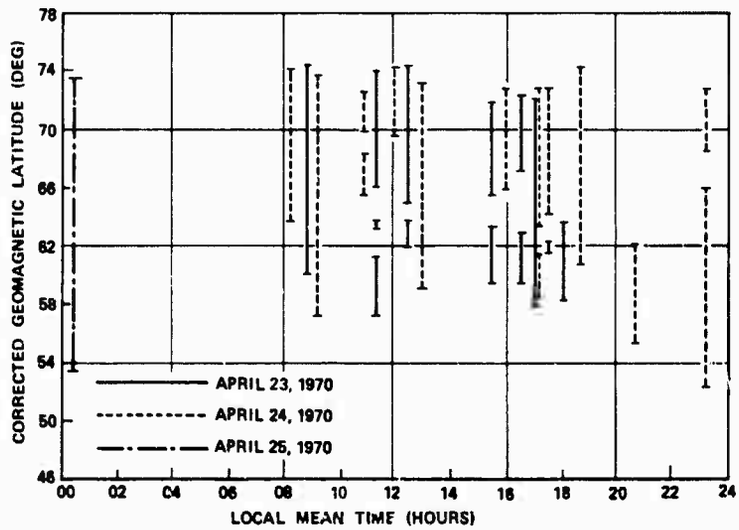


Figure 8. Corrected Geomagnetic Latitude Extent of Phase Scintillation Regions as a Function of Local Mean Time at the 350-km Altitude - Subionospheric Point

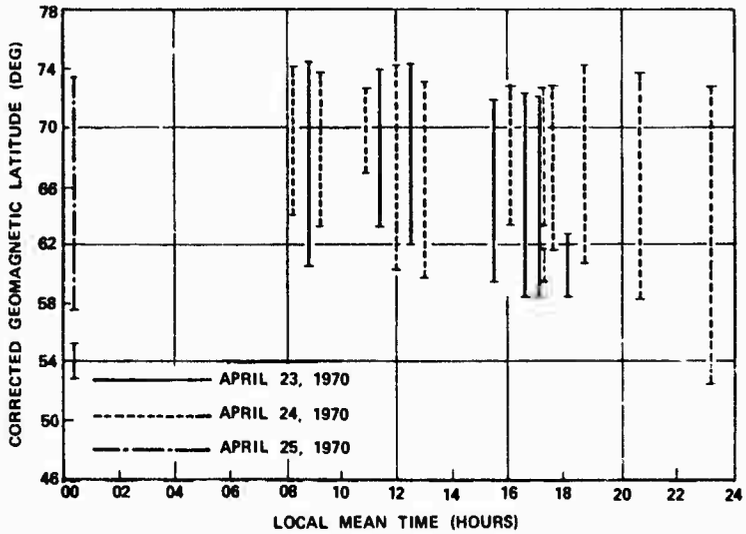


Figure 9. Corrected Geomagnetic Latitude Extent of VHF Amplitude Scintillation Regions as a Function of Local Mean Time at the 350-km Altitude - Subionospheric Point

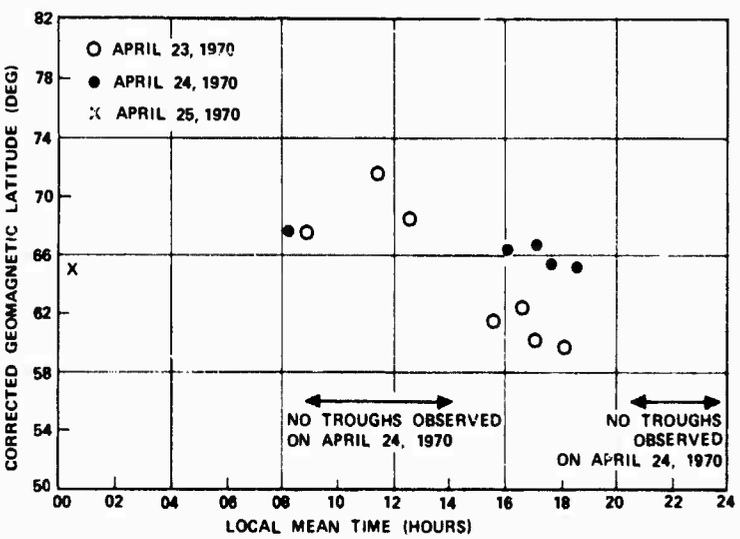


Figure 10. Equatorward Edge of the Trough as Observed from Differential Phase Measurements as a Function of Local Mean Time at the 350-km Altitude - Subionospheric Point

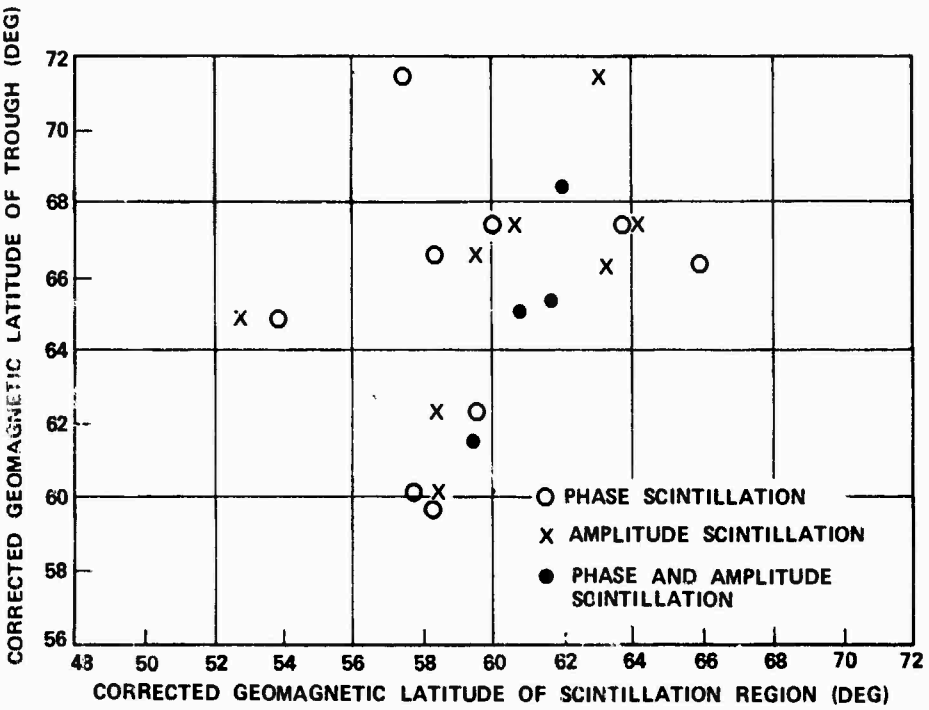


Figure 11. Corrected Geomagnetic Latitude of Equatorward Edge of the Trough as a Function of the Corrected Geomagnetic Latitude of the Equatorward Edge of the Scintillation Region

DISCUSSION

J.M.Goodman, Space Systems Division, Naval Research Laboratory, Washington D.C., USA

You indicated that amplitude scintillation occurred at times in which there was no phase scintillation. As you pointed out, theoretically this is opposite of what one would expect. Because of the steepness of the phase records it is difficult to detect small phase fluctuations. What was the criterion for ascertaining the occurrence of phase scintillation? Also, was the same time constant used in the analysis of both amplitude and phase data?

Author's Reply

The phase and amplitude scintillation regions were designated from the analog chart recordings, an example of which is shown in Figure 1. The resolution of the differential phase, as observed from the analog recordings is on the order of about 15°. The recorded data were digitized at 1 second intervals. The differential phase fluctuations about the least square fit were employed in the analysis of the nonlinear phase shifts.

PROPAGATION EFFECTS OBSERVED IN CONNECTION WITH NTS-1 OBSERVATIONS NEAR THE MAGNETIC EQUATOR

RUDOLPH R. ZIRM
and
JOHN M. GOODMAN
Space Systems Division
Naval Research Laboratory
Washington, D. C.

ABSTRACT

During a two week period in May 1975 near the magnetic equator, the Naval Research Laboratory measured amplitude and differential propagation delay scintillation of the 335 and 1580 MHz transmissions from NTS-1 satellite. NTS-1 is an experimental test vehicle in support of the NAVSTAR-GPS program and was launched on 14 JUL 1974 into a 7,277 by 7,444 nautical mile orbit having an inclination of 125°. The measurements which are described in this paper are part of Operation Antarqui, named for Inca god, and was a joint NASA Instituto Geoficisco del Peru campaign with specific experiments by the universities of Illinois, Pittsburg, Denver, and Penn State, the Dudley Observatory, and NRL.

In this paper we will discuss the results obtained through analysis of amplitude and differential propagation delay data at 335 and 1580 MHz and observed at the former mini-track station at Ancon, 25 miles north of Lima. Data were also taken during this experiment in support of the NTS-1 orbit improvement and time transfer research effort.

These results also have immediate applications in resolving such practical system problems as the contribution of propagation delay scintillation to error budgets of satellite navigation systems such as NAVSTAR/GPS, and the power margins necessary to overcome the effects of amplitude scintillation in the equatorial region. For example, the data taken on this campaign (which is only representative of the two week data span) show amplitude fading greater than 30 dB below the mean at 335 MHz, but little or no effect on 1580 MHz. These data have been scaled for SI and have been compared with the Fremouw-Pope-Rino Model. Differential propagation delays at these frequencies during time of scintillation activity were observed to be two nanoseconds peak-to-peak.

Experimental receivers designed for NTS-1 were used with special doppler converters to generate convenient analogue data displays of the differential delay and amplitude variables. A special transportable antenna system was designed and deployed and data was recorded on chart and magnetic tape recorder.

Several conclusions have been reached on the basis of our analysis.

1. Severe scintillation activity was observed at 335 MHz for approximately four days out of a two week interval and varied from a few dB to over 30dB.
2. No apparent scintillation was observed at 1580 MHz.
3. The differential delay scintillation was not observed to exceed two nanoseconds peak-to-peak. The smoothed differential delay information suggested the existence of an anomaly in the total electron content as the satellite passed over the magnetic equator. This feature is not inconsistent with currently accepted models of electron redistribution near the equator.

It is emphasized that these data were obtained over a limited span of time and the conclusions are valid for equatorial-equinox conditions only.

1. INTRODUCTION

The Navigation Technology Satellite I, NTS-1, was the latest in a series of experimental satellites designed and constructed by the Naval Research Laboratory (NRL) to investigate the feasibility of Satellite Navigation by passive ranging and the first launched as R&D support for the Tri-service NAVSTAR Global Positioning System (GPS). The next satellite, NTS-II will be launched shortly in the operational circular orbit of 12 siderial hours period, 63° inclination. The pertinent parameters of NTS-1 are shown in Table 1, a pre-launch photograph is shown in Figure 1 and an artist's conception of the on-orbit configuration in Figure 2. Figure 3 is a drawing of this spacecraft showing the antenna configuration and the location of laser retroreflectors, solar cell experiments and gravity gradient statilization system.

It became apparent shortly after achieving the nominal orbit, that the spacecraft was not becoming stabilized as shown in Figure 3 but was tumbling rather slowly at an increasing rate due to solar torques. The difficulty was found to be in the damper (the sphere on the end of the boom shown in Figure 3), which was not dissipating the energy of rotation as designed. This malfunction has since been traced to poor manufacturing techniques and has been corrected at the vendors assembly line.

The electronics was performing as designed,. however some experiments were per-

formed at a limited level as described in this paper.

1.1 NTS-I Payload Description

Figure 4 is a block diagram of the Satellite navigation subsystem. A sidetone range (STR) signal format is employed at the frequencies shown in figure 5 at P and L bands. The dual band operation was chosen to eliminate ionospheric refraction errors. All frequencies are coherently derived from a precision oscillators as shown in figure 4. Two rubidium vapor oscillators and one precision crystal oscillator were carried and were selectable on ground command. The on-orbit performance of the rubidium oscillators was one of the prime experiments. Independent range measurements are obtained every minute to a precision of approximately 10% λ at 6.4 MHz or about 1.6 ft. The forthcoming NTS-2 launch will carry GPS navigation equipment which employs a PRN waveform of 10 MB/Sec for ranging in addition to the STR system described above. A summary of the electrical parameters of the satellite is shown on figure 6 and a list of some of the experiments and prime experimentors planned for this flight is shown on figure 7. A continuing experiment of great importance is the stability of high orbits when perturbed by the solar flux. The satellite also carried a PRN code that could be radiated on command from the ground when the STR system was turned off to be used by the USAF.

1.2 NTS-1 Propagation Experiments

It is beyond the scope of this paper to report on the results of most of the NTS-1 experiments, however, it is pointed out that the spacecraft (S/C) stability limited them to some extent. The primary problem was that of temperature control, since the desired orientation was not obtained (except for brief transient intervals). Thus, thermal design conditions were not met, requiring the careful monitoring of S/C temperatures, and load shedding where necessary. For example, P band ranging was usually obtained because with directional receiving antennas sufficient back radiation was obtained. S/C temperature did not allow simultaneous L and P band ranging so very precise ranging was not obtained (except for a series of laser measurements made at Goddard Space Flight Center (NASA)). However, simultaneous measurements at both frequencies using the CW tracking beacons was usually possible so that ionospheric propagation effects could be observed using a method developed at NRL (ZIRM, et al, 1972) and independently by General Electric Co.(MILLMAN and ANDERSON, 1968), and generally referred to as differential doppler. Figure 8 shows the test setup and details of the doppler converters are shown in figure 9. This equipment was used at the test sites near Washington, D. C., and Miami, Florida to observe differential delay effects at the two tracking beacon frequencies of 335 and 1580 MHz. Figure 10 is a record of several minutes of continuous data made while the satellite was rotating relatively rapidly, but is indicative of the "normal differential delay effect" expected of these frequencies; i.e., approximately linear differential delay increasing with time as the total path delay changes. When the satellite rotation axis was out of the line of sight, a periodic delay variation of approximately 2ns (2 ft.) is observed in the differential delay as seen in figure 11, because the 335 MHz and 1580 MHz antennas were on opposite faces of the spacecraft, (see figure 3). Thus, small scale ionospheric delay variations comparable to the satellite rotation effect, are not separable without detailed satellite orientation and motion information. However variations of time scale of the order of tenths of seconds to seconds (scintillation) should be observable. No evidence of scintillation activity at the latitude of the NTS-1 receiving site near NRL was observed over a years time, however one possible scintillation pass was observed from the Florida site while the S/C was low on the horizon to the south (see figure 12).

During 1975 an opportunity to set up the receiving equipment near the geomagnetic equator presented itself with an invitation from NASA to participate in a joint campaign with the Instituto Geoficisco del Peru, to measure upper atmosphere and lower ionospheric effects from ground based sensors, balloons, and rockets. Portable receiving equipment, figure 13, was installed at the former minitrack site at Ancon, Peru and data taken for about two weeks in late May. Even before the system was calibrated and the differential doppler equipment installed, severe amplitude scintillation was observed at 335 MHz (see figure 14). However, after the equipment was properly set up, amplitude and phase scintillation was observed on four nights for varying periods of time (a few minutes to hours) with the 335 MHz amplitude scintillation down to phase lock loss level in the receivers (> 35dB). In these instances no scintillation activity was observed on the 1580 MHz record, and the phase effect was always less than 2ns peak-to-peak or a contribution to CW ranging error of $\pm$2 ft. (See Figures 15-17).

2. DATA ANALYSIS

The data have been recovered by an HP 3960 magnetic tape recorder with chart recorders as backup. Data reduction equipment problems have precluded the processing of significant quantities of the tape data into meaningful statistics as of this writing. However, playback of the data through the 100ns phase meter allowed complete curves of the differential delay variation vs. time to be obtained. Figure 18 shows a typical differential delay curve extracted from a twilight pass near the southern horizon and figure 19 depicts the geometry associated with this pass. Figure 20 is a typical delay curve obtained during a high elevation early morning NE-SW pass and figure 21 is the corresponding geometry. For this high elevation, the delay function is smaller than expected at low elevation angles which implies an electron content (TEC) anomaly symmetrical about the magnetic equator. It is noteworthy that the mid-latitude trough also produces such an effect for non-equatorial observation stations. This effect is obviously important to ranging and positioning systems that rely on simple models to remove excess delay. Superimposed is a prediction of the group path delay as derived from the Ching-Chiu model, (CHING,B. K., and CHIU, T. T., 1973), taking the appropriate

geometry into account and carrying out the necessary integration along the ray path. Inputs to the Ching-Chiu model include sunspot number, day of year, and time of day, in addition to certain geographic conditions.other models, including the BENT model (LLEWELLYN, S. K., and BENT, R. B., 1973) are being examined as well.

Preliminary scintillation analysis was carried out at NRL to extract the first order statistics from selected segments of phase and amplitude data.

It is of interest to distinguish between several types of phase fluctuations exhibited in the data. The most obvious component is of very low frequency and results from global TEC variation; this component is clearly of interest for purposes of gross ionospheric ranging error. In addition there is likely in existence a hierarchy of more rapid fluctuation regimes which are related to the size of the ionospheric inhomogeneities. Large scale oscillatory perturbations such as travelling ionospheric disturbances contribute between 0.1 and 0.5 Hz and randomized irregularities contribute above 0.5 Hz. In a manner similar to FREMOUW and COUSINS (1975), we have separated the data into a focussing (or regular) component and a scattering (or quasi-random) component through use of digital filtering. The first filter is a detrender (a digital high pass filer with f_c = 0.1 Hz)which eliminatesthe so-called D. C. component or global TEC variations in the data. The second software filter is a band-pass operation which passes frequency components between 0.1 and 0.4 Hz. Finally we pass the data through a second detrender (a digital high pass filter with f_c = 0.4 Hz) which delivers the random component at its output port. We called the output of the band-pass operation the regular component and we call the out-put of the second detrender the quasi-random component.

For purposes of illustration, we have examined a small segment of data of 100 seconds duration at 0125 zulu during which the S_4 index was essential 1.0. The in-phase components for (a) the unfiltered data, (b) the regular plus quasi-random component, (c) the regular component and, (d) the quasi-random component are given in figure 22. Figure 23 shows the phase quadrative components. Figures 24 and 25 show the polar diagrams of the regular and quasi-random components respectively. Clearly the statistics of the quasi-random component are not simple since there is likely to be multiple scattering in evidence. Nevertheless the power spectrum of the quasi-random component (figure 26) indicates that an f^{-3} law is obeyed as would be expected.

The analytical effort outlined above is continuing with emphasis being placed upon separating the ranging error due to the various categories of irregularities. We also plan to exercize codes which exhibit the phase-error distributions as well as the average interval between phase errors, and the average duration of phase errors as a function of phase-error magnitude. Separate phase and amplitude power spectra will be deduced.

3. FUTURE PLANS

The launch of NTS-2 is scheduled for launch in early 1977 into the NAVSTAR-GPS orbit of 12 sidereal hours period with a 63° inclination. It will comprise the same equipment as NTS-1 with an additional GPS NAV package, which will allow experimental investigation of the influence of ionospheric effects on signals in addition to the type data described in this paper. The data will be recorded in a digital format on tape cartridges, so that with proper processing, correlation functions, power spectra and certain relevant statistics may be generated without the manual labor necessary with analogue data. It is hoped to collect data from a variety of sites for periods long enough to show seasonal effects, and hopefully continued long enough so that the effects of increasing solar activity will be observed.

4. REFERENCES

1. Ching, B. K., and Y. T. Chiu, 1973, J. Atmos. Terrest. Phys., 35, 1635.
2. Fremouw, E. J., and M. D. Cousins, 21 Nov 1975, SRI Quarterly Technical Report #2.
3. Llewellyn, S. K., and R. B. Bent, AFCRL-TR-73-0657
4. Millman, G. H., and R. E. Anderson, July 1968, "Ionospheric Phase Fluctuations of Satellite Transmissions", Journal of Geophysical Research, Space Physics, Vol. 73, No. 13.
5. Zirm, R. R., R. E. Brescia, and C. Morris, 1972, "Transionospheric Differential Delay Measurements from Timation II at 150 MHz and 400 MHz", NRL Report 7419.

TABLE I

Physical Parameters of NTS-1

VEHICLE	ATLAS-F, FAIRCHILD TRANSFER STAGE
LAUNCH TIME/DATE	0455Z, 14 JULY 1974, WTR
INTERNATIONAL DESIGNATOR	1974-54A
SPACE DEFENCE CENTER NUMBER	7369
PERIOD – MINUTES	468.73
APOGEE HEIGHT – N. MILES	7423
PERIGEE HEIGHT – N. MILES	7274
ECCENTRICITY	0.00688
INCLINATION – DEG.	125.12
SPACECRAFT WEIGHT IN ORBIT – Kg.	282.6
STABILIZATION	TWO SIXTY FOOT GRAVITY GRADIENT BOOMS, ONE G.E. DAMPER

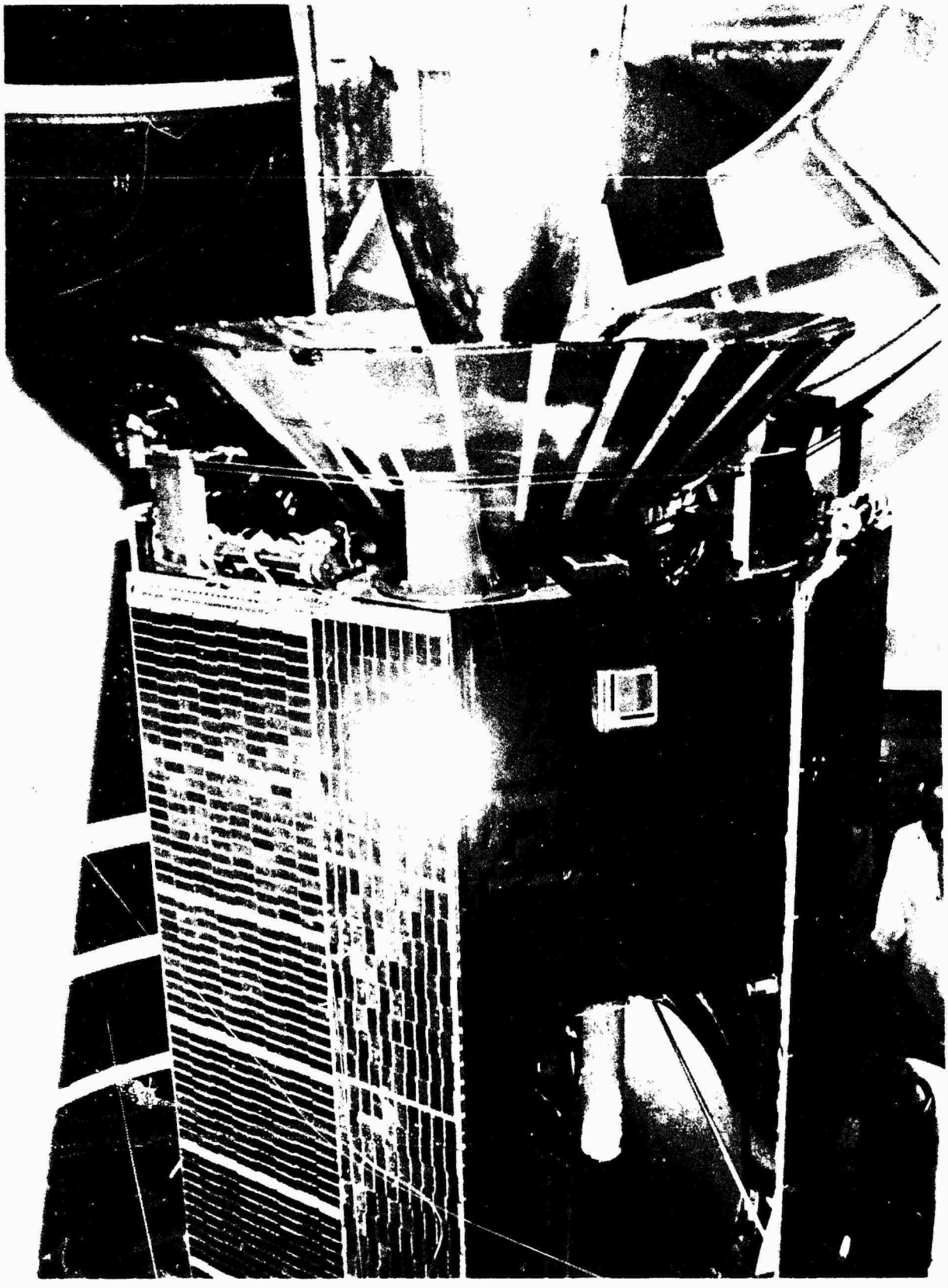


Fig. 1 - Pre-launch photograph of NTS-1

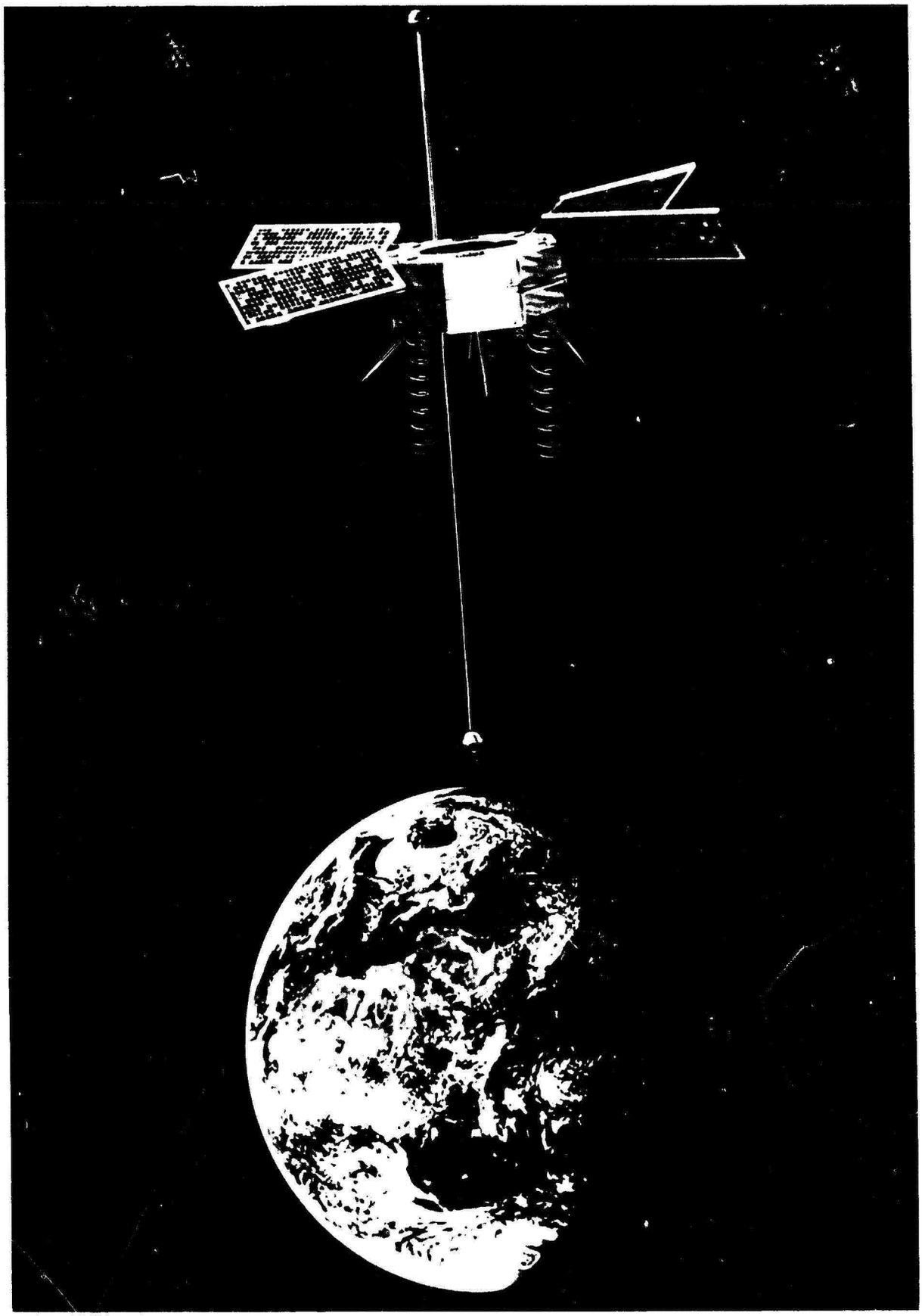


Fig. 2 - Artist conception of the NTS-1 on-orbit configuration

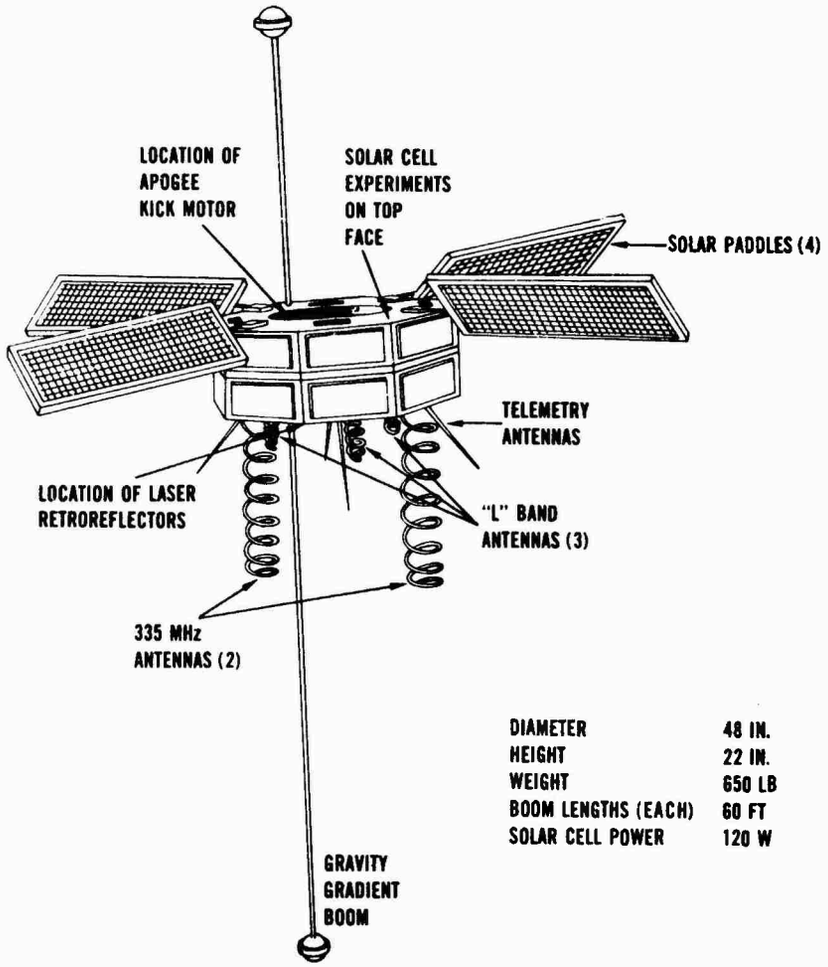
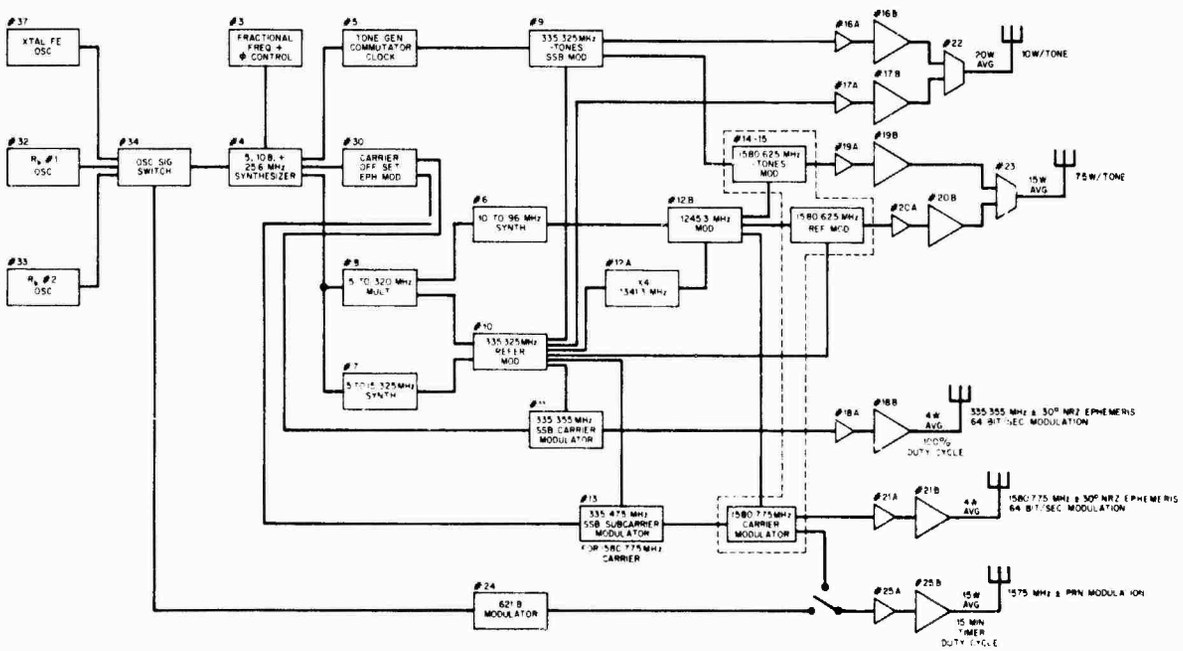


Fig. 3 - Spacecraft details



* THESE TONES ARE COMMUTATED DECREASING FREQUENCY FOR 0.5 SEC RESPECTIVELY EXCEPT THE LOWEST FREQUENCY IS ON FOR 2.5 SEC WHEN THE TOTAL DUTY CYCLE IS 7.5 SEC

Fig. 4 - Satellite navigation subsystem

DESIGNATION	LOW BAND FREQ. (MHz)	HIGH BAND FREQ. (MHz)	TONE (kHz)
RANGE REFERENCE	335.325	1580.625	—
RTF1	335.3249	1580.6249	.100
RTF2	335.32475	1580.62475	.250
RTF3	335.324	1580.624	1
RTF4	335.321	1580.621	4
RTF5	335.309	1580.609	16
RTF6	335.261	1580.561	64
RTF7	335.225	1580.525	100
RTF8	334.925	1580.225	400
RTF9	333.725	1579.025	1,600
RTF10	328.925	1574.225	6,400

Fig. 5 - Timation III (NTS-1) frequencies

POWER SUPPLY: 120W NICAD, 4 SOLAR CELL ARRAYS

FREQUENCY STANDARDS: FEL, 5 MHz DOUBLE OVEN, A FEW PARTS IN 10^{12}; TWO EFRATOM RUBIDIUM VAPOR (EXPERIMENTAL)

CARRIER SYNTHESIS: COHERENT, ALL FREQUENCIES

TRACKING BEACONS: 335.355 MHz, 4 W., 10 db HELIX, ON CONTINUOUSLY; 1580.775 MHz, 4 W., 10 db HELIX, ON CONTINUOUSLY

RANGING SIGNAL: 335.325 MHz, 8 W., 10 db HELIX, ON 5.5 OR 7.5 SEC. ON THE MINUTE; 1580.625 MHz, 8 W., 10 db HELIX, ON 5.5 OR 7.5 SEC. ON THE MINUTE

PRN SIGNAL: 1575 MHz, 10 MBS, 15 W. AVG., 10 db HELIX, ON COMMAND

Fig. 6 - Electrical parameters of NTS-1

1. ORBIT STABILITY USING LASER TRACKING. 420 FUSED SILICA CUBE CORNER RETROREFLECTORS WITH 15 MM HEXOGONAL ENTRANCE PUPIL AND ONE 5 IN. HOLLOW CUBE CORNER USING THREE FRONT-SILVERED MIRRORS. (NRL, NWL, NASA)

2. EVALUATION OF QUARTZ CRYSTAL AND RUBIDIUM VAPOR FREQUENCY STANDARDS. (NRL)

3. SOLAR CELL RADIATION DAMAGE. COMPARE VIOLET, HELIOS, RADIAL GRID, LITHIUM DIFFUSED AND THIN WRAP-AROUND CONTACT TYPES. ALSO COMPARE FZ AND CZ SILICON CELLS WITH 1 AND 10 OHM-CM AND CORRELATE WITH DOSIMETER LEVEL AND RATE. (NRL, RAE, NASA, COMSAT)

4. PRN MULTIPATH AND FOLIAGE PROPAGATION. (US ARMY ETL)

5. TRANSIONOSPHERIC DIFFERENTIAL DELAY AND AMPLITUDE SCINTILLATION. (NRL)

6. NAVIGATION AND TIME TRANSFER. (NRL, RGO, NMD)

Fig. 7 - Experiments on NTS-1

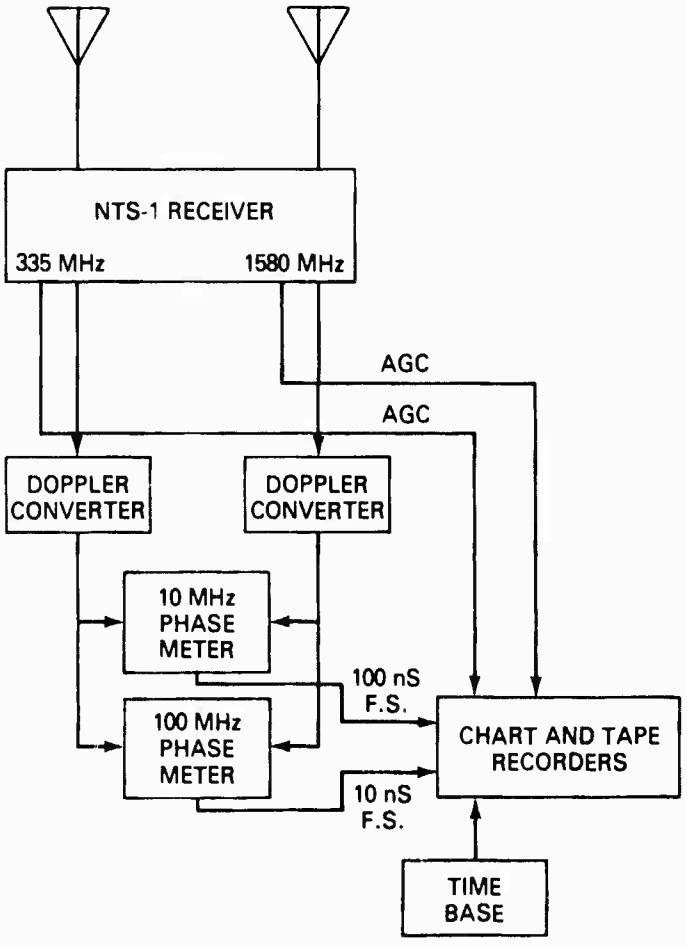


Fig. 8 - Experimental equipment

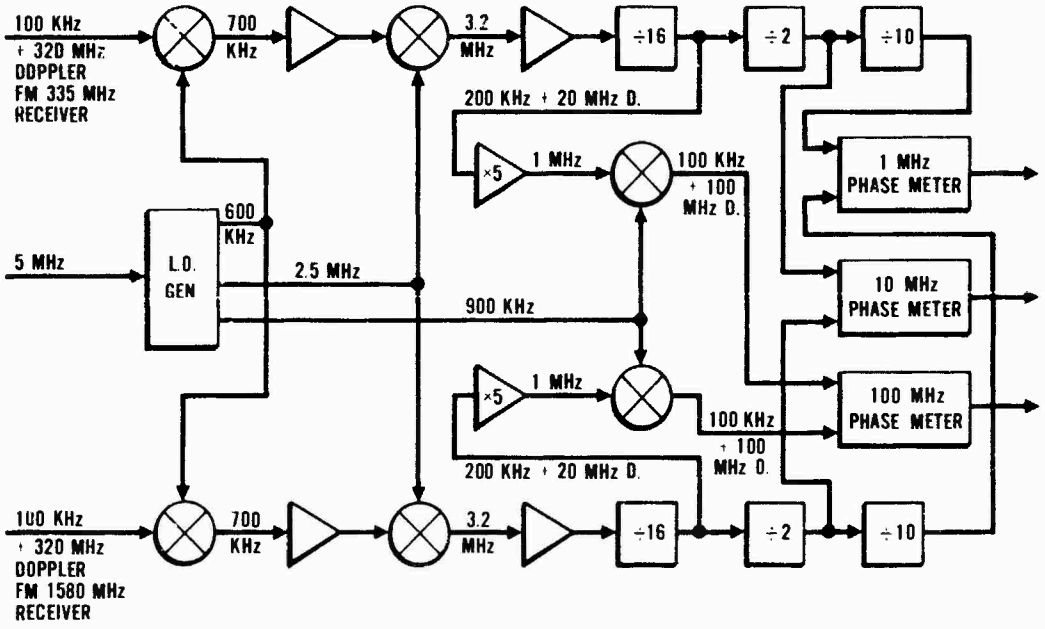


Fig. 9 - Doppler converter for the Doppler Receiver

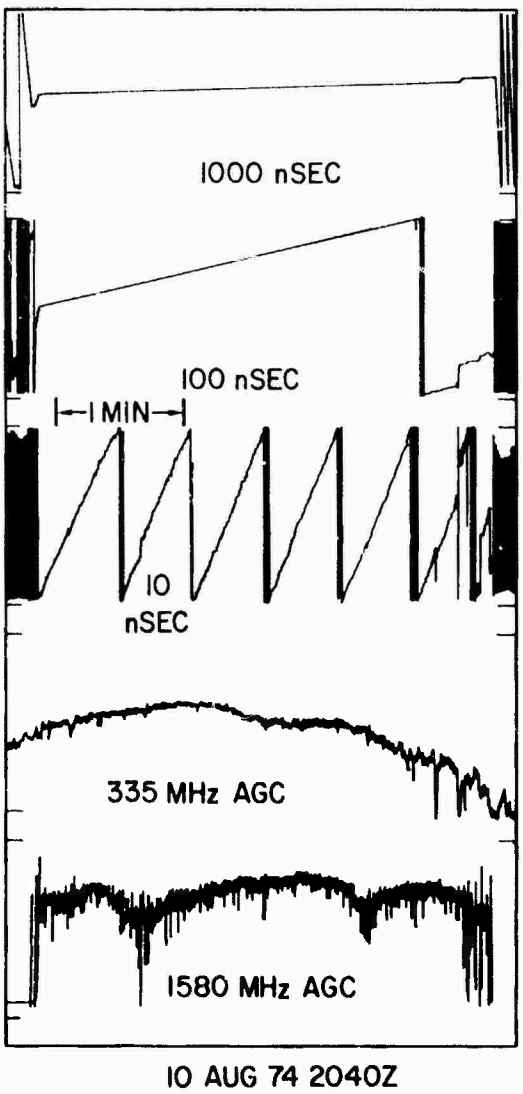


Fig. 10 - Data from phase meters and AGC output at 335 MHz on 10 AUG 1974 at a mid-latitude site.

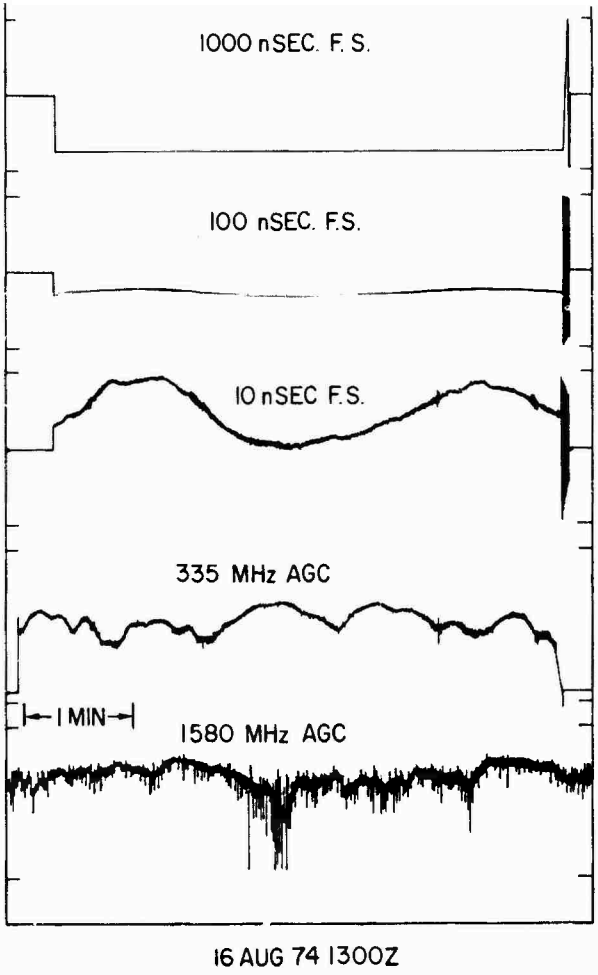


Fig. 11 - Data obtained from a mid-latitude site on 16 AUG 1974 illustrating systematic

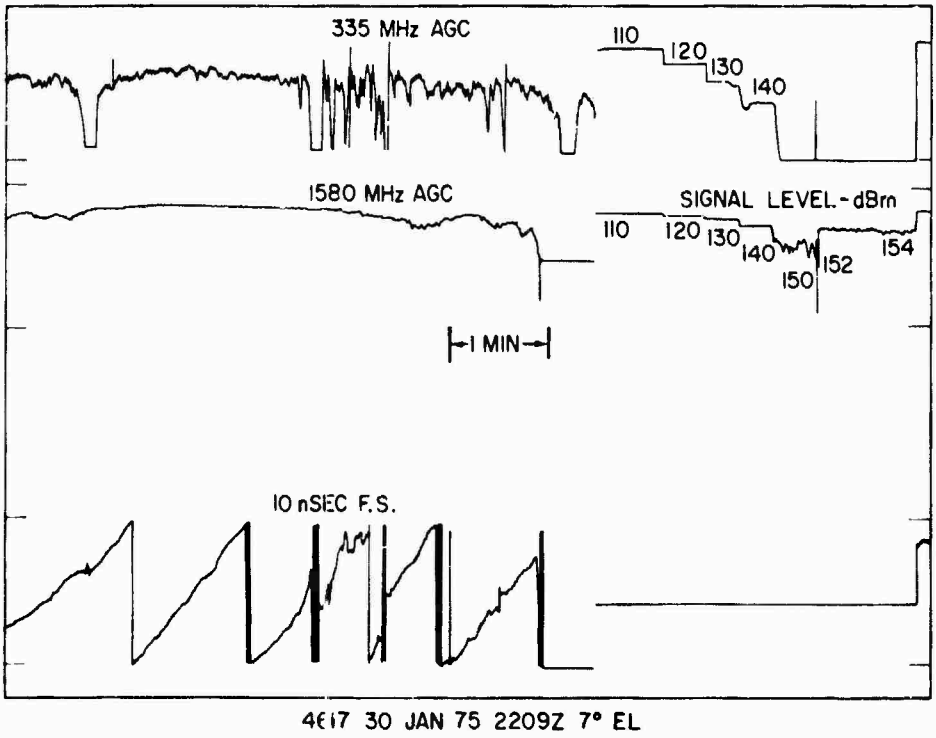


Fig. 12 - Scintillation data obtained at the Florida site on 30 JAN 1975.

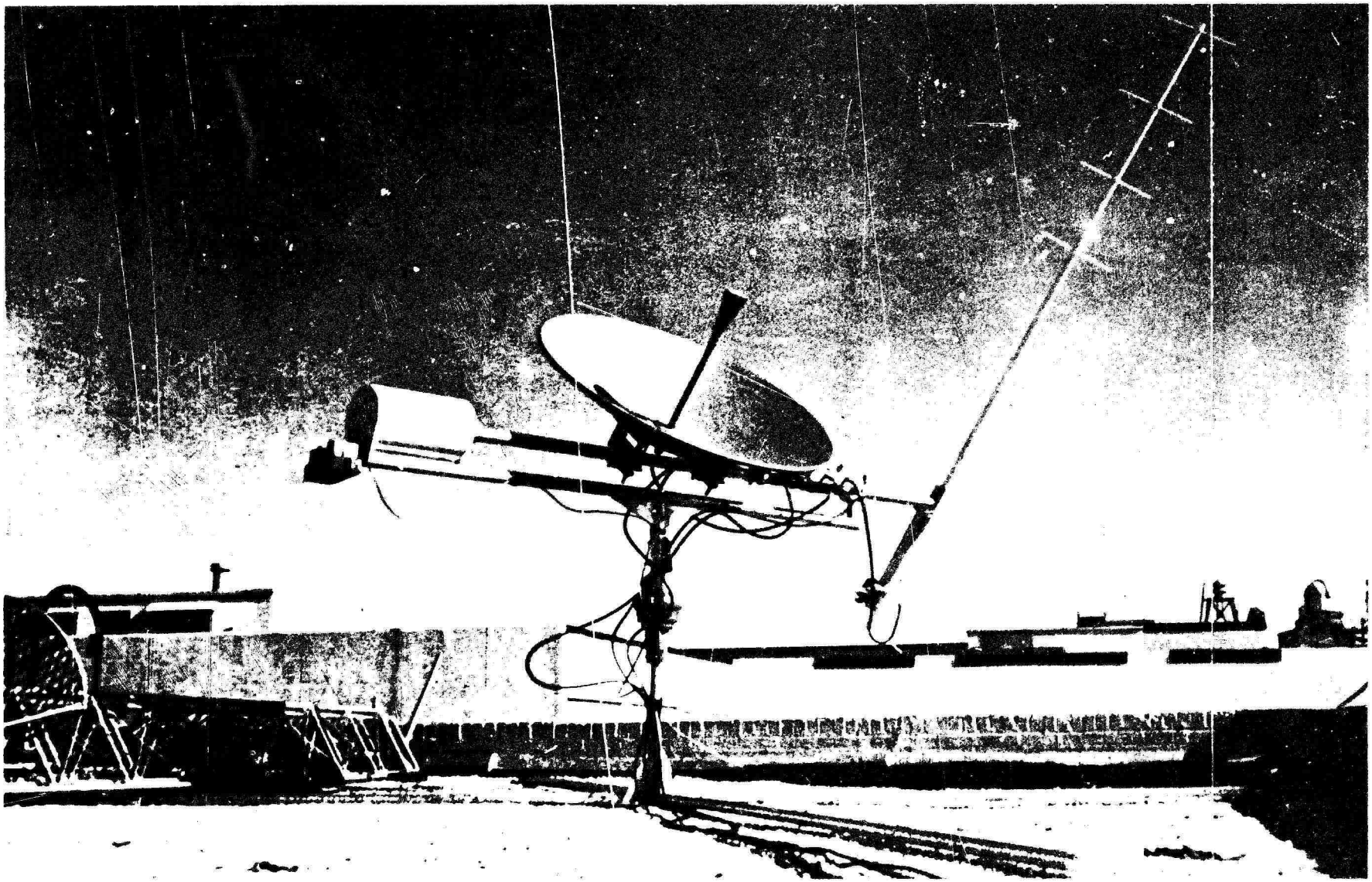


Fig. 13 - Receiver site near Ancon, Peru.

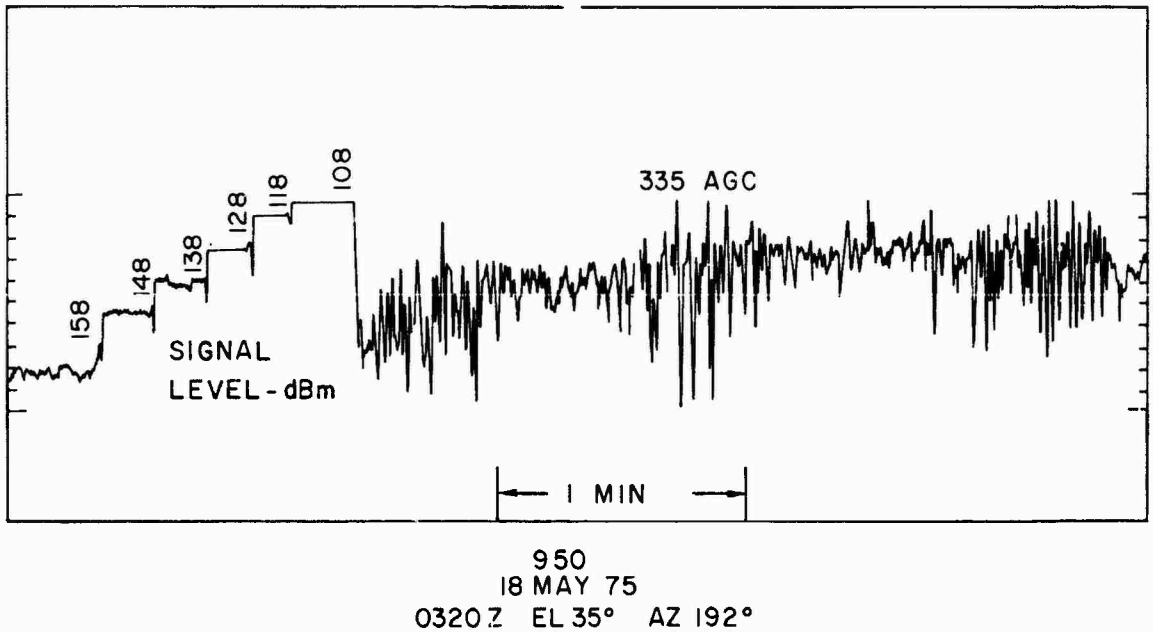


Fig. 14 - Example of amplitude scintillation at ANCON on 18 MAY 1975.

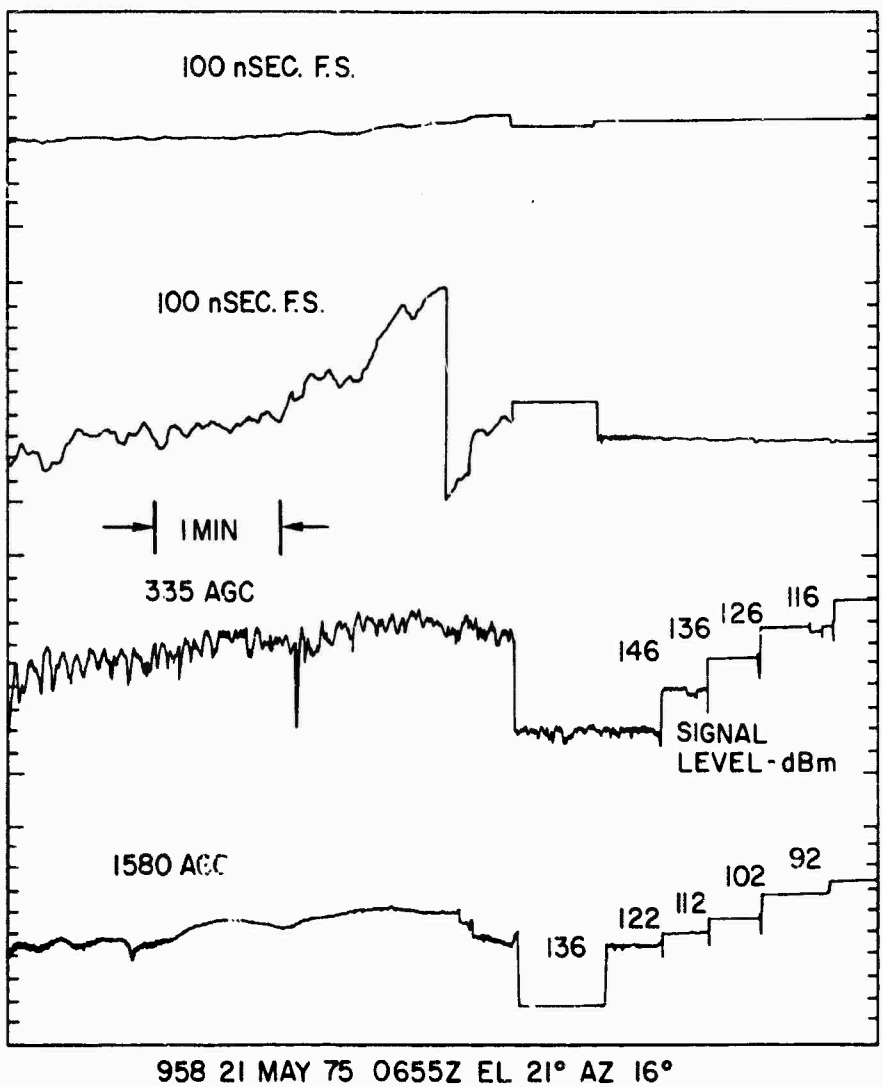


Fig. 15 - Amplitude scintillation and phase meter data obtained on 21 MAY 1975 during pass number 958.

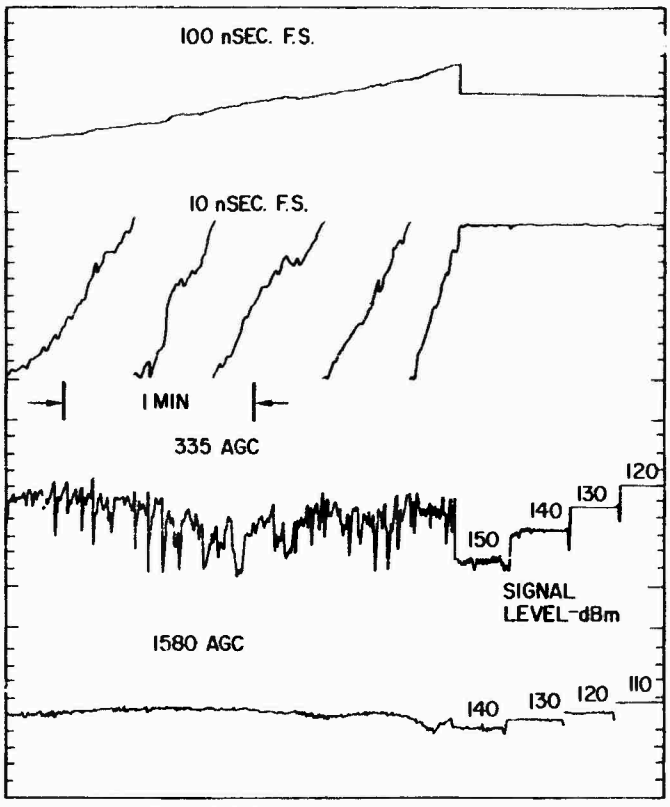


Fig. 16 - Amplitude scintillation and phase meter data obtained on 26 MAY 1975 during pass number 974.

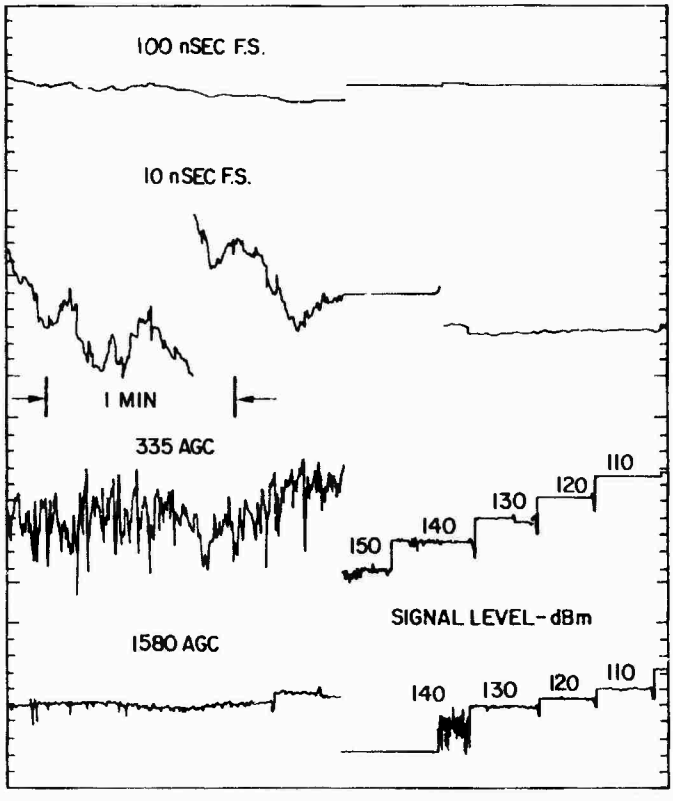


Fig. 17 - Amplitude scintillation and phase meter data obtained on 21 MAY 1975 during pass number 957.

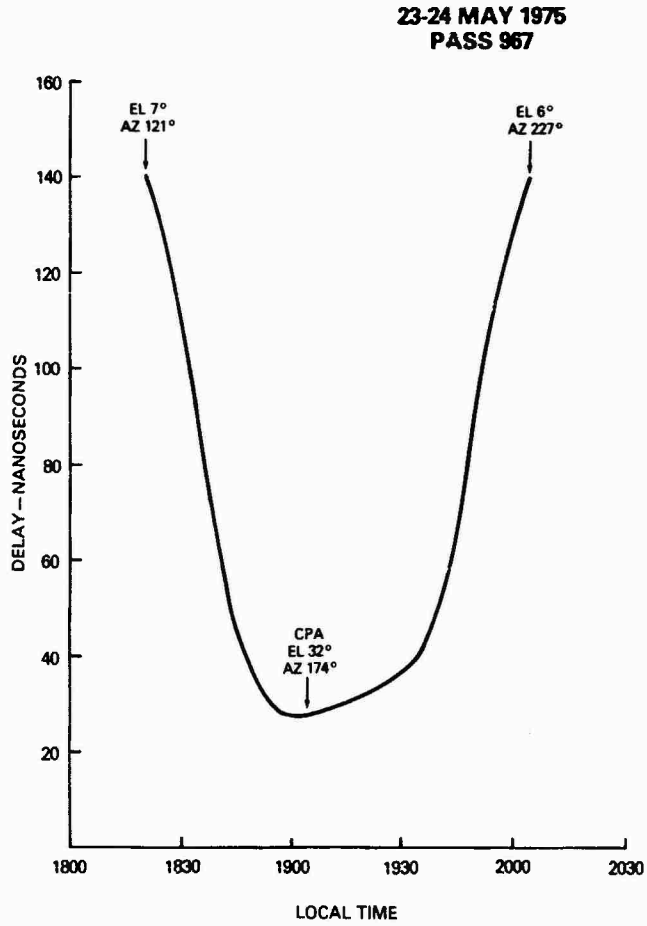


Fig. 18 - Typical differential delay curve obtained on 23-24 MAY during pass number 967.

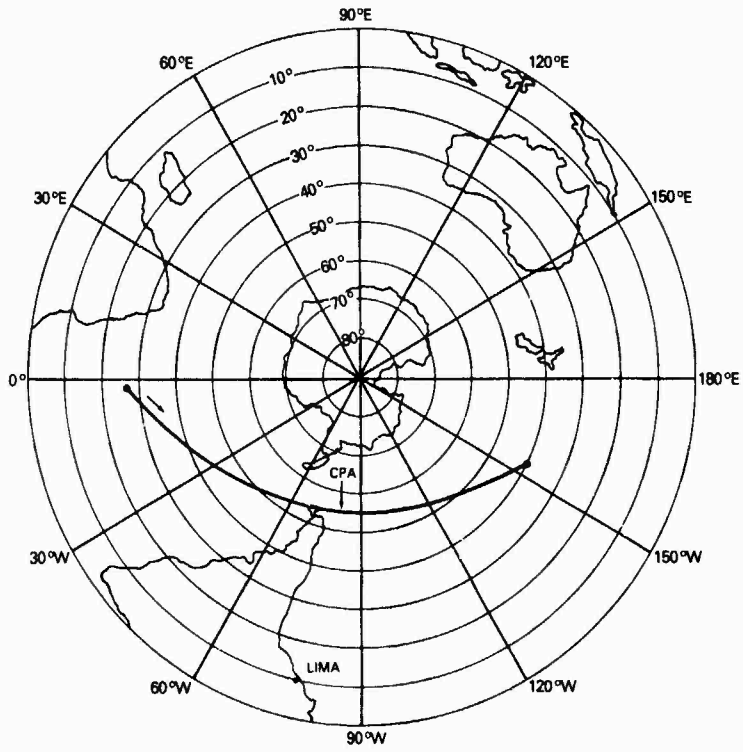


Fig. 19 - Geometry for pass number 967.

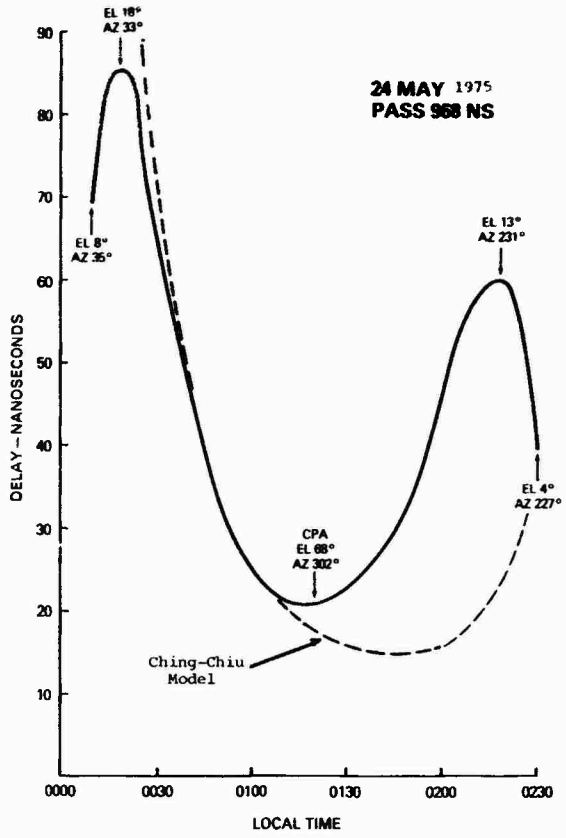


Fig. 20 - Differential delay curve obtained on 24 MAY 1975 showing anomalous pattern for pass number 968.

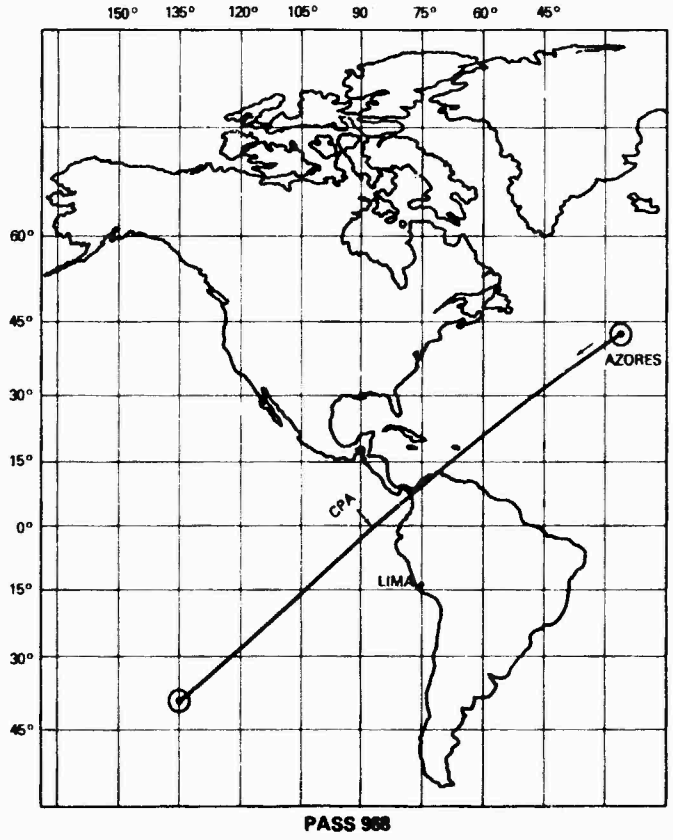


Fig. 21 - Geometry for pass number 968.

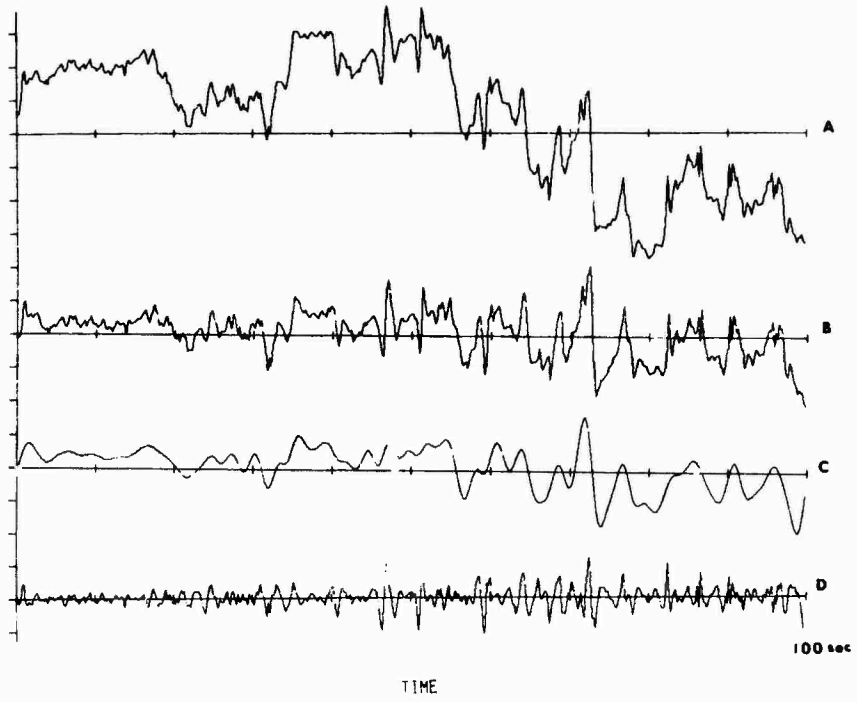


Fig. 22 - In phase components of the scintillating 335 MHz signal.
(a) unfiltered data
(b) regular plus quasi-random components
(c) regular component
(d) quasi-random component

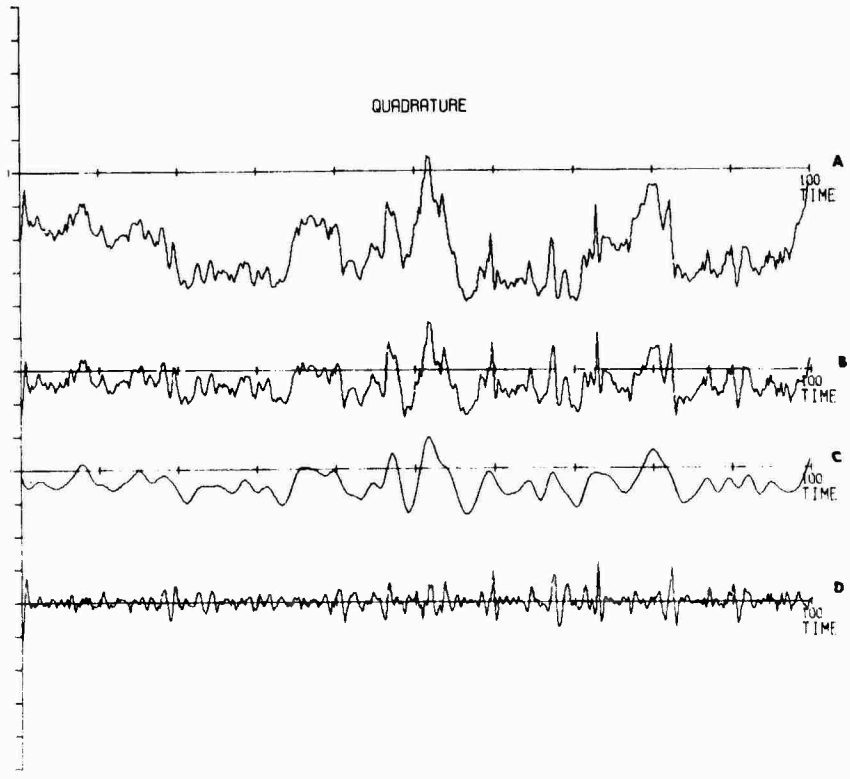


Fig. 23 - Quadrature components of the scintillating 335 MHz signal.
(a) unfiltered data
(b) regular plus quasi-random components
(c) regular component
(d) quasi-random component

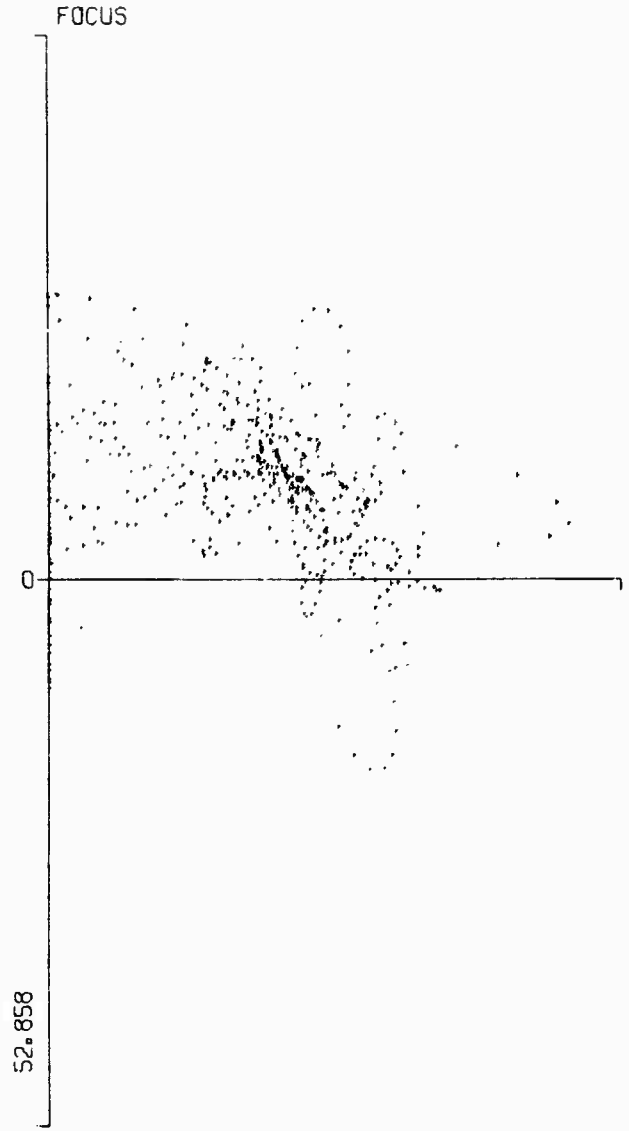


Fig. 24 - Polar diagram of the regular component

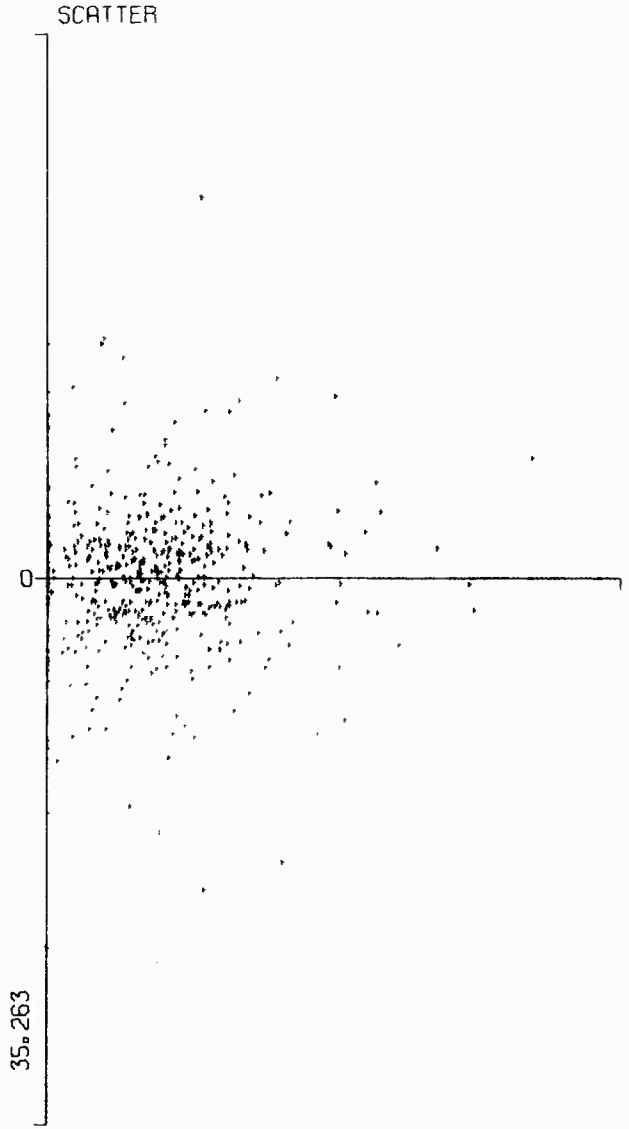


Fig. 25 - Polar diagram of the quasi-random component

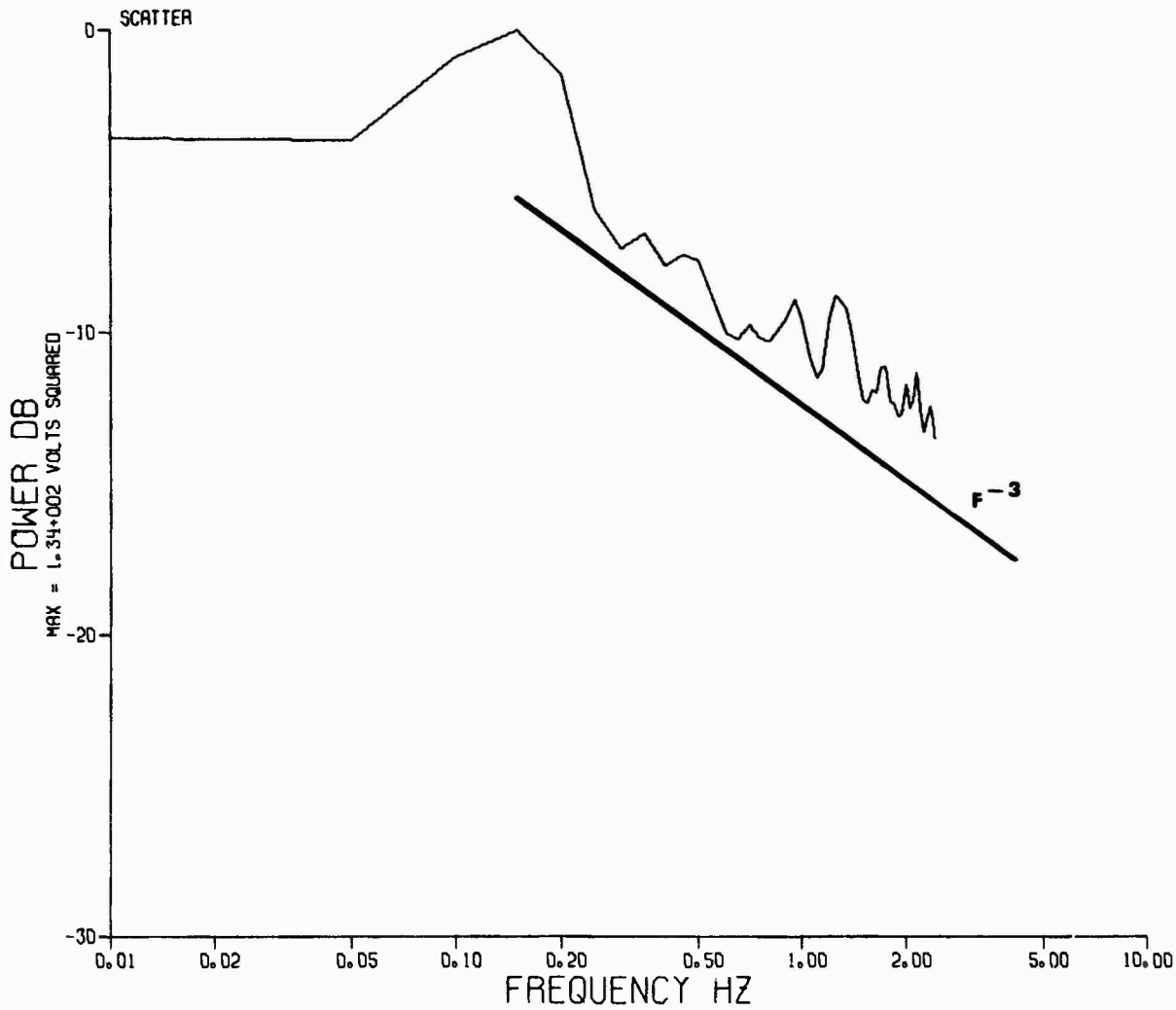


Fig. 26 - Power spectrum of scintillating signal.

J.S.Belrose

The paper concentrated on the short term phase perturbation and frequently throughout the paper both authors state that these scintillations were less than ±4 ft. No mention was made of absolute position accuracy. My comment relates to the obvious, but perhaps I should nevertheless make it. Your latitudinal (or longitudinal) US doppler data show marked asymmetries, and so clearly the positional inaccuracies are much greater than ±4 ft. This, while obvious, should perhaps be stated.

Author's Reply

The point which we made was that the so-called scatter or random component corresponds to a ranging error of ~± 4 feet, an amount which is unlikely to be removed because of its stochastic nature. The focused components and the large TID effects and indeed the zenith angle effects have a promise of being eliminated either by modeling, dual frequency schemes, or by adaptive measures similar to those mentioned by Dr Grossi in a previous paper. Although we did emphasize the so-called scatter component of ranging error, you are correct in pointing out that the total system accuracy is dependent upon other gross ranging error components (presumably regular ones) which are much larger unless sophisticated techniques are employed such as those alluded to above.

The second point we wished to discuss in the paper is that amplitude scintillation at the magnetic equator may reduce the S/N ratio to such an extent that the system may momentarily lose phase lock. This and other statistical properties of the data will be discussed in the final manuscript.

The scope of this paper has precluded discussion of these effects on system navigational accuracy per se, but obviously must be accounted for in system error budget considerations.

H.Soicher, US Army Electronic Command, Fort Monmouth, NJ 07703, USA

If the purpose of dual-frequency satellite transmission is to eliminate ionospheric errors, should not the frequency separation be wide? (as in NTS, as opposed to GPS)

Author's Reply

Obviously a greater separation between frequencies allows the removal of ionospheric refraction and differential delay effects more accurately, however, practical system constraints have dictated the GPS choice.

H.Soicher

I was pleased to see that your TEC results for the geomagnetic anomaly indicate that the crests maximize in the summer hemisphere. This agrees with our own Sao Paulo results, as well as from topside sounder results.

Author's Reply

Yes. Several examples of the type shown in the slide to which you refer have been noted. We are now in the process of comparing the magnitude of the observed anomaly with accepted models such as the BENT and the Ching-Chiu models.

Funding limitations have precluded data reduction of large numbers of passes of this type, but the raw data exists.

IONOSPHERIC RANGE ERROR CORRECTION IN PRECISION RADAR SYSTEMS BY ADAPTIVE PROBING OF THE PROPAGATION MEDIUM(*)

R. S. Allen
Air Force Geophysics Laboratory
Hanscom Air Force Base, Massachusetts 01731

D. D. DuLong
Regis College
Weston, Massachusetts 02193

M. D. Grossi and A. H. Katz
Raytheon Company
Sudbury, Massachusetts 01776

ABSTRACT

A real-time adaptive scheme developed by the authors for ionospheric range error correction in precision UHF radars appears capable of reducing this error to below 10 ft. The scheme is based on the use of a model of the monthly median ionosphere constructed from existing worldwide climatology (and capable of providing by itself a residual range error not larger than 35 ft) and on updating and correcting this model with real-time dual-frequency measurements (accurate in themselves to about 4 ft) of the columnar electron content. This is done for directions and at times for which targets of opportunity (such as satellites and target vehicles embedded in the ionosphere) become available. The space-time cell in which the dual-frequency correction maintains its validity is so large (even when the sunspot number grows from low to high) that a single dual-frequency sample taken along a direction that cuts across the volume monitored by the radar as rarely as once every half hour still suffices. This requires that the adopted ionospheric model, while the solar cycle progresses, correspondingly becomes more and more sophisticated in such features as its reproduction of horizontal gradients. When strong impulsive traveling ionospheric disturbances propagate through the volume covered by the radar, the size of this space-time cell reduces and more frequent dual-frequency probing may become necessary.

1. INTRODUCTION AND SUMMARY

Precision radar systems require corrections for time or range errors caused by the ionosphere, when that medium is crossed in all or in part by the propagation path. To first order such errors are directly proportional to the integrated electron content along the path to the target. In real time these radar errors may be substantially reduced by predictions of the expected propagation effects using a model of the monthly median ionosphere constructed from existing worldwide climatologies. Such median corrections are shown to have a residual day-to-day r.m.s. variability about the median on the order of 20-25 percent of the median value.

For instance, the expected worst median case (daytime, equinox, sunspot maximum) for range errors experienced by an L-band radar at mid-latitudes are shown in Table 1.

Table 1
Worst Case Monthly Median Uncorrected Range Errors (ft) at 1235 MHz and Their Variability (1 Sigma)

Target Altitude (km)	Radar Elevation 0° Range Error (ft)	0° Sigma	20° Range Error (ft)	20° Sigma	90° Range Error (ft)	90° Sigma
1000	147	30	102	20	47	9
400	107	20	74	16	34	7
300	57	11	40	8	18	4

(*) This research was supported by the U.S. Air Force, in part under Contract F19628-76-C-0094 to Raytheon Company.

By use of such a model, a goal of 35 ft for the 1-sigma residual range error can be achieved at L-band even during expected worst median conditions. In precision radar systems more stringent requirements exist and the ionospheric induced error must be typically reduced to a few feet.

To satisfy such requirements, a real time adaptive scheme has been developed consisting of a model of the median ionosphere for the particular radar coverage region (derived from more complicated models of worldwide climatology) and of a set of algorithms to provide error corrections for targets at specified coordinates, which together attempt only to deliver an r.m.s. error on the order of 25% of the expected median correction at that location. This estimate is then specified by a dual-frequency dispersive probing of the ionosphere in the target area by the radar itself, using satellites of opportunity or even target vehicles embedded in the ionosphere.

The dual-frequency measurement, with accuracy on the order of 4 feet, is used to derive a normalization factor for the median ionospheric model in the surveillance volume, dependent on the radar range, elevation and altitude of the radar target.

The interval between the model's adaptive updatings by the dual-frequency probing depends on the space and time characteristics of the ionosphere.

At times of low sunspot number, a 10-ft range error specification can be met at mid-latitude by using a spherically symmetric monthly median ionospheric model, stored in the radar processor memory and adapted with dual-frequency measurements made at half hourly intervals. This correction is effective even in the presence of a strong Traveling Ionospheric Disturbance (TID).

At times of high sunspot number, the radar algorithms must consider realistic horizontal gradients modeled by the median ionosphere. The adaptive dual-frequency updating is still sufficient even at these times to meet the 10-ft range error.

During sporadic disturbed periods, such as when impulsive TIDs, generated in auroral substorms, pass over narrow regions of the radar coverage volume, the space-time cell over which the 10-ft specification can be met by a single normalization will shrink in both space and time, therefore more frequent (in time and space) dual-frequency probing may become necessary.

2. COMPUTER SIMULATION OF IONOSPHERIC RANGE ERROR CORRECTION PROCEDURE (ADAPTIVE APPROACH)

Figure 1 provides the simplified block diagram of the computer simulation of the range error correction procedure.

Block (2) "Stored Ionospheric Model" is a median model with realistic horizontal gradients, climatologically adjusted to the month and the hour under investigation. The adjustments are provided by Block (1).

Block (3) "Simulated Radar Data" has been constructed with a ray tracing program that computes the apparent range (group path range) for directions and geometric ranges of interest in the volume covered by the radar. In these computations several ionospheric models have been adopted to treat the cases listed in Table 2.

The "spherically symmetric median" model has been obtained from the model in Block (2) by arbitrarily setting to zero the horizontal gradients.

The "spherically symmetric anomalous" model has been obtained from the model in Block (2) by moving its f_oF_2 and F_2H_{max} 2σ away from median values. In the other two models (both with horizontal gradients) similar 1σ and 2σ alterations have been applied.

In Block (4) the "adaptive model updating" is done based on simulated ionospheric measurements, with the pulse pair (double frequency) method applied only once on a satellite of opportunity. More precisely, the "stored ionospheric model" is adaptively updated by resetting to zero the ionosphere-induced range error (by multiplicative correction) at the point in the three-dimensional space that corresponds to the position of the observed satellite and by extrapolating this correction (with the same multiplicative factor) to all other predictions of ionospheric range errors everywhere else in the volume under observation. For instance, let's assume the following situation:

a) The real-time sampling of the ionosphere by the pulse pair method indicates that at 0.6° above horizon and for a satellite down range distance of 3200 km the ionospheric range error is +40 ft.

Table 2

Cases Used In Evaluating Range Error Correction Procedure

MODELS USED IN SIMULATING DUAL-FREQUENCY RADAR MEASUREMENTS OF THE IONOSPHERE	RADAR PROPAGATION PATH WHERE PULSE MEASUREMENT IS PERFORMED			
	A) RADAR-TO-SATELLITE PATH CROSSES VOLUME AT OBSERVATION TIME	B) RADAR-TO-SATELLITE PATH CROSSES VOLUME 1/2 HOUR OUT OF TIME	C) RADAR-TO-SATELLITE PATH 40° ABOVE HORIZON. PASS OCCURS AT OBSERVATION TIME.	D) RADAR-TO-SATELLITE PATH 40° ABOVE HORIZON. PASS OCCURS 1/2 HOUR OUT OF TIME.
1. SPHERICALLY SYMMETRIC - MEDIAN	1A	1B	1C	1D
2. SPHERICALLY SYMMETRIC - ANOMALOUS	2A	2B	2C	2D
3. MODEL WITH HORIZONTAL GRADIENTS, 1σ FROM MEDIAN	3A	3B	3C	3D
4. MODEL WITH HORIZONTAL GRADIENTS, 2σ FROM MEDIAN	4A	4B	4C	4D
NOTE: THE "STORED IONOSPHERIC MODEL" IS A MEDIAN MODEL WITH HORIZONTAL GRADIENTS.				

b) The error prediction contained in the "stored ionospheric model" indicates for the same elevation and for a down range distance as above we have an error of +60 ft.

c) We conclude that the "stored ionospheric model" is too dense and we multiply the table of all predicted ionospheric range errors by 0.67.

d) We reset to zero (40 - 0.67 x 60 = 0) the error residual for the elevation and down range distance mentioned above and we adopt as the "updated estimate" of the ionospheric-induced errors in the entire radar volume a new table obtained from the old one with all the predictions multiplied by 0.67.

e) From all radar-measured ranges (apparent ranges) we now subtract the "updated" estimate of the ionospheric error in order to obtain the estimate of the true geometric range. In the computer simulation we use the same procedure to evaluate the range error residuals (difference between simulated pulse pair measurement of the ionospheric range error and the error prediction obtained from the updated ionospheric model) and we check whether these residuals are within the maximum allowed value (say, nominally a few feet).

Figures 2 through 6 provide some examples of the calculations that we have performed. Figure 2 is a Block (7) ionospheric error prediction printout computed with the "median model with realistic horizontal gradients" at an azimuth located at the center of the radar coverage volume for December 1980, Local Time 1200 Hours, and for the frequency of 1275 MHz. The numbers in Figure 2 represent the range errors due to the ionosphere, when no correction procedure is adopted. Figures 3 and 4 are two sets of Block (5) ionospheric error examples obtained from simulated radar data (Case 3A in Figure 3 and Case 3B in Figure 4). Figures 5 and 6 give the residual error after adaptive correction (Blocks 8 and 9 of Figure 1) for these two cases. It can be concluded that one adaptive updating every half hour done on satellites of opportunity, located at such points that the propagation path crosses the radar coverage volume, suffices to provide acceptable range error residuals for an ionosphere at a peak of the solar flare activity (SSN = 100) and 1σ away from median conditions.

3. MODEL'S SOFTWARE MECHANIZATION

3.1 General

The range correction model consists of three 16-element vectors [Allen, 1974]. Let's assume, as an example, a radar azimuth span from 308° to 325°T, an elevation angle span from 0.5° to 15°, and a height span from 0 to 1000 km. Ionospheric range corrections (let's call them ΔR) are required at each location in the volume. The value of ΔR is generated by multiplication of the three 16-element vectors producing 4,096 values of ΔR_{ijk}, with i, j, k indices representing 16-element matrices (i for height, j for elevation angle, and k for azimuth). ΔR_{ijk} is generated by the following equation:

$$\Delta R_{ijk} = (H_i \; E_j \; A_k) \; C_0(U)$$

where

H_i is a 16-element vector which represents the variation in height referenced to 1000 km at a constant elevation angle (0.5°) and constant azimuth (319°T);

E_j is a 16-element vector which represents the variation in elevation angle referenced to 0.5° at a constant azimuth (319°) and constant height (1000 km);

A_k is a 16-element vector which represents the variation in azimuth referenced to 319°T at a constant height (1000 km) and constant elevation angle (0.5°);

C_0 is a normalization constant which makes the largest value in the H_i vector (which occurs at 1000 km) equal to 1;

U is a factor which updates the model either from a real time (every half hour) pulse pair measurement from a satellite within the radar coverage or from an ionospheric update service such as the Air Weather Service (AWS) Prediction Service.

3.2 Height Vector

At 319°T (assumed to be boresite, $\Delta Az = 0$) and 0.5° elevation angle the range corrections at a frequency of 1275 MHz are determined in 16 steps in height. If we adopt a function I defined as

$$I = K\left[(1 + h/R)^2 - 1\right]$$

it can be shown that

$$I = K\left[(r/R)^2 + 2(r/R)\sin E\right]$$

where h = height of target above earth in kilometers, R is the radius of the earth in kilometers, K is the scaling constant such that I = 16 at h = 1000 km, r is the target range in kilometers, and E is the elevation angle. As r and sin E are normally available in a real time system, it is relatively straightforward to calculate I. Thus, storing the ionospheric corrections in I space reduces computational requirements. Table 3 shows the value of i, I and the corresponding height for 16 steps along an 0.5° elevation angle. The constant K = 47.25. Note that at h = 0 km, I = 0 and the ionospheric correction is zero. If in the calculation of I we have I = 5.4, the value of H at i = 5.4 is determined by linear interpolation between the value of H at i = 5 and i = 6. The constant C_0 is equal to $1/H_{16}$ and all values of H are divided by H_{16} before entry into the H_i vector.

Table 3

Height Vector H_i

i	I	h (km)	H_i*
	0	0	
1	1	67.05	
2	2	133.4	
3	3	199.1	
4	4	264.2	
5	5	328.6	
6	6	392.4	
7	7	455.6	
8	8	518.2	
9	9	580.2	
10	10	641.7	
11	11	702.7	
12	12	763.2	
13	13	823.1	
14	14	882.5	
15	15	941.5	
16	16	1000.0	

*H_i generated by model.

3.3 Elevation Vector

The elevation vector E are arrayed in (sin E) space where Sin E = v cos E_0 + w sin E_0 (E_0 is the tilt of the array from zenith). The 16 elements of the vector are generated in equal increments of sin space from 0.5 (j = 1) to 15 (j = 16) degrees. Table 4 lists the element number "j", the elevation and sin E for the 16 elements. If

the calculated sin E from the radar data falls between vector elements, linear interpolation between vector elements are used.

Table 4

Elevation Vector E_j

(j) Vector Element	E	sin E	E_j*
1	0.50	0.00873	
2	1.46	0.02540	
3	2.41	0.04207	
4	3.37	0.05875	
5	4.33	0.07542	
6	5.28	0.09209	
7	6.24	0.10876	
8	7.21	0.12544	
9	8.17	0.14211	
10	9.14	0.15878	
11	10.11	0.17545	
12	11.08	0.19213	
13	12.05	0.20880	
14	13.03	0.22547	
15	14.01	0.24215	
16	15.00	0.25882	

*E_j generated by model.

3.4 Azimuth Vector

The azimuth vector A_K [Kahrilas, 1976] is arrayed in (Tan ΔAz) space where

$$\text{Tan } \Delta Az = \frac{u}{w \cos E_0 - v \sin E_0}$$

and

$$\Delta Az = Az_0 - Az \quad (Az_0 = 319^\circ T)$$

The 16-element azimuth vector is generated in equal steps of Tan ΔAz at 0.5° elevation angle (j = 1) and h = 1000 km (i = 16). Table 5 lists the 16 elements of the azimuthal vector including the element number "k", the true azimuth, the difference between true and boresite azimuth, and the tan ΔAz. If the calculated Tan ΔAz from the radar data falls between vector elements, linear interpolation between vector elements is used.

Table 5

Azimuth Vector A_K

K	Az (°T)	ΔAz (°)	Tan ΔAz	A_K*
1	324.6	-5.6	-0.09719	
2	323.4	-4.4	-0.07775	
3	322.3	-3.3	-0.05831	
4	321.3	-2.2	-0.03888	
5	320.1	-1.1	-0.01944	
6	319.0	0	0	
7	317.9	1.1	0.01944	
8	316.8	2.2	0.03888	
9	315.7	3.3	0.05831	
10	314.6	4.4	0.07775	
11	313.4	5.6	0.09719	
12	312.3	6.7	0.11663	
13	311.3	7.7	0.13607	
14	310.2	8.8	0.15550	
15	309.1	9.9	0.17494	
16	308.0	11.0	0.19438	

*A_K generated by model

3.5 Update Procedure

In our study we have considered two update procedures, one adaptive and one non-adaptive. In both cases a single value multiplicatively updates the values of ΔR_{ijk} for all i, j and k. The first update procedure is based on the use of a pulse pair (dual frequency) measurement of the ionosphere by using a satellite of opportunity as the target. This update procedure must be repeated every half-hour. The second procedure is a non-adaptive update approach based on data provided by an ionospheric update service such as AWS. If the dual-frequency measurement is not available or otherwise unacceptable for whatever reason, the AWS update is used instead. In addition, the AWS update is used to ensure that the pulse pair (dual frequency) update does not generate unreasonable corrections.

The update algorithm for the pulse pair (Upp) is

$$Upp = \frac{\Delta b}{\Delta R_{ijk}}$$

where

$$\Delta b = \frac{(R_L - R_H)\, f_H^{\,2}\, f_L^{\,2}}{f_{OP}^{\,2}(f_H^{\,2} - f_L^{\,2})}$$

R_L is the apparent range to satellite at f_L

R_H is the apparent range to satellite at f_H

f_{OP} is the operating frequency of the radar (1275 MHz)

Δb is the ionospheric correction factor obtained directly from the pulse pair measurements and ΔR_{ijk} is the prediction of the model at that location.

The Air Weather Service (AWS) provides an update U_{AWS} on a three-hour basis. This number is used in the same manner as the number generated from the dual-frequency measurement except that it requires an operator-initiated action. The update is defined as the ratio U_{AWS} where

$$U_{AWS} = \frac{\text{AWS Predicted TEC}}{\text{Median TEC}}$$

where TEC is the vertical Total Electron Content. The update U_{AWS} is provided for 210°W longitude and 63°N latitude, which is in the example that we have chosen the "nominal ionospheric characteristic location" for the radar coverage.

3.6 Software Structure for the Ionospheric Range Correction Procedure

For each hour of the day (24 hours per day) seven computer cards form the set of inputs required by the model. The 24 sets of seven computer cards are valid for each day of the month. Each new month requires a new set of cards. Figure 7 shows a listing of the seven computer cards for April 1976 at 0000Z. The first card is a header card which specifies the month, year, hour (in universal time) and the constant C_0 in feet. The last 8 columns of each card (columns 73-80) contain the year, month, hour and card number with the header card, Card No. 00. The next two cards contain the height vector H_i. The location of H_i for i = 1 to 16 is noted on Figure 7. In like manner cards 3 and 4 contain the elevation vector E_j in sin space and cards 5 and 6 contain the azimuth vector A_K in tan space.

Figure 8 is a listing of the computer cards which provide the ionospheric range error model for the month of June 1976. In addition to the deck of computer cards represented by Figure 8, a plot of range error versus down range and height above the earth is also provided. One such plot is provided for each hour, or a total of 24 plots for each month.

4. EXPERIMENTAL EVIDENCE ON CELL STRUCTURE SIZE

The residual error after using the adaptive update of the model will have both a spatial and a temporal variability. Since the radar range error is directly related to the total electron content (TEC) encountered along the radar line of sight, that error can be estimated from measurements of TEC from archive sources.

4.1 Spatial Variability

An example of the potential of this specification technique to remove spatial variability is shown in Figure 9. Differential Doppler records taken at Millstone Hill, Mass. during 1972 using one of the TRANSIT satellites of the Navy Navigation Satellite System were reduced to equivalent vertical electron content (Nicholas Tomjanovich, preliminary data). A standard Air Force prediction program, used to estimate the median value of TEC along the satellite path, and separate measurements of TEC from Hamilton, Mass. and from Goose Bay, Labrador, using Faraday rotation measurements of ATS-3 signals at the time of the pass, were compared with the observations. There is a general fit of the latitude gradients along the path for both the median prediction and the two sets of observations.

From the TRANSIT pass, the value of TEC at closest approach was used to scale the median prediction at this point, and this corrective factor was applied to the median prediction along the entire path. The resultant normalized prediction has a spatial variability of less than 5 percent within a few degrees of the central scaling region. The maximum deviation increases with distance along the path, to the order of 10 percent at the extremes. This is a consistent daytime result (Figure 10) for a representative sample of five TRANSIT passes from five different months selected by computer to fall within ±0.5 degrees of longitude of the subionospheric point of the ATS-3 measurements made from Hamilton, Mass.

These initial results for specification of the daytime ionosphere indicate that the proposed technique of adaptive modeling of the median range error by normalization to the central region of a satellite pass-track can produce a tenfold reduction of the variability expected from monthly median observations.

The median range errors expected during night hours are about a magnitude less than those expected during the day. Although the day-to-day percentage variability about the median is only slightly worse for night hours (30 percent) than for day hours (20 percent), sharply localized features in the ionosphere may make it just as necessary to use the adaptive model during the night.

Figure 11 illustrates the significant local features which frequently occur in the night ionosphere. The steep gradient of electron content found over small latitude intervals along the TRANSIT path is confirmed by the gradient measured simultaneously with the ATS-3 beacon at Hamilton, Mass. and at Goose Bay, Labrador. The prediction of the monthly median has not be normalized to the observations. The model predicts only a slight gradient at this time, but the absolute error at nighttime is of the same order as the absolute daytime error using the scaled median prediction. We have not had sufficient data to determine whether scaling at nighttime will improve the prediction. If the observed gradients are a consistent nighttime feature, then scaling could introduce new errors and an improvement in prediction would have to result from improved modeling of these features.

4.2 Temporal Variability

In parallel with the spatial variability, the temporal variability, using a similar scheme for scaling the median prediction, is being examined. The values of TEC, as determined from measurements of ATS-3 signals taken at Hamilton, Mass. in 15-minute intervals for the year 1972, have been used for the local observations in the initial phase of this study. The effectiveness of the median 10-day prediction was examined by comparing the standard deviation of the observations from the monthly mean of the observations with the standard deviation of the observations from the median prediction. The error is equivalent in both cases, one of the indications that the observed mean is being successfully modeled.

To simulate adaptive updating of a model ionosphere, the ATS-3 measurements in TEC units and the median prediction of TEC were used to determine a scaling factor at each hourly observation. The resultant scaling factor was then used to scale the prediction at 15-minute intervals for the succeeding 12 hours. The monthly mean of this scaled prediction was determined, and the standard deviation of the observations from the new prediction was calculated. Figure 12 shows the error, in TEC units, for the hourly-based 12-hour predictions for 2-hour intervals for the month of September 1972, with the error from the monthly mean of the observations and the median prediction. It is apparent that throughout most of the day and nighttime hours a local measurement will improve the median prediction for several hours, with reductions in error of 50% or more still possible two or three hours after normalizing the model. A constraint in this scheme occurs in the pre-dawn period when the error for a prediction based on any measurement, including one taken in the period near sunrise, rapidly exceeds the error using a median prediction. Except for this brief period, the error using a scaled

prediction is usually less than, and does not exceed, the median prediction error even 12 hours after normalization.

The potential effectiveness of this technique depends on the time interval between local TEC measurements. Figures 13 and 14 summarize the results of predictions based on 30-minute and 2-hour observations respectively, for the year 1972. The dashed lines approximate sunrise and sunset, the contours are the rms percent error. To consistently reduce the rms error to less than 5% for most of the daytime, measurements would have to be made in intervals no greater than 30 minutes. The error rarely exceeds 10% at any time other than the period near sunrise, but even at these times, the absolute error does not exceed the maximum daytime error.

For a prediction based on a 2-hour observation, there is still a significant reduction in error. The rms error is less than 15% for most of the daytime hours and near sunrise it exceeds the median prediction error, but for most of the year it is less than the 20 to 25% expected rms error using a median climatology.

The comparison of the TRANSIT passes with the scaled median prediction has shown that a reduction in error to 5 percent is possible in daytime, within a few degrees latitude of the normalization point. This is equivalent to the daytime reduction in rms error using 30-minute observations to scale the median prediction. The daytime error in spatial variation increases to about 10% over 15° latitude, which is comparable to the 10% rms error in daytime using a median prediction scaled with hourly observations. Initial indications are that spatial and temporal variations are comparable in magnitude when considered on a mean basis. Also, since the TRANSIT passes are effectively north-south, and the temporal variations can be considered east-west, it appears that direction is not a primary consideration in determining variation on a mean basis.

5. CONCLUSIONS

The problem of correcting the range errors induced by the propagation in the ionosphere of electromagnetic waves used by navigation and positioning systems is acutely felt by numerous communities of investigators and users.

The dispersivity of the ionosphere is of substantial help in solving this problem, when dual-frequency observations are feasible and practical. Models of the ionosphere are also of some use when only single-frequency observations are available.

When possible, an adaptive combination of dual-frequency measurements and models based on climatology represents a solution with relevant practical advantages. Time commitment of the radar to perform ionospheric measurements can be reduced to a minimum so that most of the radar time can be devoted to its fundamental mission of target observation. A stored model of the ionosphere provides the basic data for continuous range error correction. The dual-frequency observations update the model, possibly and preferably performing this function at times that are not at premium (when no important targets are in sight). Our research has shown that residual range errors of a few feet are attainable by performing dual-frequency model updating, only once every half-hour or so.

An approach of adaptive model updating such as the one described in this paper appears to these authors to show a good potential not only for the correction of ionospheric errors, but for the correction as well of all sorts of propagation errors. This is especially true for cases in which the model of the medium would not be able to provide by itself alone the required accuracy, while the radar itself could not devote any sizeable portion of its operation time to medium sampling.

6. REFERENCES

ALLEN, R.S., 1974, An Ionospheric Correction Model for the Cobra Dane Program, AFCRL Letter LIR/x-2943, 16 December.

KAHRILAS, P.J., 1976, Electronic Scanning Radar Systems (ESRS) Design Handbook, Chapter 7, ARTECH House, Inc.

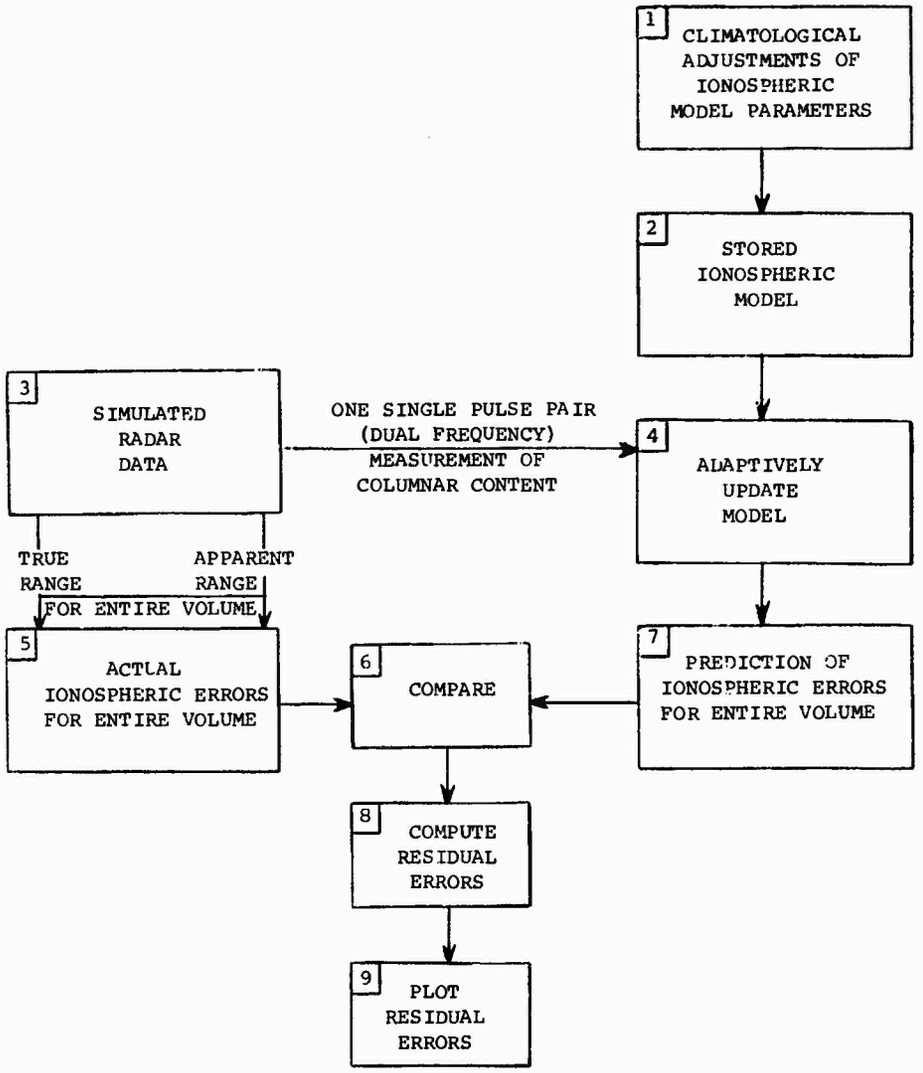


Figure 1 Ionospheric Error Correction Simulation - Simplified Flow Chart.

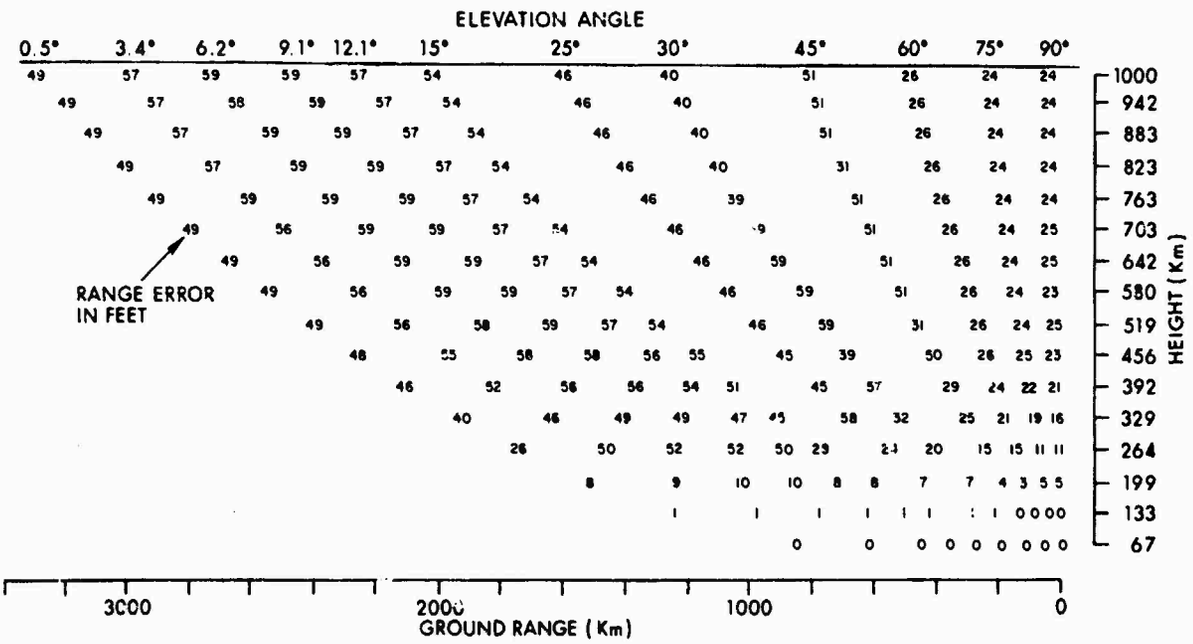


Figure 2 Stored Ionospheric Model (Median Model for Ionospheric Range Errors, LT 1200, December 1980, Including Realistic Horizontal Gradients).

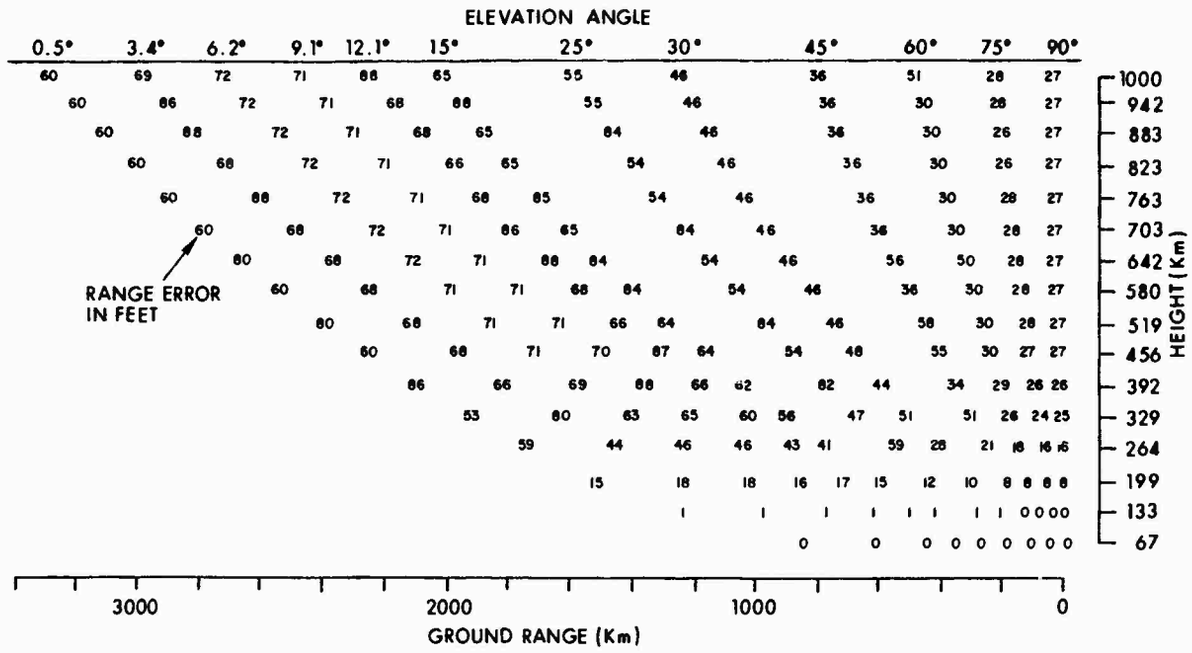


Figure 3 Simulated Radar Data (Ionospheric Range Error) for Case 3A (see Table 1).

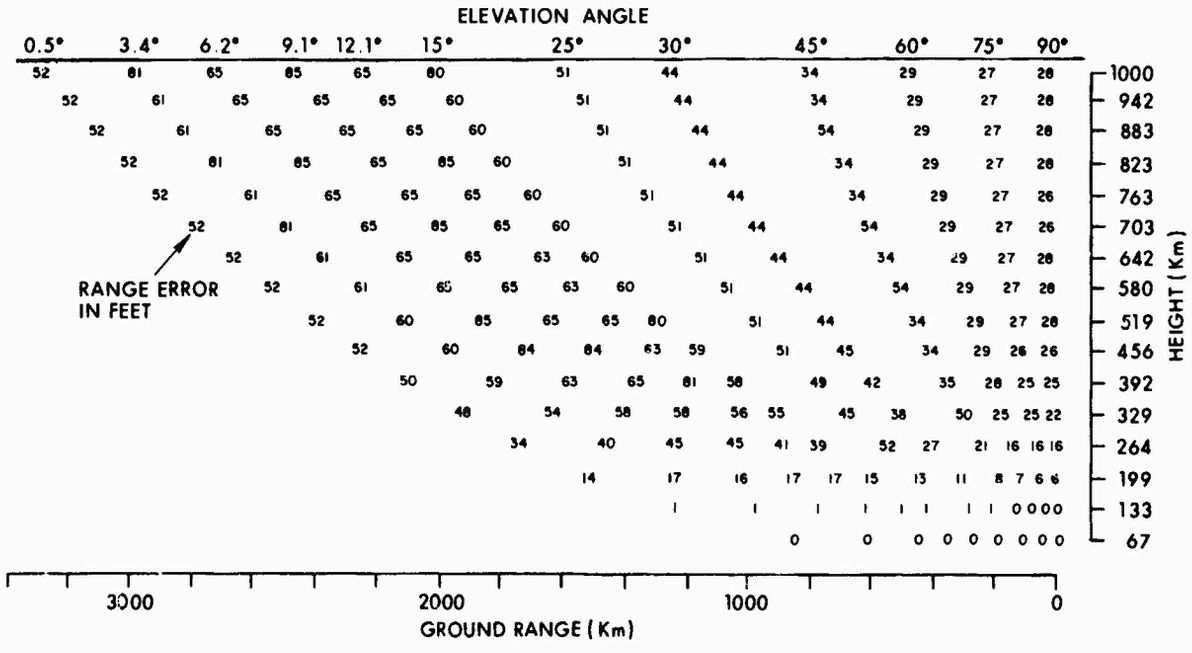


Figure 4 Simulated Radar Data (Ionospheric Range Error) for Case 3B (see Table 1).

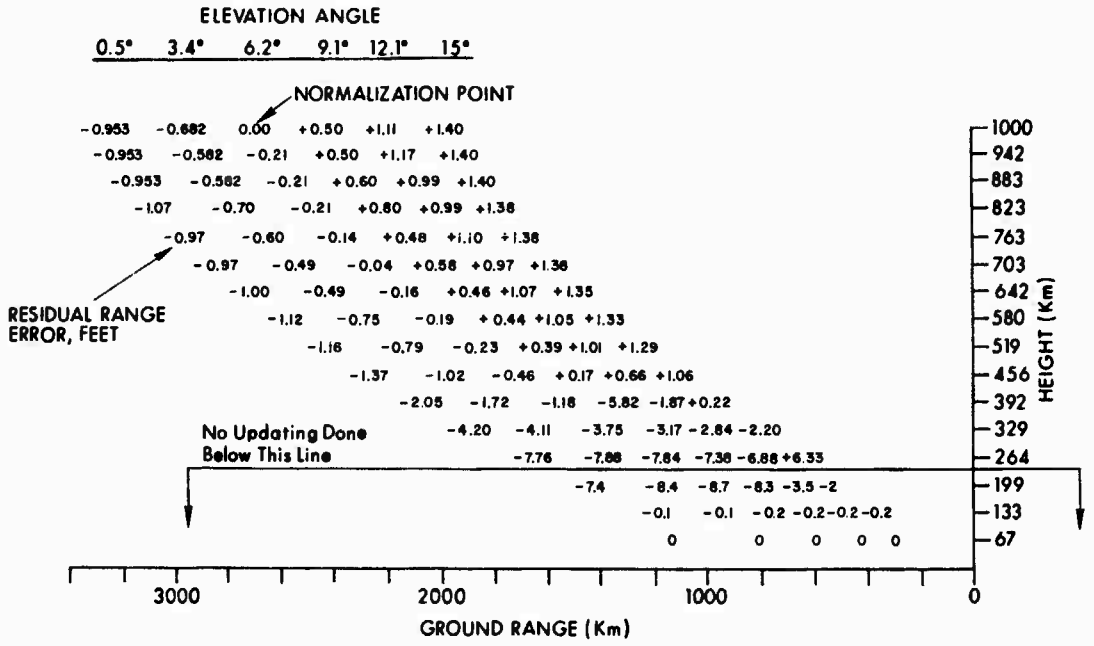


Figure 5 Residual Ionospheric Range Errors for Case 3A (see Table 1) After Adaptive Correction (by simulation of radar dual-frequency measurements).

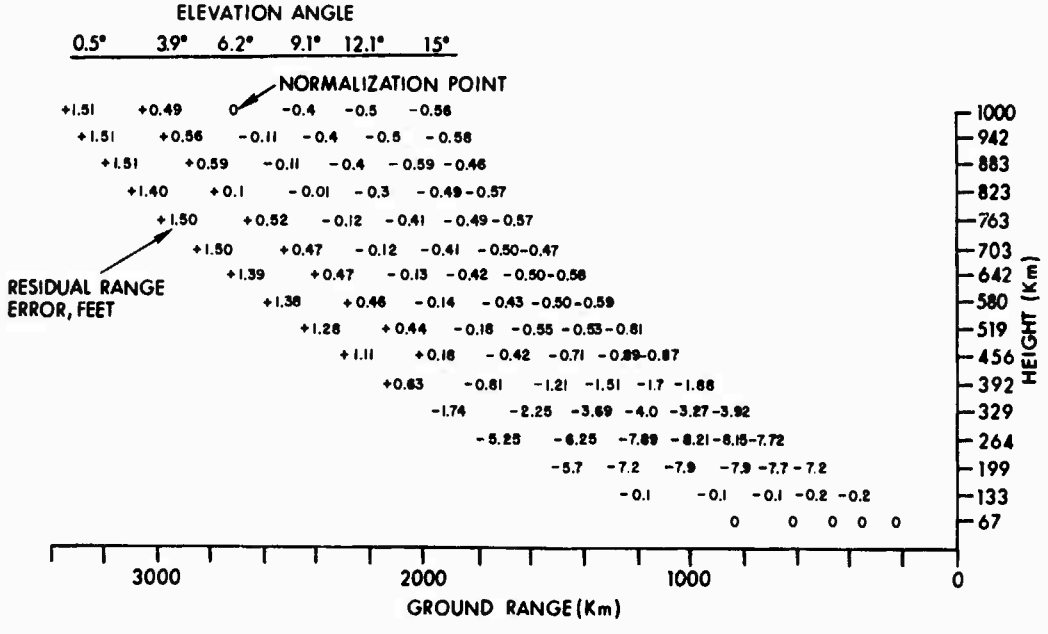


Figure 6 Residual Ionospheric Range Errors for Case 3B (see Table 1) After Adaptive Correction (by simulation of radar dual-frequency measurements).

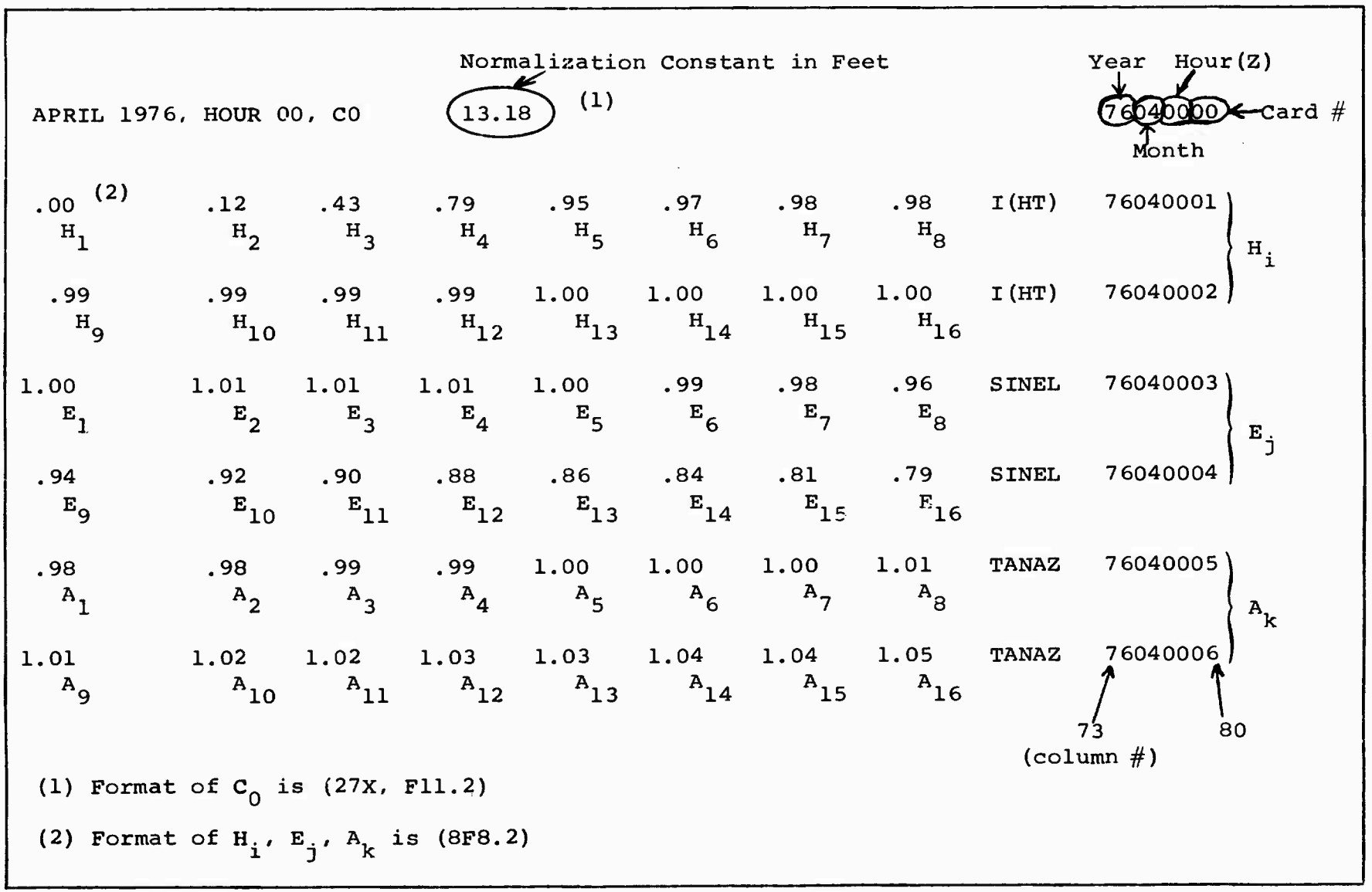
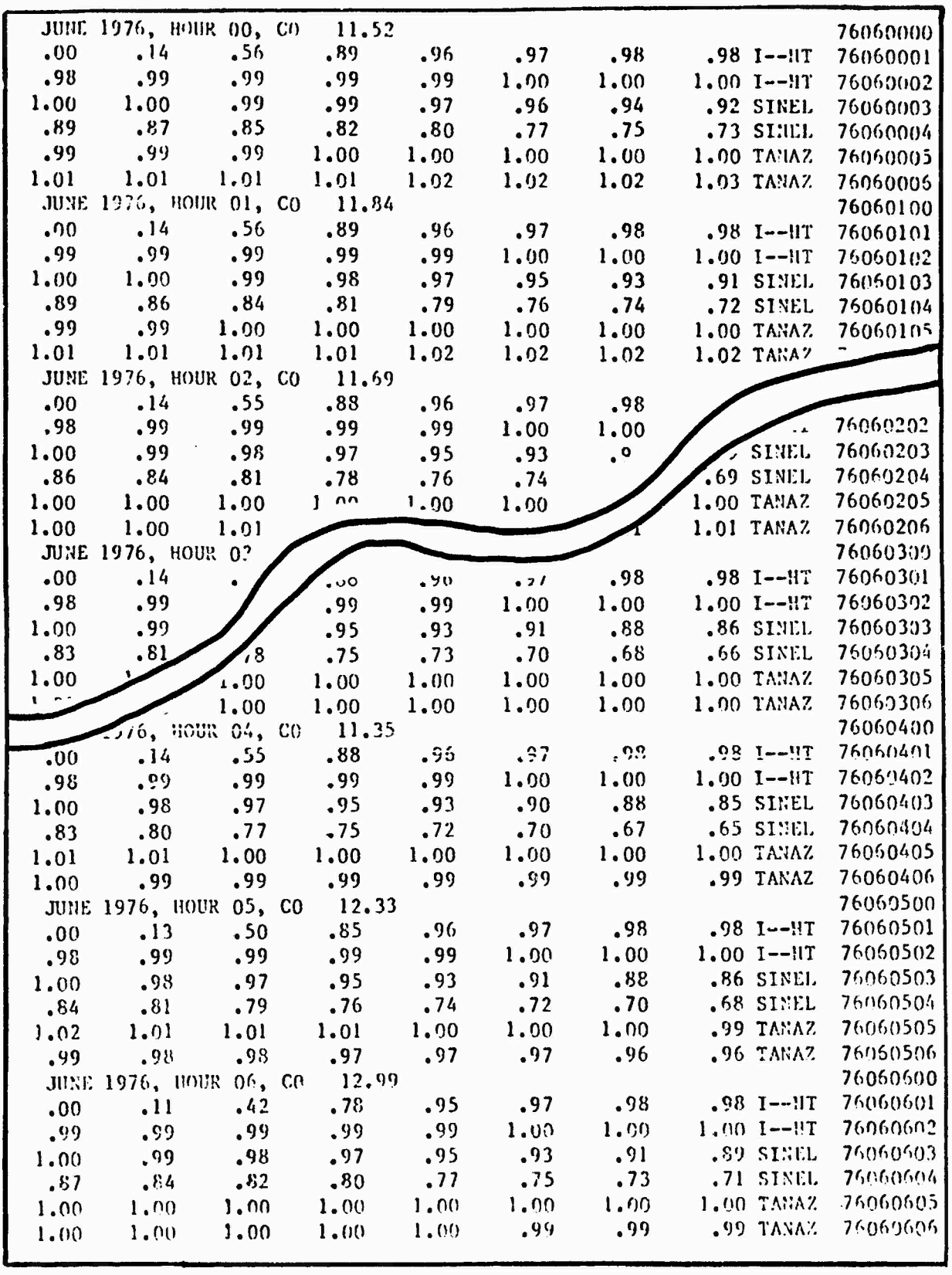


Figure 7 Listing of Computer Cards for One Hour in April 1976.



```
 JUNE 1976, HOUR 00, CO   11.52                                          76060000
 .00     .14     .56     .89     .96     .97     .98     .98 I--HT       76060001
 .98     .99     .99     .99     .99    1.00    1.00    1.00 I--HT       76060002
1.00    1.00     .99     .99     .97     .96     .94     .92 SINEL       76060003
 .89     .87     .85     .82     .80     .77     .75     .73 SINEL       76060004
 .99     .99     .99    1.00    1.00    1.00    1.00    1.00 TANAZ       76060005
1.01    1.01    1.01    1.01    1.02    1.02    1.02    1.03 TANAZ       76060006
 JUNE 1976, HOUR 01, CO   11.84                                          76060100
 .00     .14     .56     .89     .96     .97     .98     .98 I--HT       76060101
 .99     .99     .99     .99     .99    1.00    1.00    1.00 I--HT       76060102
1.00    1.00     .99     .98     .97     .95     .93     .91 SINEL       76060103
 .89     .86     .84     .81     .79     .76     .74     .72 SINEL       76060104
 .99     .99    1.00    1.00    1.00    1.00    1.00    1.00 TANAZ       76060105
1.01    1.01    1.01    1.01    1.02    1.02    1.02    1.02 TANAZ       7
 JUNE 1976, HOUR 02, CO   11.69
 .00     .14     .55     .88     .96     .97     .98
 .98     .99     .99     .99     .99    1.00    1.00                     76060202
1.00     .99     .98     .97     .95     .93     .9      SINEL           76060203
 .86     .84     .81     .78     .76     .74             .69 SINEL       76060204
1.00    1.00    1.00    1.00    1.00    1.00            1.00 TANAZ       76060205
1.00    1.00    1.01                    1               1.01 TANAZ       76060206
 JUNE 1976, HOUR 0?                                                      76060300
 .00     .14     .                .00     .90     .97     .98     .98 I--HT       76060301
 .98     .99                     .99     .99    1.00    1.00    1.00 I--HT       76060302
1.00     .99                     .95     .93     .91     .88     .86 SINEL       76060303
 .83     .81             .78     .75     .73     .70     .68     .66 SINEL       76060304
1.00             1.00    1.00    1.00    1.00    1.00    1.00    1.00 TANAZ       76060305
                 1.00    1.00    1.00    1.00    1.00    1.00    1.00 TANAZ       76060306
          976, HOUR 04, CO   11.35                                          76060400
 .00     .14     .55     .88     .96     .97     .98     .98 I--HT       76060401
 .98     .99     .99     .99     .99    1.00    1.00    1.00 I--HT       76060402
1.00     .98     .97     .95     .93     .90     .88     .85 SINEL       76060403
 .83     .80     .77     .75     .72     .70     .67     .65 SINEL       76060404
1.01    1.01    1.00    1.00    1.00    1.00    1.00    1.00 TANAZ       76060405
1.00     .99     .99     .99     .99     .99     .99     .99 TANAZ       76060406
 JUNE 1976, HOUR 05, CO   12.33                                          76060500
 .00     .13     .50     .85     .96     .97     .98     .98 I--HT       76060501
 .98     .99     .99     .99     .99    1.00    1.00    1.00 I--HT       76060502
1.00     .98     .97     .95     .93     .91     .88     .86 SINEL       76060503
 .84     .81     .79     .76     .74     .72     .70     .68 SINEL       76060504
1.02    1.01    1.01    1.01    1.00    1.00    1.00     .99 TANAZ       76060505
 .99     .98     .98     .97     .97     .97     .96     .96 TANAZ       76060506
 JUNE 1976, HOUR 06, CO   12.99                                          76060600
 .00     .11     .42     .78     .95     .97     .98     .98 I--HT       76060601
 .99     .99     .99     .99     .99    1.00    1.00    1.00 I--HT       76060602
1.00     .99     .98     .97     .95     .93     .91     .89 SINEL       76060603
 .87     .84     .82     .80     .77     .75     .73     .71 SINEL       76060604
1.00    1.00    1.00    1.00    1.00    1.00    1.00    1.00 TANAZ       76060605
1.00    1.00    1.00    1.00    1.00     .99     .99     .99 TANAZ       76060606
```

Figure 8 Example of Listing of Computer Deck for Ionospheric Correction (June 1976).

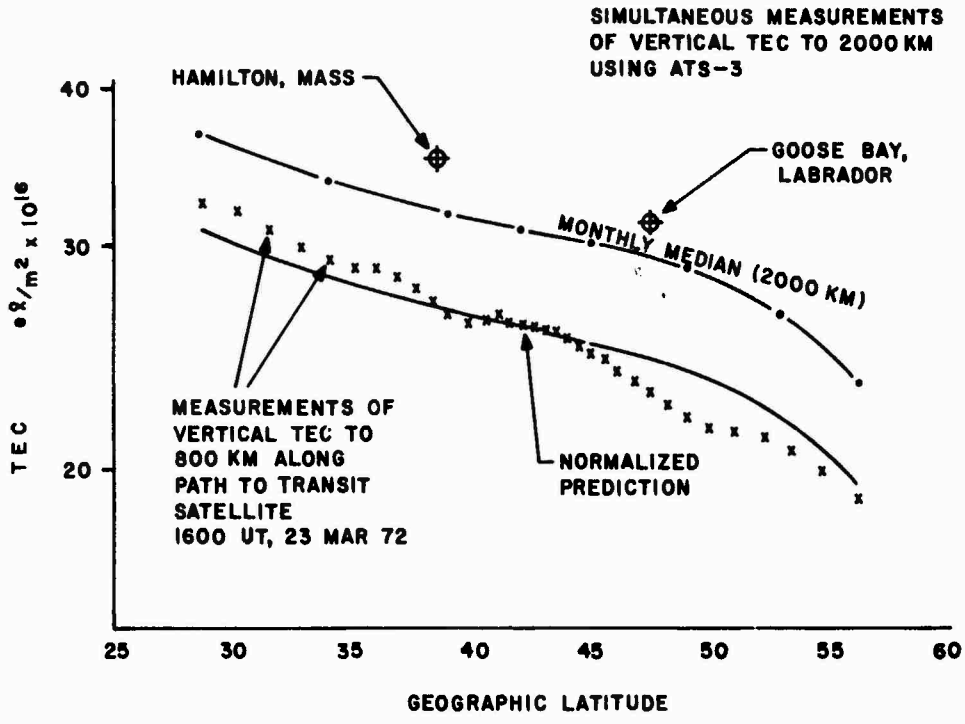


Figure 9 Specification of Median Latitude Variation of Vertical Content by a Local Measurement.

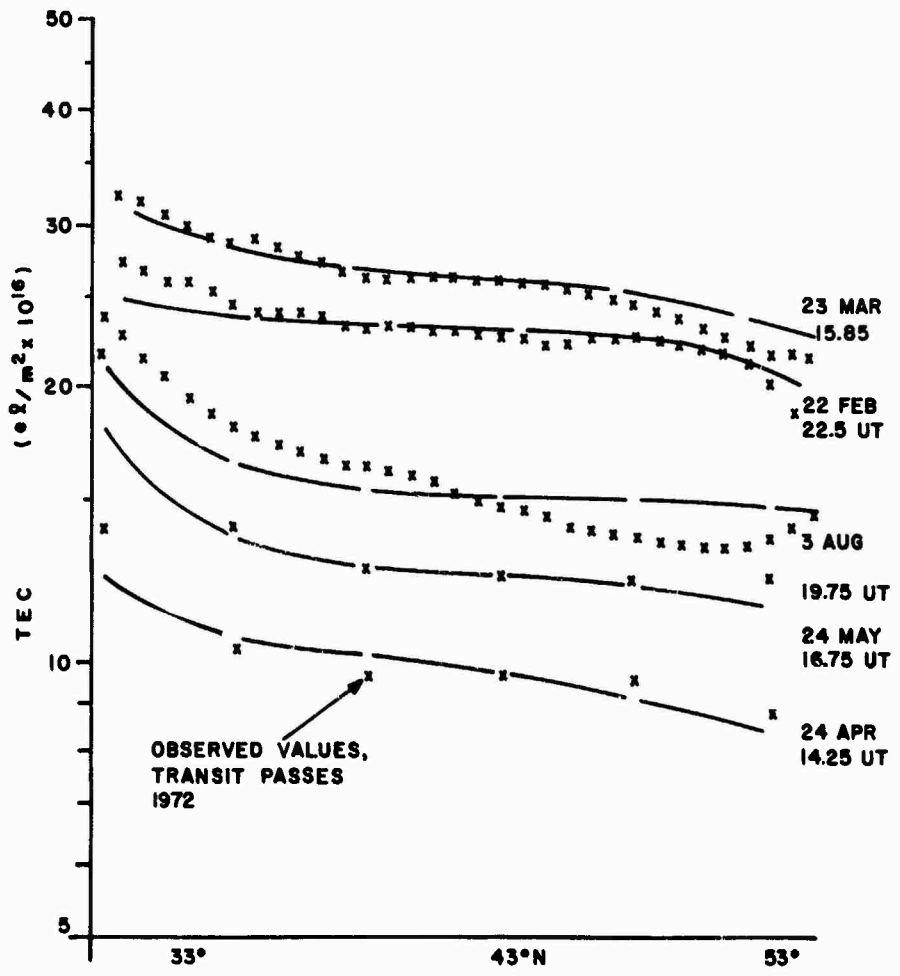


Figure 10 Comparison of Normalized Prediction of Median Gradient with Observed Total Content to 800 km.

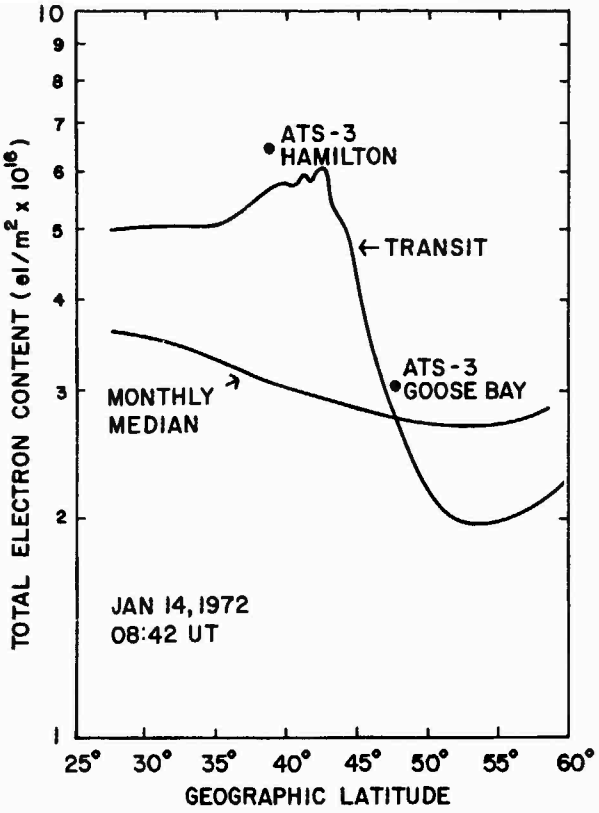


Figure 11 Comparison of Predicted Median Gradient with Observed Nighttime Gradient in Total Content

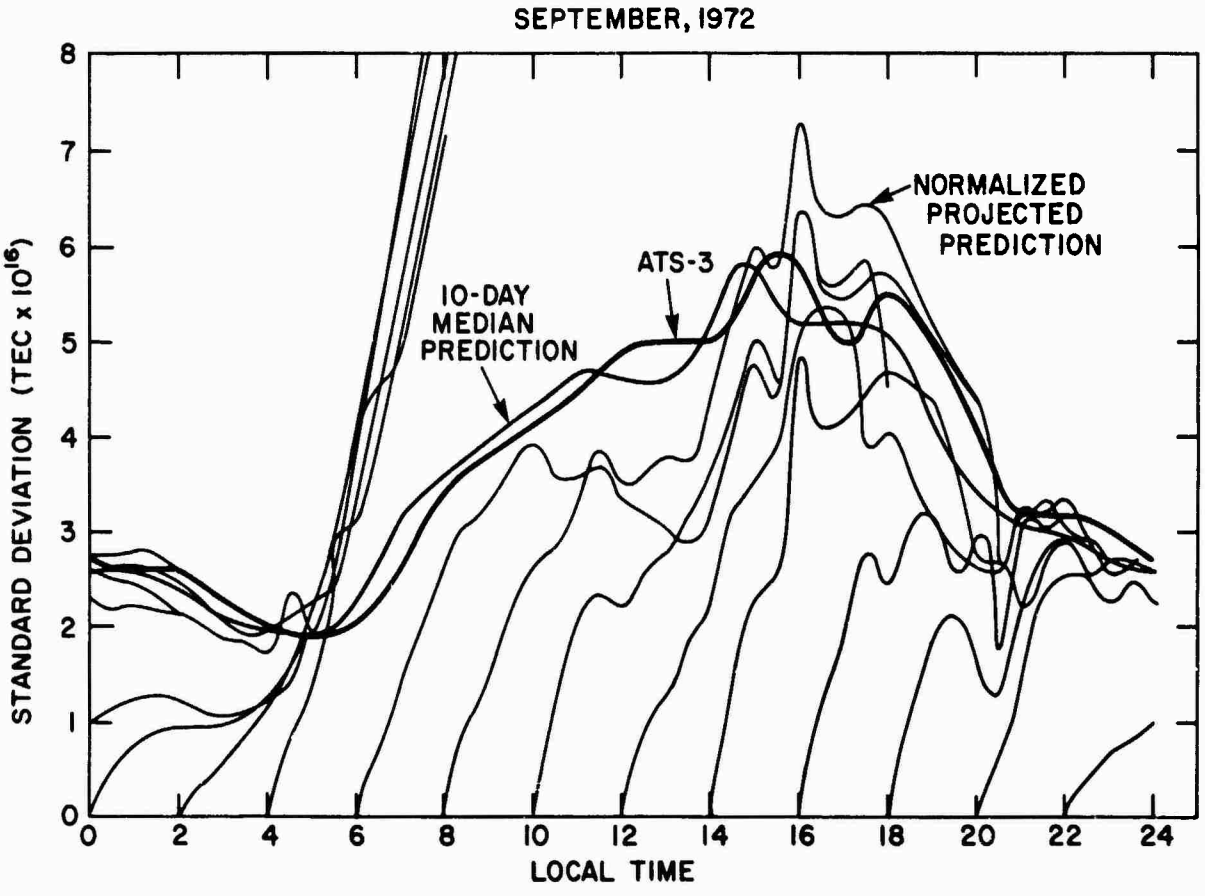


Figure 12 Standard Deviation from Observed Values of Total Electron Content (ATS-3 Observations, September 1972).

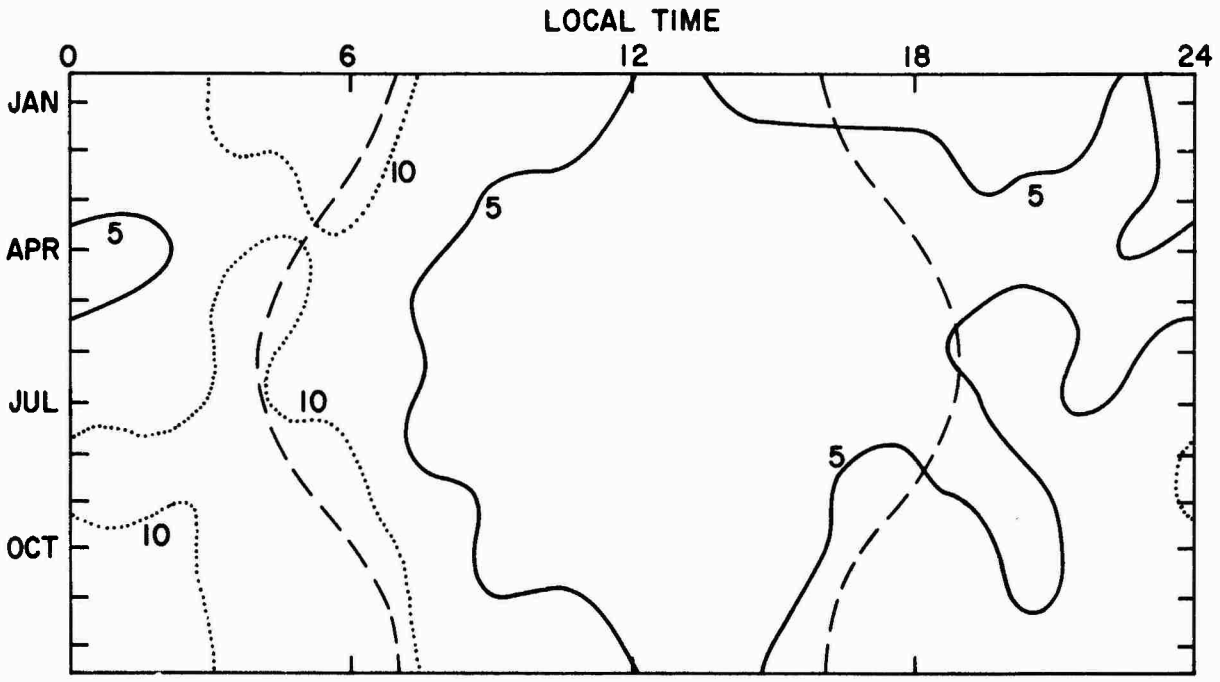


Figure 13 Map of the RMS Percentage Error in Predicting Ionospheric Range Effects Based On Observations Made 30 Minutes Ahead of Time (scaled 10-day median).

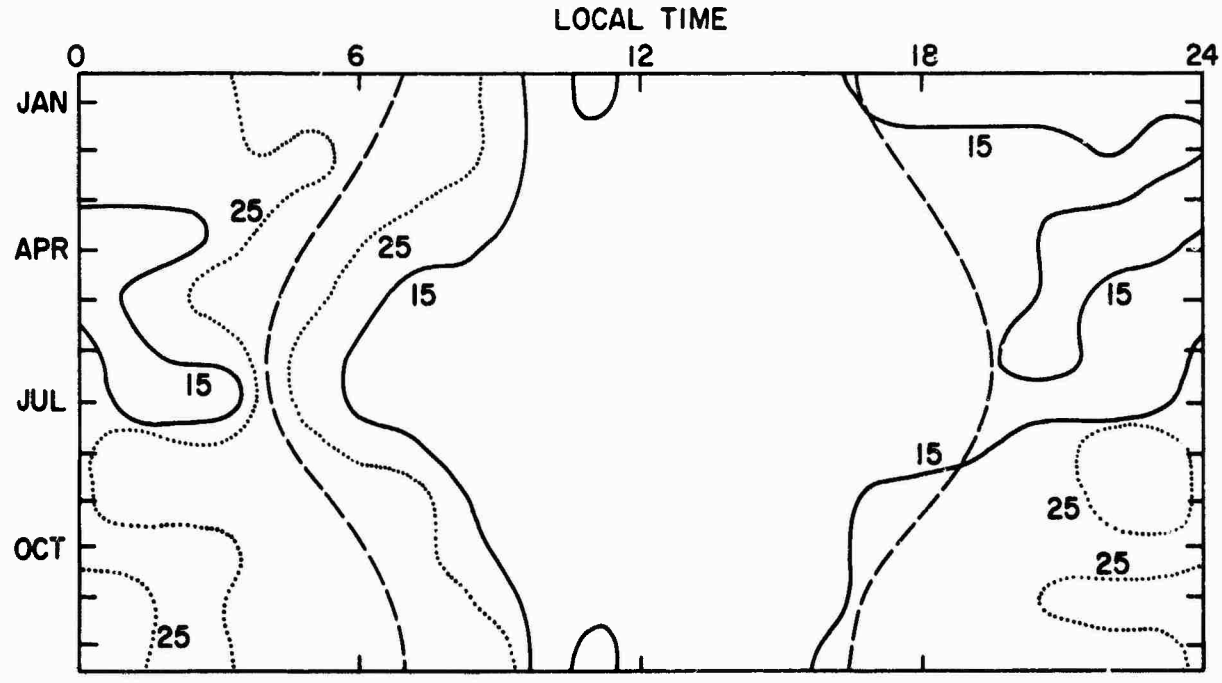


Figure 14 Map of the RMS Percentage Error in Predicting Ionospheric Range Effects Based On Observations Made 2 Hours Ahead of Time (scaled 10-day median).

THE EFFECTS OF RADIO LENSES IN THE IONOSPHERE ON THE SCINTILLATION OF SATELLITE-TO-GROUND RADIO SIGNALS

Kenneth Davies*
Space Environment Laboratory
National Oceanic and Atmospheric Administration
Boulder, Colorado 80302, U.S.A.

and

J. D. Whitehead
Department of Physics
University of Queensland
St. Lucia, 4067, Australia

ABSTRACT

Radio signals from geostationary satellites are observed to scintillate in amplitude and phase as the result of traversing the ionosphere. These scintillations affect the operation of radio systems at very high and ultra high frequencies.

The radio beacon on board the geostationary satellite ATS6 emits phase coherent signals on frequencies near 40, 140 and 360 MHz and observations of the beacon (amplitude and phase) were made at Boulder from June 1974 to May 1975. A particularly interesting phenomenon was observed characterized by the following features:

(1) An almost total fade out of the amplitudes of the 40, 140 and 360 MHz signals.

(2) Asymmetric Fresnel-type fading of the signals on both sides of the deep fades, with periods of the order of 1 sec.

(3) A distinctive modulation of the diffraction (fading) pattern.

The peculiar diffraction patterns have been simulated by the use of cylindrical lenses in the ionosphere which give Gaussian distributions of phase advance. To produce the observed modulations and deep phases the lenses must have cross sections with radii of the order of 100 m and intense core densities. The core of the lens may be overdense for the 40 MHz signal.

The asymmetric fading pattern may be caused by one of two effects:

(a) An asymmetric cross section.

(b) A cylinder expanding in time.

To examine the effects of the lens on gigahertz frequencies fading patterns have been calculated for radio frequencies of 720, 1440 and 2880 MHz using appropriate scaling rules.

Calculations have also been made with a focussing (positive) lens on 40 MHz and 140 MHz.

1. INTRODUCTION

The amplitude and phase fluctuations of radio signals passing through the ionosphere imposes limitations on the available bandwidth. These fluctuations are produced by ionospheric irregularities which modify the shape of the wavefront. When the deformation of the wavefront is small the ionosphere can be treated as a weak diffraction screen. When the phase variation across the wavefront exceeds about one radian the ionosphere can no longer be treated as a weak diffraction screen and the mathematical analysis in this case becomes extremely complicated, especially in the general case of an irregular screen. The present paper deals with a strongly diffracting ionospheric irregularity.

In recent years many instances have been reported of regular (or quasi-periodic) amplitude fluctuations produced by isolated ionospheric plasma blobs or lenses. For example, Singleton (1964) has investigated the focussing of radio-star signals while Ireland and Preddy (1967), Elkins and Slack (1969) and Slack (1972) observed quasi-periodic fading of radio signals from both orbiting and geostationary satellites. These authors have explained the fading phenomenon as the interference of refracted signals. Titheridge (1971) has used the Kirchoff integral to calculate the diffraction pattern on the ground, caused by cylindrical lenses that produce Gaussian distributions of phase across the wavefront.

In the work reported here we discuss a particular feature sometimes observed with the quasi-periodic fading, namely the modulation of the amplitude pattern. We shall show that such modulated patterns can be produced by small scale (less than one Fresnel radius) irregularities. A review of small scale structure of the ionosphere has recently been published by Pope and Rufenach (1975) with special emphasis on the effects on the scintillation of SHF waves. We shall calculate appropriate diffraction patterns for some super-high frequencies in section 4.

---

*This work was carried out while the author held the H. C. Webster Fellowship, Department of Physics, University of Queensland, St. Lucia, Qld., Australia, 4067.

2. THE ATS6 RADIO BEACON EXPERIMENT

The particular observation to be discussed here is illustrated in Figure 1 which contains the amplitude records of 40, 140 and 360 MHz signals emitted by the ATS6 geostationary satellite (0°N, 94°W) and received at Boulder, Colorado. Further details of the radio experiment have been published elsewhere (see Davies et al., 1975 a,b). The recorded signals are circularly polarized and, therefore, free of Faraday rotation. The quasi-periodic patterns are characterized by a deep central minimum and a somewhat asymmetric pattern in which the fading frequency increases away from the centre. The event shown in Figure 1 is rather unusual in that there is almost a complete fade-out even on 360 MHz. The presence of the modulated envelopes can be seen in other published records as for example Ireland and Preddy (1967, Fig. 1(d)), Kelleher and Martin (1975, Fig. 1), and Slack (1972, Fig. 1) but they have not been emphasized. A simulation of some of these records has recently been published by Herron (1976). The particular event considered here has also been studied by Pope and Rufenach (1975).

3. LENS MODELS

In our studies of the modulation of the fading pattern we have used an extension of the classical Cornu spiral (see Jenkins and White, 1957, Chapter 18). In particular, in their Figure 18V ($\Delta v = 0.5$), Jenkins and White show that a modulated diffraction pattern is produced by a very narrow opaque strip. It is important to realize that, whereas the regular (unmodulated) fading is produced by the interference of two signals, three signals are required to produce the modulation.

The cylindrical lens used in the present study is shown in Figure 2 and produces a Gaussian distribution of phase ϕ across the emergent wavefront, i.e.

$$\phi = \phi_o \exp \left\{ - \frac{(R - R_o)^2}{2\sigma^2} \right\} \tag{1}$$

where ϕ_o is the phase advance for the wave passing through the centre of the lens located at R from the perpendicular from the observing point P to the incident wavefront. R is the distance along the emerging wavefront and σ is the size of the lens. R, R_o and σ are measured in terms of the radius of the first Fresnel zone (see eq. 3). Provided there are no discontinuities in the emergent wavefront there will be an odd number of rays arriving at O. To fit the observations the unrefracted ray (1) dominates except near the centre of the pattern. Modulation of the pattern is the result of phase differences between rays (2) and (3).

In the case of the cylindrical lens the in-phase and quadrature amplitude components X and Y are obtained from the modified Fresnel integrals:

$$X = \int_{-\infty}^{+\infty} \cos \left[\frac{\pi}{2} R^2 + \phi_o \exp \left\{ - \frac{(R - R_o)^2}{2\sigma^2} \right\} \right] dR \qquad \text{(a)}$$

$$Y = \int_{-\infty}^{+\infty} \sin \left[\frac{\pi}{2} R^2 + \phi_o \exp \left\{ - \frac{(R - R_o)^2}{2\sigma^2} \right\} \right] dR \qquad \text{(b)} \tag{2}$$

Hence, the amplitude A is given by $\sqrt{X^2 + Y^2}$.

The integration of equation (2) was carried out by computer. To relate the parameters ϕ, σ and R to ionospheric and radio parameters, we note that :

$$R = S \sqrt{\frac{2}{d\lambda}} \tag{3}$$

where S is distance along the wavefront, d is the distance from the ionosphere to the ground and λ is the radio wavelength. Thus in scaling from one radio frequency f to another we have, since σ is in the same units as R,

$$\sigma \propto \sqrt{f} \tag{4}$$

Furthermore, when the refractive index μ of the lens is close to unity and the effect of the earth's magnetic field on μ is negligible

$$1 - \mu \propto 1/f^2$$

$$\therefore \quad \phi_o \propto 1/f \tag{5}$$

This latter assumption gets progressively poorer as the frequency decreases but should hold on 140 MHz and 360 MHz. The distance S must be related to the time t by allowing the lens to move with velocity u making

$$S = ut \cos \theta \tag{6}$$

where θ is the elevation of the ray at O.

Thus a test of the usefulness of the present models is to scale σ and ϕ_o according to equations (4) and (5) respectively, determine an appropriate value for u and see whether the resulting amplitude patterns simulate those in Figure 1.

The calculation involves some trial and error. First of all, the phase difference between rays (2) and (3) at a point P just far enough away from 0 (Fig. 2) to allow three rays is found by drawing the corresponding modified Cornu spiral (Fig. 3). $R_o = 2$ is about the right value for a wide range of σ and ϕ_o. In Figure 3, it is easy to recognize the three geometrical rays as points of inflexion and measure the phase difference between rays (2) and (3).

To match the diffraction pattern at 140 MHz, this phase difference has to be about 6: this requires that σ should be of the order 0.3.

The next requirement was that the amplitude at 0 (Fig. 2) be small. Amplitude at 0 as a function of σ and ϕ_o is plotted in Figure 4 and it is seen that again σ has to be of the order of 0.3. Suitable values of ϕ_o are near -3 radians repeating at approximately 6 radian intervals.

It is rather odd that the amplitude is reduced greatly only for σ close to 0.3 (except for much larger lenses). It seems to be because the waves which pass through the centre of such a lens are of the same amplitude as the waves diffracted round its edges and if the direct and diffracted waves are out of phase a deep minimum occurs.

The total change in the phase difference between ray (1) and either (2) or (3) as the observing point is moved from 0 to large distance gives a *rough* idea of the total number of fades.

Much larger lenses cause defocussing and reduced amplitude but the phase change through the centre of larger lenses requires an unacceptable amount of ionization, nor do the diffraction patterns resemble the observations. Since deep minima are required at all three frequencies, our choice was therefore limited. Finally, the velocity u is chosen to give the best match between the calculated and observed diffraction patterns.

During the observations, the receiver output is filtered by a CR filter ($\tau = 1$ sec.) to increase the signal/noise ratio. We have calculated diffraction patterns with and without similar filtering. Our filtered diffraction patterns do not reproduce the higher frequency fading which is present.

The values finally arrived at were:

$$\sigma_{360} = 0.4 \qquad \phi_{360} = -15.6$$

$$\sigma_{140} = 0.25 \qquad \phi_{140} = -40$$

$$\sigma_{40} = 0.13 \qquad \phi_{40} = -140$$

The observed time separations of the first few peaks following the central minimum (obtained from magnetic tape recordings at 1 sec intervals) are matched to the calculated positions of the peaks. The distance separations obtained from our calculations depend on the distance d from lens to receiver. Taking into account the obliquity of the raypath ($\simeq 45°$) we have $d \simeq 140$ km and $\simeq 420$ km for lens heights of 100 km and 300 km respectively. The corresponding values of u (m sec^{-1}) are given below:

f (MHz)	u (m sec^{-1})	
	For d = 140 km	For d = 420 km
140	50	86
360	51	89

Thus, the values of the velocities obtained from observations at the two frequencies are in good agreement.

Figures 5(a), (b) and (c) give filtered amplitude versus time patterns for cylindrical lenses for conditions representing 40 MHz, 140 MHz and 360 MHz respectively. Included in Figures 5(b) and 5(c) are the 1 sec data points. Figures 6(a), (b) and (c) are the corresponding unfiltered amplitude patterns. The effect of filtering is particularly noticeable on the higher radio frequencies and on the wings of the patterns. In Figure 6 an average value of u = 50 m sec^{-1} has been used.

One of the remarkable features of the diffraction pattern is that the width of the central minimum is much greater than the width of the cylindrical Gaussian lens. The width of the lens (corresponding to σ) is equal to 97 m (if the lens is E-region height) or 168 m (at F-region height), whereas the distance along the ground from the centre of the diffraction pattern to the first maximum is 1100 m (E-region) or 1900 m (F-region). Our model then represents a very small, very intense lens in the ionosphere. Figure 5(c) shows that the calculated and observed amplitudes are in remarkably good agreement with details such as the small subsidiary peak near the middle of the central minimum and the modulation envelope being reproduced. The agreement on 140 MHz (Fig. 5(b)) is quite reasonable, though the phase of the modulation is not correct. The agreement on 40 MHz is poor.

4. EXTENSION TO HIGHER FREQUENCIES

With the introduction of super high frequencies for satellite-to-ground communications it was expected that scintillation effects caused by the ionosphere would be eliminated. Such was not to be the case. Christiansen (1971) observed scintillations of signals from the lunar ALSEP transmitter received at Ascension

Island on 2278 MHz. Somewhat similar scintillations on frequencies near 4 GHz and 6 GHz of signals from the geostationary satellite INTELSAT have been observed by Craft and Westerlund (1972). It is, therefore, of interest to extend the amplitude calculations using the above lens to higher frequencies to see whether significant scintillations are possible. This has been done using the scaling rules given in equations (4) and (5) for frequencies of 720 MHz, 1440 MHz and 2880 MHz for which:

$$\phi_{720} = -7.8 \qquad \sigma_{720} = 0.57$$

$$\phi_{1440} = -4.0 \qquad \sigma_{1440} = 0.8$$

$$\phi_{2880} = -2.0 \qquad \sigma_{2880} = 1.14$$

The amplitude versus time (or distance) patterns are shown in Figures 7(a), (b) and (c). Here the RC filtering has been eliminated. On 720 MHz the minimum signal is 0.27 of the undisturbed value (1.414) and the maximum is 2.0 times the undisturbed value. On 1440 MHz the minimum and maximum values are 0.58 and 1.50 times the undisturbed value. The minimum and maximum values on 2880 MHz are 0.82 and 1.125 respectively. With parameters appropriate to a frequency of 6 GHz the minimum and maximum values are 0.94 and 1.03 respectively of the undisturbed value. The patterns are still larger on the ground than is the size of the lens, e.g. 570 m on 720 MHz, 370 m on 1440 MHz and 367 m on 2880 MHz for d = 140 km. The corresponding values for d = 420 km are 990 m, 640 m and 636 m compared with the lens radius of 97 m.

5. POSITIVE LENSES

In the foregoing models the phase shift ϕ_o has been negative corresponding to an increased electron density and, therefore, a defocussing lens. It is quite possible that focussing lenses, with electron densities below the ambient values, can exist. Indeed focussing of the signals can be seen in Figure 1 marked B, where the peak amplitude is about 2.2 times the undisturbed value on 40 MHz and 1.13 times the undisturbed value on 140 MHz. The change in amplitude on 360 MHz while noticeable is not measurable. Some calculated patterns for frequencies of 40 MHz and 140 MHz are given in Figures 8(a) and 7(b). Here again there is no RC filtering. With a positive lens there is a limit to the phase shift since the electron density cannot be negative: for example for a background plasma frequency of 8 MHz the maximum value of ϕ_o on 40 MHz is 17 radians and on 140 MHz it is 5 radians.

6. DISCUSSION

The cylindrical Gaussian model gives the required modulation of the diffraction pattern provided that the maximum phase shift is sufficiently large *and* that the width (σ) is sufficiently small. On the assumption that the cross section of the cylinder was circular and that the refractive index is near unity for the 140 MHz signal and neglecting the background ionization, the maximum electron density in the lens is about 2.4×10^{13} el m^{-3} corresponding to a plasma frequency of about 43 MHz. This being the case the core of the lens would be overdense to the 40 MHz signal. It is, therefore, not surprising that the fading pattern for $\sigma = 0.15$, $\phi_o = -140$ radians does not correspond in detail with the 40 MHz record of Figure 1.

The asymmetric patterns could be produced by cylindrical lenses with an unsymmetrical cross section. Alternatively, it may be the result of temporal changes, e.g. the expansion of the lens. The latter idea, together with the intense plasma concentration, suggests a meteor trail. However, from data presented by Sugar (1964), we find that the probability of observing a trail with the required concentration is about one in 10^7 days. Furthermore, with diffusion coefficients of 1 m^2/sec to 140 m^2/sec in the height range 85 to 115 km the ionized column could not last for the several minutes required to explain the fading shown in Figure 1. Hence a meteor-trail source is unlikely. Another possible cause is the ionized trail produced by re-entry of a satellite into the earth's atmosphere.

From our data, it is not possible to determine the height of the lens. We must, however, rule out certain man-made objects: the low velocity rules out aircraft, the modulation of the fading and the asymmetry of the patterns rules out balloons. We cannot however make a choice between E- and F-regions of the ionosphere. The lens is certainly elongated: we attempted to model a spherical lens but failed to reproduce the required depth of modulation. The cylinder need not have circular cross section, but rather could be in the form of a sheet. On the grounds that small intense blobs of ionization are required to explain some of the radio observations of sporadic E, we favour an E-region origin. The phase shifts required may rule out a lens above F-region. The ray path to the satellite was not tangent to the geomagnetic lines anywhere so that a weak perturbation over a long path through the magnetosphere (Booker, 1975) is unlikely to provide an explanation.

Hajkowicz (1974) has pointed out the quasi-periodic scintillations may be closely linked with an interference effect of waves simultaneously transmitted by two satellites moving within the beamwidth of the receiving antenna. This could not be the cause of the scintillations reported here since these occur on all three carrier frequencies and, as far as we know, no other satellite transmits simultaneously on these precise frequencies.

7. CONCLUSION

The observation requires an ionospheric lens approximately 100 m across, but elongated in one direction to explain the diffracted pattern. Such a lens can produce appreciable amplitude fluctuations on frequencies above 1 GHz. Lenses caused by a deficiency of electrons also produce diffraction patterns. Thus a single ionospheric irregularity can produce several maxima and minima in the amplitude pattern observed on the ground.

8. REFERENCES

Booker, H. G., 1975, The role of the magnetosphere in satellite and radio-star scintillation, *J. Atmos. Terr. Phys.* 37, 1089.

Christiansen, R. M., 1971, Preliminary Report of S-band propagation disturbance during ALSEP mission support (Nov. 19, 1969 - June 30, 1970), NASA Report No. X-861-71-239, National Aeronautic and Space Administration, Goddard Space Flight Center, Greenbelt, Maryland 20771.

Craft, H. D. and L. H. Westerlund, 1972, Scintillations at 4 and 6 GHz caused by the ionosphere, AIAA Paper No. 72-179, 10th Aerospace Science Meeting.

Davies, K., Fritz, R. B., Grubb, R. N., and Jones, J. E., 1975a, Some early results from the ATS-6 radio beacon experiment, *Radio Sci.* 10, 785.

Davies, K., Fritz, R. B., Grubb, R. N., and Jones, J. E., 1975b, ATS-6 radio beacon experiment: The first year, *IEEE Trans.* AES-11, 1103.

Elkins, T. J. and Slack, F. F., 1969, Observations of travelling ionospheric disturbances using stationary satellites, *J. Atmos. Terr. Phys.* 31, 421.

Hajkowicz, A., 1974, Radio transmissions from two satellites as a possible cause of quasiperiodic scintillations in amplitude recordings, *J. Atmos. Terr. Phys.* 36, 1689.

Herron, M., 1976, *J. Atmos. Terr. Phys.* (submitted).

Ireland, W. and Preddy, G. F., 1967, Regular fading of satellite transmissions, *J. Atmos. Terr. Phys.* 29, 137.

Jenkins, F. A. and White J. E., 1957, *Fundamentals of Optics*, 3rd Edition, McGraw-Hill, New York.

Kelleher, R. F. and Martin, P., 1975, Fresnel-type fading on satellite records at low latitudes, *J. Atmos. Terr. Phys.* 37, 1109.

Pope, J. H. and Rufenach, C., 1975, Review of ionospheric small scale structure with emphasis on equatorial radio scintillation in the SHF band, *NOAA Tech. Rept.* ERL343-SEL34, U. S. Government Printing Office, Washington, DC 20402.

Singleton, D. G., 1964, Broadband radio-star scintillations, II. Interpretation, *Radio Sci.* 68D, 1095.

Slack, F. F., 1972, Quasiperiodic scintillation in the ionosphere, *J. Atmos. Terr. Phys.* 34, 927.

Sugar, G. R., 1964, Radio propagation by reflection from the meteor trails, *Proc. IEEE* 52, 116.

Titheridge, J. E., 1971, The deflection of satellite signals by isolated ionospheric irregularities, *J. Atmos. Terr. Phys.* 33, 47.

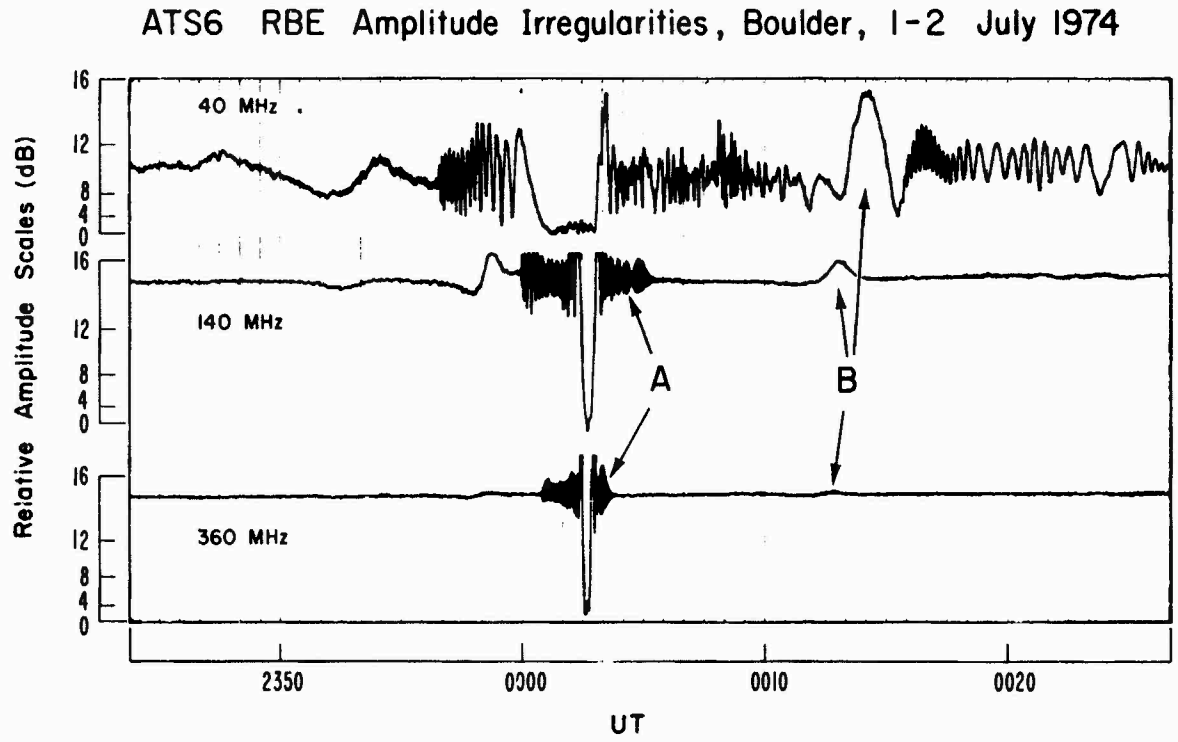


Fig. 1 Amplitude variations of radio signals from the ATS6 radio beacon. Note the modulations of the fading patterns marked A and the amplitude enhancements marked B.

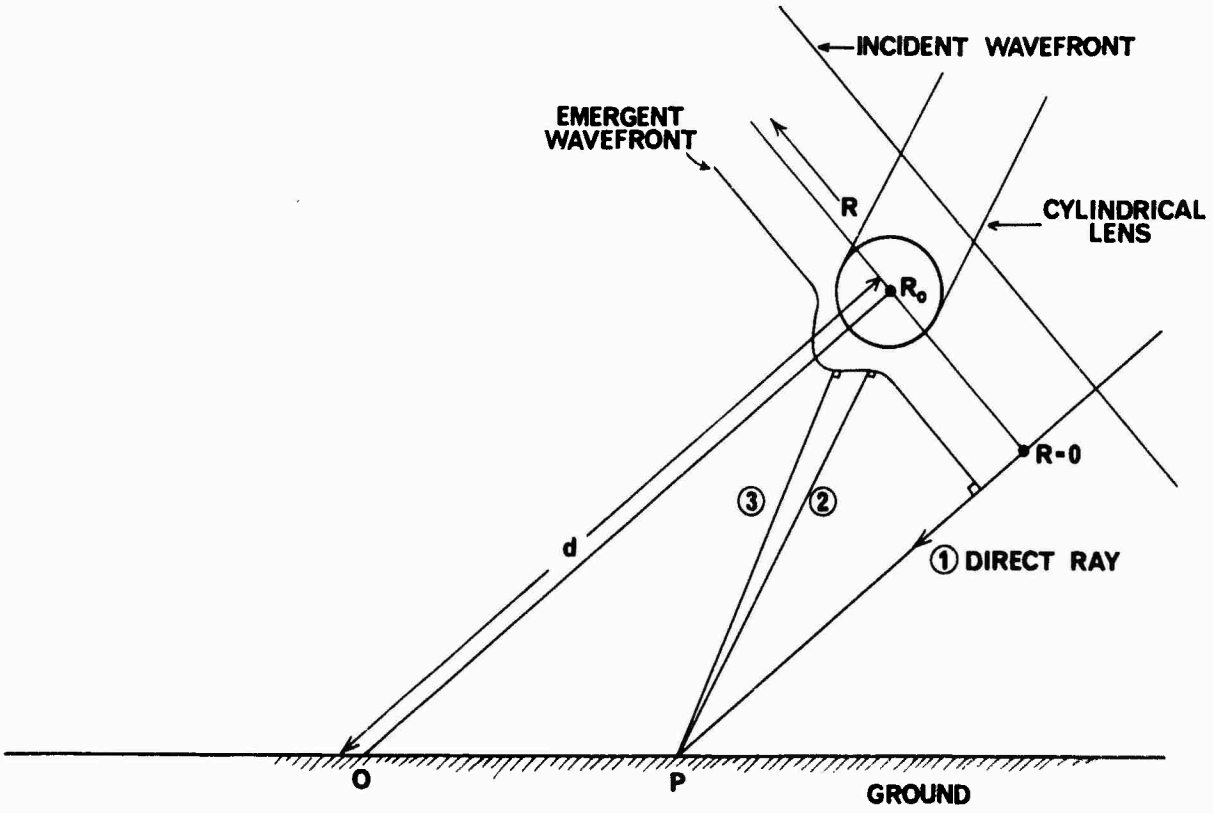
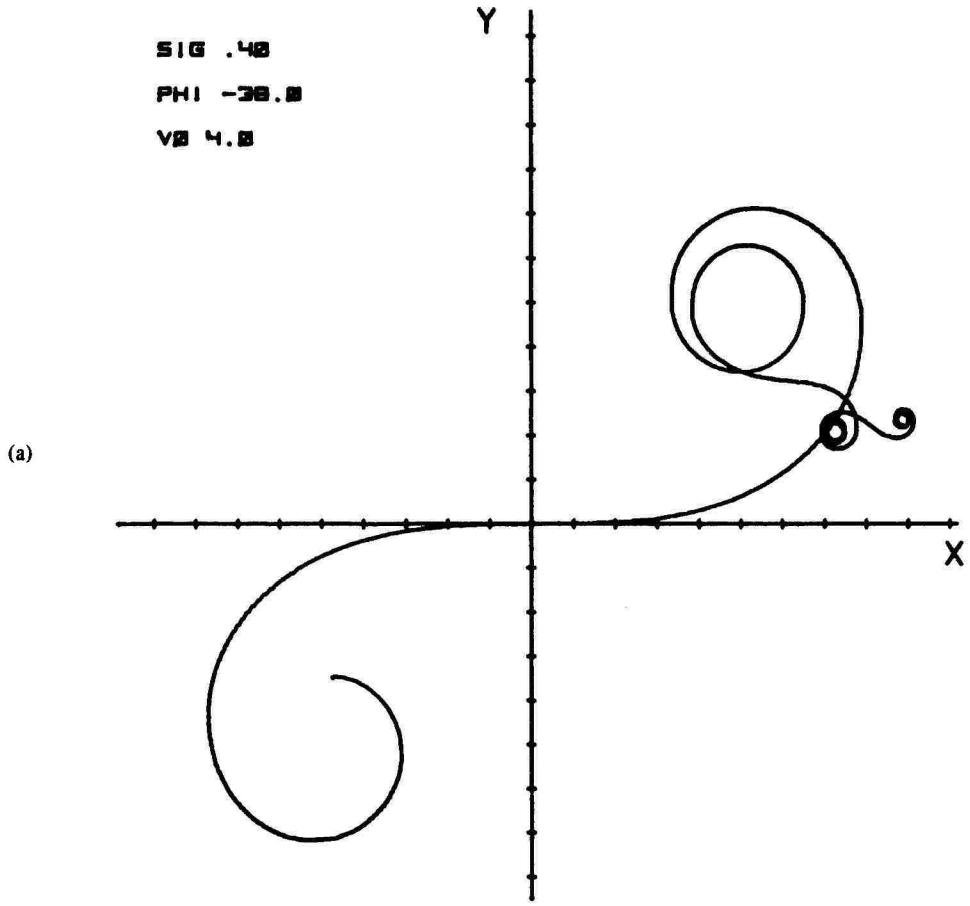
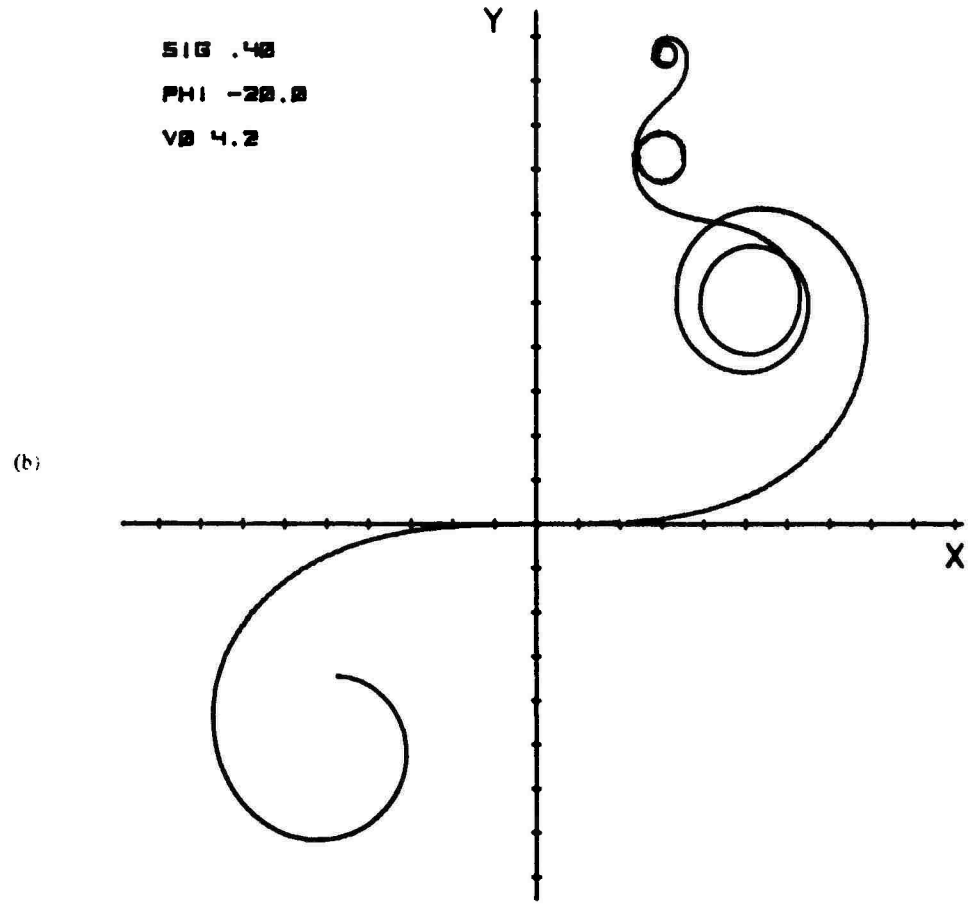


Fig. 2 Geometry of cylindrical lens.

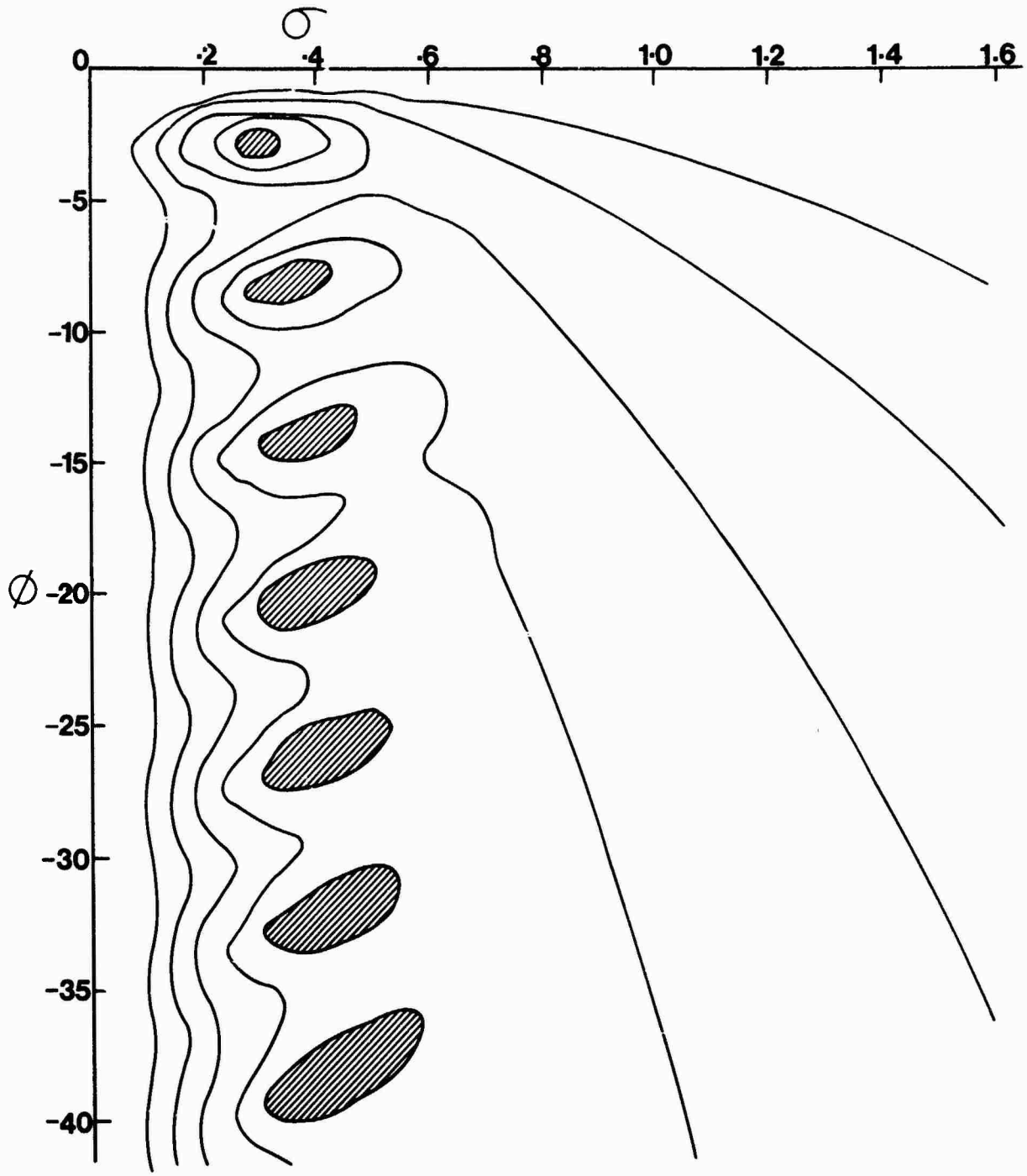


(a)



(b)

Fig. 3 Examples of modified Cornu spirals caused by cylindrical Gaussian lenses.



CONTOURS OF EQUAL AMPLITUDE AS A FUNCTION OF σ AND Ø. SHADED AREAS INDICATE AMPLITUDE LESS THAN 14 % OF UNPERTURBED AMPLITUDE.

Fig. 4 Variation with σ and ϕ of amplitude at the centre.

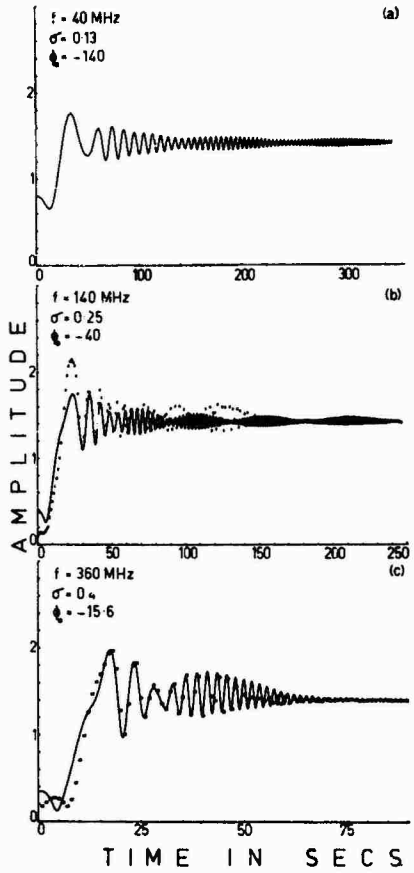


Fig. 5 Calculated amplitude patterns on (a) 40 MHz, (b) 140 MHz, and (c) 360 MHz, with 1 sec smoothing.

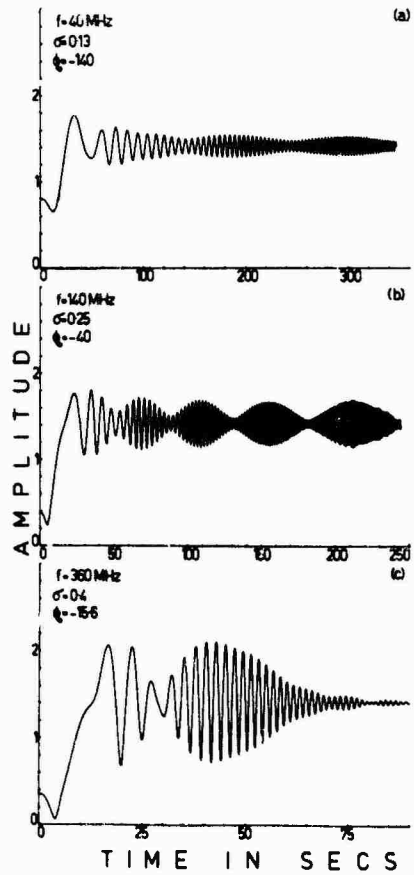


Fig. 6 Calculated amplitude patterns on (a) 40 MHz, (b) 140 MHz, (c) 360 MHz, without smoothing.

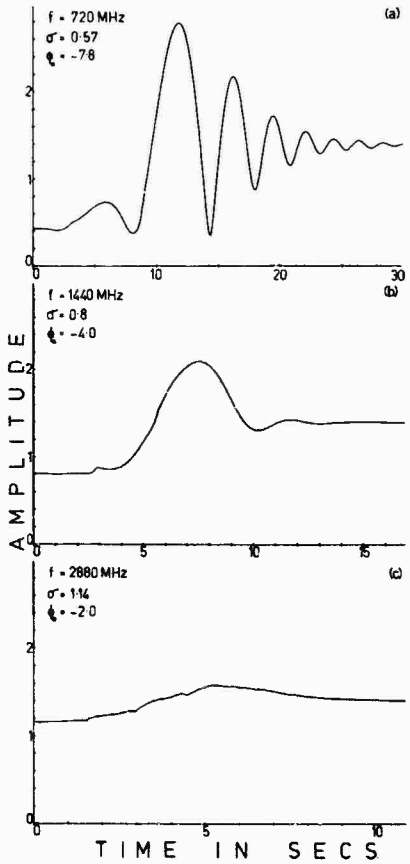


Fig. 7 Calculated amplitude patterns on (a) 720 MHz, (b) 1440 MHz, and (c) 2880 MHz.

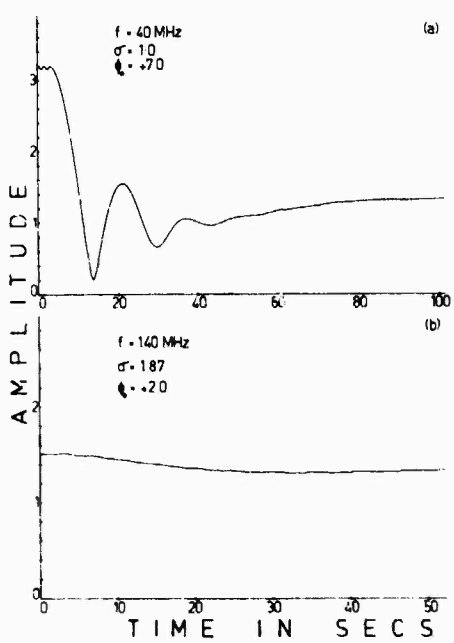


Fig. 8

DISCUSSION

J.Röttger, Max-Planck-Institut für Aeronomie, D-3411 Katlenburg-Lindau 3, F.R.G.
Is there any idea on the origin of these over dense drifting radio lenses?

Author's Reply

We have no explanation of the cause of this irregularity. Meteor trails have been ruled out. The Space Data Center of NASA/Goddard have told me that there was no reentry of a space vehicle at this time and place. It may be of interest to note that ATS-6 observations at Lindau, West Germany, showed a similar but less intense pattern on May 12, 1976 near 1836 UT i.e., the phenomenon occurs at other places at other times.

LOW ANGLE EFFECTS ON VHF AND UHF PROPAGATION DUE TO IONOSPHERE AND TROPOSPHERE

(Review)

G.K. Hartmann

Max-Planck-Institut für Aeronomie - IfL
D 3411 Katlenburg-Lindau 3, P.O. Box 80
FRG

ABSTRACT

A literature survey is attempted for low angle effects on VHF and UHF propagation due to ionosphere and troposphere. Main emphasis is given to effects which occur for satellite elevation angles of less than 30° degrees. The most important effect is due to scintillation. In many cases it was not possible to find out the exact elevation angle under which the reported scintillation occurred. Thus we restricted ourselves to those cases where it was unambiguously clear that the observed effects occurred below 30° and as far as possible to relevant already existing review papers.

Propagation of electromagnetic waves through an inhomogeneous ionosphere and troposphere usually results in variations or fluctuations in the amplitude, phase, and angle of arrival of the waves as received at the ground. These fluctuations are collectively known as scintillations. It is quite clear that these scintillations affect "Radio Navigation and Positioning Systems".

In addition to the general aspects of low angle effects as an example a specific effect that occurs very often is discussed in more detail. It is the effect that is due to the antenna beamwidth and due to small changes of the medium within that 3 dB cone.

A. INTRODUCTION

The literature on scintillation ... ow very extensive and a complete review is impossible in the space and time available. (Aarons, ., Whitney, H.E., Allen, R.S., 1971; Aarons, J., 1973; Hartmann, G.K., 1975; Hartmann, G.K., 1972; Warnik, W.A., Liu, C.H., Youakim, M.Y., Yeh, K.C., 1973).
Thus we will only emphasize those effects which occurr for elevation angles less than 30 degrees. Propagation of electromagnetic waves through an inhomogeneous ionosphere and troposphere usually results in variations or fluctuations in the amplitude, phase, and angle of arrival of the waves as received at the ground. These fluctuations were formerly collectively known as scintillations. However, since several years the word scintillation is used in a closer sense to describe a phenomen whereby radio signals experience random fading in amplitude and phase after propagating through a transmission medium. Since all types of amplitude and phase fluctuations can affect "Radio Navigation and Positioning Systems" we will use the word scintillation in its earlier broader meaning. (AGARD, 1970; AGARD, 1971; Crane, R.K., 1976).

B. SCINTILLATIONS

Scintillation occurrence and scintillation intensity depend upon frequency, location, propagation path geometry, geophysical conditions, receiving equipment characteristics, - such like the receiving antenna beamwidth, the receiver resolution and stability, and the data acquisition facilities -, and the measure used to describe scintillation. This causes many problems by comparing data from different sources. (Fremouw, E.J., 1974).

Since amplitude measurements require in general only a fairly simple and inexpensive equipment, roughly 85 % of the now available scintillation data were obtained by amplitude measurements. This is the reason that fairly often, for the sake of brevity, the terminus technicus "scintillation" is exclusively related to measured amplitude fluctuations. Phase measurements, strictly speaking relative phase measurements, require a more sophisticated and a much more expensive equipment. Therefore only approximately 10 % of the available scintillation data were obtained by phase measurements. (Hartmann, G.K., 1972). Angle-of-arrival measurements require a still more complicated equipment, especially an extremely accurate antennasystem. Therefore only about 5 % of the data were obtained by angle-of-arrival measurements.

C. LOW ANGLE EFFECTS

C. 1 General aspects

The length of a satellite to ground radio link increases with decreasing elevation angle of the satellite. This gives rise to various problems that can be almost completely neglected if the satellite is observable under high elevation angles. Besides grazing angle and diffraction effects in the ionosphere and troposphere we have to consider those which are due to refraction and due to the antenna beamwidths. These fairly "normal" effects will be briefly discussed using the observation of the geostationary US satellite ATS-6 as an example.
The ATS-6 geostationary satellite was launched on 30 May 1974 and was positioned for about 12 months at 94° W before it was moved to 35° E where ATS-6 was positioned for another 12 months. During this period the satellite was observable from Europe. The signals of the ATS-6 Radio Beacon Experiment (RBE) where monitored from the field site Gillersheim (51.65° N and 10.13° E, about 7 km distant from the main institute at Lindau) of the Max-Planck-Institut für Aeronomie, Lindau, FRG. ATS-6 was positioned at 35° E, 36 000 km above the surface of the Earth. Its topocentric coordinates with respect to the receiving site Lindau were azimuth: 30° from south towards east, elevation: 28°. The line of sight between ATS-6 and Gillersheim intersected between 50° and 40° geogr. latitude the ionosphere in altitudes between 100 km and 1000 km and had a

relevant length of 1600 km. Thus the results (Degenhardt, W., Hartmann, G.K., 1976) - from an ionospheric and tropospheric point of view - were due to medium latitudes, due to the fact that in our case geographic and geomagnetic coordinates are nearly identical. Ionospheric and tropospheric low angle effects were reported (Degenhardt, W., Hartmann, G.K., 1976).
Figure 1 shows the geometry using a Short Backfire (SBF) antenna with a beamwidth of $\pm$ 15° for observing the ATS-6 RBE signals at Gillersheim. The 3 dB cone of the SBF antenna is displayed, defined by Li P_1 P_2. The effect due to refraction in the ionosphere implies that no rays originating from the upper shaded area (RA_1) will be received within the 3 dB beamwidth and that all rays originating from the lower shaded area will be received within the 3 dB beamwidth. Refractions due to the troposphere can be treated similarly. Figure 2 shows the same geometry however with more details however disregarding refraction effects i.e. assuming straight line propagation. The signal at the receiving antenna is to a first approximation the vector sum of all rays incident on the plane - defined by P_1 and P_2 - of figure 1. Amplitude and phase of the received signal remain constant - assuming the transmitter and receiver characteristics remain constant - as long as the medium within the 3 dB cone remains constant. However this is neither a true statement for the ionosphere nor for the troposphere. Even if there are only weak changes within the cone that does not significantly affect the amplitude of each single ray they affect in any case the phase of the rays and this may lead - due to the vectorial summing of the rays at the receiving antenna - to fairly strong amplitude and phase changes of the resulting signal. It is obvious how dependent they are from the antenna beamwidth. Thus they are different using antennas with different beamwidths. Effects of this nature were observed with ATS-6 RBE signals (Degenhardt, W., Hartmann, G.K., 1976).
Using a simple model the magnitude of the just mentioned phase variations will be calculated. Fig. 3 shows a portion of the volume within the 3 dB cone of the antenna, displayed in fig. 1. For the sake of simplicity it is assumed to be plane. However, more rigorous calculations of cylindrically or conically shaped intersections lead to the same results. The geometry of figure 3 was used to derive the following formula (Degenhardt, W., Hartmann, G.K., 1976).

Two rays traversing a medium with a refractive index n arrive at A and C where they penetrate into a medium with refractive index n_1 and n_2 having a thickness D. At D' and E they leave the medium and start again traversing a medium with a refractive index n. After having left the medium with the thickness D they have a phase difference $\Delta\varphi$ given by:

$$\Delta\varphi = \frac{2\pi f}{c} D \left[\frac{1 - \sin\alpha \sin\beta_1}{\cos\beta_1} - \frac{1 - \sin\alpha \sin\beta_2}{\cos\beta_2} \right] \qquad (1)$$

where

$$\sin\beta_1 = \frac{n}{n_1} \sin\alpha$$

$$\sin\beta_2 = \frac{n}{n_2} \sin\alpha$$

f: signal frequency [Hz]
c: velocity of the light [m/sec]
D: layer thickness [m]

Using signal frequencies of 40 MHz (λ_0 = 7.5 m), 140 MHz (λ_0 = 2.14 m), and 360 MHz (λ_0 = 0.83 m) we get at 360 MHz a phase difference $\Delta\varphi$ for the troposphere that is 9 times larger that at 40 MHz and 2.6 times larger than at 140 MHz for a given angle of incidence and a given layer thickness D. Refraction effects in the ionosphere show an opposite frequency behaviour.

$\Delta\varphi \sim f^{-1}$ (f = signal frequency).

We can see that $\Delta\varphi$ for the troposphere is proportional to the frequency f and to the layer thickness D and furthermore depends on α, β_1 and β_2.

Fig. 4 shows $|\Delta\varphi|$ for the troposphere in degrees as a function of the angle of incidence α calculated for f = 360 MHz (λ = 0.833 m), a layer-thickness D (8000 m), and

(x) n = 1, n_1 = 1.000005, n_2 = 1.000010

(o) n = 1, n_1 = 1.000295, n_2 = 1.000300

The calculation was carried out for tropospheric conditions and for angles α between 0° and 80°. The angles between 80° and 90° were skipped since in this domain the grazing angle effects start to be much stronger than the effects described by equation (1).

The more the frequency f increases and the longer the path through the troposphere gets the more also the effects of the water vapour in the atmosphere have to be considered (Hartmann, G.K., Oberländer, H., 1976). This is especially true for frequencies higher than 10^9 Hz.

Fig. 5 shows a good approximation of the effect of the phase shift $\Delta\varphi$ of the two rays on the resulting normalized signal amplitude A_r disregarding reflection and absorption effects. Be F the area of the volume - layer with thickness D - intersected from the atmosphere by the antenna beam, then we define two areas FA_1 and FA_2 which represent the regions where refractive index is n_1 and n_2.
We get $F = FA_1 + FA_2$ and $n_1 \neq n_2$. The resulting signal amplitude A_r is now dependent from the phase shift $\Delta\varphi$ and the ratio of two areas FA_1/FA_2.
It is quite obvious that for $FA_1/FA_2 = 1$ - i.e. the signal amplitude of the ray passing through the portion of the volume with refractive index n_1 is identical with that passing through the region with n_2 - and for $\Delta\varphi = 180°$ the resulting signal amplitude is zero. We see that at low elevation angles already weak refractive index variations can cause fairly drastic amplitude changes which are due to phase changes along the ray path of the two rays. The calculations can easily be carried out for more than two rays, however this would not yield new informations unless we know more about the layer thickness D and the refractive index characteristics within the antenna beam volume.

C. 2 Results

1) Since many of the data which were recorded with signals from orbiting - non geostationary - satellites are not directly given as a function of elevation angle, it is not possible to calculate the percentage of data that is due to elevation angles less than 30°.

2) At 137 MHz scintillation with fades in excess of 6 dB occurs on zenith paths for less than 20 % of the time near the geomagnetic equator, less than 2 % of the time in the auroral regions, and less than 0.1 % of the time at middle latitudes. (Crane, R.K., 1976). However, this latter figure may increase by several orders of magnitude if we consider those mid latitude regions where often the troposphere exhibits irregular structures and if simultaneously we only consider elevation angles less than 30 degrees. (Tyagi, 1975).

3) The effect of the three major and different ionospheric scintillation areas - a) Equatorial regions, b) Middle latitude regions, and c) High latitude regions - are getting less prominent with increasing frequency. (Aarons, J., Whitney, H.E., Allen, R.S., 1971; Golden, T.S., 1970; Hartmann, G.K., 1972). For elevation angles less than 30° we have to consider additional effects that are due to irregularities in the troposphere and which increase with frequency and decreasing elevation angle. (Badillo, V.L., 1970; Barrett, L., Hartmann, G.K., Leo, R.L., Weiss, W., Zwick, R., 1975; Degenhardt, W. Hartmann, G.K., 1976; Fengler, C., 1976). Especially in many mid latitude regions these effects can be much larger and more frequent than the above mentioned ionospheric effects. (Tyagi, 1975).

4) Almost all now available model calculations, describing the influence of the lower atmosphere on VHF and UHF waves are very useful in cases of a "normal, mean" behaviour of the atmosphere. (Fluess, H.W., Eckl, W., 1975; AGARD, 1968). However, they do not take into account scintillation effects caused by gradients and/or irregularities in the troposphere which are the actual limiting factors for radio navigation systems. (Fengler, C., 1976; Hartmann, G.K., 1969).

5) The number of scintillation events at mid latitude regions increases drastically for the whole VHF and UHF frequency range if the elevation decreases below 30°. No clear statements about the seasonal and daytime behaviour of the scintillation occurrence and the scintillation intensity for these cases are possible till now.

6) The data available thus far for these low elevation angles seem to be insufficient even for more detailed morphological studies like those which were carried out for higher elevation angle scintillations caused in the ionosphere and which yielded a fairly good overall picture of the main irregularity regions of the ionosphere.

7) More detailed investigations of low angle scintillations require a better cooperation between scientists whose field is ionosphere physics, troposphere physics and hydrosphere physics, e.g. Commissions F and G of URSI.

8) Some general discussions on the usefulness of such scintillation data in geophysics rather than for only the most important application purposes in the design of radio navigations systems are lacking. (Mendillo, M., editor, 1976). This should be considered in the context with the information analysis centers (IACs), which start to investigate some general data evaluation problems, e.g. CODATA founded in 1975 an Advisory Panel on the Geosciences consisting of seven individuals from different disciplines.

9) It is evident that scintillation is a fact of life for a number of communication or radio navigation systems - like radar - that have to operate through the auroral or equatorial ionosphere. Considering elevation angles of less than 30° for the incident rays we have to extend this statement also to all those mid latitude regions were often the troposphere exhibits irregular structures. Space diversity or time diversity coding schemes are required to mitigate the effects of scintillation. Using the weak scintillation model, the parameters for the design of diversity systems can be determined if the height, axial ratio, and, for time diversity schemes, the drift velocity are known. Adequate statistical data are only available for scintillation intensity. Adequate statistical data for the design of diversity systems are still lacking and await either further experimentation or better models for the production and movement of the irregularities. Thus it seems to be difficult in the moment to move much beyond morphological studies, which are just the first step in the course of more rigorous attempts to interpret the geophysical phenomena and to obtain more reliable models for the prediction of scintillation effects especially for elevation angles less than 30°. For better geophysical interpretations also a higher number of standardized transmitting and receiving systems and evaluation methods is required.

REFERENCES

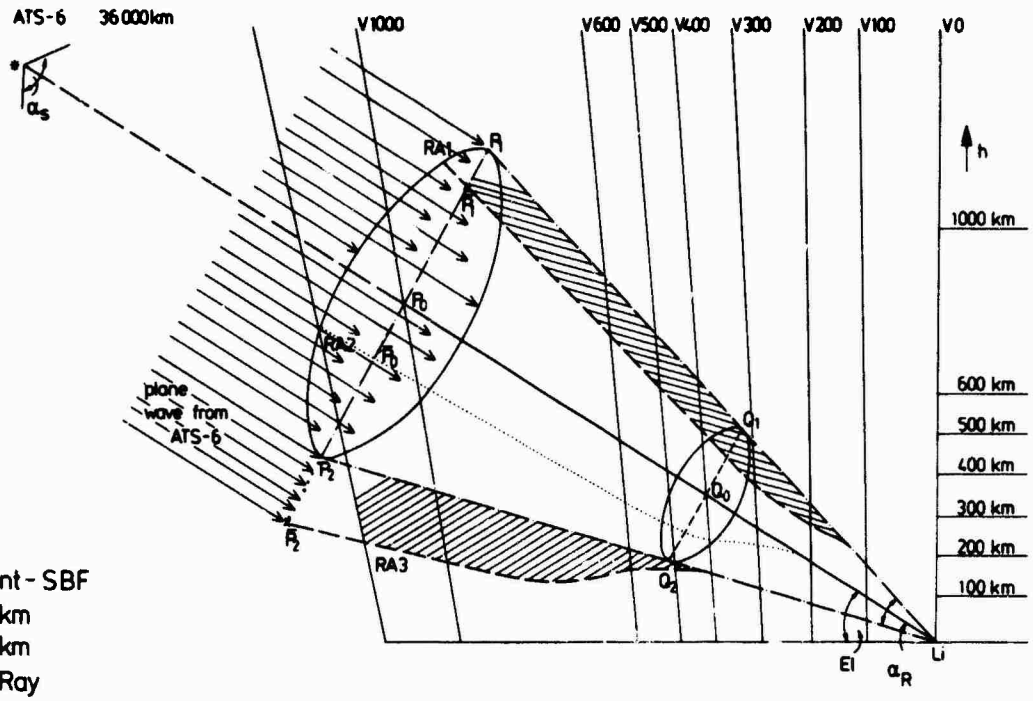
Aarons, J. Whitney, H.E. Allen, R.S.	1971	Global morphology of ionospheric scintillations Proc. IEEE 59, 159-172
Aarons, J.	1973	Total electron content and scintillation studies of the ionosphere AGAR Dograph No 166
AGARD	1968	Radio Wave Propagation AGARD Lecture Series XXIX
AGARD	1970	Application of Propagation Data of VHF Satellite Communication and Navigation Systems AGARD Lecture Series No 41
AGARD	1971	Tropospheric Radio Wave Propagation Part I, Part II AGARD Conference Proceedings No 70
Badillo, V.L.	1970	Solar microwave scintillations Radio science, vol 5, no 6, p 979-982

Barratt, L. Hartmann, G.K. Leo, R.L. Weiss, W. Zwick, R.	1975	Tropospheric irregularity effects on the 360 MHz ATS-6 Radio Beacon Submitted for publication to Radio Science
Crane, R.K.	1976	Ionospheric scintillations Submitted to Proceedings of IEEE
Degenhardt, W. Hartmann, G.K.	1976	Periodic amplitude and phase fluctuations ranging from 1 sec to 30 min observed with the ATS-6 RBE at Lindau Paper presented at the COSPAR Beacon Satellite Symposium at Boston, USA, June 2-4, 1976; In: Mendillo 1976
Fengler, C.	1976	Tropospheric effects on Satellite-Earth links In: Mendillo 1976
Fluess, H.W. Eckl, W.	1975	A comprehensive description of the atmosphere on microwaves in the 0.5 to 100 GHz regions by means of a computer program Messerschmitt-Bölkow-Blohm GmbH, 8000 München 80, P.O.Box 801160, FRG
Golden, T.S.	1970	A note on correlation distance of VHF fading from irregularities in the equatorial ionosphere Radio science, vol 5, no 6, p 943
Hartmann, G.K.	1969	Tropospheric diffraction phenomena of radio signals from the beacon satellite Explorer 22 J. Atmos. Terr. Phys. vol 31, p 663-669
Hartmann, G.K.	1975	HF and UHF propagation studies of the mid-latitude ionosphere Ann. Geophys. vol 31, no 1, p 39-52 (Review)
Hartmann, G.K.	1972	Brief review of scintillation studies Space Research XII, 1221-1228
Hartmann, G.K.	1972	Satellite scintillations between 43° and 66° northern latitude from 1964 to 1969 In: Radar propagation in the Arctic. AGARD Conference Proceedings No 97, p 20-1 to 20-12
Hartmann, G.K. Oberländer, H.	1976	Physikalische Wasseruntersuchungsmethoden. Ein Übersichtsartikel - Kap. 6 II. Süßwasser - (Physical methods for water research. A review) Published by: Institut für Dokumentationswesen, D-6000 Frankfurt/Main 71, Herriotstr. 5, FRG, Nov. 1976
Mendillo, M. editor	1976	Proceedings of the COSPAR Beacon Satellite Symposium at Boston, USA, June 2-4, 1976, in press
Warnik, W.A. Liu, C.H. Youakim, M.Y. Yeh, K.C.	1973	A theoretical study of scintillation of transionospheric radio signals Final report NASA GSFC NGR 14005[illegible]9, Technical Report 50
Tuhi Ram Tyagi Somayajulu, V.V. Ghosh, A.B.	1975	Satellite Radio Beacon Monitoring of the Troposphere J.Atmos.Terr.Phys. 37, pp 1603 - 1608

Addendum

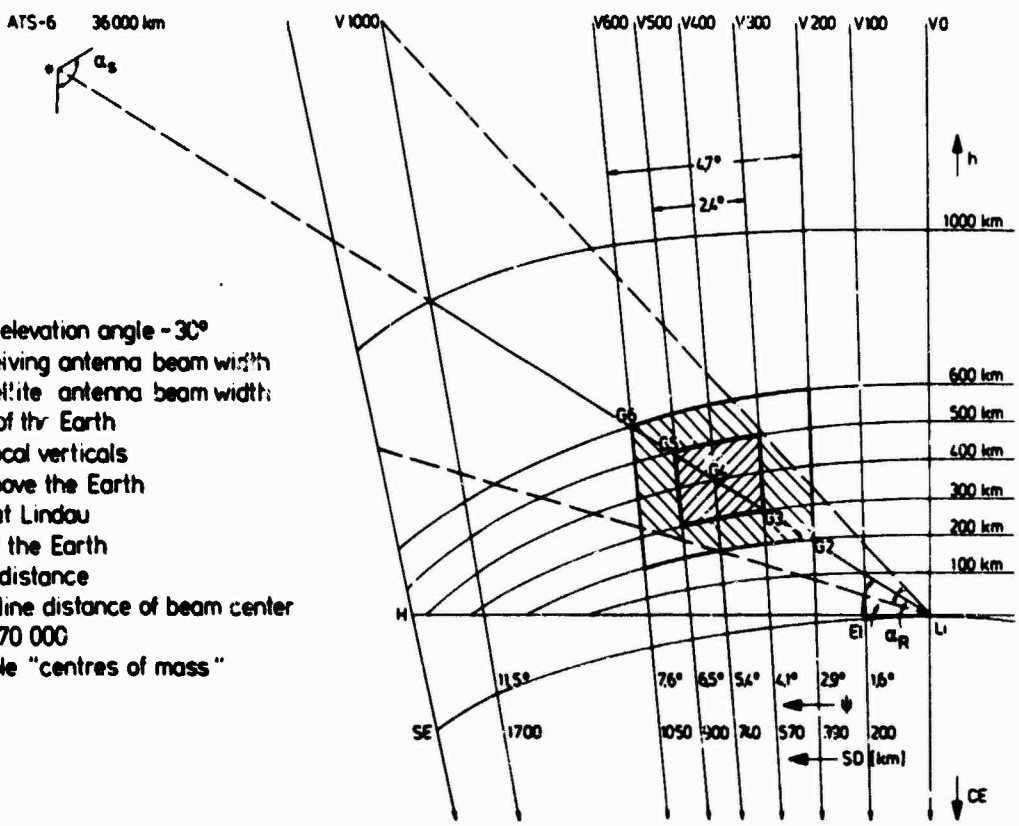
Leitinger, R. • Hartmann, G.K.	1976	Zeit- und Breitenabhängigkeit von transionosphärischen Ausbreitungsfehlern. Kleinheubacher Berichte, Nr.20, in press
Davies, K. • Hartmann, G.K. Leitinger, R.	1976	A comparison of several methods of estimating the columnar electron content of the plasmasphere. To be published J.Atmos.Terr.Phys.

• Available at Dr. R. Leitinger, Institut f. Meteorologie u. Geophysik der Universität Graz, Halbärthg.1, A 8010 Graz, Austria



one-element-SBF
$\overline{Q_0Q_1} \approx 190$ km
$\overline{P_0P_1} \approx 450$ km
RA1-RA3 Ray
bending due to refraction

Fig. 1: Geometry for observations of the ATS-6 at the MPI f. Aeronomie at Lindau FRG considering refraction effects



Li : Lindau
El : Satellite elevation angle - 30°
α_R : 3dB receiving antenna beam width
α_S : 3dB satellite antenna beam width
SE : Surface of the Earth
V0 - V1000 local verticals
h : height above the Earth
H : horizon at Lindau
CE : Center of the Earth
ψ : angular distance
SD : Straight line distance of beam center
Scale : 1:6 670 000
G2 - G6 possible "centres of mass"

Fig. 2: Detailed geometry for observations of the ATS-6 at Lindau using Single Element Short Backfire Antennas

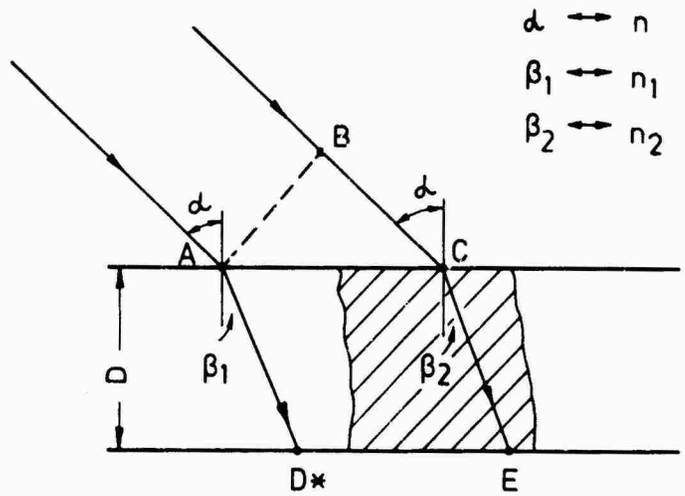


Fig. 3: Two rays incident from a medium with refractive index n on a layer with two different refractive indices n_1 and n_2.

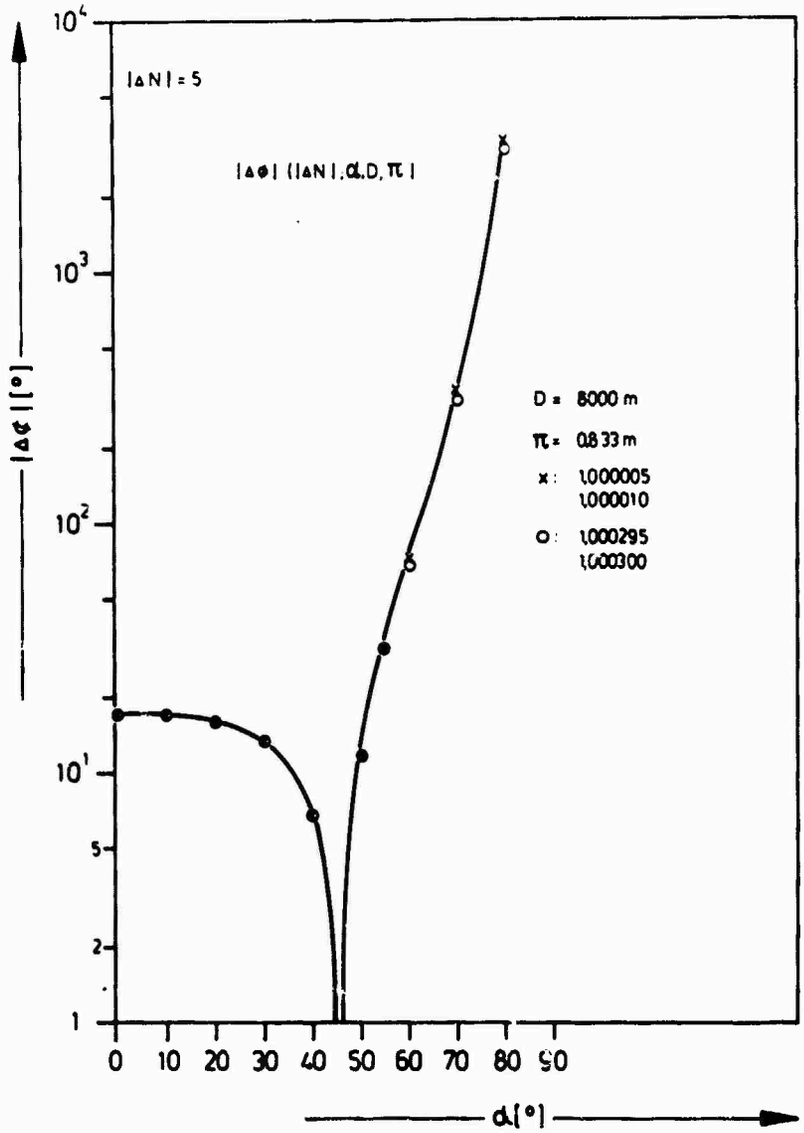


Fig. 4: Phase shift $|\Delta\varphi|$ between the two rays of fig. 3 as a function of the angle of incidence α after passing the layer

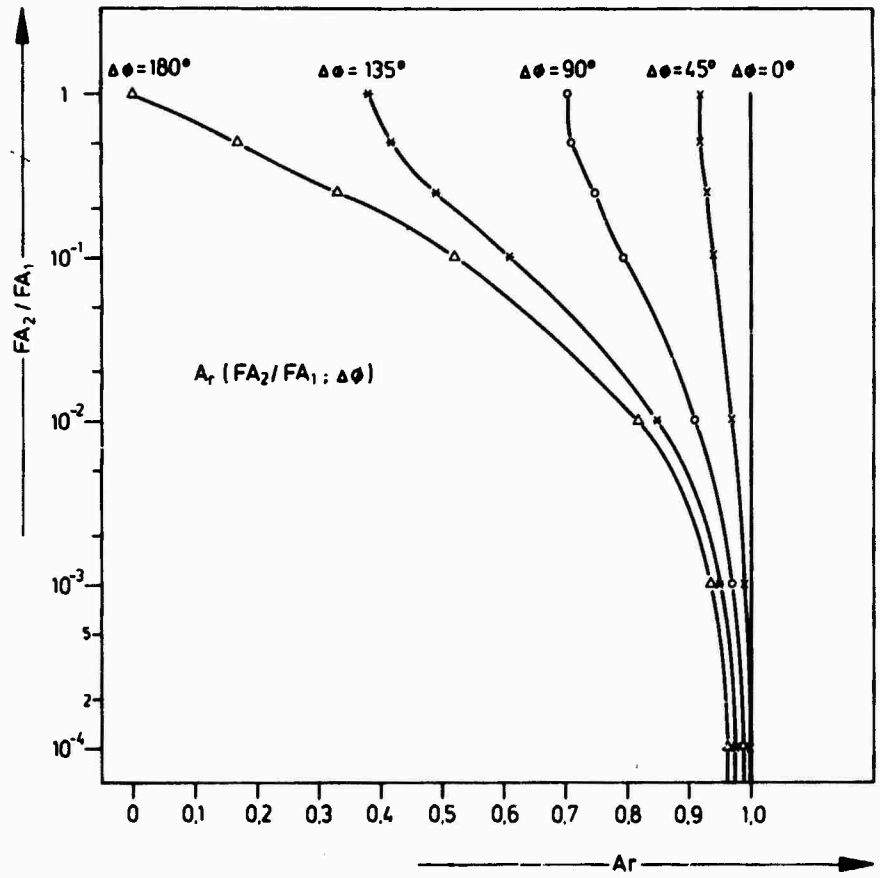


Fig. 5: Resulting normalized signal amplitude A_r as a function of the ratio FA_2/FA_1 for various phase shifts $\Delta\phi$

DISCUSSION

J.S.Belrose

With regard to the deep minimum at 45° elevation angle in your calculated phase data. Is this a Brewster angle effect for an assumed vertically polarized down coming wave?

Author's Reply

We only applied Snell's law which is valid for both polarizations. Thus, we do not have to worry about Brewster angle phenomena which has to be considered only in cases of reflection.

H.Soicher, US Army Electronics Command, Fort Monmouth, NJ 07703, USA

Are you aware of time or space diversity schemes that have been used to offset the effects of scintillations?

Author's Reply

We know of some schemes. I doubt whether simple and inexpensive schemes can be developed for low elevation, or for strong polar and equatorial scintillations.

J..\arons, Air Force Geophysics Lab., Bedford, Mass., USA

CRC studies are in progress at 1° of elevation on the ANIK satellite (4 and 6 GHz) on space diversity gains.

Author's Reply

Studies of this nature should be supported as much as possible.

J.M.Goodman, Space Systems Division, Naval Research Laboratory, Washington D.C., USA

Your paper indicates that tropospheric scintillation is an important factor below 30° elevation and can also be important at UHF and above. Low elevation studies at NRL using microwave frequencies show periodicities in angle-of-arrival (in the vertical) which are attributed to buoyancy oscillations in the atmospheric boundary layer. Your study makes me more confident in the NRL microwave studies results, and you are certainly correct in pointing out this very important potential limitation to which too little attention has been given.

Author's Reply

It is good to get additional support for an idea of tropospheric low angle effects which was, till now, subject of strong controversies.

A REVIEW OF LF/VLF RADIO NAVIGATION SYSTEMS AND SOME RELATED PROPAGATION INFLUENCES

Dr. B. Burgess
Royal Aircraft Establishment
FARNBOROUGH, Hants
UK

SUMMARY

A review is given of the important radio navigational aids that have been developed in the LF/VLF bands since the 1940s. Attention is paid to the operational requirement that could be met by this type of aid and the effects that the propagation characteristics of the radio waves at these frequencies have on the performance of the aids. This paper is not intended to be a comprehensive review but mainly to set the scene to specialist papers in the sessions on "LF/VLF systems".

1. INTRODUCTION

The LF/VLF radio navigational aids that are currently in use or will be in operation in the near future are situated either in the 10-14 kHz VLF band or between the low frequencies (LF) 70 and 130 kHz. The LF band came into use for radio navigation just after the end of World War II with the introduction of the British Decca system. The US Loran system, developed during World War II, and operating at a frequency of about 2 MHz, was in the 1950s developed for use at LF and eventually came into being in the late 1950s as a pulsed system operating at 100 kHz and using a bandwidth of $\pm$10 kHz. The HF version of Loran has become known as Loran-A while that operating at 100 kHz is known as Loran-C.

In the late 1940s and the 1950s a number of other radio navigational aids were conceived and various tests undertaken. Among systems proposed during these times were the British Delrac and Dectra systems, operating in the LF and VLF bands respectively. Flight tests of the Dectra system were performed in the 1960s. The 1960s saw the steady development of the Omega system and the expansion of Decca and Loran-C on a world wide basis. Omega uses frequencies in the internationally assigned frequency band of 10-14 kHz set apart for navigational use, but proposals and indeed systems have been made utilising transmissions from VLF communications and frequency standard transmitters in the band 14-30 kHz. Note should also be made that the USSR are developing a VLF navigation system on a similar basis to Omega and also operate two Loran-C chains that are compatible with the US system (Ref 1). A variant of Loran-C for tactical military use was also developed at this time and is known as Loran-D (Ref 2).

Nearly all the above mentioned systems operate as hyperbolic navigational aids, a term used to represent the geometry made by the measurement of phase or time difference of the received signals. Omega and Decca are examples of phase comparison systems while Loran measures the difference in time of arrival of radiated pulses of energy. In this paper I would like to review the principle features of these systems and the role that radio wave propagation plays in determining the accuracy and reliability of these systems.

Before doing this, however, attention will be paid to the requirements of users of these types of systems to give a perspective on the role of LF/VLF radio navigation systems.

2. OPERATIONAL REQUIREMENTS

LF/VLF radio navigational aids, because of their ability to be used at distances greater than line of sight, have applications ranging from use on a world wide basis to that in harbour entrance zones; large high speed aircraft to small fishing ships with a spectrum of accuracy and environmental requirement.

For maritime users one can differentiate between three categories of requirements: oceanic, coastal/confluence zones and harbour areas. For oceanic or high seas operation it is currently estimated (Ref 3) that position accuracy is in the range 2-5 nm and that improved accuracy to 1-2 nm would be desirable though not essential. For coastal/confluence zones which cover areas where the distance from the coast is up to 50 nm or to the edge of continental shelf a general guide to the requirement is a repeatable rms accuracy of 0.25 nm, though in certain areas a more stringent requirement will exist. Harbour and harbour entrance areas requires the highest degree of accuracy to avoid collisions and stranding. Accuracies in this case are now possibly limited by the size of the vessels.

Turning to aviation users (Ref 4) one may similarly define requirements according to categories: oceanic and area navigation (RNAV). This latter category can be subdivided into en route, terminal and approach. Taking as our example of oceanic routes the North Atlantic area, current lateral separation of aircraft flying in the same direction and at the same height is nominally 120 n miles, which is reflected

in a 10 nm rms accuracy in position determination for the aircraft navigation system. Studies of future traffic in this area points to a need for an increased navigation accuracy and a reasonable goal for future requirements is an accuracy of the order 1–2 nm (1σ).

Norwood (Ref 4) in his paper on the development of air navigation system requirements gives the following as the rms crosstrack errors in nm (metres) applicable to RNAV systems.

	Present	Future
En route	1.5 (3000)	0.25 (500)
Terminal	1.0 (2000)	0.10 (200)
Approach	0.5 (1000)	0.05 (100)

Other papers (Refs 5, 6) give somewhat different figures, especially for the future requirements, but the above accuracy figures for present requirements are typical, with improvements for future requirements varying by a factor of 2 or more.

Summarizing it is evident that there are varying requirements in accuracies of the position of airborne and marine vehicles that are used on long or relatively long distance journeys. These range from some 10 nm for long range commercial aircraft down to something of the order of 100 metres or less for harbour or airport approaches, or fishing vessels. In stating these performance figures, this usually applies to the system as a whole, so that the errors that can be allowed for propagation variations must be somewhat less than these overall figures.

3. LF/VLF SYSTEMS

3.1 Decca (70–130 kHz)

Decca is an LF continuous wave hyperbolic system which is in current commercial use and over the past 30 years some 50 Decca chains have been installed throughout the world. Decca is basically a high accuracy, short to medium range aid intended for coastal and land navigation. A Decca chain consists of a master station and two or three slave stations, situated about 100 to 200 km from the master station. The master and slave stations radiate on different frequencies, normally near to 70, 85, 113 and 126 kHz, but having a fixed relationship to one another; the phase of the signal radiated by the slave stations being referenced to that received from the master station. A receiver converts the different received frequencies to a common frequency for phase comparison and a lattice of hyperbolae are formed with the master and each slave as focii in turn.

The Decca system accuracy predictions are based on some 30 years data gathering from operational systems. Daytime accuracies are limited by signal to noise considerations, the signal being essentially ground wave only. At night time, with the disappearance of the D region, the skywave signal becomes marked and performance consequently degrades since the phase comparison system cannot differentiate skywave from ground wave. Accuracies better than 0.05 nm (rms) out to ranges of 100 to 150 nm are achieved during daytime, the accuracy degrading with distance from the transmitting stations. At night-time this degrades by about a factor of 5 (Ref 7).

Ambiguity resolution is achieved by momentarily interrupting the regular chain transmissions, so that combinations of frequencies can be radiated from a given transmitter. This then enables a coarse hyperbolic pattern to be set up to identify the correct lane.

From a propagation point of view it is worth noting that one is comparing the propagation delay on different frequencies over differing paths.

3.2 Dectra

Dectra is an area coverage navigation and operating on the 70 kHz frequencies associated with Decca. Its aim is to permit close lateral track separations to be flown on transoceanic routes such as those across the North Atlantic. In the system that has been tested a pair of master–slave stations was located in Newfoundland and another pair in UK. In contrast to the Decca system both the master and slave radiate the same frequency thus minimising the dispersion characteristics of the propagation which in this system is mainly by skywave. These transmissions give a tracking pattern. The master station at the other end of the oceanic route radiates a frequency close to that of the other pair, and this difference frequency generates a hyperbolic ranging pattern with focii at either end of the oceanic chain.

Good accuracy has been claimed for this system from both airborne and maritime trials that have taken place in the 1960s, typical values being 2.5 to 5 nm (1σ) over a large area of the North Atlantic. (Ref 8)

3.3 Loran–C

Loran–C is a high accuracy, long range pulsed ground wave radio navigation system with an operating frequency of 100 kHz. It has been developed from Loran–A and has been in operation for some 20 years. While being a hyperbolic system like Decca and also employing master and slave stations which radiate a signal synchronised to the master transmission, it differs fundamentally in employing only one carrier frequency and relies on differential time delay as opposed to phase measurements to give the hyperbolic lines of position. Its great virtue is that due to the different propagation delay of skywave and groundwave, the Loran–C system is able to obtain navigation information from only the groundwave component of the transmitted signal which (1) propagates to ranges in excess of 1000 nm and (2) has highly stable propagation characteristics.

At present there are nine chains in operation, which cover a substantial part of the Northern hemisphere. Typical accuracy figures (Ref 6) are 50 m (rms) at 200 nm, increasing in a linear fashion to some 200 m at 1000 nm ranges.

3.4 Loran-D

Loran-D is a development of Loran-C for tactical military use (Ref 2). It is compatible with Loran-C and is no different in basic principle. In achieving tactical mobility, the radiated power from Loran-D is significantly less; however due to the propagation properties of LF groundwave a 500 nm Loran system only needs one hundredth of the radiated power of a 1000 nm system. Another advantage accrues in virtue of the greater difference in time delay between the ground and skywave. This allows the sampling of the groundwave at a later time in the received pulse and hence at a higher signal level. Otherwise from a propagation viewpoint there is no difference between the two systems Loran-C and D.

Loran-D has basically achieved the objectives of the Cytac System, the predecessor of Loran-C.

3.5 Omega

Omega is a world-wide radio navigational aid (Ref 9), using frequencies in the 10-14 kHz band, which should be fully operational in the late 1970s. A pair of transmitting stations provides the navigator with a family of hyperbolic lines-of-position (lop), and eight transmitting stations with 5000-6000 nm baselines will give a global coverage. The eight transmitters radiate the various frequencies in a time division multiplex format. All transmitters will be phase synchronised and repeat their individual radiated format every 10 seconds.

The 10.2, 11⅓ and 13.6 kHz transmissions are the basic navigation frequencies; other radiated frequencies, unique to each transmitter, may be used for navigation by distance measurement based on a cycle counting technique.

The 10.2 kHz transmissions generate lops which have ambiguities spaced by 15 km on the baseline joining two transmitters; for marine use, where the vessel is slow moving, the resolution of these 15 km ambiguities can be achieved by dead reckoning means. However if necessary, the 13.6 kHz transmissions could be utilised in conjunction with the 10.2 kHz transmissions to give ambiguous lops separated by some 45 km on the baseline. Similarly the 10.2 kHz and 11⅓ kHz transmissions can give ambiguous lops separated by 135km.

The exact position of these lops on a chart, relative to known geographic coordinates are governed by the value of the propagation velocity, v, that is used for the appropriate frequency. For 10.2 kHz, the standard value used is given by $c/v = 0.9974$ where c is the velocity of propagation in free space (299792.5 m/s). This velocity of propagation is a function of many geophysical parameters such as ground conductivity, direction of propagation, effective height of ionospheric reflection, which cause the lops to vary about the charted values, temporally and spatially. Corrections can be applied to minimise these variations and these effects are discussed in a paper by Swanson at this Conference (Ref 10).

In assessing the accuracy that is possible using Omega, attention must be paid to the way it is used. Three ways in which Omega may be used are (1) using only a single frequency, (2) using a difference frequency, and (3) using it in the differential mode.

In using a single frequency an accuracy in the range 0.5 to 2.0 nm may be achieved. This is probably the simplest use of the system but suffers from the possibility of lane ambiguity and the necessity to use 'skywave corrections' (SWC). When use is made of the difference frequency techniques, lane identification problem are eased, and SWCs are not so necessary if one can accept the degraded accuracy which will be in the range 1-5 nm. The third techniuqe called 'Differential Omega' will give a much improved accuracy in the range 0.25-0.5 nm, but needs the creation of a system of monitor stations. These monitor stations situated within say 100-200 miles of the user can be used to correct the Omega measurements at the user, since phase variations due to propagation are highly correlated over these separation distances (Ref 11).

4. PROPAGATION CHARACTERISTICS OF LF/VLF

4.1 Introduction

A considerable literature has been published dealing with both the theoretical and experimental aspects of VLF and LF propagation, especially over the past twenty five years. In discussing propagation at these frequencies it is desirable at the VLF end of the frequency scale to talk in terms of waveguide modes, while at the LF end of the spectrum to consider propagation as the summation of a series of wave-hops. These terms have arisen from a theoretical approach to the subject, but we must note that Johler (Ref 12) has developed a theory which covers the frequency range from VLF to LF which considers propagation in terms of spherical wave functions. This approach has the virtue that with the use of high speed computers, it is a more viable approach, economically, to producing propagation data, theoretically, at these frequencies.

4.2 LF Propagation

In the wavehop theory, (Ref 13) the full wave solution for propagation between a spherical earth and a concentric ionosphere can be expanded into a series of complex integrals. If these integrals are replaced by their saddle-point approximations the series can be identified as the ray hop series of geometric optics, and thus these integrals are called wave-hops. The saddle point approximation is inadequate near and beyond the caustic and in this area the wavehops can be evaluated by numerical integration or by summing a residue series. Berry and Chisman (Ref 13) represent the series in the form:

$$E = E_o + \sum_{j=1.} R_j I_j$$

where E_o is the groundwave and R_j is the ionospheric reflection coefficient, computed for any model of the ionosphere that is desired. The path integral I_j takes into account ground conductivity, path length, reflection height and earth curvature.

An indication of the propagation characteristics of LF signals can be obtained from this theory if one computes the signal level as a function of distance, up to say 2000 km, for both day and night conditions. Under daytime conditions the propagation is very stable, with a stable interference pattern. At the lower frequencies, the skywave becomes important at distances of the order of 500 km, but even so the skywave shows good stability in both phase and amplitude. This has been bourne out by experimental results on frequencies in the range 45 to 80 kHz (Ref 14).

Under night time conditions, the skywave component is much stronger, due to the disappearance of the ionospheric D region, and strong interference between ground and skywave occurs at ranges starting from around 200 km from the transmitter. Under these conditions the most severe interference takes place in the range 200 to 1000 km where signal strength patterns can show fades of up to 40 dB. Beyond a 1000 km range, the propagation tends to stabilise. Again these results are confirmed by experimental values, (Ref 14) but at LF frequencies there is a lack of experimental data, compared with VLF, to substantiate theory.

4.3 VLF Propagation

At VLF frequencies for long range propagation the earth behaves like a good electrical conductor with a reflection coefficient approaching +1 while the ionosphere has a reflection coefficient approaching −1. It is convenient thus to treat the earth and the ionosphere as the two boundaries of a spherical waveguide and to consider propagation to great distances in terms of waveguide modes (Ref 15,16). For each mode there are three parameters which govern the characteristics of the mode, ie the attenuation rate, α dB/Mm, the phase velocity of the mode, usually quoted relative to the velocity of light, v/c, and the excitation factor A which is approximately the ratio of the power launched into the earth ionosphere waveguide to that launched into a flat waveguide with perfectly conducting boundaries.

From the experimental and theoretical considerations of VLF propagation the lower frequency band of 10–14 kHz has definite advantages over the higher frequencies for navigational purposes. The main advantage is that due to the frequency dependence of the attenuation and excitation factors of the modes, the second mode perturbing effects are much smaller in this lower frequency band.

At frequencies below 10 kHz the attenuation rates increase rapidly to high values. This coupled with the economics of designing efficient ground transmitting aerials at this end of the band effectively puts a lower limit of about 8 kHz to the optimum band for navigation use.

These considerations together with the maxim of using as low a frequency as possible for lane resolution has led to the OMEGA World-Wide Navigational System choosing its lowest frequency of transmission as 10.2 kHz.

Propagation effects as they affect the Omega system will be discussed in detail by Swanson (Ref 10), but since the higher VLF frequencies of 14–30 kHz have been proposed for navigational uses, it is relevant to note briefly some differences in propagation characteristics between these frequencies.

Of particular importance for navigational purposes is the overall stability of propagation. This stability is demonstrated by the consistent diurnal pattern of the phase delay of VLF transmission observed at a distant point. At 10.2 kHz, the daytime, night-time and transitional dawn/dusk periods give a variation in stability of about 2 to 1, with no evidence of cycle slipping. However at frequencies somewhat above 16 kHz, at long ranges, the transitional dawn/dusk period is noted for the phenomena of cycle slippage, due to the low attenuation of higher order mode propagation under night conditions (Ref 14).

This strong miltimode nature of propagation at the higher frequencies will also mean that lane resolution techniques will be more unreliable than at the Omega frequencies at night, and also that non-reciprocal propagation effects and anomalies in position fixing due to the proximity of transmitters will be enhanced. Thus particular care will have to be used in choosing the appropriate transmissions at these higher frequencies.

5. CONCLUDING REMARKS

The paper has attempted to give an overview of LF and VLF radio navigational systems as they have developed since the end of World War II. Indications have been given of the accuracy that can be obtained from these aids together with the underlying principles of operation of the systems. Attention has also been drawn to the requirements of navigators who look to this type of navigational aid to satisfy their needs. It is seen that to a large degree the aids are complementary. Finally a brief discussion is given of the propagation characteristics that govern LF and VLF transmissions and their relevance to the performance of the navigation systems.

REFERENCES

1 BEUKERS, J. M., 1974, "A Review and Applications of VLF and LF Transmissions for Navigation and Tracking", J. Inst. Nav, Vol. 21 No. 2 p. 117-133, (Summer).

2 FRANK, R. L., 1974, "Current Developments in Loran-D", J. Inst. Nav, Vol. 21, No. 3, p. 234-241, (Fall).

3 HAISLIP, D. T., GOLDSMITH, A., 1974, "Meeting the Maritime Requirements in United States Waters", J. Inst. Nav, Vol. 21, No. 2, pp. 159-164, (Summer).

4 NORWOOD, A., 1975, "Development of Navigation Systems for Advanced Commercial Transports", J. Inst. Nav, Vol. 22, No. 1, pp 68-75, (Spring).

5 TYLER, J. S., BRANDEWIE, D. M., HEINE, W., ADAMS, R., 1975, "Area Navigation Systems: Present Performance and Future Requirements", J. Inst. Nav, Vol. 22, pp. 95-111, (Summer).

6 HOLLISTER, W. M., DODGE, S. M., 1975, "An Evaluation of Differential Omega for General Aviation Area Navigation", J. Inst. Nav, Vol. 22, No. 3, pp. 259-273, (Fall).

7 BLOOD, E. B., 1973, "Navigation aid Requirement for the US Coastal/Confluence Region", J. Inst. Nav, Vol. 20, No. 1, pp 68-76, (Spring).

8 The Decca Navigator Company, "Decca 2".

9 PIERCE, J. A., 1975, "Omega", IEEE Trans. on Aerospace and Electronic Systems, AES-1, pp. 206-15.

10 SWANSON, E. R., 1976, "Propagation Effects on Omega", Agard Conference: "Propagation Limitations of Navigation and Positioning Systems" Instanbul, (October).

11 KASPER, J. F., CREEKMORE, E. E., "Omega Utilization by Non-Military Subscribers", J. Inst. Nav, Vol. 19, No. 3, pp 215-225, Fall 1972.

12 JOHLER, J. R., 1970, "Spherical wave Theory for MF, LF and VLF Propagation", Radio Science, Vol. 5, No. 12, pp. 1429-1444.

13 BERRY, L. A., CHRISMAN, M. E., 1965, "The Path Integrals of LF/VLF Wave-Hop Theory", Radio Science, Vol. 69D, No. 11, pp. 1469-80.

14 BURGESS, B., JONES, T. B., 1975, "The Propagation of LF and VLF Radio Waves with Reference to Some System Applications", The Radio and Electronic Engineer, Vol. 45, No. 1/2, pp. 47-61.

15 BUDDEN, K. G., 1961, "The Waveguide Mode Theory of Wave Propagation", Logos Press, London.

16 WAIT, J. R., "Electromagnetic Waves in Stratified Media", McMillan, New York, 1962.

DISCUSSION

A.Biggs

Katsafrakis at SRI made phase measurements over the Antarctic ice pack – and Guy etc made phase measurements over Greenland – and we made them over the floating Arctic ice pack. How do you reconcile the phase change over Greenland? With respect to the floating ice pack, do you expect any observable changes in propagation parameters with the thin (5–10 meters) Arctic ice pack?

Author's Reply

There is definitely a change in the phase velocity of propagation of VLF radio waves as the waves cross from a sea or land mass to an ice-cap area. This has been predicted by theory and confirmed by experimental measurements. In order for this change in phase velocity to be of a significant magnitude, the depth of the ice must be comparable to the skin depth, at these frequencies, of ice. Under these considerations, the ice cap of Greenland, which is some 3000 meters thick, greatly affects VLF propagation, while the Arctic ice pack has little effect.

Effects of Irregular Media on Navigation and Positioning Systems--Full Wave Solutions*

E. Bahar
Electrical Engineering Department, University of Nebraska
Lincoln, Nebraska 68588

ABSTRACT

The steady state magnitude and phase of radio waves are affected by the medium through which they propagate. Since these effects are frequency dependent, pulsed signals will undergo distortions and delays that depend on the medium of propagation. Thus when CW radio signals are used for navigation (Omega System), it is necessary to determine the deviations in the phase of received signals (phase anomalies) due to medium effects. Similarly, when pulsed signals are used for navigation (Loran C system), it is necessary to determine the signal delays due to medium effects.

Radio wave propagation over several irregular propagation paths is examined in detail. The phase anomalies, the delays and the distortions for pulsed radio signals transmitted across hills or valleys on the earth's surface are determined using a full wave approach.

1. INTRODUCTION

When the medium of propagation is represented reasonably well by simple models such as horizontal or concentric spherical or cylindrical dielectric layers it is possible to use standard separable solutions to the propagation problem to determine the corresponding phase anomalies and pulse delays.

As the demands on the navigation systems become more stringent, it is necessary to obtain more accurate predictions of the received signals. Thus for many relevant propagation problems it is not sufficient to use such idealized models of the propagation medium. More realistic models of the propagation medium consist of irregular boundaries and stratified media with layers of nonuniform thickness and varying electromagnetic parameters. For these models it is not possible to obtain separable solutions to the propagation problem. In these cases the fields are expanded in terms of complete sets of basis functions that reflect the local features of the medium, and Maxwell's equations are transformed into coupled sets of first-order differential equations for the complex wave amplitudes. Thus, the fields at the surface of the earth are expressed completely as a superposition of vertically and horizontally polarized surface waves, lateral waves and radiation fields.

In this work propagation over irregular models of the earth's crust is examined for LORAN C pulse excitations. Expressions for the envelope of the transient response and the effects of the ground parameters on the pulse delays and the phase anomalies are determined.

2. FORMULATION OF THE PROBLEM

The scattering of vertically polarized waves by a one dimensionally rough surface is considered here in detail. The interface between two semi-infinite media of complex permittivity ϵ and permeability μ is given by $h(x)$ (Fig. 1). Thus

$$\epsilon = \epsilon_R - i\epsilon_I = \epsilon_k - \frac{i\sigma}{\omega} = \begin{cases} \epsilon_o, & y > h(x) \\ \epsilon_1, & y < h(x) \end{cases}, \quad \mu = \begin{cases} \mu_o \\ \mu_1 \end{cases} \tag{2.1}$$

The vertically polarized waves are excited by a magnetic line source $\bar{J}_m$ (analogous to the electric line current $\bar{J}$) and to obtain the steady state solutions an $\exp(i\omega t)$ time dependence is assumed.

Using a full wave approach, the fields are expressed in terms of generalized Fourier transforms that are suitable for the treetment of rough surface scattering. Thus (Bahar 1972a) the magnetic field component H_z is expressed as

$$H_z(x,y) = H_o(x,y) + H_1(x,y) + H_s(x,y) = \Sigma H(x,u)\psi(u,y)$$

$$= \int_0^\infty H_o(x,y)\psi_o(u,y)du_o + \int_0^o H_1(x,y)\psi_1(u,y)du_1 + H_s(x,y)\psi_s(x,y) \tag{2.2a}$$

where the field transforms ere

$$H_m(x,u) = \int_{-\infty}^{\infty} H_z(x,y)Z(u,y)\Psi_m(u,y)dy, \quad m = 0,1 \text{ or } s \tag{2.2b}$$

and H_o, H_1 and H_s ere the radiation field, the lateral wave end the surfece wave terms respectively. The corresponding basis functions ere

$$\Psi_o(u,y) = R_{oh}(u)\psi_o(u,y) = \begin{cases} [\exp(iu_o y) + R_{oh}(u)\exp(-iu_o y)]/[2\pi Z_o(u)]^{\frac{1}{2}}, & y \geq h(x) \\ \exp[iu_1(y-h)\exp(iu_o h)T_o(u)/[2\pi Z_o(u)]^{\frac{1}{2}}, & y \leq h(x) \end{cases} \tag{2.3a}$$

$$\Psi_1(u,y) = R_{1h}(u)\psi_1(u,y) = \begin{cases} \exp -iu_o(y-h)\ \exp(-iu_1 h)T_1(u)/[2\pi Z_1(u)]^{\frac{1}{2}}, & y \geq h(x) \\ [\exp(-iu_1 y) + R_{1h}(u)\exp(iu_1 y)]/[2\pi Z_1(u)]^{\frac{1}{2}}, & y \leq h(x) \end{cases} \tag{2.3b}$$

---

*This work was supported by the U.S. Army Research Office, under Grant No. DAHC04-74-G-0074.

and

$$\Psi_s(u,y) = \psi_s(u,y) = \Psi_s(u,h) \begin{cases} \exp[-iu_{os}(y-h)], & y \geq h(x) \\ \exp[iu_{1s}(y-h)], & y \leq h(x) \end{cases} \tag{2.3c}$$

where

$$[\Psi_s(u,h)]^2 = 2iu_{os}/Z_o(u_s)[1-(\varepsilon_o/\varepsilon_1)^2] = -2i\omega\varepsilon_o/n(1-1/\varepsilon_r^2) \tag{2.3d}$$

and n is the refractive index

$$n^2 = \varepsilon_1/\varepsilon_o = \varepsilon_r \tag{2.4}$$

The transverse wave impedance is

$$Z(u,y) = \beta/\omega\varepsilon = \begin{cases} \beta/\omega\varepsilon_o = Z_o(u), & y > h(x) \\ \beta/\omega\varepsilon_1 = Z_1(u), & y < h(x) \end{cases} \tag{2.5}$$

where

$$\beta = (k_{o,1}^2 - u_{o,1}^2)^{\frac{1}{2}}, \quad \mathrm{Im}(\beta) \leq 0 \tag{2.6}$$

and $k_{o,1} = \omega(\mu_o,\varepsilon_{o,1})^{\frac{1}{2}}$ is the wave number.

The reflection coefficients referred to the surface y = 0 are

$$R_{oh}(u) = R_o(u)\exp(i2u_o h), \; R_{1h}(u) = R_1(u)\exp(-i2u_1 h) \tag{2.7a}$$

and the Fresnel reflection and transmission coefficients (referred to the boundary y = h) are

$$R_o(u) = -R_1(u) = (u_o\varepsilon_1 - u_1\varepsilon_o)/(u_o\varepsilon_1 + u_1\varepsilon_o) \text{ and } T_{o,1}(u) = 1 + R_{o,1}(u) \tag{2.7b}$$

For the surface wave $1/R_o(u_s) = 0$, thus

$$u_{os}\varepsilon_1 + u_{1s}\varepsilon_o = 0, \quad \mathrm{Im}(u_{o,1}) < 0 \tag{2.8a}$$

Thus assuming principal square roots

$$\beta_s = k_o\sqrt{\frac{\varepsilon_r}{1+\varepsilon_r}}, \; u_{os} = \frac{-k_o}{\sqrt{1+\varepsilon_r}} \text{ and } u_{1s} = \frac{k_o\varepsilon_r}{\sqrt{1+\varepsilon_r}} \tag{2.8b}$$

E_y is

$$E_y(x,y) = \Sigma E(x,y)Z(u,y)\psi(u,y) \tag{2.9a}$$

where

$$E_m(x,y) = \int_{-\infty}^{\infty} E_y(x,y)\Psi_m(u,y)dy, \quad m = 0,1 \text{ or } s \tag{2.9b}$$

For the full wave analysis to the problem, the fields are subjected to the exact boundary conditions at the interface y = h(x). Thus (Bahar 1972b)

$$H_z(x,h^+) = H_z(x,h^-) \tag{2.10a}$$

and

$$\left[E_y(x,h^+) - E_y(x,h^-)\right]\frac{dh}{dx} = \frac{i}{\omega}\frac{\partial}{\partial y}\left[\frac{1}{\varepsilon_o}H_z(x,h^+) - \frac{1}{\varepsilon_1}H_z(x,h^-)\right] \tag{2.10b}$$

The differential equations for the traneverse field componente E_y and H_z are

$$-\partial E_y/\partial x = \frac{i}{\omega\varepsilon}[k^2 H_z + \partial^2 H_z/\partial y^2] + J_m \tag{2.11a}$$

$$-\partial H_z/\partial x = i\omega\varepsilon E_y \tag{2.11b}$$

On substituting the field traneforms (2.2) and (2.9) into (2.11), subject to the boundary conditione (2.10), Maxwell'e equations are converted into generalized telegraphist's equations for the field transforms. Due to the irregularities of the boundary y = h(x), the first order differential equations for the field transforms are coupled. Thus the incident radiation fielde are scattered at the irregular boundary in non-specular directiona and converted into surface and lateral waves. Similarly, the incident surface and lateral waves are coupled into the other full wave components of the field.

The dual problem of scattering of horizontally polarized waves can be analyzed in a similar manner by replacing the magnetic line source with an electric line source and interchanging the electric with the magnetic fields and the dielectric coefficients with the magnetic permeability. The more general case of scattering by two dimensionally rough surface, where effects of depolarization are important, can also be solved using the full wave approach (Bahar 1974). When it is neceesary to account for the earth curvature effects, either irregular cylindrical or epherical modele of the earth can be used (Bahar 1975, 1976).

3. STEADY STATE FULL WAVE SOLUTIONS FOR THE SCATTERED FIELDS

For a line source at a large distance from the nonuniform boundary, the scattered radiation field is given by (Bahar 1972b)

$$H_z^r(x,y) = |H_{sp}|\exp[-ik_o(\rho_o+\rho)]\left(-\frac{i\omega}{\omega_c}\right) F(C^f,C^i)I(C^f,C^i,h,L) \tag{3.1a}$$

where $|H_{sp}|$ is the magnitude of the specularly reflected field for a flat perfectly conducting surface at the carrier frequency ω_c.

$$|H_{sp}| = K\, k_{oc}\, LC_o^i/2\pi\eta_o(\rho_o\rho)^{\frac{1}{2}} = |H_o^i|2k_{oc}\, LC_o^i/(2k_{oc}\rho)^{\frac{1}{2}} \tag{3.1b}$$

The magnitude of the unperturbed incident wave at the origin is $|H_o^i|$; $\rho_o = (x_o^2+y^2)^{\frac{1}{2}}$, $\rho = (x^2+y^2)^{\frac{1}{2}}$ are the distances from the source and field point to the origin (see Fig. 1) and $k_{oc} = \omega_c(\mu_o\varepsilon_o)^{\frac{1}{2}}$ is the free space wave number at the carrier frequency.

$$F(C^f,C^i) = \frac{2C_o^f(C_1^fC_1^i-S_o^fS_o^i)(1-1/\varepsilon_r)}{[C_o^i+C_1^i/n][C_o^f+C_1^f/n][C_o^i+C_o^f]} \tag{3.1c}$$

and

$$I(C^f,C^i,h,L) = \frac{1}{2L}\int_{-L}^{L}\exp\left\{i(C_o^i+C_o^f)k_oh + i(S_o^f-S_o^i)k_ox\right\}dx \tag{3.1d}$$

where C_o^i and S_o^i are the cosine and sine of the angle of incidence, θ_o^i in the medium $y > h$ and C_o^f and S_o^f are the cosine and sine of the scattering angle θ_o^f. The corresponding expressions for the medium $y < h$ are denoted by symbols with subscript 1 instead of 0. Thus

$$u_{o,1}^{i,f} = k_{o,1}\, C_{o,1}^{i,f} = k_{o,1}\cos\theta_{o,1}^{i,f} \tag{3.2a}$$

and

$$\beta_{o,1}^{i,f} = k_{o,1}\, S_{o,1}^{i,f} = k_{o,1}\sin\theta_{o,1}^{i,f} \tag{3.2b}$$

The intrinsic impedance is

$$\eta = \begin{cases}\eta_o = \sqrt{\mu_o/\varepsilon_o}\\ \eta_1 = \sqrt{\mu_1/\varepsilon_1}\end{cases} \tag{3.3}$$

and the nonuniform boundary is given by

$$y = h(x),\ -L \le x \le L \tag{3.4}$$

when medium 1 ($y < h$) is highly conducting and the fields in the region $y < h$ are of no practical interest, it is convenient to analyze the problem by characterizing the interface $y = h(x)$ by an approximate surface impedance for waves of grazing incidence (Bahar 1972c)

$$Z_s \to \eta_o(\varepsilon_r-1)^{\frac{1}{2}}/\varepsilon_r \tag{3.5}$$

The full wave solutions for the scattered radiation fields (3.1) satisfy the reciprocity relationships in electromagnetic theory.

The scattered radiation fields (3.1) vanish near the boundary $\theta_o^f \to \frac{\pi}{2}$. This agrees with the empirically well known result that for any surface, regardless of roughness, a scattered radiation field will not exist at the surface except for grazing incidence (Beckmann and Spizzichino, 1963). However, due to surface irregularities, the incident radiation fields excited by line sources at large distances, are coupled into guided surface waves and lateral waves (see Fig. 2). The scattered surface waves and lateral waves could therefore be the predominant contributions to the total field near irregular boundaries.

The scattered surface waves due to line sources at large distances from the boundary is (Bahar 1972b).

$$H_{zo}^s(x,y) = |H_o^i|\exp[-ik_o\rho_o]\left(\frac{i\omega}{\omega_c}\right)^{\frac{1}{2}} 4ik_oLF(C_s,C^i)I(C_s,C^i,h,L)\cdot\exp[-i(\beta_s x + u_{os}y)] \tag{3.6e}$$

where

$$F(C_s,C^i) = (C_1^iC_{1s}-S_o^iS_{os})/(C_o^i+C_1^i/n)n(1+1/\varepsilon_r) \tag{3.6b}$$

and

$$I(C_s,C^i,h,L) = \frac{1}{2L}\int_{-L}^{L}\exp[iu_o^ih + i(\beta_s-\beta^i)x]dx \tag{3.6c}$$

This solution for the scattered surface wave is also consistent with reciprocity relationships. For highly conducting media 1, the surface impedance concept (3.5) can be used to characterize the boundary $y = h(x)$ to obtain approximate expressions for the scattered surface wave.

The scattered lateral waves due to line sources at large distances from the boundary is (Bahar 1972b)

$$H_{zo}^\ell(x,y) = |H_o^i|\exp\left[-ik_o\rho_o\right]\exp\left[-i(k_1L_1+k_oL_\rho)\right]\frac{(2L/\lambda_c)}{2\pi(L_1/\lambda_c)^{3/2}}\ F(C^\delta,C^i)\cdot I(C^\delta,C^i,h,L) \tag{3.7a}$$

where

$$F(C^{\delta},C^{i}) = (nC_1^iC_o^{\delta} + S_o^iS_o^{\delta})/(C_o^i + C_1^i/n)n^{3/2}\varepsilon_r \tag{3.7b}$$

and

$$I(C^{\delta},C^{i},h,L) = \frac{1}{2L}\int_{-L}^{L}\exp[iu_o^ih + i(\beta^{\delta}-\beta^i)x]dx \tag{3.7c}$$

in which λ_c is the free space wavelength at the carrier frequency ω_c

$$\beta^{\delta} = k_oS_o^{\delta} = k_1S_1^{\delta} = k_1 = k_on,\ u_o^{\delta} = k_oC_o^{\delta} = k_o\sqrt{1-n^2},\ u_1^{\delta} = k_1C_1^{\delta} = 0 \tag{3.7d}$$

The lateral waves are of practical significance only for low loss medium 1 where

$$n = n'-in'' = S_o^{\delta} = \sin\delta = \sin(\delta'-i\delta'') \simeq \sin\delta'-i\delta''\cos\delta' \tag{3.8}$$

Thus

$$u_o^{\delta}y + \beta^{\delta}x \simeq k_oL_{\delta} + k_1L_1 \tag{3.9a}$$

and the distances L_{δ} and L_1 are (Fig. 2)

$$L_{\delta} = y/\cos\delta' \text{ and } L_1 = x-y\tan\delta' \tag{3.9b}$$

where δ' is the critical angle for total internal reflection. This solution for the scattered lateral wave is consistent with reciprocity relationships. The lateral wave contribution cannot be darived if the approximate surface impedance concept (3.5) is used.

When the line source is near the boundary h(x) it excites surface waves and lataral waves in addition to the radiation term. In this case it is also necessary to consider tha coupling between the lateral waves and the surface wave at the irregular boundary. Tha surface wave generated at the irregular boundary h(x) by an incident lateral wave is

$$H_{\ell}^s(x,y) = \frac{Ki\omega\varepsilon ik_oL}{(2\pi)^{\frac{1}{2}}[-ik_1x_o]^{3/2}}\exp[-i(u_o^{\delta}y_o-\beta^{\delta}x_o)]\exp[-i(u_{os}y+\beta_{os}x)]$$

$$\cdot F(C_s,C^{\delta})I_1(u_s,u^{\delta},h,L) \tag{3.10a}$$

where

$$F(C_s,C_o^{\delta}) = [C_o^{\delta}+1][\varepsilon_r/(1-\varepsilon_i^2)]^{3/2} \tag{3.10b}$$

$$I(C_s,C^{\delta},h,L) = \frac{1}{2L}\int_{-L}^{L}\exp[-iu_o^{\delta}h-i(\beta^{\delta}-\beta_s)x]dx \tag{3.10c}$$

and for $-x_o \gg L$ the amplitude of the unperturbed latersl wava at the origin is

$$H_{\ell}(o,o) = \frac{-Ki\omega\varepsilon_o}{(1-\varepsilon_r)2(2\pi)^{\frac{1}{2}}[-ik_1x_o]^{3/2}}\exp[-i(u_o^{\delta}y_o-\beta^{\delta}x_o)] \tag{3.10d}$$

This solution for tha coupling betwean the surface wave and the incidant lateral wave by irragular surfaces is also consistent with raciprocity relationships in electromagnetic theory.

It should also ba pointed out that if the exact boundary conditions (2.10) at the irregular surfaca are raplaced by the approximate Kirchoff type boundsry conditions (Beckmann 1963), both the surface wave and lateral wave contributions vanish.

4. FULL WAVE SOLUTIONS FOR TRANSIENT EXCITATIONS

To obtain the full wave solutions for transient excitations, analytical and numarical Fourier techniquas are employed. For arbitrary tima varying excitations the instantaneous exprassion for the magnetic line source $\bar{J}_m$ is

$$\bar{J}_m(\bar{r},t) = \text{Re } f_s(t)\delta(x-x_o)\delta(y-y_o)\bar{a}_z \tag{4.1a}$$

where for convenience the excitation function $f_s(t)$ is taken to ba complex. Tha Fourier transform of $f_s(t)$ is

$$F(\omega) = F[f_s(t)] = \int_{-i\infty}^{i\infty} f_s(t)\exp(-i\omega t)dt \tag{4.1b}$$

The transfar function $H_z(\omega)$ for the scattared fialds satisfias the ralationship

$$H_z(\omega) = H_z^*(-\omega) \tag{4.2}$$

whera * is tha symbol for the complex conjugate. However, in genaral $F(\omega) \neq F^*(-\omega)$ sinca $f_s(t)$ is complex.

The transient response can be shown to be given by

$$h(\bar{r},t) = \mathrm{Re}\ F^{-1}\left[H_z(\omega)F(\omega)\right] = \mathrm{Re}\ \frac{1}{2\pi}\int_{-\infty}^{\infty} H_z(\omega)F(\omega)\exp(i\omega t)dw$$

$$\equiv \mathrm{Re}\ h_s(\bar{r},t) \tag{4.3}$$

In this work LORAN C pulse excitations are considered in detail. This pulse can be represented as a sum of three damped sinusoids (Johler and Horowitz, 1973).

$$f_s(t) = \sum_{j=1}^{3} A_j \exp(-\Gamma_j t)u(t) \tag{4.4a}$$

in which u(t) is the unit step function

$$\begin{aligned} A_1 &= i/2 \quad & \Gamma_1 &= c + i\omega_1 \quad & \omega_1 &= \omega_c \\ A_2 &= -i/4 \quad & \Gamma_2 &= c + i\omega_2 \quad & \omega_2 &= \omega_c + 2\omega_p \\ A_3 &= -i/4 \quad & \Gamma_3 &= c + i\omega_3 \quad & \omega_3 &= \omega_c - 2\omega_p \end{aligned} \tag{4.4b}$$

Thus

$$F(\omega) = \Sigma\, A_j/(s+\Gamma_j)$$

$$f_s(t) = i\exp(-ct)\sin^2 \omega_p t \exp(-i\omega_o t) \tag{4.4c}$$

$$f(t) = |f_s(t)|\sin \omega_o t \tag{4.4d}$$

and the envelope is

$$|f_s(t)| = \exp(-ct)\sin^2\omega_p t \tag{4.4e}$$

In general for dissipative media the transfer functions $H_z(\omega)$ are non-analytic. In these cases the transient response $h_s(r,t)$ is expressed as follows:

$$h_s(t) = h_L(t) + h_H(t) \tag{4.5a}$$

where

$$h_L(t) = \frac{1}{2\pi i}\sum_{j=1}^{3}\int_{-i\infty}^{i\infty} H^{(1)}(-\Gamma_j)H^{(2)}(s,-\Gamma_j)\exp\left[s(t-T_o)\right]A_i\, ds/(s+\Gamma_j)$$

$$\equiv h_z(t)\exp(-i\omega_o t) \tag{4.5b}$$

$$h_H(t) = \frac{1}{2\pi i}\sum_{j=1}^{3}\int_{-i\infty}^{i\infty}\left[H_z(s)-H^{(1)}(-\Gamma_j)H^{(2)}(s,-\Gamma_j)\exp(-sT_o)\right]\exp(st)A_j\, ds/(s+\Gamma_j) \tag{4.5c}$$

$$H_z(-\Gamma_j) = H^{(1)}(-\Gamma_j)H^{(2)}(s,-\Gamma_j)\exp(-sT_o)\Big|_{s=-\Gamma_j} \tag{4.5d}$$

and $s = i\omega$. For the scattered radiation fields (3.1) we set $|H_{sp}| = 1$ and

$$H^{(1)}(s) = F(C^f,C^i) \tag{4.6a}$$

$$H^{(2)}(s,-\Gamma_j) = \frac{1}{2\pi}\int_{-L}^{L}\exp\left[s\left\{\frac{(C_o^i+C_o^f)h + (S_o^f-S_o^i)x}{v_o}\right\}\right]_{-\Gamma_j}\exp\,(s+\Gamma_j)\tau\, dx\left(\frac{-s}{\omega_c}\right) \tag{4.6b}$$

where

$$\tau = \lim_{s\to i\infty}\left[(C_o^i+C_o^f)h + (S_o^f-S_o^i)x\right]/v_o \tag{4.6c}$$

$$T_o = \lim_{s\to i\infty}(\rho_o+\rho)/v_o \tag{4.6d}$$

and

$$v_o = 1/(\mu_o\varepsilon_o)^{\frac{1}{2}} \tag{4.6e}$$

Thus T_o is the time it takes a specularly reflected wave from a flat perfectly conducting surface to reach the observation point.

In (4.5a) the term $h_L(t)$ is due solely to the residues of the poles of the excitation transfer function F(s) and $h_H(t)$ is due to the singularities of $H_z(s)$. Thus for nondissipative media $h_H(t) \to 0$ and $|h_s(t)|$ (4.5b) is the envelope of the transient response. In general however, the envelope of the total response $h_T(t)$ can be regarded as the superposition of $|h_L(t)|$ on the function Re $h_H(t)$. Thus

$$h_T(t) = \mathrm{Re}\ h_H(t) \pm |h_z(t)| \tag{4.7}$$

Since, in this work, non-specular scattering is considered, in general, T_o is not the arrival time of the earliest signal at the observation point. The earliest signal at the observation point arrives at time

$$T_A = T_o - T_M \tag{4.8a}$$

where T_M is the largest value of τ in (4.6c). Thus if

$$d\tau/dx\Big|_{x=x_m} = 0 \text{ for } -L < x_m < L \tag{4.8b}$$

$$T_M = \begin{cases} \tau(x_m) & \text{, for } \tau(x_m) > T_1 \\ T_1 = \lim\limits_{s\to i\infty} \left|S_o^f - S_o^i\right| L/v_o, & \text{for } T_1 > \tau(x_m) \end{cases} \tag{4.8c}$$

and if (4.8b) is not satisfied $T_M = T_1$. It is convenient in this work to employ a transformation in time $t' = t - T_A$. To this end the integrands in (4.5b) and (4.5c) are multiplied by $\exp(sT_A)$ and $t' = o$, $(t = T_A)$ corresponds to the time of arrival of the earliest signal at the observation point. The time it takes the signal scattered at the stationary point $x = x_m$ to reach the observer is $T_p = T_o - \tau(x_m)$ and in general $T_p \geq T_A$. The signal scattered at the stationary point $x = x_m$ (the largest scattered signal) is not necessarily the earliest signal to arrive at the observation point.

For the scattered surface wave at $x = L$, $y = 0$ (3.6) we set $|H_o^i|4k_{oc}L = 1$ and

$$H^{(1)}(s) = \left(\frac{s}{\omega_c}\right)^{1/2} F(C_s, C^i) \tag{4.9a}$$

$$H^{(2)}(s,-\Gamma_j) = \frac{1}{2L}\int_{-L}^{L} \exp\left[s\left\{\frac{C_o^i h + (S_{os}-S_o^i)x}{v_o}\right\}\right]_{-\Gamma_j} \exp(s+\Gamma_j)\tau dx \left(\frac{s}{\omega_c}\right)$$
$$\cdot \exp\left[-i\beta_s L\right]_{-\Gamma_j} \exp\left[-T_s\Gamma_j\right] \tag{4.9b}$$

where

$$\tau = \lim_{s\to i\infty} \left[C_o^i h + (S_{os}-S_o^i)x\right]/v_o \tag{4.9c}$$

and T_s is the time it takes the surfsce wave to travel the distance L.

$$T_s = \lim_{s\to i\infty} S_{os} L/v_o \tag{4.9d}$$

Thus for the scattered surface wave

$$T_o = T_s + \lim_{s\to i\infty} \rho_o/v_o \tag{4.9e}$$

and T_A (arrival time of earliest signal) is given by (4.8) with S_o^f raplaced by S_{os} in (4.8c). Similarly for the scattered lateral waves at $x = L$ and $y = 0$, (3.7) we sat

$$|H_o^i|(2L/\lambda_c)/2\pi(L_1/\lambda_c)^{3/2} = 1 \text{ and}$$

$$H^{(1)}(S) = F(C^\delta, C^i)\left(\frac{\omega_c}{s}\right) \tag{4.10a}$$

$$H^{(2)}(s-\Gamma_j) = \frac{1}{2L}\int \exp\left[s\left\{\frac{C_o^i h + (S_o^\delta - S_o^i)x}{v_o}\right\}\right]_{-\Gamma_j} \exp(s+\Gamma_j)\tau dx \left(\frac{s}{\omega_c}\right)$$
$$\cdot \exp\left[-ik_1 L\right]_{-\Gamma_j} \exp\left[-T_L\Gamma_j\right] \tag{4.10b}$$

where

$$\tau = \lim_{s\to i\infty} \left[C_o^i h + (S_o^\delta - S_o^i)x\right]/v_o \tag{4.10c}$$

and T_L is the time it takes the lateral wave to travel the distance L

$$T_L = \lim_{s\to i\infty} S_o^\delta L/v_o \tag{4.10d}$$

Thus

$$T_o = T_L + \lim_{s\to i\infty} \rho_o/v_o \tag{4.10e}$$

and T_A is given by (4.8) with S_o^f replaced by S_o^δ

For a surface wave at $x = L$ and $y = 0$ generated at the irregular boundary by an incident lateral wave (3.10) we set $K\omega_c k_{oc} L/2\pi)^{\frac{1}{2}}(-k_{oc}x_o)^{3/2} = 1$ and

$$H^{(1)}(s) = \left(\frac{\omega_c}{s}\right)^{\frac{1}{2}} F(C_s, C^{\delta})/(S_o^{\delta})^{3/2} \tag{4.11a}$$

$$H^{(2)}(s-\Gamma_j) = \frac{1}{2L}\int_{-L}^{L} \exp\Gamma\left[-s\left\{\frac{C_o^{\delta}h + (S_o^{\delta}-S_{os})x}{v_o}\right\}\right]_{\Gamma_j} \exp(s+\Gamma_j)\tau dx \left(\frac{s}{\omega_c}\right)$$

$$\cdot \exp\left[-i\beta_s L\right]_{-\Gamma_j} \exp\left[-T_s\Gamma_j\right] \tag{4.11b}$$

where

$$\tau = \lim_{s\to i\infty} -\left[C_o^{\delta}h + (S_o^{\delta}-S_{os})x\right]/v_o \tag{4.11c}$$

and T_s is given by (4.9d). Thus in this case

$$T_o = T_s + \lim_{s\to i\infty}\left[C_o^{\delta}\, y_o - S_o^{\delta}\, x_o\right] \tag{4.11d}$$

and T_A is given by (4.8) with $(S_o^f - S_o^i)$ replaced by $(S_{os} - S_o^{\delta})$ in (4.8c)

The term $h_H(t)$ in (4.5) which is attributed to the singularities of the transfer function $H_z(\omega)$ is significant only at the leading edge of the response. However, any distortion of the leading edge of the pulse due to medium effects could result in an effective delay in arrival time of the radio signal.

The instantaneous phase $\phi(t)$ of the transient response is given by the phase of the complex response $h_s(t)$. Thus

$$\phi(t) = \arctan[\mathrm{Im}\{f_s(t)\}/\mathrm{Re}\{f_s(t)\}] \tag{4.12a}$$

For the undistorted LORAN C pulse $\phi(t) \simeq -\omega_o t$. However, in general the phase deviation (or phase anomaly) for the transient response

$$\Delta\phi(t) = \phi(t) + \omega_o t \tag{4.12b}$$

does not vanish. Very rapid variations in $\Delta\phi(t)$ occur at the leading edge of the transient response. Thus examination of the terms $h_H(t)$ and $\Delta\phi(t)$ is useful to determine medium effects on pulse delays.

5. ILLUSTRATIVE EXAMPLES

For the illustrative examples presented in this section the carrier frequency $f_c = 10^5$ Hz, $f_p = f_c/40$ and $c = 10\, f_p$. The irregular boundary between the two media is given by

$$h(x) = h_{max}[1 + \cos(\pi x/L)]/2 \qquad -L \le x \le L \tag{5.1}$$

where $2L = 10\lambda_{oc}$, $k_{oc}\, h_{max} = 2$. The free space wavelength at f_c is $\lambda_{oc} = 3.10^3$ meters.

In Fig. 3 the instantaneous, specularly scattered radiation field is plotted as a function of $t' = t - t_A$ where t_A is the arrival time of the earliest response (4.8). Medium 1 ($y < h$) is assumed to be highly conducting $\varepsilon_r = 100 - i\, 10^6\left(\frac{\omega_c}{\omega}\right)$ and $\theta_o^i = \theta_o^f = 80^o$. The envelopes of the scattered field and the undistorted LORAN C pulse are also given in Fig. 3. For convenience both have been normalized to unity at their respective peaks. Along the time axis the symbol x is marked three times. The first is at time

$$T_{x1} = T_M - \tau(x_m) \ge 0 \tag{5.2}$$

Thus T_{x1} is the difference in the time of arrival of the earliest response and the time of arrival of the response from the stationary point where $d\tau/dx = 0$ (4.8b). The second x symbol appears at the time

$$T_{x2} = T_{x1} + 1.5/f_o \tag{5.3}$$

where T_{x2} occurs 3 half periods after T_{x1}. The third zero crossing of the received signal is usually used to mark the arrival time of the LORAN C pulse. The third symbol indicates the time T_{x3} when the undistorted LORAN C pulse peaks. For the case studied in Fig. 3, the pulse scattered in the specular direction undergoes very little distortion and $T_{x1} = 0$.

In Fig. 4, the specularly scattered radiation field is plotted for the case $\varepsilon_r = 3\left(1 - i\frac{\omega_c}{\omega}\right)$. The rest of the data is as in Fig. 3. Here, too, the envelope of the scattered pulse undergoes very little distortion, however since the Fresnel reflection coefficient is not equal to unity as in the previous example, there is a marked shift in the third zero crossing. In Figs. 4b and 4c, the phase anomaly $\Delta(\phi)$, (4.12b) and Re $h_H(t)$, (4.5) are plotted as functions of time for the case considered in Figs. 4.

The 2π change in $\Delta(\phi)$ which occurs at time $T_p/2 = 1/2\ f_p = 2.10^4$ sec when the envelope (4.4e) of the LORAN pulse vanishes, can be ignored. Thus, the effects of finite ground conductivity on the response are significant only near the leading edge of the pulse. In Figures 5a, 5b and 5c the non-specularly scattered radiation field is considered. For this case $\varepsilon_r = 10(1 - i\frac{\omega_c}{\omega})$, $\theta^i = 80^o$ and $\theta_o^f = 50.8^o$. This value of θ_o^f corresponds to the direction where the scattered field is at a minimum. In Fig. 5a in which the scattered instantaneous response and envelope are plotted, we note that the pulse is distorted and that $T_{x1} > 0$. Furthermore in Fig. 5b, the undulations in phase anomaly result from the fact that near the direction of a minima ($\theta_o^f = 50.8^o$) the scattered fields undergo distructive interference. In Fig. 5c, Re $h_H(t)$ is significant only near the leading edge of the pulse.

In Figures 6a and 6b, the scattered surface wave due to the incident radiation field is considered. For the case considered $\varepsilon_r = 10(1 - i\frac{\omega_c}{\omega})$ and $\theta_o^i \to 90^o$ grazing incidence. For $\theta_o^i \to 90^o$, the coupling between the incident radiation field and the surface wave is largest in the example considered. In Fig. 6a the instantaneous response and the envelope of the scattered surface wave are plotted. Here $T_{x1} > 0$, this results in the small shift in the envelope of the response with respect to the envelope of the unperturbed LORAN C pulse. Fig. 6b also indicates that the Re $h_H(t)$ is negligible beyond the leading edge of the pulse.

In Figures 7a and 7b, the scattered lateral wave due to the incident radiation field is considered. For the case considered here, $\varepsilon_r = 1 - 0.5/(1 - 0.1\ i\frac{\omega}{\omega_c})$ and $\theta_o^i = 70^o$. At the angle $\theta_o^i = 70^o$, the coupling between the lateral wave and the incident radiation field is maximum. In Fig. 7a, the leading edge of the pulse is shown to be very distorted and $T_{x1} > 0$. For the lateral wave, which propagates parallel to the boundary in medium 1 ($y < 0$), the term Re $h_H(t)$,(Fig. 7b) is not insignificant beyond the leading edge of the pulse. This term is due to the singularities of the transfer function $H_z(s)$ (4.5).

In Fig. 8, the scattered surface wave response due to an incident lateral wave is plotted as a function of time. For this case $\varepsilon_r = 1 - 0.5/(1 - 0.1\ i\frac{\omega}{\omega_c})$, since the wave parameter β (2.6) is not the same for the surface and the lateral waves, here too, $T_{x1} > 0$ and there is a shift in the enveloped of the response with respect to the unperturbed LORAN C envelope.

6. CONCLUDING REMARKS

Analytical and numerical solutions for low frequency radio waves propagation over rough, finite conducting surfaces are presented. These solutions are based upon a full wave approach to the problem. Thus consideration is not only given to non-specular scattering of the radiation fields, but also to coupling between the radiation fields, the lateral and the surface waves. For the case of scattering in the specular direction, the time of arrival of the response from the stationary points of the irregular surface is determined by $\tau(x_m)$(4.8b). Furthermore since the time of arrival of the LORAN C pulse is usually considered to be the third zero crossing, the phase of the reflection coefficient and therefore the ground permittivity determines the effective time of arrival of the pulse. For non-specular scattering, the leading edge of the pulse would be considerably distorted and $T_{1\Lambda} \neq 0$ (4.8c), adding to the inaccuracy in determining the effective arrival time of the LORAN C pulse. When the conductivity of medium 1 ($y < h(x)$) is low, the contributions from the scattered lateral waves and surface waves may be significant, particularly if the surface wave and lateral wave contributions arrive at approximately the same time. These contributions are significant only when the transmitter or receiver are near the irregular boundary.

7. REFERENCES

BAHAR, E., (1972a),"Generalized Fourier Transforms for Stratified Media," Canadian Journal of Physics, Vol. 50, No. 24, pp. 3123-3131.

BAHAR, E., (1972b),"Radio Wave Propagation in Stratified Media with Non-uniform Boundaries and Varying Electromagnetic Parameters, Full Wave Analysis," Canadian Journal of Physics, Vol. 50, No. 24, pp. 3132-3142.

BAHAR, E., (1972c), "Radio Wave Propagation Over a Rough Variable Impedance Boundary Part I-Full Wave Analysis," IEEE Transactions on Antennas and Propagation, Vol. AP-20, No. 3, pp. 36 -368.

BAHAR, E., (1974), "Depolarization in Nonuniform Multilayered Structures," Journal of Mathematical Physics, Vol. 15, No. 2, pp. 202-208.

BAHAR, E., (1975), "Propagation in Irregular Multilayered Cylindrical Structures of Finite Conductivity-Full Wave Solutions," Canadian Journal of Physics, Vol. 53, No. 11, pp. 1088-1096.

BAHAR, E., (1976), "Electromagnetic Waves in Irregular Multilayered Spheroidal Structures of Finite Conductivity-Full Wave Solutions," Radio Science, Vol. 11, No. 2, pp. 137-147.

BECKMANN, P. and A. SPIZZICHINO, (1963), "The Scattering of Electromagnetic Waves from Rough Surfaces," McMillan Co., New York.

JOHLER, J. R. and S. HOROWITZ, (1973), Propagation of Loran Ground and Ionospheric Wave Pulses, Office of Telecommunications Report 73-20, Superintendent of Documents, U. S. Government Printing Office, Washington, D. C. 20402.

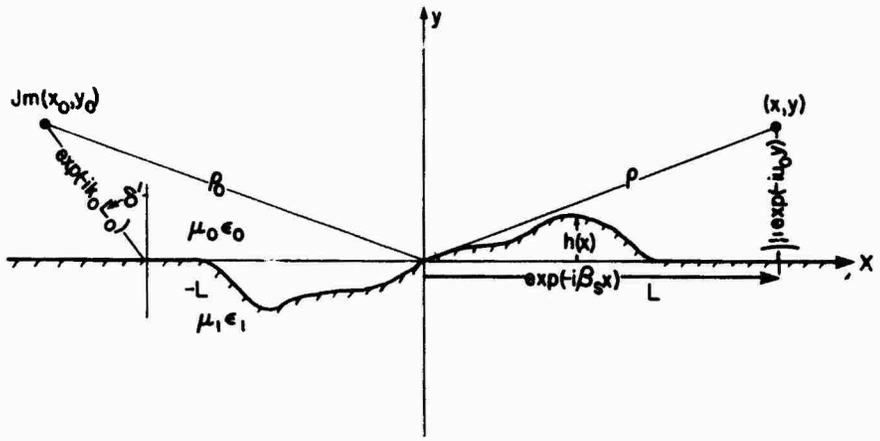


Fig. 1. The scattered radiation fields due to incident plane waves over rough surfaces.

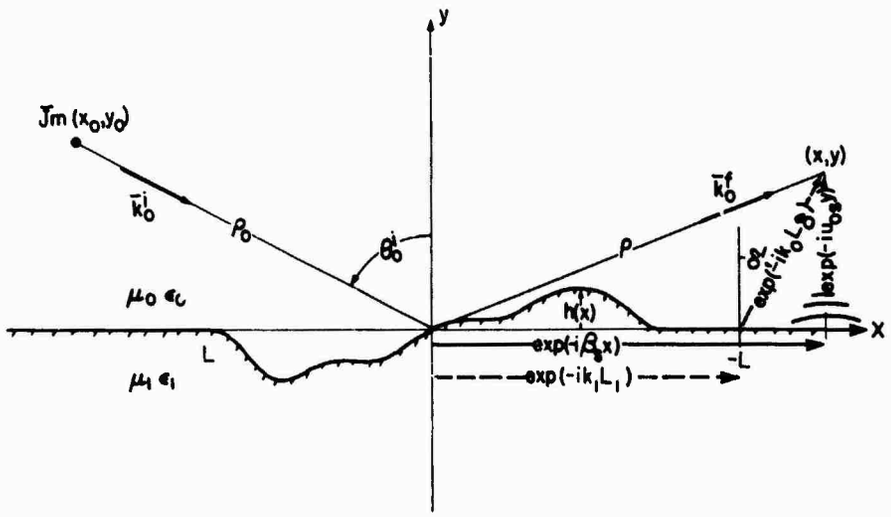


Fig. 2. The scattered surface waves and lateral waves due to incident plane waves over rough surfaces.

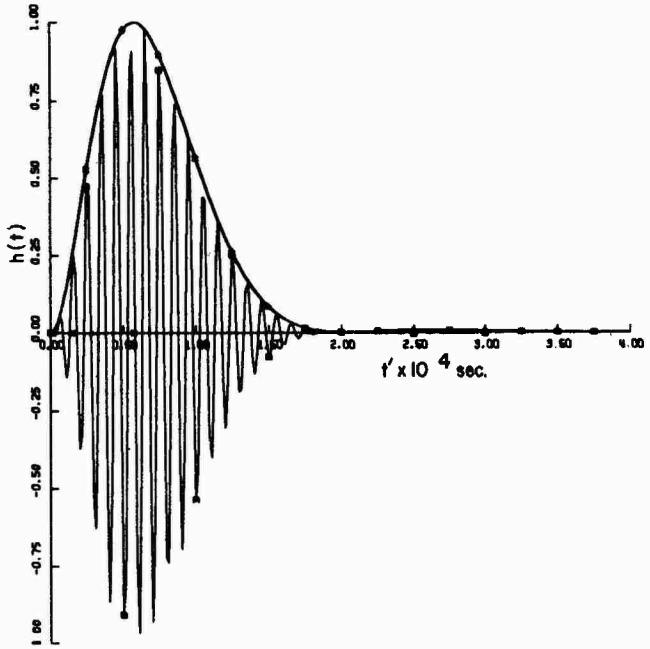


Fig. 3. The scattered radiation field for $\varepsilon_r = 100 - i\,10^6 \frac{\omega_c}{\omega}$ and $\theta_o^i = \theta_o^f = 80^o$. The instantaneous response and the envelopes.

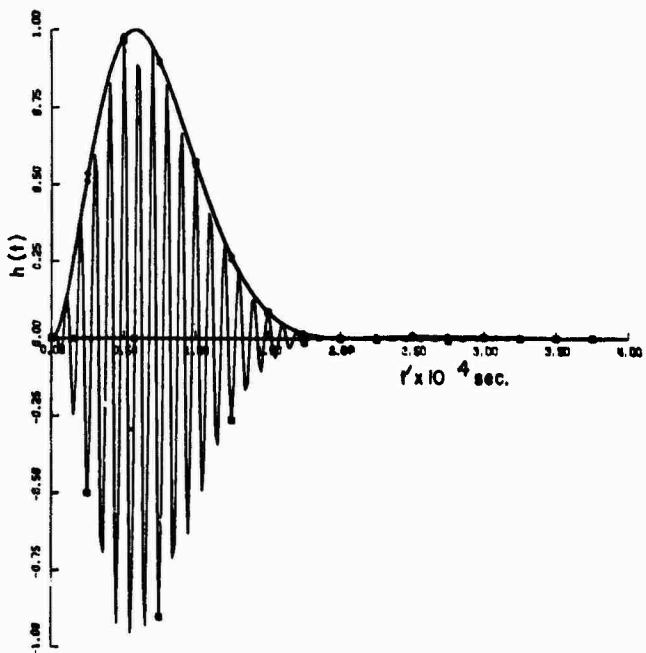


Fig. 4. The scattered radiation field for $\varepsilon_r = 3(1 - i\frac{\omega_c}{\omega})$ and $\theta_o^i = \theta_o^f = 80^o$.

(a) The instantaneous response and the envelopes.

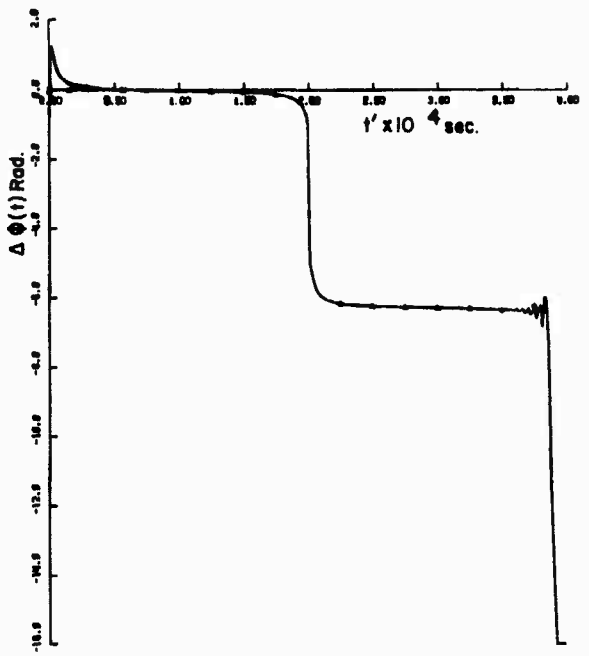


Fig. 4. The scattered radiation field for $\varepsilon_r = 3(1 - i\frac{\omega_c}{\omega})$ and $\theta_o^i = \theta_o^f = 80^o$.
(b) The phase anomaly $\Delta(\phi)$.

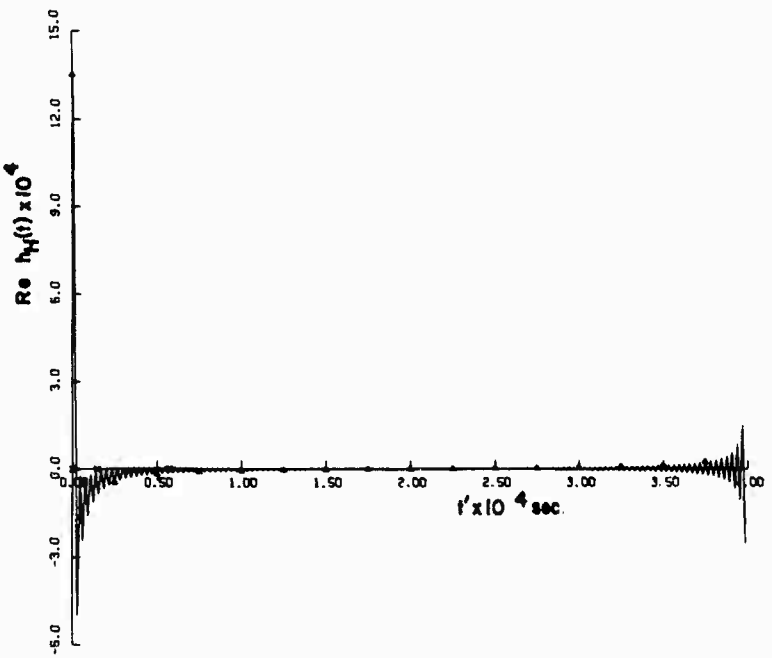


Fig. 4. The scattered radiation field for $\varepsilon_r = 3(1 - i\frac{\omega_c}{\omega})$ and $\theta_o^i = \theta_o^f = 80^o$.
(c) Re $h_H(t)$.

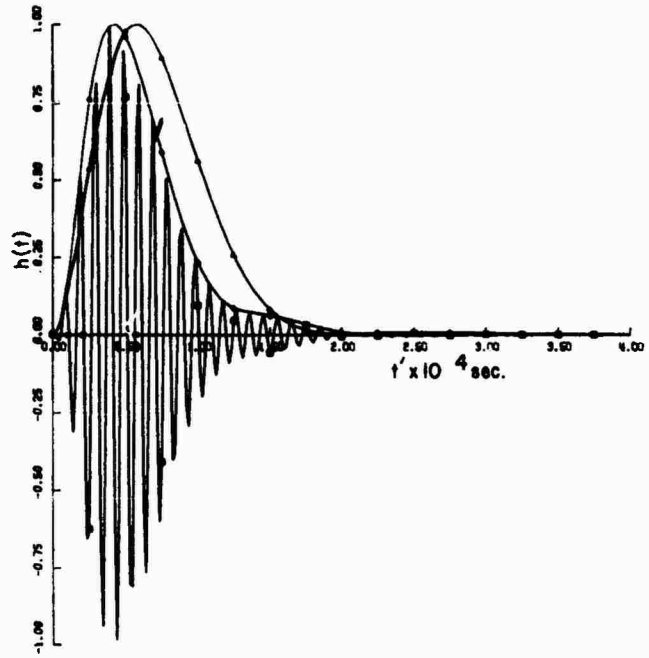


Fig. 5. The scattered radiation field for $\varepsilon_r = 10(1 - i\frac{\omega}{\omega_c})$, $\theta_o^i = 80^o$ and $\theta_o^f = 50.8^o$

(a) The instantaneous response and the envelopes.

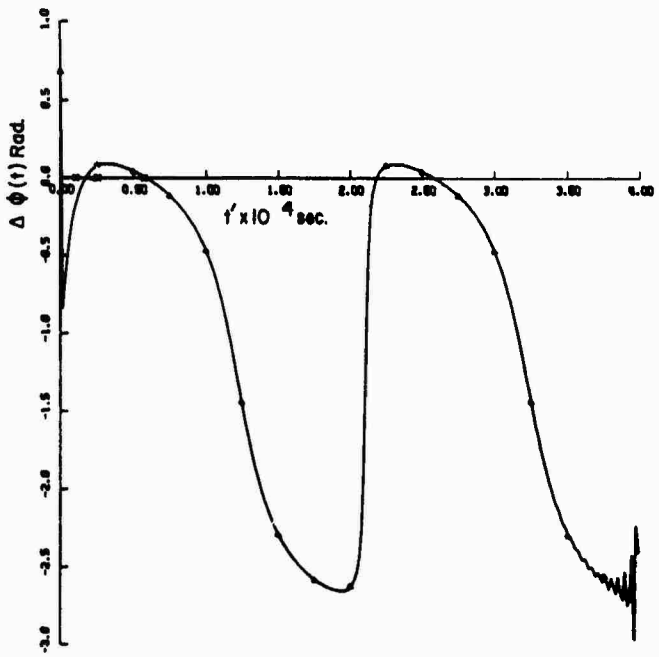


Fig. 5. The scattered radiation field for $\varepsilon_r = 10(1 - i\frac{\omega}{\omega_c})$, $\theta_o^i = 80^o$ and $\theta_o^f = 50.8^o$

(b) The phase anomaly $\Delta(\phi)$.

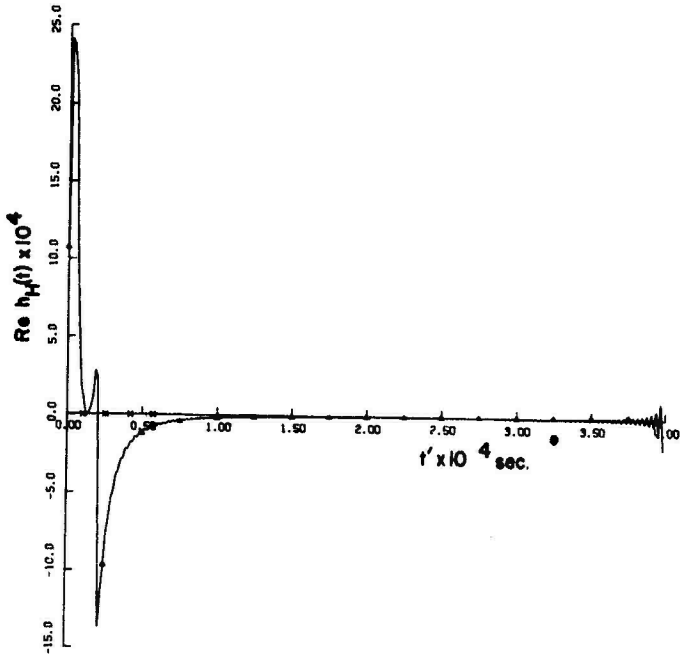


Fig. 5. The scattered radiation field for $\epsilon_r = 10(1 - i\frac{\omega_c}{\omega})$, $\theta_o^i = 80^o$ and $\theta_o^f = 50.8^o$
(c) Re $h_H(t)$.

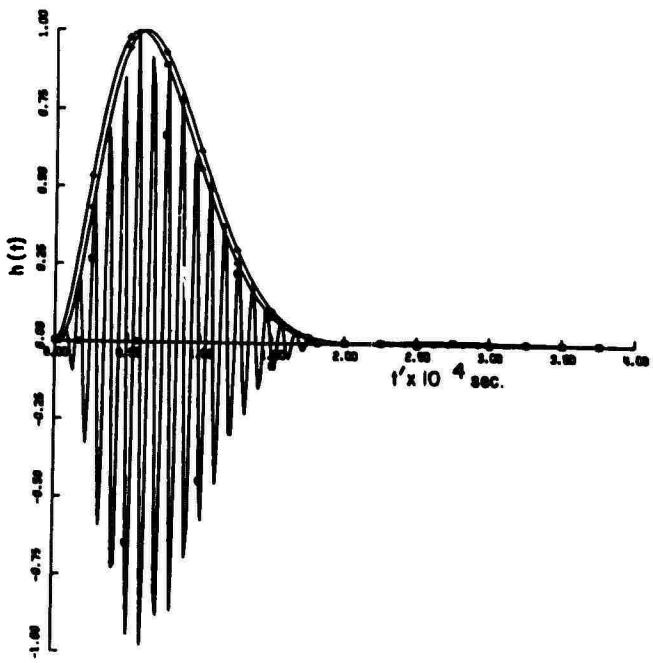


Fig. 6. The scattered surface wave for $\epsilon_r = 10(1 - i\frac{\omega_c}{\omega})$ and $\theta_o^i \to 90^o$
(a) The instantaneous response and the envelopes.

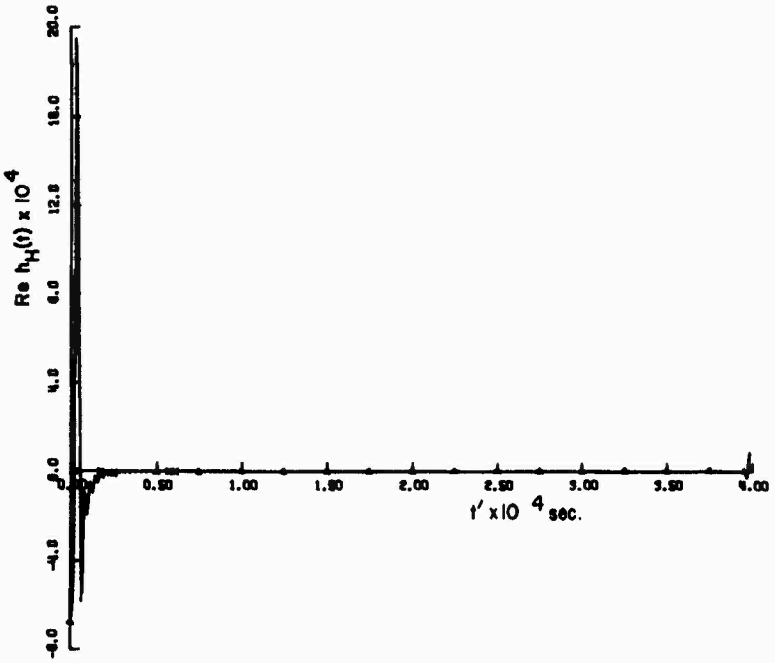


Fig. 6. The scattered surface wave for $\varepsilon_r = 10(1 - i\frac{\omega_c}{\omega})$ and $\theta_o^i = 90^\circ$

(b) Re $h_H(t)$

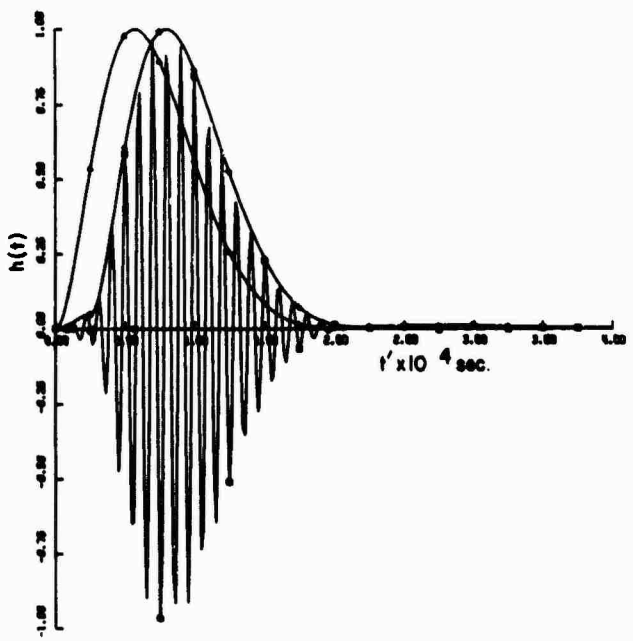


Fig. 7. The scattered lateral wave for $\varepsilon_r = 1 - 0.5/(1 - 0.1\, i\frac{\omega}{\omega_c})$ and $\theta_o^i = 70^\circ$

(a) Instantaneous response and the enveloped.

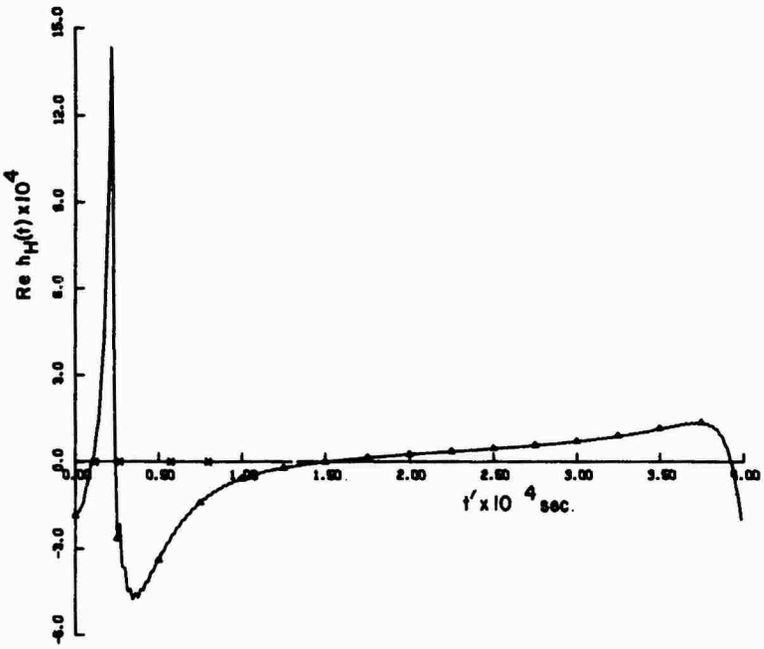


Fig. 7. The scattered lateral wave for $\varepsilon_r = 1 - 0.5/(1 - 0.1\ i\ \frac{\omega}{\omega_c})$ and $\theta_o^i = 70^\circ$

(b) Re $h_H(t)$

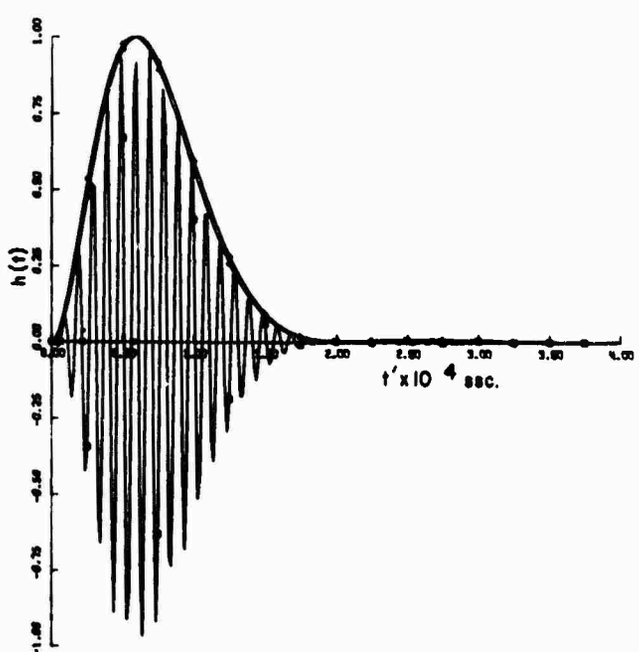


Fig. 8. The scattered surface wave due to an incident lateral wave for $\varepsilon_r = 1 - 0.5/(1 - 0.1\ i\ \frac{\omega}{\omega_c})$.

The instantaneous response and the envelopes.

DISCUSSION

Dr Langenberg

I did not really understand how you came from your steady-state solution to the direct transient response. Was it by means of a Fourier-transform or did you separate the contribution of the singularities of the transfer function as for example, the SEM (Singularity Expansion Method) does?

Author's reply

As I pointed out in my talk, the transient response (the product of the excitation transform and the steady state transform function) was split into two parts. The first term contains the singularities (poles) of the excitation transform only. The second term contains all the singularities (poles and branch points) of the transfer function.

The transient response due to the first term contains $\exp(-i\omega_c t)$ as a factor (ω_c is the carrier frequency). It can be calculated readily by analytical methods. The second term is in general far more difficult to calculate analytically (particularly due to the branch points). These terms are readily calculated using Fast Fourier Transform Techniques.

DISCUSSION OF REAL AND APPARENT LORAN-C PROPAGATION LIMITATIONS

by

R. H. Doherty
Colorado Research and Prediction Laboratory Inc.
$CRPL_i$
P.O. Box 1056
Boulder, Colorado 80306 U.S.A.

ABSTRACT

An attempt to determine Loran-C propagation limitations from measured data is related to the prediction model against which the data are analyzed. Inadequate models have resulted in analysis problems and reported misleading results. After measuring and analyzing Loran-C data throughout the world over a period of 20 years, it has become possible to recognize these analysis problems. Until the modeling limitations are completely understood and resolved, the subject of the propagation limitations of Loran-C cannot be adequately summarized.

Some examples of real versus apparent propagation type errors are presented to show the problems associated with certain analysis techniques. Also, an example of an unfounded assumption is used to show how errors in basic assumptions can affect the evaluation of the entire system's capability.

I. INTRODUCTION

A simplified and often used solution to the modeling problem is to empirically fit a purely mathematical prediction to the measured data and analyze the results statistically. Unfortunately, reasonable physical restraints are often not included in the model, and the results lack uniqueness. Thus, the nonunique statistical solution can yield results considerably different from a solution obtained when physical restraints are imposed. This can be illustrated by using perfect data (i.e., data that yields zero errors) as input to an empirical model fitting technique. Subsequent statistical analysis with up to 0.2 μs of random noise superimposed on the input data produces apparent propagation errors in excess of ± 10 μs. This clearly illustrates the need for physical restraints.

Nonrigorous or speculative prediction models resulting from misinterpretation of measured propagation effects may also result in a nonuniqueness problem. Over highly irregular and nonhomogeneous surfaces, the secondary phase correction is more variable than over smooth homogeneous terrain. An increase in variability also occurs in the pulse amplitude and envelope (i.e., envelope to cycle discrepancy or ECD); therefore, it has been assumed that if you were to measure one parameter you could predict the second at all locations. In the case of the secondary phase correction related to ECD, the complete theoretical analysis shows that a unique relation exists at any single point, but the relation changes with location so that no unique and singular relation exists. Rigorous mathematical theory shows that over smooth homogeneous surfaces the primary factor affecting the secondary phase correction is the surface impedance, whereas the primary factor affecting the ECD is the earth's curvature and not the surface impedance.

II. PHILOSOPHY OF PREDICTION AND RANDOMNESS OF MEASURED DATA

Loran-C prediction philosophy is based on very simple straightforward concepts, even though the rigorous mathematical solutions become quite complex. To a first approximation, a radio signal travels at the speed of light in a vacuum. The actual earth's surface with its irregularities and nonhomogeneities change the velocity of propagation by a small but important correction factor.

In the case of the Loran-C radio navigation system, one is always interested in the arrival time of the signal as denoted by the phase of the signal. Since the surface conditions always slow down the signal, the phase can be defined as the primary phase (based on the speed of light in normal atmosphere) and the secondary phase correction.

It takes all of the mathematical rigor of Maxwell's equations to predict this relatively small secondary phase correction. Figure 1 compares the magnitude of the primary phase with the secondary phase correction. Even though this calculation was made over some of the most rugged terrain in the United States, it is still impossible to see the secondary phase correction after it has been added to the primary phase delay.

It is quite obvious from this example that the very large and extremely predictable primary phase should be removed before the secondary phase correction propagation effects are studied. An attempt to model measured data without removing the primary phase may lead to a "numerical swamping" effect. Examples of this will be shown later in the applications portion of this paper.

The secondary phase correction portion of the propagation theory is the limiting factor in loran prediction capability. If the surface terrain and surface impedances are known for all points between the transmitter and the receiver, the secondary phase correction can be accurately predicted. However, for any given point within a service area or along a given path, the secondary phase correction may appear to be quite random (see fig. 1).

The apparent randomness of the secondary phase correction is completely analogous to the random number generator in a computer. Empirical fitting of any number of random samples from a computer random number generator will not allow one to predict additional future samples. However, if one knows the seed and formula for generating the random numbers, any or all of the random numbers can be determined exactly. This is exactly the same situation with the extremely complex secondary phase corrections of the Loran system. Hence, the use of a least squares fitting to a number of Loran-C measurements as an attempt to predict additional measurements is fraught with disaster. All that can be achieved is detection of the systematic portion of the propagation. But this could have been predicted from the simple and well proven homogeneous smooth earth theory, without the need for any measurements!

Now, as with the seed and formula in the random number generator, if one accounts for the path parameters from transmitter to receiver and utilizes a mathematically rigorous solution, the secondary phase variations can be predicted. So that, an apparently random phenomena, as seen from the measurement point of view, is not random at all as seen from the prediction point of view.

Attempts to explain the random phenomena in the measurements by least square fitting techniques may be further aggravated by numerical swamping. This phenomenon becomes a problem when improper mathematical procedures are used. In the Loran-C case, this can occur when the primary wave is not considered separately from the secondary phase correction; that is, when the great numerical value in its entirety, as shown in figure 1, is adjusted by a least square fitting of measured data. As can be seen from figure 1, the random secondary phase correction can not even be visually detected in the total phase that includes primary phase plus secondary phase correction. Examples of this numerical swamping problem will be presented in the next section.

III. EXAMPLES OF REAL VERSUS APPARENT PROPAGATION ERRORS

The phase of the low frequency radio signal can be expressed as

$$\phi = (\omega/c)\ \eta_a d + \phi_c \quad , \tag{1}$$

where $\omega = 2\pi f$, c is the velocity of light in a vacuum, η_a is the surface refractive index, d is the distance from the transmitter, and the ϕ_c is the secondary phase correction (in radians) caused by finite impedance and refractive index changes on either side of the irregular surface boundary along which the signal propagates. The first term in this equation is the primary wave, and all of the factors are simple well-known parameters. As can be seen from figure 1, this term in the equation is by far the largest.

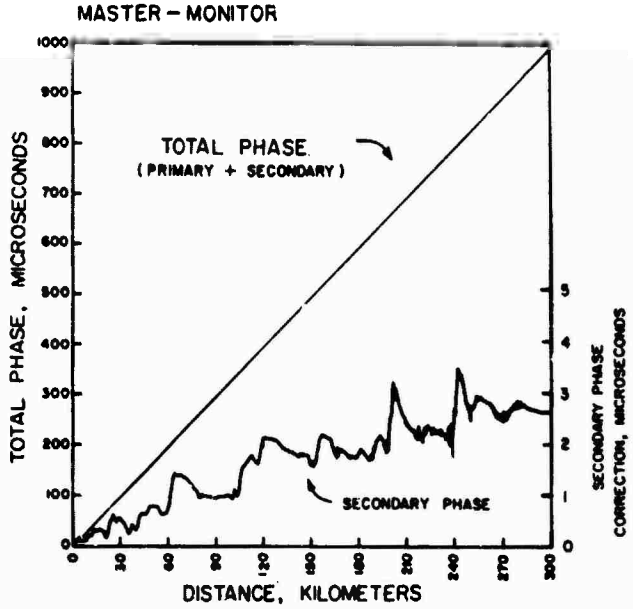


Figure 1. Primary phase and secondary phase correction computed from scaled terrain and impedance data for the Master to Monitor path in Nevada and California, including Death Valley.

Although the second term of the equation is more than an order of magnitude less than the first or primary term, it is by far the most complex (Johler, 1971). When the actual irregularity of the surface is taken into account and the surface impedance Δ_2 at all points along the path is considered, the following formulas may be obtained.

$$W[\Delta_2(0)] = \left\{1 - \frac{ik_1}{2\pi}\int_{S_0} W[\Delta_2(Q)]\ \exp(-ik_1 r) \times \left[\Delta_2(Q) + \left((1 + \frac{1}{ik_1 r_2}\right)\frac{\partial r_2}{\partial n}\right]\frac{r_o}{r_1 r_2}\ ds\right\} A, \quad (2)$$

$$W = \Pi/2\Pi_{pri}, \quad (3)$$

$$\Pi_{pri} = \exp(-ik_1 r_o)/r_o \quad , \quad (4)$$

where 0 and Q are spatial positions to be calculated and $r = r_1 + r_2 - r_0$, and A = {½ if Q is above S_0 and 1 if Q is on S_0}. Also, the r_0, r_1, r_2, and $\partial r_2/\partial n$ are all functions of α. The distance between the transmitter and scatter point under consideration is r_0. The distance from the transmitter to another scatter point is r_1, and r_2 is the distance between the two scatter points. In other words, in this propagation simulator expressed as an integral equation, the terrain is introduced by r_0, r_1, r_2, and $\partial r_2/\partial n$; whereas ground surface impedance is introduced as $\Delta_2(Q)$ and $\Delta_2(0)$. The secondary phase correction ϕ_c is then represented by the argument of W.

$$W = |W|\ \exp\{-i[\arg W]\} \quad (5)$$

or

$$W = |W|\ \exp\{-i(\phi_c)\} \quad . \quad (6)$$

Finally, in the case of the Loran-C system, the time difference between the arrival time of a signal from a secondary station with a fixed emission delay (ED) and the master station is measured. That is,

$$TD_A = (\omega/C)\ \eta_a d_A + \phi_{secondary} - (\omega/C)\ \eta_a d_M - \phi_{master} + ED_A \quad , \quad (7)$$

where the phase corrections over the path from the master transmitter subtract from the phase corrections over the path from the secondary transmitter. This is a very complex theory, and many attempts have been made to avoid it or to ignore it in favor of more simplifing data analysis techniques.

In such an attempt to statistically fit corrections to data measurements, Pearce and Walker (1974) used the following formula:

$$TD[P] = (ED + \alpha) + \beta(D_S - D_M) + \varepsilon(\theta_S, \theta_M) \quad , \quad (8)$$

where $\varepsilon(\theta_S, \theta_M)$ is to account for seawater paths at bearing angle θ from the secondary and master transmitters to the field points. Using this analysis and 61 data points in a 100 km square, they determined for the Master and Secondary Z path an $\alpha = -0.19$, $\beta = 3.339$ and an RMS deviation for the set = 0.33. For a 30-data point sub set of these data in a 60 km square, they found $\alpha = -4.11$, $\beta = 3.349$, and the RMS deviation for the set = 0.25. In this particular case, the Secondary-Z transmitter was approximately 1000 km; whereas, the Master station was approximately 600 km. Therefore, the 0.01 change in β between these two solutions times the $D_S - D_M$ of 400 km yields approximately 4 μs or the change in α.

Although a solution of this nature is a valid empirical technique, and could be used to predict points within the area where data is already available, it has no physical significance. The large α numbers of several microseconds are frequently interpreted as Loran associated errors, but in reality are only a result of the mathematical analysis. This type of solution and the next example are indicative of numerical swamping.

In a separate attempt to statistically analyze Loran-C data, measurements made in Alabama and Florida, Wieder and Washburn (1970) used a least squares fitting technique. Nine locations were measured in Alabama and five locations were measured in Florida in this experiment. They first assigned

$$\Delta TDA = \varepsilon_1 D_{SA} - \varepsilon_2 D_M + \varepsilon_4$$
$$\Delta TDB = \varepsilon_3 D_{SB} - \varepsilon_2 D_M + \varepsilon_5 \quad . \quad (9)$$

They then minimized the quantity

$$S = \sum_{i=1}^{N} W_{Ai}(\Delta TD_{Ai} - \varepsilon_1 D_{SAi} + \varepsilon_2 D_{Mi} - \varepsilon_4)^2 + W_{Bi}(\Delta TD_{Bi} - \varepsilon_3 D_{SBi} + \varepsilon_2 D_{Mi} - \varepsilon_5)^2 \quad . \quad (10)$$

Here again the small ε's multiplied by the large D's produced large microsecond corrections that were offset by ε_4 and ε_5. The results of all of their analysis showed ε_4 and ε_5 values ranging from $\pm$7.5 to $\pm$15.8 μs.

In an attempt to understand the reason for these large ε's, a separate experiment was performed. TD values were chosen for each location so that all of the ε's were a flat zero as analyzed by the computer. Subsequently, the random number generator in the computer was used to add or subtract errors from each of the TD numbers used in the analysis. When the random errors introduced into the TD values had peak values of ± 0.2 μs, the ε_4 and ε_5 values exceeded ± 10 μs. In other words, by introducing purely random noise between 0 and ± 0.2 μs into the psuedo measured data, the ε's produced by the best mathematical least squares fit exceeded fifty times the maximum random error introduced. This is an excellent example of numerical swamping in data analysis. As was pointed out by Doherty (1972), many of these problems could be avoided by introducing the physical restraints of the propagation theory.

A further example of a real versus an apparent propagation error occurs in the area of ECD (Envelope to Cycle Discrepancy). The rigorous theoretical method for predicting a point on the envelope of a Loran-C pulse involves evaluating equations (1)-(6) as a function of frequency. A complete analysis of these equations over irregular inhomogeneous ground (Johler and Horowitz, 1974) showed that both the phase and the envelope variability increased when the terrain became more rugged. However, the analysis showed there was no simple unique relation between the phase and the envelope. Again the apparent randomness could be predicted by knowing the input parameters, but could not be predicted from measurements alone.

This recent analysis, as well as an earlier one (Johler, 1963), showed that over a smooth homogeneous terrain the ECD is practically independent of impedance and is primarily dependent on geometry. This is true because the ECD is primarily related to the earth's curvature through the lapse rate of the refractive index. In the case of the secondary phase correction, the opposite is true--the impedance effect is significant and the earth's curvature effect becomes secondary.

Irregardless of these and other studies, observations that both the phase and envelope became more variable over irregular terrain, and the desire for a simple solution prompted Fehlner and McCarty (1974) to propose the following relationship:

$$D/C = TP + A/B(ECD) \quad . \tag{11}$$

This relation implied that the secondary phase correction could be made by merely multiplying A/B (constants to be determined) times the measured ECD and algebraically adding this to the speed of light velocity. They stated "Thus, it is apparent that, so long as the dispersiveness of the medium is constant, the ratio of A to B is predicted to be constant." Their major justification for this scheme was based on the effect of HF signals passing through the ionosphere. In this respect, there appears to be no theoretical justification for comparing the dispersion of low frequency ground waves with high frequency waves passing through the ionosphere.

The Fehlner and McCarty proposed simplified relation assumed that the ECD was a primary function of impedance over smooth homogeneous terrain and not primarily a function of distance. It also assumed that over irregular inhomogeneous terrain there was a unique relation between the phase and the envelope.

The rigorous solution of equations (1) - (6), as pointed out above, showed that their simplifying assumptions were not true. In spite of this, experimental money was spent in an attempt to prove these naive simplifying assumptions to be true.

IV. CONCLUSIONS

Theory based upon the solution of Maxwell's equations is singularly unique in this low frequency propagation field. Complete solutions of these theoretical equations have been rigorously derived mathematically and implemented on digital computers. Therefore, there is no longer reason for using a prediction model with less sophistication or rigor.

This author maintains that the best use of measurements is to validate and improve the theoretical predictions. A fairly common viewpoint is that measurements alone are the complete and ultimate test and evaluation of the system. This viewpoint may actually prove harmful to the system in that it hinders improvements and exaggerates measurement errors. This viewpoint also limits the progress that can be made toward separating the various sources of measurement errors.

V. REFERENCES

Doherty, R. H. (1972), A Loran-C grid calibration and prediction method, OT/TRER 25 (Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402).

Fehlner, L. F., and T. A. McCarty (1974), How to harvest the full potential of Loran-C, NAVIGATION 21, pp. 223-233.

Johler, J.R. (1963), The propagation time of a radio pulse, IEEE Trans. Ant. Prop.,AP-11, No. 6, pp. 661-668.

Johler, J. R. (1971), Loran radio navigation over irregular and inhomogeneous ground with effective ground impedance maps, OT/TRER 22 (Supt. of Documents, U.S. Government Printing Office, Washington, DC 20402).

Johler, J. R., and S. Horowitz (1974), Propagation of a Loran pulse over irregular, inhomogeneous ground (AGARD-NATO Paper No. 28, Conference of Electromagnetic Wave Propagation Involving Irregular Surface and Inhomogeneous Media, XX Technical Meeting of the Electromagnetic Wave Propagation Panel AGARD-NATO, Netherlands, March 1974).

Pearce, D. C., and J. W. Walker (1974), Ground effects of Loran-C signals, Proc. of the Sixth Annual Precise Time and Time Interval (PTTI) Planning Meeting, December 3-5, 1974, U.S. Naval Research Laboratory, p. 319.

Wieder, B., and J. S. Washburn (1970), Evaluation of the Loran tests at Anniston, Alabama, and Panama City, Florida. ESSA Technical Report ERL 166-ITS 107 (Supt. of Documents, U.S. Government Printing Office, Washington, DC 20402).

DISCUSSION

H.Soicher, US Army Electronics Command, Fort Monmouth, NJ 07703, USA

(1) With LORAN-C what accuracy may one expect in mountainous regions?

(2) It seems that in order to improve accuracy one must have detailed mapping of terrain and conductivities in the very area which includes the transmitter and the receiver. Is this feasible? Is it practical?

Author's Reply

(1) We feel that with the proper values for the input data, we can predict to 50 nano-seconds over mountainous or smooth terrain.

(2) Yes, the detailed terrain and conductivity information is required and we feel it is feasible and practical. Reference Burch, L.B., Donerty, R.H., and Johler, J.R., 1976, LORAN Calibration by Prediction, Navigation, Journal of the Institute of Navigation, Vol.23, No.3. (Fall Issue)

T.B.Jones, Physics Department, University of Leicester, UK

The LORAN Calibration by prediction methods described by Mr Doherty was most interesting. When will this method become operational and be generally available?

Author's Reply

I cannot say when this system will be fully implemented. The work that we were doing was abruptly terminated June 14, 1976, even though there were no technical difficulties and the entire program was well within a few months of successful completion. This occurred even though the program was actually a culmination of many years of our previous research work and the entire idea of calibration by prediction originated with us.

We have been isolated from any continuing work on this technique.

PREDICTION OF GROUND WAVE PROPAGATION TIME ANOMALIES IN THE LORAN-C SIGNAL TRANSMISSIONS OVER LAND

J. Ralph Johler

Colorado Research and Prediction Laboratory
CRPL$_i$
P.O. Box 1056
Boulder, Colorado 80306
U.S.A.

ABSTRACT

The ultimate accuracy to which Loran-C can be predicted is dependent firstly upon the repeatability of the loran chain grid including the measuring equipment and secondly upon the predictability. Prediction of the signal propagation time to full accuracy over land or over seawater involving land masses in the propagation paths to each of the transmitters requires special considerations. This is true both on the ground and aloft in aircraft. We have found that such propagation time may be anomalous. This means that spatial perturbations of the signal propagation time will, without more advanced and sophisticated theory, limit the prediction accuracy of the loran coordinates (time differences). This limitation arrises from the physical nature of the ground and its effect on wave propagation--land is nonhomogeneous and irregular. Thus, classical ground wave theory used in the past, although helpful in removing gross propagation errors in loran, fails to take into account with sufficient accuracy propagation errors observable on modern loran receivers. We therefore wish to emphasize that a more sophisticated theory is required to fully exploit loran receiver and chain precision.

Fortunately, ground wave propagation theory has reached the state of development such that nonhomogeneous and irregular ground in the propagation paths can be taken into account. Previously, when great spatial perturbations of the order of several microseconds were observed, we have called the signal propagation anomalous. It is argued in this paper that such anomalous propagation is a unique function of the geographic location of the propagation path and requires the introduction of terrain, soil and basement rock electrical properties and features along the propagation path. Such detail is readily introduced into the modern solution of the ground wave propagation problem using an integral equation and modern advanced computer data handling capabilities. We believe that loran coordinates can be predicted from geographic coordinates with an accuracy of 50 ns using such a technique.

1. INTRODUCTION

The recent applications of Loran-C positioning techniques over the land areas of the earth and the concomittant development of loran receivers with a resolution of the order of nanoseconds has led to the identification of ground wave pulse propagation time anomalies. We have demonstrated such anomalies theoretically and we have observed them in measurements (JOHLER and HOROWITZ, 1975; DOHERTY, 1975). These anomalies affect the loran time difference coordinates because they are spatially dependent and appear as a spatial perturbation of the secondary phase correction, t_c, or the correction in microseconds of the pulse cycle propagation time relative to the propagation time at the speed of light, c.

A loran coordinate or time difference (y) is given in microseconds by:

$$y = 10^6\, n_a(d_s - d_m)/c + t_c(d_s) - t_c(d_m) + C_s \quad , \tag{1}$$

where d_s and d_m are geodesic distances to a secondary and a master transmitter respectively, and C_s is a constant independent of spatial positions (latitude, longitude and altitude) of the loran receiver. The quantity, n_a is the index of refraction of air at the ground level, $n_a = 1.000338$. The quantity t_c or secondary phase correction (JOHLER, 1956) is the spatially dependent correction of the pulse cycle time. Thus, t_c, is quite complicated but nevertheless must be considered in a forthright manner using propagation theory if nanosecond prediction accuracy is to be achieved for loran. Actually it is the phase of the rigorous propagation function, W (JOHLER, 1971) for the ground wave:

$$W(O) = \{1 + (-ik/4\pi) \int_S W(Q)\, \exp(-ikr)[x(Q) + (ikr_2)^{-1}(\partial r_2/\partial n)] \cdot r_0(r_1 r_2)^{-1}\, dS\}\, A \quad . \tag{2}$$

Here, $A = \frac{1}{2}$ if Q is above S and $A = 1$ if Q is on S. The quantity $r = r_1 + r_2 - r_0$, where r_0 is the distance from the transmitter to the observer, O; r_1 is the distance from the transmitter to the scatter point, Q, on the integration surface, S, and r_2 is the distance from the scatter point Q to O. k is the wave number in air, $k = 2\pi n_a f/c$; $\partial r_2/\partial n$ is the normal derivative to the surface at Q. Thus, if the surface is smooth and spherical, $\partial r_2/\partial n = 0$. $x(Q)$ is the spatially dependent complex impedance at a particular point, Q, normalized to the impedance of space, $Z_0 = 377$ ohms. Thus, the secondary phase correction, t_c, in microseconds, can be found:

$$t_c = 10^6 \arg(W F_i)/\omega \quad , \tag{3}$$

where the angular frequency is ω radians per second. The function, F_i, depends upon the nature of the receiving antenna and represents the induction field of the transmitting antenna. Two types of antennas are used with loran receivers: the "whip" antenna, that is sensitive to the vertical electric, E_r, field, and the "loop" antenna, that is sensitive to the horizontal magnetic, H_ϕ, field. Thus, we shall differentiate the factors F_{ir} and F_{im} respectively. At great distance from the transmitting antenna, $F_{ir} = 1$ and $F_{im} = (Z_o)^{-1}$. Thus,

$$F_{ie} = \{[D^{-1}(-2\cos\theta) + D^{-3}(3aa'\sin^2\theta)] + i[D^{-2}(kaa'\sin^2\theta + 2k^{-1}\cos\theta) + D^{-4}(-3aa'k^{-1}\sin^2\theta)]\}(ik)^{-1} \tag{4}$$

and

$$F_{im} = (\mu_o \omega D)^{-1}(ka'\sin\theta)[1 + (ikD)^{-1}] \quad , \tag{5}$$

where

$$a' = a + h$$

$$D^2 = (a')^2 + a^2 - 2aa'\cos\theta$$

$$\theta = d/a \quad ,$$

and where a is the radius of a spherical earth, and h is the altitude of the observer above the spherical earth. As the distance, d, to the transmitter becomes small, and assuming $h = 0$, then $\cos\theta \sim 1$, $\sin\theta \sim d/a$, and $D \sim d$, then we find

$$F_{ie} = 1 + (ikd)^{-1} + (ikd)^{-2} \tag{6}$$

and

$$F_{im} = Z_o^{-1}[1 + (ikd)^{-1}] \quad . \tag{7}$$

Here $Z_o = \mu_o c = 377$ ohms, the impedance of space. Of course one can calculate the complete complex vertical electric, E_r, and horizontal magnetic, H_ϕ, fields guaranteed by Maxwell's equations:

$$E_r = F_{ie}\, WC \tag{8}$$

and

$$H_\phi = F_{im}\, WC \quad , \tag{9}$$

where $C = (4\pi)^{-1} I_o \ell\, \mu_o c$ and $Z_o = E_r/H\phi$. The quantity $I_o\ell$ has dimensions of ampere-meters and represents the antenna current moment. The antenna is assumed to be a dipole; i.e., it has a uniform current distribution over its length. Non-uniform current distributions are a comparatively simple extension of the theory (JOHLER, 1971).

In this paper we use specialized solutions of equation (2) to demonstrate, as a computer simulated propagation mechanism, various types of anomalous propagation of importance to Loran-C signal propagation time.

2. THEORETICAL CONSIDERATIONS

The general form of the integral equation (2) can be made tractable and suitable for computer programming by reduction of the integral to a one dimensional integration and transforming the integral equation into a set of algebraic equations (JOHLER and BERRY, 1967):

$$W(d_j) = 1 - B_j \sum_{\ell=1}^{j-1} \{p_k\, W(d_\ell)[F_1(d_j,d_\ell)\, x(d_\ell) + F_2(d_j,d_\ell)]\} - B_j p_j\, W(d_j)[F_1(d_j,d_j)\, x(d_j) + F_2(d_j,d_j)] \quad , \tag{10}$$

where

$$B_{j,\ell} = (r_o/d_{j,\ell})(k/4\pi)^{\frac{1}{2}} \exp(i\pi/4) \quad , \tag{11}$$

$$F_1 = \{\omega_j/[d_\ell(d_j - d_\ell)]\}^{\frac{1}{2}} \exp(-ikr) \quad , \tag{12}$$

$$F_2 = [1 + (ikr_2)^{-1}]\, F_1 \quad . \tag{13}$$

The quantity p_ℓ is the Gaussian weight used in the quadrature. The distances d_j and d_ℓ are geodesic distances along a radial from the transmitter. It is clear that the solution (10) is the set of points that represents the field along d_j. The terrain and ground impedance data enter equation (10) through $\partial r_2/\partial n$ and x respectively. For convenience in the numerical procedure, the solution can be obtained by projecting the integration surface into a plane containing the distance .

The calculation of $x(d_{j,\ell})$ should be obtained from comparatively rigorous solutions of the boundary conditions $^{j,\ell}$ at the surface of the ground. Thus, some account should be taken of soil conductivity and thickness, bedrock conductivity and water table level, if such data are available for the particular land area under consideration. Solutions of this type are well documented in the literature and the following formula was obtained from JOHLER and HARPER (1962) as specialized to a three layer model boundary condition:

$$x(d_{j,\ell}) = Z/Z_o = [(1 - T)/(1 + T)] \cos \phi \quad , \tag{14}$$

where Z is defined in a cartesian u,v,w coordinate system as,

$$Z = E_v/H_u \quad , \tag{15}$$

and at the surface of which the impedance is calculated in the u-v plane. The angle of incidence of the wave, ϕ, may be complex or real, but is not critical. In fact, Z is almost independent of ϕ. The quantity T is calculated from the boundary conditions,

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & & & \\ a_{21} & a_{22} & a_{23} & & & \\ & a_{32} & a_{33} & a_{34} & a_{35} & \\ & a_{42} & a_{43} & a_{44} & a_{45} & \\ & & & a_{54} & a_{55} & a_{56} \\ & & & a_{64} & a_{65} & a_{66} \end{bmatrix} \cdot \begin{bmatrix} T \\ U_i^{(1)} \\ U_r^{(1)} \\ U_i^{(2)} \\ U_r^{(2)} \\ U_i^{(3)} \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \\ \\ \\ \\ \end{bmatrix} \quad , \tag{16}$$

where $a_{10} = a_{11} - \cos\phi$; $a_{20} = a_{12} = a_{13} = -a_{21} = a_{34} = a_{35} = -a_{56} = -1$; $a_{23} = -a_{22} = \sin^2\phi/\delta_1 + \delta_1^2$; $a_{32} = \exp(-ik\, w_1\delta_1)$; $a_{33} = \exp(ik\, w_1\delta_1)$; $a_{42} = a_{22}\, a_{32}$; $a_{43} = a_{23}\, a_{33}$; $a_{44} = \sin^2\phi/\delta_2 + \delta_2^2$; $a_{45} = -a_{44}$; $a_{54} = \exp(-ik\, w_2\delta_2)$; $a_{55} = \exp(ik\, w_2\delta_2)$; $a_{64} = a_{44}\, a_{54}$; $a_{65} = a_{45}\, a_{55}$; $a_{66} = \sin^2\phi/\delta_3 + \delta_3^2$; $\delta_q = \varepsilon_q - i\sigma_q(\varepsilon_o\omega)^{-1} \sin^2\phi$, where q = 1, 2, 3, refers to the three layer model of the ground starting with the top layer. The dielectric constant of each layer is ε_q, and σ_q is the conductivity in mhos/m of the q-th layer. The thickness of each layer in meters is given by w_q. The solution of equation (16) is straightforward algebra and it obviously can be extended to any number of layers if the geological and soil data for such stratification becomes available. The reduction of the multilayer matrix problem has been discussed by Johler (1967) and it is quite tractable on computers. Thus, for purposes of this paper, we identify the top layer of soil with q = 1. The second layer is then q = 2. Finally, the bedrock is identified as q = 3.

We shall introduce, for purposes of demonstrating anomalous propagation, a distortion of the smooth spherical ground in the form of a mathematical hill:

$$f(d_{j,\ell}) = H \exp[-9(d_{j,\ell} - d_m)^2/b^2] \quad . \tag{17}$$

This is a ridge of maximum height, H, and a width, b, on each side of the hill crest emplaced at a distance, d_m, from the transmitter.

3. PROPAGATION ANOMALIES

We shall now consider a smooth, homogeneous propagation path and compare it with a similar propagation path in which the ground has become nonhomogeneous and an irregularity in the form of a Gaussian hill given by equation (17). The maximum height, H, is 1 km, and the width, b, is 8 km on each side of the crest that has been emplaced at a distance, d_m, equal to 100 km from the transmitter. As some additional physical realism for the

anomaly, we shall introduce a coast line at 125 km such that the ground conductivity changes from land, σ = 0.005 mhos/m, to seawater, σ = 5 mhos/m. The resultant secondary phase correction for smooth, homogeneous land and irregular, nonhomogeneous ground is depicted in figure 1. The receiver is assumed to be located on the surface of the ground. The propagation time relative to the propagation time at the speed of light is shown in microseconds as a function of distance from the transmitter. The curve that increases monotonically with distance represents the smooth, homogeneous case. It is clear that the ground anomaly that we have simulated on the computer with the theory of electromagnetic wave propagation not only produces a disturbance of the phase in the immediate vicinity of the irregularity in the ground, but also at distances greater than 75 km at sea! Since the total phase at 100 km is approximately 210 radians or 334 μs, the perturbation of the secondary phase correction depicted in figure 1 is 0.3 percent of the total phase. However, such anomalous phase perturbations are significant if nanosecond predictability and calibration accuracy as proposed by BURCH, DOHERTY, and JOHLER (1976) is to be accomplished. Thus, to correct the loran signal propagation time to the speed of light with nanosecond precision, it is necessary to introduce soil and geological structure together with terrain along each propagation path involved in the generation of loran coordinates. This means that we must introduce soil conductivity, thickness, and the terrain elevations into the electromagnetic wave propagation simulator to get a unique propagation time for each propagation path.

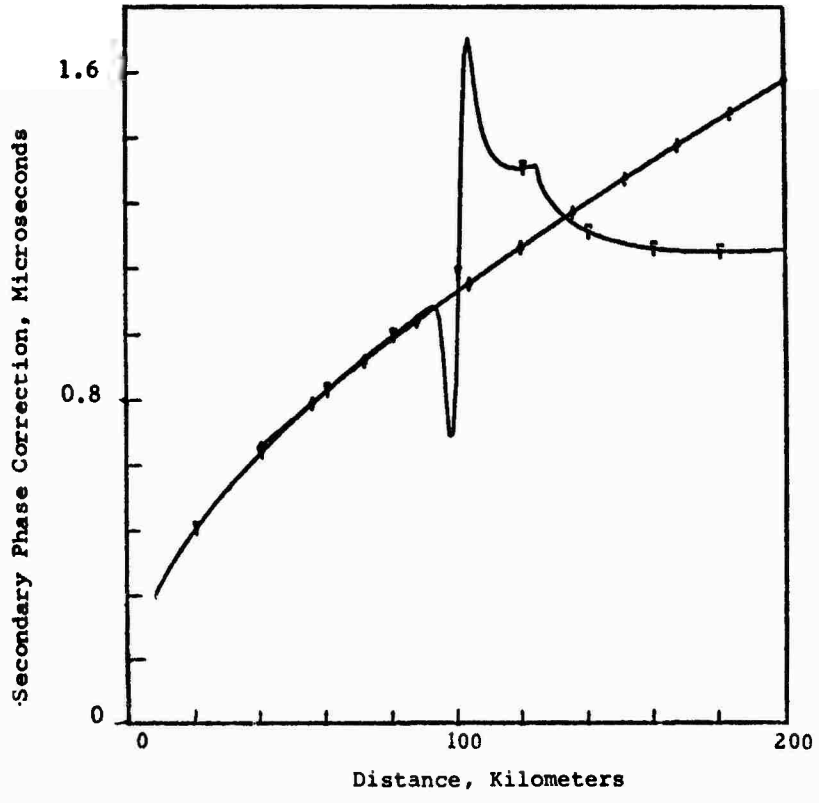


Figure 1. Secondary phase correction as a function of distance from the transmitter depicting the effect of a hill at 100 km and a seawater expanse at distances greater than 125 km. Also given is the secondary phase correction over smooth homogeneous land.

If one employs aircraft outfitted with loran receivers, the three dimensional nature of the problem is depicted in figure 2. Here we have assumed the ground to be homogeneous with an impedance of 0.03336 exp(i0.7756) corresponding to a conductivity of 0.005 mhos/m, assuming homogeneous or nonstratified ground. The altitude of the aircraft, h, is assumed to be 0.3 and 10 km. A hill is introduced at 85 km distance from the transmitter with a maximum height, H, of 0, 0.5, and 1 km. The width, b, is assumed to be 8 km each side of the crest emplaced at d_m. We have assumed vertical electric polarization sensitivity of the receiving antenna and hence the induction field given by equation (4) is evident at short distances from the transmitter. More important, the secondary phase correction manifests the influence of the induction field to even greater distance as the aircraft altitude is increased. Thus, at 10 km altitude, the induction field is important to distances as great as 100 km, if nanosecond precision is a requirement for loran navigation. This also means that we must apply a different secondary phase correction to a receiver with a loop antenna as compared with a receiver with a vertical whip antenna at distances as great as 100 km and altitudes of 10 km.

The effect of the 0.5 km hill and the 1 km hill also propagates to the aircraft altitudes of 3 and 10 km in significant amounts. In fact the perturbation becomes decorrelated with the precise anomaly (hill) on the ground as the altitude is increased. Thus, to obtain nanosecond prediction accuracy, it is necessary to take account of the ground anomalies irregardless of the aircraft altitude.

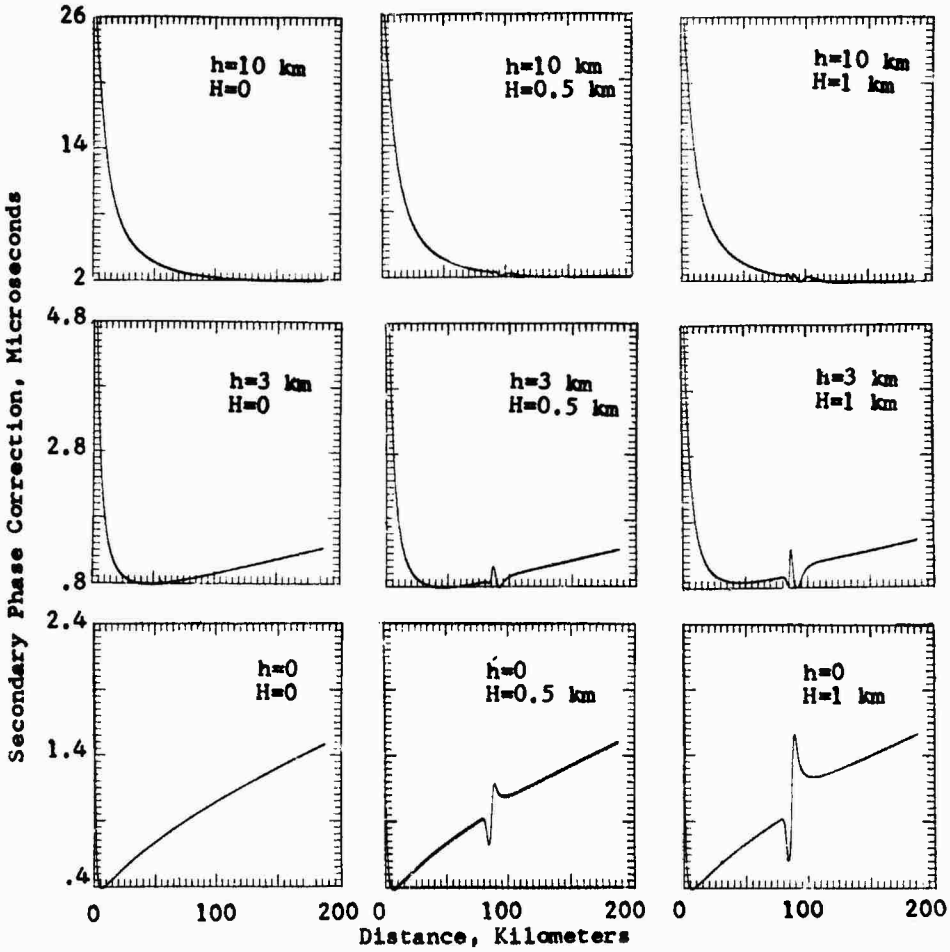


Figure 2. Illustrating the effect on the secondary phase correction of a hill in the propagation path of maximum height, H, equal to 0, 0.5, and 1 km at 85 km from the transmitter and a width of 8 km each side of the crest. The ground is homogeneous with a conductivity of 0.005 mhos/m. The effect of the induction field at shorter distance is also shown.

In figure 3, the same ground irregularity as figure 2 has been employed, but the additional complication of a coast line has been introduced. Thus, at distances greater than 100 km, we have introduced seawater. Furthermore, we have introduced some geological structure into the region of the hill between 75 and 100 km. The impedance, x, given by equation (14) has been introduced as an inhomogeneity in the form of a three layer model. The soil thickness is assumed to be 20 m and is composed of two layers each 10 m in thickness with a soil conductivity of 0.001 mhos/m for the top layer and 0.005 mhos/m for the second layer. The geological basement rock is assumed to be poorly conducting, with a value of 0.00001 mhos/m. This results in an impedance at the surface of the ground of $0.04590 \exp(i0.54242)$. The resultant spatial perturbation propagates to altitudes of 10 km or greater and to distances at sea greater than several hundred kilometers. Thus, if we wish to calibrate and predict our navigation system with nanosecond precision, we must introduce the technology of propagation over nonhomogeneous irregular ground described in this paper and in the papers of JOHLER and HOROWITZ (1975), DOHERTY (1975), and BURCH, DOHERTY, and JOHLER (1976), where we have given great detail for implementation of this concept on a wide geographical basis.

In figure 4 an anomaly inducing patch of ground is imbedded in an average smooth spherical ground of $\sigma = 0.005$ mhos/m conductivity. The ground anomaly could be a poorly conducting granite rock protruding a maximum of 500 m above the countryside and extending over the distance shown. The classical change in phase with altitude near 150 km is depicted as a small phase advance followed by a phase retardation at altitudes above 2500 m. The phase-altitude signature is traced at 2 km intervals through the perturbation caused by the anomaly. Notice that the phase advance relative to the surface value increases until the altitude-phase signature shows a great advance followed by a retardation. Finally, long after the cause of the anomalous propagation is past, the phase change with altitude (at 200 km) comprises a large phase advance followed by a small phase retardation with altitude--the opposite of the signature exhibited before the wave passed the anomaly. Thus, there is a certain permanent effect on the propagated wave at great distances from the cause of the perturbation.

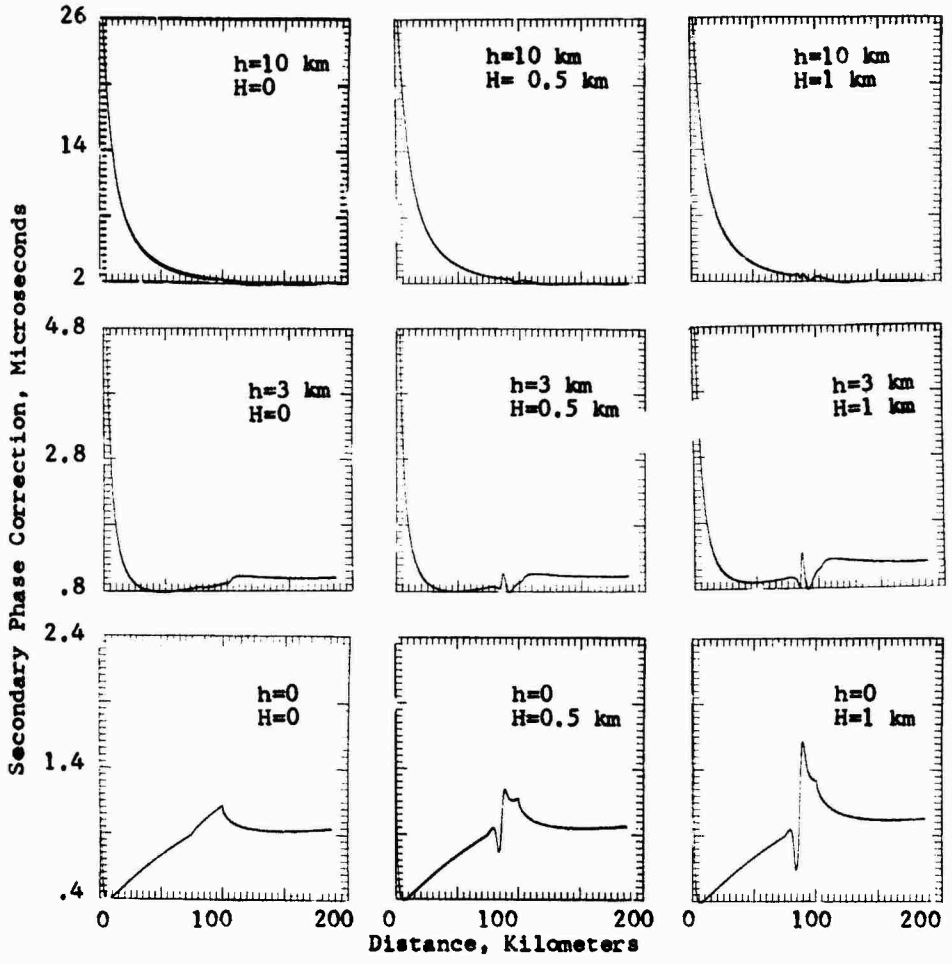


Figure 3. Illustrating a hill effect similar to figure 2, with a conductivity of sea water or 5 mhos/m at distances greater than 100 km. The ground impedance of the hill, $x = 0.04590\ exp(i.54242)$, corresponds to a three layer model with two top layers 10 m each in thickness and three conductivities, $\sigma_1 = 0.001$, $\sigma_2 = 0.005$, $\sigma_3 = 0.00001$.

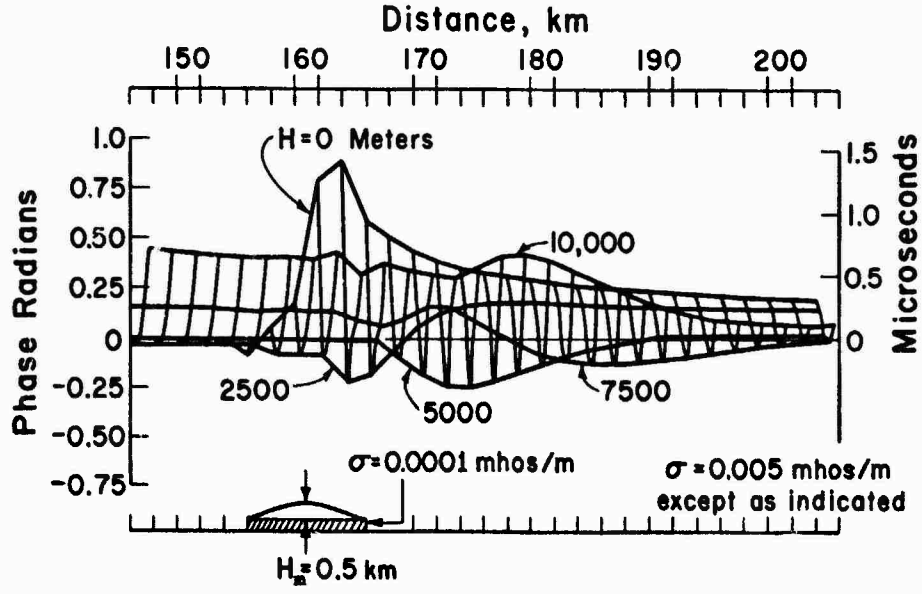


Figure 4. Phase correction perturbation caused by an anomaly inducing patch of ground imbedded in an average smooth spherical ground of 0.005 mhos/m conductivity. The height-gain or altitude effect curve is traced through perturbations in phase caused by an anomaly centered about d = 160 km.

4. CONCLUSIONS

It is concluded from the rigorous theory of propagation and from measurement that anomalous propagation not only exists, but is important to loran operation on ground wave if nanosecond precision is to be obtained in the use of Loran-C for navigation. It is further concluded that such propagation anomalies are significant not only on the ground in the immediate vicinity of the anomaly, but also aloft and at great distance from the anomaly. It is also concluded that the type of antenna used by aircraft navigating on loran must be considered as to its effect on the secondary phase correction at shorter distance from the transmitter. It is also noted that the type of antenna is significant to even greater distance as the altitude of the aircraft is increased.

It has also been demonstrated that the perturbations from a localized anomaly may propagate to great distances from the anomaly producing a permanent change in the ground wave propagation phase with distance. This manifests itself in the secondary phase correction employed in loran to correct the signal propagation time to the speed of light and it is quite predictable.

5. REFERENCES

BURCH, L. B., DOHERTY, R. H., and JOHLER, J. R., 1976, Loran calibration by prediction, Navigation, Journal of the Institute of Navigation, 23, No. 3 (Fall Issue).

DOHERTY, R. H., 1975, Spatial and temporal electrical properties derived from LF pulse ground wave propagation measurements, AGARD-CP-144 (The Hague, Netherlands conference, March 1974).

JOHLER, J. R., 1956, Phase of the low radio frequency ground wave, Natl. Bur. of Stds. Circular 573 (Suptd. of Doc., U.S. Gov. Print. Off., Washington, D.C.).

JOHLER, J. R., 1967, Theory of propagation of LF terrestrial radio waves--mathematical techniques for the interpretation of D-region propagation studies, Proceedings of the Conference on Ground-Based Radio Wave Propagation Studies of the Lower Ionosphere (Defense Research Board, Department of National Defense, Canada, Vol. 2, 399-422).

JOHLER, J. R., 1971, Loran radio navigation over irregular, inhomogeneous ground with effective ground impedance maps, Telecommunications Research and Engineering Report OT/TRER 22 (Suptd. of Doc., U.S. Gov. Print. Off., Washington, D.C., 20402).

JOHLER, J. R., and HARPER, J. D., 1962, Reflection and transmission of radio waves at a continuously stratified plasma with arbitrary magnetic induction, Jour. of Res. of NBS, 66D (Radio Prop.), No. 1, 81-89.

JOHLER, J. R., and BERRY, L. A., 1967, Loran-D phase corrections over inhomogeneous, irregular terrain, ESSA Technical Report IER 59-ITSA 56 (Suptd. of Doc., U.S. Gov. Print. Off., Washington, D.C., 20402).

JOHLER, J. R., and HOROWITZ, S., 1975, Propagation of a loran pulse over irregular, inhomogeneous ground, AGARD-CP-144 (The Hague, Netherlands Conference, March 1974).

DISCUSSION

J.Aarons, Air Force Geophysics Lab, Bedford, Mass., USA

What validation of the computer studies exists?

Author's reply

I refer you to my reference: Doherty, R.H., 1975, Spatial and Temporal Electrical Properties Derived from LF Pulse Ground Wave Propagation Measurements, AGARD-CP-144 (The Hague, Netherlands, March 1974) which has reported validation of the technique. This work comprises our initial and very limited validation of the general concept given in the Paper: Burch, Z.B., Doherty, R.H., and Johler, J.R., 1976, LORAN Calibration by Prediction, Navigation, Journal of the Institute of Navigation 23, No.3 (Fall Issue).

The comprehensive validation of our concept was terminated June 14, 1976 for non-technical reasons.

H.J.Albrecht, FGAN

Commenting on the predominance of tropospheric refractive index or of other effects upon LF propagation, such as electrical ground parameters, attention is drawn to the fact that the changes of the refractive index gradient and its extreme conditions can only be analyzed with difficulty for the accuracies required for standard transmissions. In this respect, information given with so-called weather-map accuracy may not be adequate unless it can be supplemented using data gathered by balloon-borne refractometers or, at least, radiosonde measurements of refractive index elements, throughout the relevant altitude region and along the propagation path of interest. Such more precise data should be taken for each test period.

LORAN C/D COORDINATE PREDICTION DEPENDENCE ON GROUND ELECTRICAL PROPERTIES

S. Horowitz
Ionospheric Radio Physics Branch
Electromagnetic Sciences Division
L G HANSCOM AFB MA 01731, USA

and

J.R. Johler
CRPL, Inc.
Boulder, CO, USA

SUMMARY

Radio navigation using LF ground wave propagation has now been a successful endeavour throughout the world in the form of the Loran C/D system. The use of ground wave to the exclusion of the skywave to distances of thousands of kilometers permits orders of magnitude greater navigation accuracy than is possible with ordinary radio systems. As both military and civilian requirements for even greater navigation precision are set forth, it becomes appropriate to inquire about the ultimate accuracy and precision imposed by nature on such a navigation system. With the current improved Loran C/D receiver instrumentation, the error due to the pulse propagation through the medium is emerging as one of great importance. The need to understand and reduce these propagation errors to the minimum is becoming increasingly important. Thus, the prediction accuracy of the system depends on the knowledge of the propagation characteristics of the medium.

The time of arrival of a loran pulse depends on the electrical properties of the earth's surface over which these signals propagate. These electrical properties include the impedance or conductivity of the ground, the roughness or terrain variations of the surface, the refractive index of the atmosphere at the surface, and the lapse rate or rate of change of refractive index with altitude above the surface. Spatial variations of the transmitted loran signal are primarily influenced by the nonhomogeneous surface impedance and by variations in the terrain. Temporal effects may be produced by time changes on these spatial features but are more easily influenced by the surface refractive index and the lapse rate of the refractive index of the earth's atmosphere, which are known to change diurnally and with changing weather conditions.

Loran groundwave time of arrival calculations have been made in the presence of irregular and nonhomogeneous ground. The irregularities are represented by elevation data as a function of range from the ground transmitters to the receiver. The conductivity along this path transformed into a surface impedance accounts for the nonhomogeneities as seen by the propagating loran pulse. This paper describes the results of such calculations. The ultimate limitation in the reduction of the propagation error as a function of quality of data for different types of land paths will be illustrated.

1. INTRODUCTION

A computer program package for predictions of Loran C/D grid time difference coordinates over land or in the vicinity of land masses is under development. The basic technique has been described by Johler (1971) and Johler and Horowitz (1974) for predicting the hyperbolic Loran time difference reading. This time difference measured at a fixed receiver is expressed as:

$$TD = \frac{n}{c} d_s + t_c(d_s) + ED - \frac{n}{c} d_m + t_c(d_m)$$

where n = atmospheric index of refraction = 1.000338.

c = velocity of light = 2.997925×10^8 meters/second.

d_s = length of geodetic path from slave transmitter to receiver.

d_m = length of geodetic path from master transmitter to receiver.

t_c = time correction or secondary phase factor for propagation over a given path length.

ED = emission delay in microseconds.

The distance from the transmitter to the receiver can be easily calculated with an accuracy depending on the precision of the given latitude and longitude of these stations. From this, the free space or primary wave delay can be calculated with great accuracy. Emission delays to remove line ambiguities are set experimentally for a fixed time difference reading at the ground area monitor receiver. These can be calibrated to $\pm$ 10 nanoseconds.

The time correction for propagation over a path length, t_c, expressed in microseconds, accounts for the disturbing influence of the earth. Depending on the electrical properties of the earth, t_c can have values of up to 10 microseconds at 500 km and thus becomes the largest contributor to the grid prediction error budget over land. Therefore, the calculation of the time correction due to variation of signal propagation velocity from uniform speed that occurs in free space is of primary importance in determining the accuracy of Loran grid predictions.

2. GROUND ELECTRICAL CHARACTERISTICS

The electrical properties of the ground that effect the propagation of these 100 kHz pulses are terrain elevation, conductivity, dielectric constant and relative permeability. The terrain elevation data is derived from topographic maps which are digitized and stored in the form of a large matrix. A similar matrix is generated for the soil and bedrock characteristics. For practical purposes and availability of data, this matrix has been initially designed for a 30-second arc grid or a 0.5 nautical mile spacing of data points. The optimum grid increment is still to be selected. For Loran C/D frequencies, an average dielectric value of 15 over land and 80 over seawater is assumed. The relative permeability can normally be regraded as unity. The integral equation for calculation of the secondary phase or time correction factor requires an input of the surface impedance and terrain elevation. The effective value of ground impedance is determined not only by the nature and depth of the soil but also its moisture content, geological structure, depth of wave penetration, temperature and frequency. Therefore, a three layer ground model has been proposed with a top layer called the unsaturated soil, a layer of wet or saturated soil and an underlying strata of bedrock to determine the local surface impedance from Johler and Harper (1962).

Temporal fluctuations have also been found to be significant by Doherty (1974). Both temperature and seasonal effects appear to be present in most recorded Loran data, but the magnitudes of the variations are highly dependent on the location of the receiver. At present, there is no generally accepted model which accurately predicts Loran temporal variations and none of the existing grid prediction programs include temporal modelling. Loran monitor stations control the temporal drifts by varying the emission delays resulting in relatively stable time differences only in the area of the monitor. To implement weather effects into the grid prediction algorithm would require an extensive meteorological capability.

3. TIME CORRECTION CALCULATIONS

a. Soil

Tables 1 and 2 are derived from calculations of ground impedance for representative ground saturated and unsaturated soils with a bedrock conductivity of 0.005 and 0.00001 mhos per meter respectively. The effect on the complex impedance of varying the unsaturated soil in conductivity and depth are illustrated. The conductivity of the top layer is varied from 5 to 0.0001 mhos per meter and the soil depth is varied from zero to 20 meters. The conductivity of the second layer is fixed at 0.005 mhos/meter. The amplitude of the impedance undergoes a large change as the depth of the top soil increases and the conductivity decreases. This is due to the increase wave penetration depth as the conductivity is decreased. Over land there is a ten-to-one (from .02 to .2) variation in impedance amplitude in going from good to poor soil. Tables 3 and 4 illustrate the corresponding phase changes of the complex impedance.

Figure 1 shows the resulting change in the secondary phase factor for a variation of the ground impedance. A soil conductivity of 0.005 mhos/meters translates from the above tables into an impedance with an amplitude of 0.03335 and a phase of 0.77651 radians. The calculated time correction factor for a 180 kilometer path is indicated by curve A. If the conductivity along the path is altered, that is reduced by a factor of 10 or to 0.0005 mhos per meter with a corresponding amplitude and phase of 0.04864 and 1.02041 midway along the path, the secondary phase factor is shown by curve C. Increasing the conductivity by 10 to 0.05 mhos per meter results in a time correction of 250 nanoseconds. Calculations of secondary phase factor over a typical ground path are shown in Figure 2. Over a 200 km path changes in ground conductivity of 30 and 60 percent will require time corrections of 60 and 100 nanoseconds respectively. Unfortunately, conductivity is usually known to within a factor greater than three and therefore time of arrival (TOA) measurements will have large differences from the computed TOA. To reduce these errors, more accurate data on ground conductivity is required and techniques developed to refine predictions as measured time difference values at known geographic locations become available. Figure 3 illustrates the conversion of soil and rock data into surface impedance. Current effort is on the modification and updating of the look up table.

b. Terrain Elevation

The model chosen to determine the effect of an elevation perturbation on the secondary phase factor is ground with a soil conductivity of 0.005 mhos per meter as in the case above. A hill with a base of 20 km is placed at 90 km in the 200 km path and the height of the hill is increased from 250 meters to 500 meters. The results of the calculation are shown in Figure 4. In the vicinity of the hill, a perturbation from smooth earth is encountered even for hills as small as 250 meters. After passing the hill and continuing on the geodetic path, the secondary phase factor attempts to recover to the smooth value. At a distance of 100 km from the hill, the difference between the two curves is negligible. A 500 meter hill produces a much larger secondary phase factor perturbation in the area of the hill and at a distance of 100 km from the hill, the time correction is 25 nanoseconds above the previous smooth value. Figure 5 illustrates the variation of secondary phase factor for an inhomogeneous irregular path. At a distance of 130 km the contribution of a 1.5 km perturbation in terrain, will produce a 100 nanosecond delay.

c. Calculations

Utilizing the techniques described in Johler and Horowitz (1974), and available maps to determine elevation, soil type, soil depth and subsurface rock structure over a Loran-covered area, time differences have been computed, without the aid of measured time differences at known geographic coordinates. The mean discrepancy between the calculated and measured Loran time difference for ground geographic coordinates is between 250 and 300 nanoseconds. This figure appears large especially when one considers that in a time difference system like Loran, if there is a correlation between the ground conductivity in the two paths, there will be a cancellation in the error contributed by each TOA resulting in an improved time difference reading. The large discrepancy only emphasizes the need for more accurate data on ground conductivity if 50 to 100 nanosecond prediction accuracy is to be obtained.

4. CONCLUSIONS

A goal of $\pm$ 50 nanosecond prediction along each Loran propagation path has been recommended as feasible by the Institute for Telecommunications Services (1975). In order to achieve accuracy, available ground conductivity data must be improved. Elkins (1976) is developing a technique to update the ground conductivities as measured time difference data at known locations become available.

In the area of hills of 250 meters or greater in elevation, the secondary phase factor undergoes a significant fluctuation. At large distances from the hill the effect becomes negligible for terrain perturbations of 250 meters or less. However, for larger terrain perturbations, the effects are significant to great ranges and therefore all terrain variations must be included in the secondary phase factor calculation. Fortunately, much of this data is accurate and not too difficult to obtain.

It is also recommended that some form of temporal-compensation be included in the modelling algorithm for the Loran grid prediction computer program to increase its accuracy.

A tool now exists where the Loran C/D prediction accuracy dependence on the variation of ground electrical properties can be studied in great detail. It is planned to exploit this technique by a series of calculations so that the detail propagation limitations of this system can be determined.

REFERENCES

DOHERTY, R.H., 1974, "Spatial and Temporal Electrical Properties Derived from LF Ground Wave Propagation Measurements", AGARD Proceedings No. 144 SHAPE Technical Centra, P.O. Box 174, Hagua, Netherlands.

ELKINS, T.E., 1976, "Empirical Correction of Soil Conductivity Model", Privata Communication.

INSTITUTE FOR TELECOMMUNICATIONS SERVICES, 1975, "Proposal to Develop an Automatic Prediction/ Calibration Computer Package", Boulder, CO.

JOHLER, J.R., 1971, "Loran Radio Navigation Over Irregular, Inhomogeneous Ground With Effective Ground Impedance Maps", TRE Report OT-22, Bouldar, CO.

JOHLER, J.R. and S. HOROWITZ, 1974, "Propagation of a Loran Pulse over Irregular, Inhomogeneous Ground", AGARD-NATO Paper No. 28, Conferanca Proceedings No. 144, SHAPE Technical Centre, P.O. Box 174, The Hague, Netherlands.

JOHLER, J.R. and J.D. HARPER, 1962, "Reflection and Transmission of Radio Waves at a Continuously Stratified Plasma with Arbitrary Magnetic Induction", Jour. of Res. of NBS 66D (Radio Prop), No. 1, 81-89.

TABLES

1. Variation of ground impedance amplitude for a three layer modal with a variable depth and conductivity for the top unsaturated soil and a constant conductivity for tha saturated (.005 mhos/meter) layar and bedrock (.005 mhos/meter).

2. Variation of ground impedance amplitude for a three layar modal with a variable depth and conductivity for the top unsaturated soil and a constant conductivity for the saturated (.005 mhos/metar) layer and bedrock (.00001 mhos/meter).

3. Variation of ground impedanca phase for a three layer modal with a variable dapth and conductivity for the top unsaturated soil and a constant conductivity for tha saturatad (.005 mhos/meter) layer and bedrock (.005 mhos/meter).

4. Variation of ground impedance phase for a three layer wodel with a variable depth and conductivity for the top unsaturated soil and a constant conductivity for tha saturated (.005 mhos/metar) layar and bedrock (.00001 mhos/meter).

FIGURES

1. Effect of change in ground conductivity at 75 km from uniform 0.005 (A) to 0.05 (B) and 0.0005 (C) mhos/meter on secondary phasa factor.

2. Changa in sacondary phase factor from ambiant path (A) for a 30% (B) and 60% (C) increase in ground conductivity.

3. Data base of ground electrical propertias.

4. Effact of changa in tarrain elevation from smooth aarth (A) to 250 metar hill (B) and 500 metar hill (C) on secondary phase factor. Conductivity is assumed at 0.005 mhos/meter over path.

5. Variation of secondary phasa factor for an inhomogeneous irragular path.

TABLE 1

Variation of Ground Impedance Amplitude for a Three Layer Model with a Variable Depth and Conductivity for the Top Unsaturated Soil and a Constant Conductivity for the Saturated (.005 mhos/meter) Layer and Bedrock (.005 mhos/meter)

GROUND IMPEDANCE
FREQUENCY = 100. KILOHERTZ
ANGLE OF INCIDENCE = 80. DEGREES
TOTAL SOIL DEPTH = 20. METERS

ELECTRICAL CONSTANTS

	UNSATURATED SOIL	SATURATED SOIL	BEDROCK	
CONDUCTIVITY (MHOS/METER) =	VARIABLE	.00500	.00500	
DIELECTRIC CONSTANT =	15.	15.	15.	(SEAWATER -- 80. IS USED)
MAGNETIC PERMEABILITY =	1.	1.	1.	

- - - AMPLITUDE - OHMS RELATIVE TO FREE SPACE -(377 OHMS) - - -

UNSATURATED SOIL DEPTH METERS	SATURATED SOIL DEPTH METERS	UNSATURATED SOIL CONDUCTIVITY (MHOS/METER) .01000	.00500	.00200	.00100	.00050	.00020	.00010
0.000	20.000	.03335	.03335	.03335	.03335	.03335	.03335	.03335
.200	19.800	.03306	.03335	.03353	.03360	.03363	.03365	.03365
.400	19.600	.03276	.03335	.03372	.03384	.03390	.03394	.03395
.600	19.400	.03247	.03335	.03390	.03408	.03418	.03424	.03425
.800	19.200	.03218	.03335	.03408	.03433	.03446	.03454	.03455
1.000	19.000	.03190	.03335	.03426	.03458	.03474	.03484	.03485
1.200	18.800	.03162	.03335	.03445	.03483	.03502	.03514	.03516
1.400	18.600	.03135	.03335	.03463	.03508	.03530	.03544	.03547
1.600	18.400	.03108	.03335	.03482	.03533	.03559	.03575	.03578
1.800	18.200	.03081	.03335	.03501	.03558	.03588	.03606	.03609
2.000	18.000	.03056	.03335	.03519	.03583	.03617	.03637	.03641
4.000	16.000	.02826	.03335	.03707	.03843	.03914	.03957	.03965
6.000	14.000	.02649	.03335	.03896	.04110	.04223	.04292	.04303
8.000	12.000	.02518	.03335	.04080	.04380	.04541	.04637	.04653
10.000	10.000	.02427	.03335	.04258	.04651	.04864	.04991	.05012
12.000	8.000	.02367	.03335	.04427	.04920	.05190	.05352	.05380
14.000	6.000	.02330	.03335	.04584	.05184	.05519	.05719	.05754
16.000	4.000	.02311	.03335	.04729	.05443	.05847	.06091	.06135
18.000	2.000	.02304	.03335	.04860	.05693	.06175	.06466	.06520
20.000	.000	.02305	.03335	.04977	.05933	.06501	.06845	.06911

TABLE 2

Variation of Ground Impedance Amplitude for a Three Layer Model with a Variable Depth and Conductivity for the Top Unsaturated Soil and a Constant Conductivity for the Saturated (.005 mhos/meter) Layer and Bedrock (.00001 mhos/meter)

GROUND IMPEDANCE
FREQUENCY = 100. KILOHERTZ
ANGLE OF INCIDENCE = 80. DEGREES
TOTAL SOIL DEPTH = 20. METERS

ELECTRICAL CONSTANTS

	UNSATURATED SOIL	SATURATED SOIL	BEDROCK	
CONDUCTIVITY (MHOS/METER) =	VARIABLE	.00500	.00001	
DIELECTRIC CONSTANT =	15.	15.	15.	(SEAWATER -- 80. IS USED)
MAGNETIC PERMEABILITY =	1.	1.	1.	

- - - AMPLITUDE - OHMS RELATIVE TO FREE SPACE -(377 OHMS) - - -

UNSATURATED SOIL DEPTH METERS	SATURATED SOIL DEPTH METERS	UNSATURATED SOIL CONDUCTIVITY (MHOS/METER) .01000	.00500	.00200	.00100	.00050	.00020	.00010
0.000	20.000	.02996	.02996	.02996	.02996	.02996	.02996	.02996
.200	19.800	.02967	.02996	.03014	.03020	.03023	.03025	.03026
.400	19.600	.02938	.02996	.03032	.03044	.03051	.03054	.03056
.600	19.400	.02910	.02996	.03050	.03068	.03078	.03084	.03086
.800	19.200	.02882	.02996	.03068	.03093	.03106	.03115	.03117
1.000	19.000	.02855	.02996	.03086	.03118	.03135	.03145	.03148
1.200	18.800	.02829	.02996	.03105	.03143	.03163	.03177	.03180
1.400	18.600	.02803	.02996	.03124	.03169	.03193	.03208	.03213
1.600	18.400	.02777	.02996	.03143	.03195	.03222	.03241	.03245
1.800	18.200	.02753	.02996	.03162	.03221	.03252	.03273	.03279
2.000	18.000	.02728	.02996	.03181	.03244	.03283	.03306	.03312
4.000	16.000	.02515	.02996	.03379	.03529	.03610	.03664	.03677
6.000	14.000	.02354	.02996	.03588	.03842	.03985	.04080	.04107
8.000	12.000	.02236	.02996	.03808	.04192	.04418	.04572	.04619
10.000	10.000	.02156	.02996	.04038	.04590	.04935	.05176	.05253
12.000	8.000	.02105	.02996	.04281	.05056	.05575	.05953	.06083
14.000	6.000	.02077	.02996	.04546	.05625	.06417	.07030	.07255
16.000	4.000	.02067	.02996	.04842	.06362	.07618	.08682	.09103
18.000	2.000	.02070	.02996	.05191	.07393	.09541	.11637	.12565
20.000	.000	.02081	.02996	.05628	.08991	.13197	.18546	.21475

TABLE 3

Variation of Ground Impedance Phase for a Three Layer Model with a Variable Depth and Conductivity for the Top Unsaturated Soil and a Constant Conductivity for the Saturated (.005 mhos/meter) Layer and Bedrock (.005 mhos/meter)

GROUND IMPEDANCE
FREQUENCY = 100. KILOHERTZ
ANGLE OF INCIDENCE = 80. DEGREES
TOTAL SOIL DEPTH = 20. METERS

ELECTRICAL CONSTANTS

	UNSATURATED SOIL	SATURATED SOIL	BEDROCK	
CONDUCTIVITY (MHOS/METER) =	VARIABLE	.00500	.00500	
DIELECTRIC CONSTANT =	15.	15.	15.	(SEAWATER -- 80. IS USED)
MAGNETIC PERMEABILITY =	1.	1.	1.	

- - - PHASE, RADIANS - - -

UNSATURATED SOIL DEPTH METERS	SATURATED SOIL DEPTH METERS	UNSATURATED SOIL CONDUCTIVITY (MHOS/METER) .01000	.00500	.00200	.00100	.00050	.00020	.00010
0.000	20.000	.77651	.77651	.77651	.77651	.77651	.77651	.77651
.200	19.800	.76795	.77651	.78172	.78344	.78427	.78462	.78456
.400	19.600	.75984	.77651	.78679	.79023	.79186	.79256	.79247
.600	19.400	.75217	.77651	.79174	.79686	.79931	.80036	.80023
.800	19.200	.74494	.77651	.79655	.80335	.80661	.80802	.80784
1.000	19.000	.73813	.77651	.80125	.80970	.81377	.81553	.81532
1.200	18.800	.73172	.77651	.80582	.81590	.82078	.82289	.82266
1.400	18.600	.72570	.77651	.81026	.82197	.82765	.83012	.82987
1.600	18.400	.72006	.77651	.81459	.82791	.83438	.83722	.83695
1.800	18.200	.71478	.77651	.81881	.83371	.84097	.84418	.84390
2.000	18.000	.70986	.77651	.82290	.83938	.84744	.85100	.85072
4.000	16.000	.67737	.77651	.85791	.88934	.90525	.91265	.91255
6.000	14.000	.66804	.77651	.88331	.92845	.95217	.96375	.96427
8.000	12.000	.67343	.77651	.90066	.95847	.99005	1.00623	1.00774
10.000	10.000	.68719	.77651	.91134	.98084	1.02041	1.04161	1.04446
12.000	8.000	.70471	.77651	.91653	.99677	1.04445	1.07111	1.07561
14.000	6.000	.72281	.77651	.91724	1.00725	1.06316	1.09570	1.10213
16.000	4.000	.73951	.77651	.91439	1.01311	1.07731	1.11615	1.12477
18.000	2.000	.75371	.77651	.90873	1.01504	1.08754	1.13306	1.14412
20.000	.000	.76500	.77651	.90095	1.01363	1.09437	1.14693	1.16066

TABLE 4

Variation of Ground Impedance Phase for a Three Layer Model with a Variable Depth and Conductivity for the Top Unsaturated Soil and a Constant Conductivity for the Saturated (.005 mhos/meter) Layer and Bedrock (.00001 mhos/meter)

GROUND IMPEDANCE
FREQUENCY = 100. KILOHERTZ
ANGLE OF INCIDENCE = 80. DEGREES
TOTAL SOIL DEPTH = 20. METERS

ELECTRICAL CONSTANTS

	UNSATURATED SOIL	SATURATED SOIL	BEDROCK	
CONDUCTIVITY (MHOS/METER) =	VARIABLE	.00500	.00001	
DIELECTRIC CONSTANT =	15.	15.	15.	(SEAWATER -- 80. IS USED)
MAGNETIC PERMEABILITY =	1.	1.	1.	

- - - PHASE, RADIANS - - -

UNSATURATED SOIL DEPTH METERS	SATURATED SOIL DEPTH METERS	UNSATURATED SOIL CONDUCTIVITY (MHOS/METER) .01000	.00500	.00200	.00100	.00050	.00020	.00010
0.000	20.000	.51873	.51873	.51873	.51873	.51873	.51873	.51873
.200	19.800	.51335	.51873	.52200	.52308	.52359	.52374	.52358
.400	19.600	.50838	.51873	.52513	.52725	.52825	.52855	.52824
.600	19.400	.50380	.51873	.52810	.53125	.53278	.53316	.53271
.800	19.200	.49961	.51873	.53093	.53506	.53700	.53758	.53689
1.000	19.000	.49579	.51873	.53362	.53870	.54109	.54181	.54098
1.200	18.800	.49233	.51873	.53616	.54217	.54500	.54585	.54498
1.400	18.600	.48922	.51873	.53857	.54546	.54872	.54971	.54870
1.600	18.400	.48644	.51873	.54083	.54859	.55226	.55337	.55224
1.800	18.200	.48398	.51873	.54296	.55154	.55563	.55686	.5556.
2.000	18.000	.48183	.51873	.54495	.55433	.55881	.56016	.55877
4.000	16.000	.47510	.51873	.55770	.57337	.58117	.58356	.58110
6.000	14.000	.48883	.51873	.55837	.57721	.58719	.59034	.58708
8.000	12.000	.51558	.51873	.54812	.56677	.57764	.58118	.57733
10.000	10.000	.54957	.51873	.52780	.54242	.55253	.55591	.55162
12.000	8.000	.58640	.51873	.49803	.50396	.51108	.51341	.50973
14.000	6.000	.62288	.51873	.45921	.45068	.45164	.45140	.44623
16.000	4.000	.65694	.51873	.41169	.38157	.37168	.36617	.36037
18.000	2.000	.68739	.51873	.35601	.29579	.26793	.25170	.24311
20.000	.000	.71375	.51873	.29351	.19422	.13675	.09496	.07640

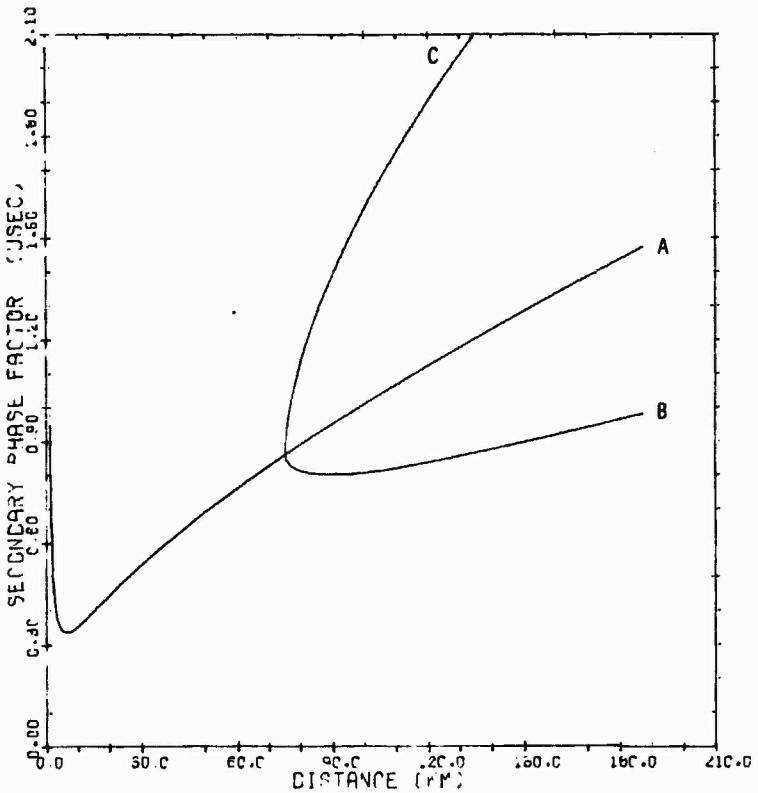


Fig.1 Effect of change in ground conductivity at 75 km from uniform 0.005 (A) to 0.05 (B) and 0.0005 (C) mhos/meter on secondary phase factor

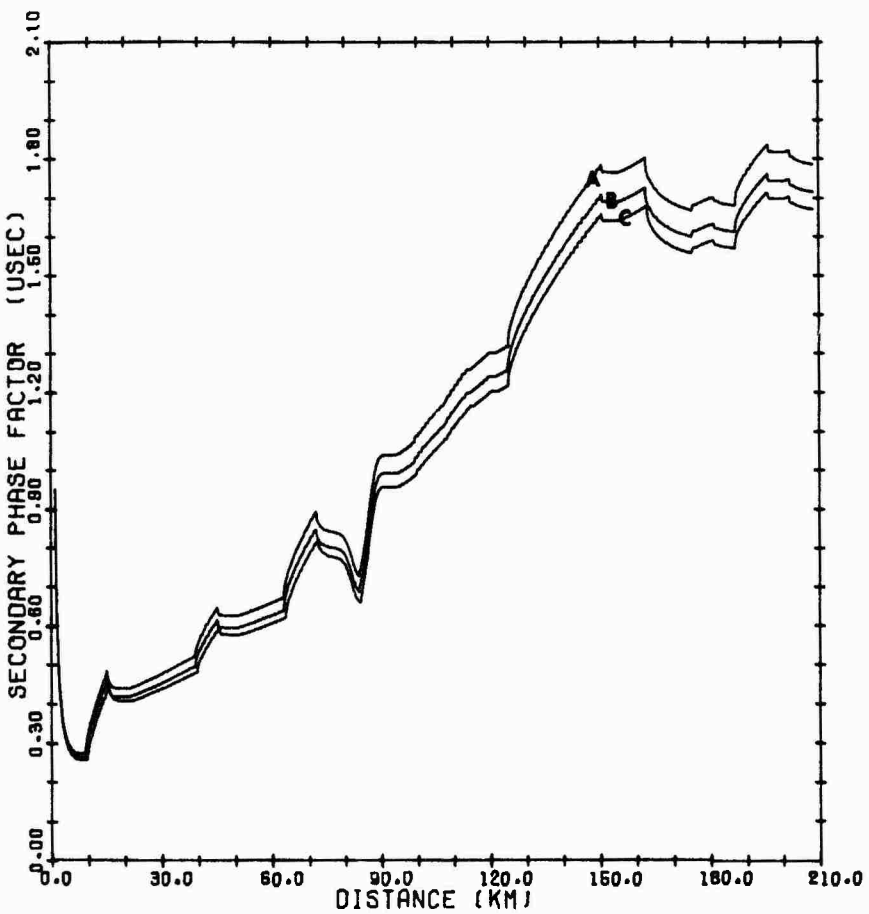


Fig.2 Change in secondary phase factor from ambient path (A) for a 30% (B) and 60% (C) increase in ground conductivity

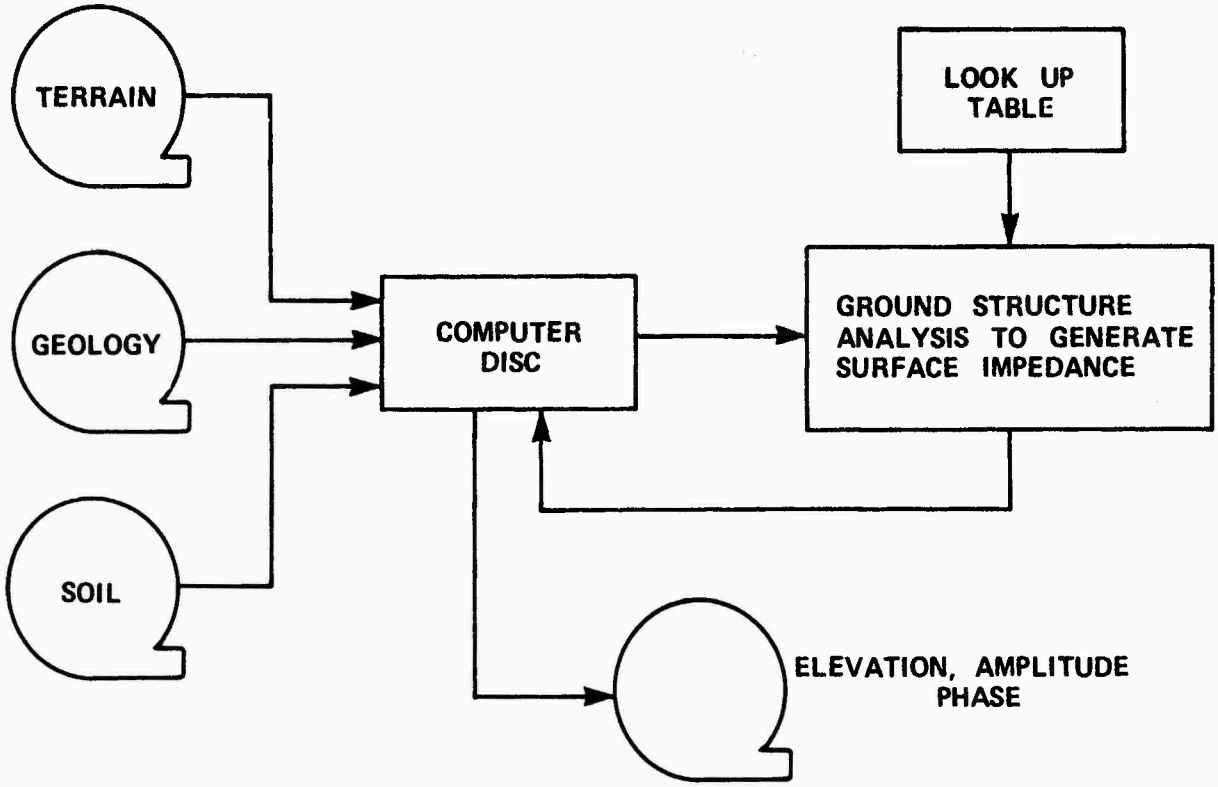


Fig.3 Data base of ground electrical properties

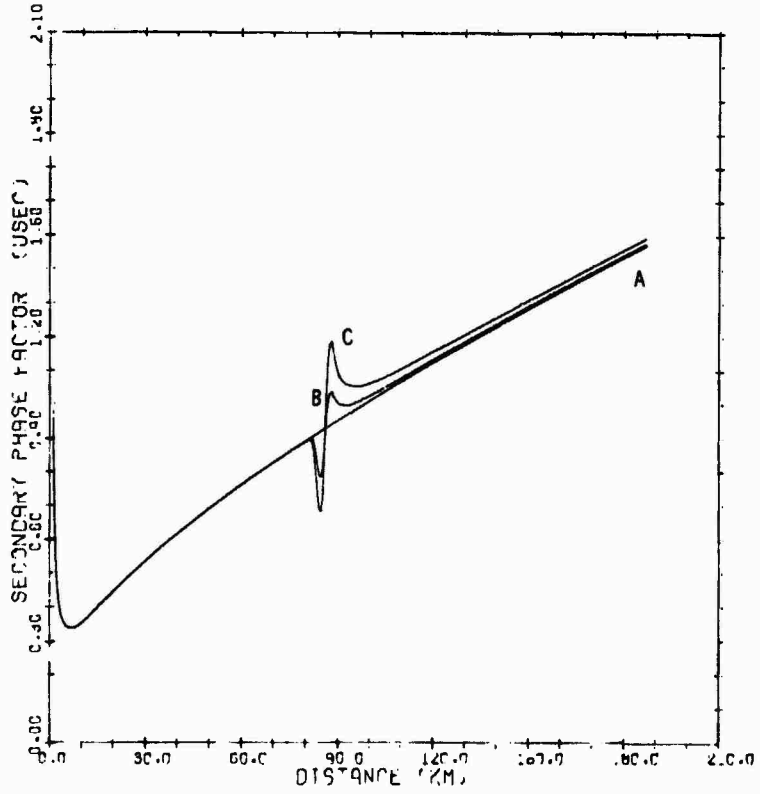
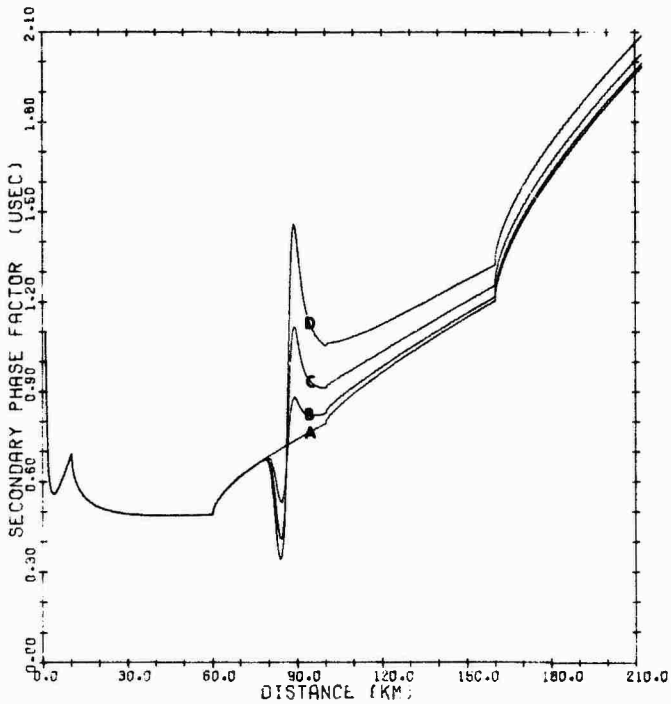


Fig.4 Effect of change in terrain elevation from smooth earth (A) to 250 meter hill (B) and 500 meter hill (C) on secondary phase factor. Conductivity is assumed at 0.005 mhos/meter over path



0-10 km, .001 Mhos/meter; 10-60 km, Seawater

69-100 km, .01 Mhos/meter; 100-160 km, .05 Mhos/meter

160-210 km, .001 Mhos/meter

A = Smooth, B = 0.5 km C = 1 km D = 1.5 km Hill

Fig.5 Variation of secondary phase factor for an inhomogeneous irregular path

DISCUSSION

A.Shuval, National Committee for Space Research, Radio Observatory, P.O.B. 4655, Haifa, Israel

What is the resolution of the grid of points entered into the computer program?

Author's Reply

Elevation, rock age, soil type and depth are digitized for a thirty arc second grid.

IONOSPHERIC EFFECTS ON LORAN-C IN POLAR REGIONS

T R Larsen and E V Thrane
Norwegian Defence Research Establishment
Kjeller, Norway

SUMMARY

Loran-C is a radio navigation system using pulsed signals at 100 kHz. Interference between ground wave and waves reflected from the ionosphere can reduce the accuracy of this system. The reflecting properties of the ionosphere are variable, and irregular disturbances affecting these properties are particularly frequent in polar regions. The paper describes computations of amplitude and time delay of a 100 kHz radio wave that is reflected from different ionospheric models. These models simulate day and night conditions and conditions during a severe ionospheric disturbance. The computations are made using a "full wave solution" and indicate at which distances from a Loran-C transmitter the skywave can interfere with the ground wave (time delay <30 μs between groundwave and skywave pulses). For undisturbed nighttime conditions interference should never occur, whereas for propagation over land during moderately and severely disturbed conditions interference may occur beyond distances of ≈800 and ≈600 km, respectively. For propagation over sea the corresponding distances are ≈1400 and ≈800 km.

1. INTRODUCTION

1.1. The principle of Loran-C navigation

Loran-C is a very accurate navigation system using pulsed radio signals with a carrier frequency of 100 kHz. A chain of stations may consist of a "master" and two or more "slaves" and may cover a geographical region with a radius of up to 2000 km. An example is the existing chain in the North Atlantic with a master station at Færøy Islands and slaves at Jan Mayen, Iceland and North-Norway. The pulses transmitted from the stations in a chain are synchronized with a great precision and the user determines his position by measuring the time delay between signals received from different stations. Accuracies of position determination of 150 m up to distances of 1800 km are quoted as typical in the literature (Potts, 1972). Each station transmits pulses in a coded pattern that ensures correct identification and prevents interference. We shall briefly discuss the principle of operation and a few important sources of error.

Figure 1 shows schematically the propagation of a Loran-C pulse from transmitter to receiver. The received signal will consist of a pulse that has travelled along the earth's surface and signals reflected one or more times from the ionosphere. At frequencies near 100 kHz the ionosphere is normally a good reflector, and we shall investigate under which conditions ground wave and skywave may interfere.

Figure 2 shows the transmitted and received pulses in more detail. The transmitted pulse is 250 μs wide, it has steep leading edge and reaches maximum amplitude after 70 μs. In the receiver the time of reception is measured at the zero crossing after the third swing, that is 30 μs after the start of the pulse. This time can be measured with an accuracy of 0.1 μs. The effect of noise is minimized through a phase coherent detection of the signal. In Figure 2 two skywaves are indicated arriving after one and two reflections from the upper atmosphere.

It is obvious that interference can only occur if the time delay between ground wave and skywave is less than 30 μs, and if the skywave has sufficient signal strength.

Our task is therefore to calculate the propagation velocity of the ground wave and the path length and reflection coefficient of the skywave.

2. GROUND WAVE PROPAGATION

Away from the near field at the transmitting antenna the ground wave will mainly be a surface wave whose velocity and attenuation depends upon the properties of the earth's surface. Computations of phase velocity and attenuation for low frequency waves over surfaces with different conductivities are available in the literature (Watts, 1967; Wait and Howe, 1956).

For a 100 kHz wave propagating over sea with conductivity $\sigma = 4$ mho m^{-1}, the phase velocity 1000 km away from the transmitter is $v = 0.9995\ c$, where c is the free space velocity of light. The attenuation is $A = 10$ dB in addition to the spatial attenuation. For soil with relatively poor conductivity $\sigma = 10^{-3}$ mho m^{-1}, the corresponding numbers are $v = 0.9977\ c$ and $A = 26$ dB. The computations are valid for homogeneous earth with a smooth surface. Rough terrain with varying conductivity may change these parameters, and poorly conducting soil dispersion may cause distortion of the pulse shape, but such effects have been neglected in the present work.

3. THE REFLECTING PROPERTIES OF THE IONOSPHERE

A 100 kHz radio wave (free space wavelength = 3 km) will be reflected from the lowest part of the ionosphere, the D-region, and for our purposes it is important to find an accurate method to determine the reflection coefficient and effective reflection height for a particular ionospheric situation and propagation path. The ionospheric properties that affect the radio wave, such as the electron density and collision frequency, can change significantly within a 3 km height interval in the D-region, and we cannot

therefore directly apply the principles of geometric optics to the propagation problem. We have instead used a "full wave" integration method developed by Pitteway (1965) and adapted by Bain and May (1967) to find reflection coefficient and height of reflection as a function of distance from the transmitter for selected ionospheric models. The method allows the specification of a plane, horizontally stratified ionosphere with arbitrary height variations of the electron density and collision frequency, penetrated by the earth's magnetic field, at an arbitrary dip angle. The complex wave fields and reflection coefficients of a plane wave incident upon the model ionosphere are then computed for different angles of incidence. The stationary phase principle is then used to find the angle of incidence corresponding to a particular distance along a curved earth from transmitter to receiver. Once this angle of incidence (I) is determined, a triangulation height is found (Piggott, Pitteway and Thrane, 1965) which may be used as an effective reflection height h_{eff} (see Figure 1).

The lower ionosphere has a systematic diurnal and seasonal variation. In addition, rapid and irregular variations may occur. These are most pronounced and frequent in polar regions, and may often disturb radio communications. We have chosen three different ionospheric models as a basis for our computations. One represents undisturbed nighttime conditions, one a weakly disturbed day and one a severe disturbance. The electron density profiles are shown in Figure 3. They have all been determined experimentally. The nighttime profile was measured by Smith et al (1966) using a ground based wave interaction technique, the others have been measured by sounding rockets in North Norway during an auroral disturbance 5 February 1969 and a Solar Proton Event on 25 February 1969 (Haug, 1972). The collision frequency ν_M has been specified using the relation $\nu_M = 8 \cdot 10^5\ p$ (Thrane, 1968), where p is the standard atmosphere pressure in Nm^{-2} for the appropriate season and latitude (CIRA, 1972).

The effective reflection heights resulting from our computations are indicated in Figure 3. They represent values obtained for a distance of 300 km from the transmitter and vary very slowly beyond this distance. We note that the reflection height is about 80 km at night, about 55-60 km during weakly disturbed daytime and may be well below 50 km during a severe disturbance. The reflection coefficients will normally increase with increasing distance. Beyond 300 km from the transmitter typical values for the three models are $|R| \approx 0.5$ (undisturbed night), $|R| \approx 0.005$ (weakly disturbed day) and $|R| \approx 0.2$ (severely disturbed day). The night model gives strong reflection because the electron densities at night are small below the reflection level. In a severe disturbance the wave may be reflected in a strong gradient of electron density at the very bottom of the D-region. Its path through the absorbing region caused by the disturbance may therefore be short. The net result is a strong reflection even though the electron density is large.

4. INTERFERENCE BETWEEN GROUND WAVE AND SKYWAVE

Based on the method and data we have discussed, we may now estimate the time delay of the skywave relative to the ground wave. Figures 4 and 5 show time delay Δt in microseconds as a function of distance from a transmitter for different effective ionospheric heights. The geometrical horizon for 1-hop-transmission is indicated on the figures. Transmission over sea and over land with conductivity $\sigma = 10^{-3}$ mho m^{-1} are chosen as examples. We note from Figure 4 that for propagation over sea the time delay will never be less than 30 μs for effective ionospheric heights greater than about 55 km. For propagation over land the corresponding height limit is 68 km. We also note that a very small change in ionospheric height may cause a large change in the distance at which interference can occur. For undisturbed nighttime conditions, interference should never occur.

The next step is to consider the relative amplitudes of ground wave and skywave. The correct procedure would be to compute the amplitudes of both waves near the sampling point (the 3rd swing in the ground wave pulse) as a function of time delay, and then use an experimentally determined criterion for when interference becomes a problem. In this preliminary study we have chosen an approximate method. Close to the transmitter the ground wave is, of course, stronger than the skywave. However, the amplitude of the skywave will decrease more slowly with distance than the amplitude of the ground wave, because the ionosphere becomes a more efficient reflector at oblique incidence. We have therefore, for each ionospheric model, and for two ground conductivities, computed the distance at which skywave and ground wave pulses have equal amplitudes. We then assume that if a) the time delay between ground wave and skywave is less than 30 μs and b) the skywave is stronger than the ground wave, then interference may occur. Table 1 shows the results. For each ionospheric model and for each type of ground, the largest distance listed indicates the radius of the zone near the transmitter that should be free from interference. Although our estimates are approximate, we feel that they nevertheless give interesting information.

Ground conductivities of about 10^{-3} mho m^{-1} are typical for North Norway, but rough and mountainous terrain may cause effective values smaller than this. In the zone where the ground wave dominates and interference cannot occur, the accuracy of the Loran-C system will be better than 0.5 μs corresponding to 150 m in position. In the zone where skywave and ground wave can interfere it is more difficult to estimate the uncertainties, but it seems clear that problems can arise. If, for example, the skywave at the sampling point has the same amplitude as the ground wave, but arrives in antiphase, the error in the time determination may correspond to several cycles.

At greater distances the skywave will dominate and the system accuracy will depend upon the stability of the ionosphere. At 2000 km a height change of 10 km in the ionosphere may give a change of 20 μs in time delay corresponding to a position error of 6 km.

Experience shows that at low and middle latitudes the ionosphere is normally very stable. Phase stability of ±1 μs during daytime and from ±4 μs to ±8 μs at night have been reported for skywave out to 3000 km from the transmitter (Doherty, 1963; Shapiro, 1968). There is little data available in the literature from polar regions. However, our calculations indicate that we must expect considerably poorer accuracy during ionospheric disturbances in and north of the auroral zone, where the propagation paths from the different Loran-C stations to the user may suffer disturbances in various degrees. If we assume that a 10 km height depression of the ionosphere represents an extreme disturbance, the uncertainty in the position is estimated to reach 10-15 km.

5. CONCLUSION

In certain geographic areas where the Loran C availability is reduced, effects like the ones discussed may contribute to a deterioration of the accuracy of the position fixes. For a geometry like the one shown in Figure 6 reception at Vadsø is at times marginal. Figures 7 and 8 give examples of Loran C pulses (integrated over 12 min) received at Vadsø. The variability of the amplitude of the skywave as well as the arrival time w.r.t. the groundwave are noticeable.

6. ACKNOWLEDGEMENT

The authors are grateful to Chief engineer Chr A Gloersen of the Norwegian Defence Communications Administration for providing the Loran C data and for many discussions.

REFERENCES

Bain, W.C., B.R. May, 1967, "D-region electron density distributions from propagation data", Proc IEE, 114, 1593-1597.

COSPAR International Reference Atmosphere, 1972.

Doherty, R.H., 1963, "Experimental observations and theoretical calculations leading to a model of the lower ionosphere", Proc Int Conf on the Ionosphere, The Institute of Physics and the Physical Society, 428-434.

Haug, A., 1972, "Electron density observations in the lower ionosphere", In: Ed. G. Bjøntegaard, VLF Propagation Symposium, Norwegian Institute of Cosmic Physics, Report No 7201, Oslo 3.

Pitteway, M.L.V. (Part I), Piggot, W.R., M.L.V. Pitteway, E.V. Thrane (Part II), 1965, "The numerical calculation of wave fields, reflexion coefficients and polarization of long radio waves in the lower ionosphere", Phil Trans Roy Soc, 251, 219-271.

Potts, C.E., 1972, "Precise time and frequency dissemination via the Loran-C system", Proc IEEE, 60, 530-539.

Shapiro, L.D., 1968, "Loran-C skywave delay measurements", IEEE, Trans Instrum Meas, IM-17, 366-372.

Smith, R.A., T.N.R. Coyne, R.G. Loch, I.A. Bourne, 1966, "Small perturbation wave interaction in the lower ionosphere", In: Ground based radio wave propagation studies of the lower ionosphere, p. 335-358, DRTE, Shirley Bay, Ottawa, Canada.

Thrane, E.V., 1968, "Collision frequency in the high latitude D-region", In: Ed. K. Folkestad, Ionospheric Radio Communications, p. 63-72, Plenum Press.

Wait, J.R., H.H. Howe, 1956, "Amplitude and phase curves for ground wave propagation in the band 200 cycles per second to 500 kilocycles", National Bureau of Standards, Circular 574, Boulder, Colo, USA.

Watt, A.D., 1967, "VLF Radio Engineering", Pergamon Press, Oxford.

Table 1.

The table gives the distances for which the propagation time delay between the groundwave and skywave equals 30 μs; the computations are made for three ionospheric electron density profiles and two values of ground conductivity. Also given are the distances for which the groundwave and skywave have equal amplitude.

		SEA ($\sigma = 4$ mho m^{-1})		LAND ($\sigma = 10^{-3}$ mho m^{-1})	
CONDITIONS	EFFECTIVE REFLECTION HEIGHT	DISTANCE WHERE $\Delta t = 30\ \mu s$	DISTANCE WHERE $E_{skw} = E_{grw}$	DISTANCE WHERE $\Delta t = 30\ \mu s$	DISTANCE WHERE $E_{skw} = E_{grw}$
UNDISTURBED NIGHT	83 km	∞	400 km	∞	200 km
WEAKLY DISTURBED DAY	56 km	1400 km	1050 km	800 km	500 km
SEVERELY DISTURBED DAY	50 km	770 km	950 km	500 km	350 km

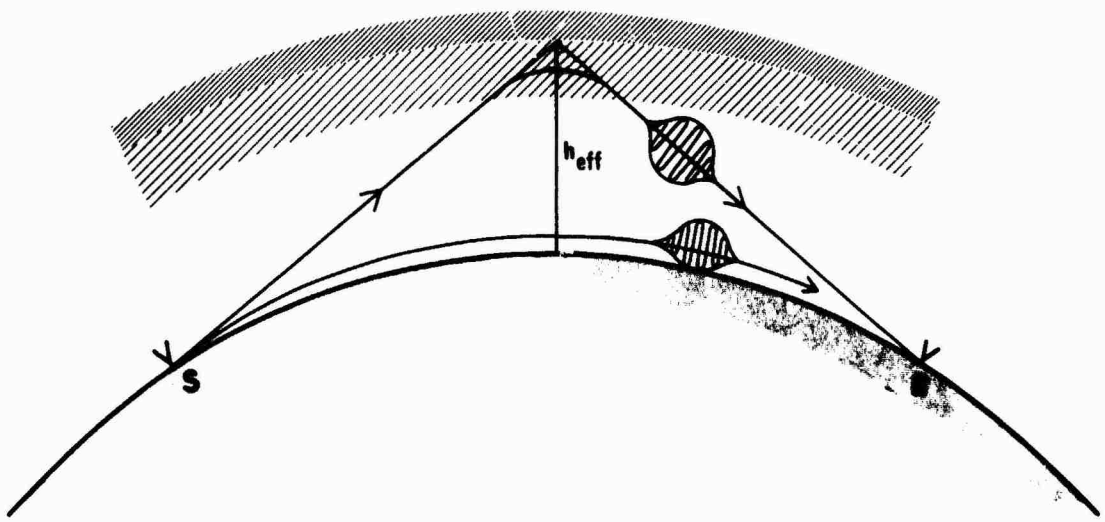


Figure 1. Schematic representation of groundwave and skywave propagation between sender (S) and receiver (R). The effective ionospheric reflecting heights (h_{eff}) is indicated. (Dimensions are not drawn to proper scale.)

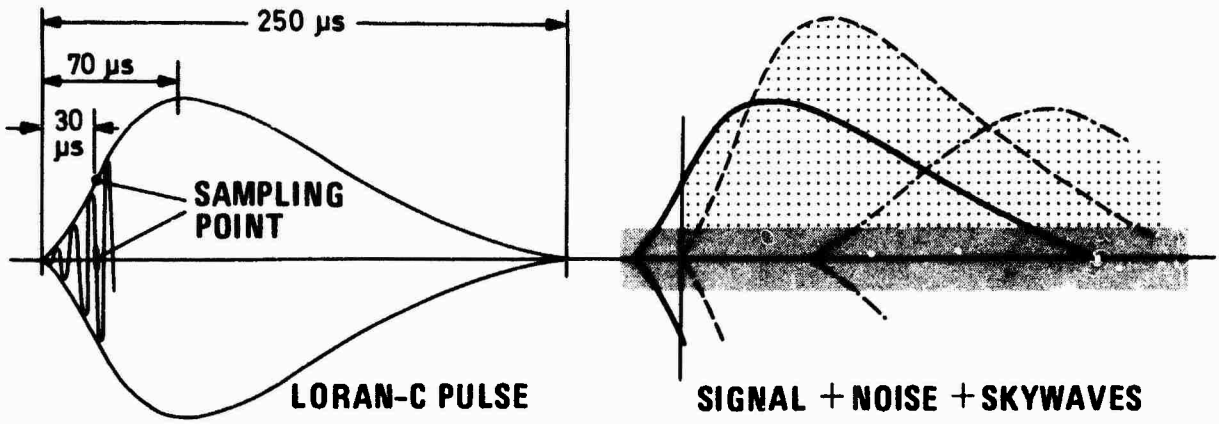


Figure 2. Schematic illustration of a LORAN C pulse (left). The received signal + noise + skywave is also illustrated (right).

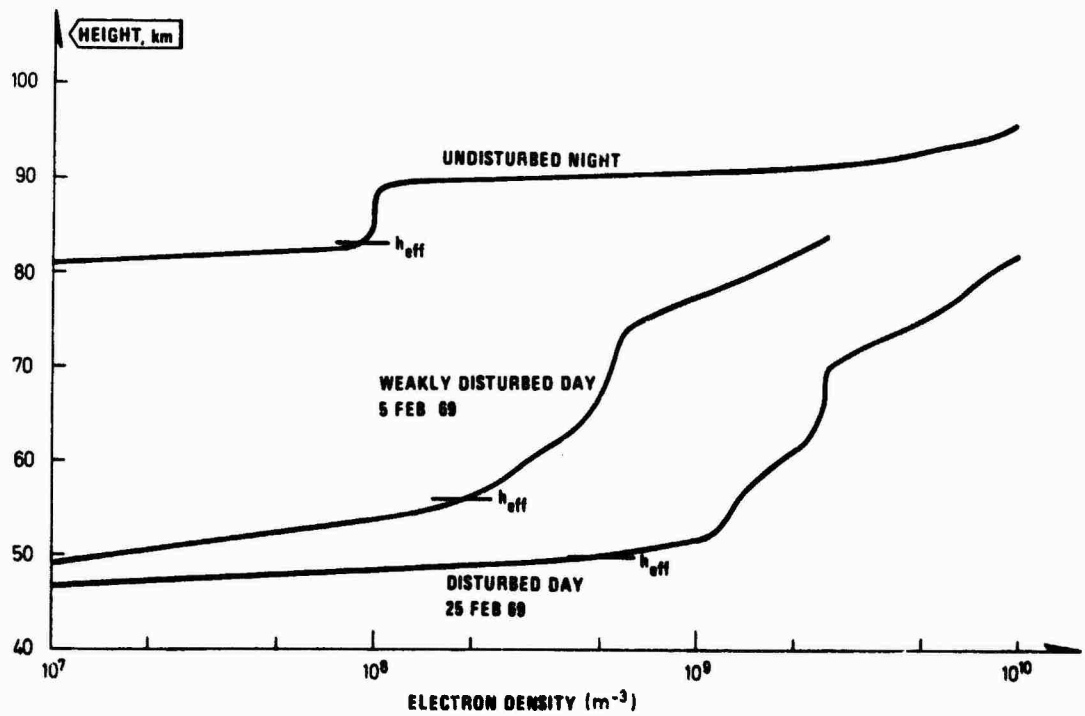


Figure 3. Ionospheric electron density height profiles for the cases considered in this study. The computed effective reflection height is marked on each profile.

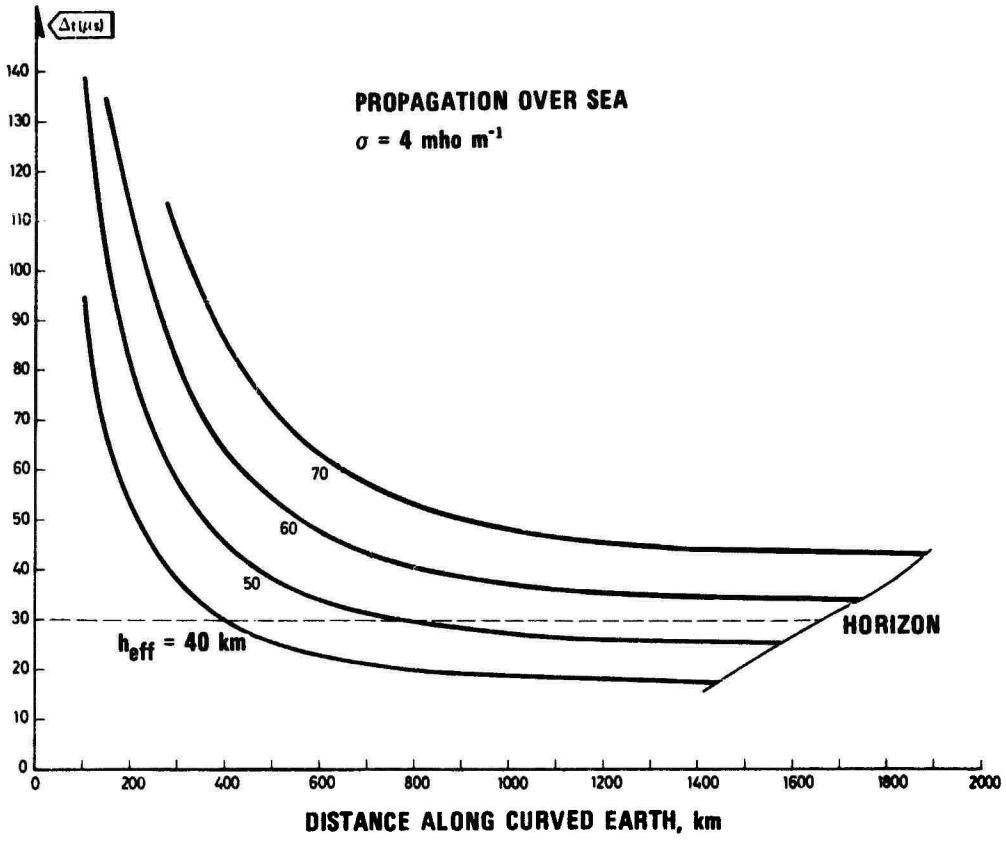


Figure 4. Difference in propagation time (in μs) between groundwave and skywave vs propagated distance along the Earth for various values of the effective reflection height. Propagation above sea.

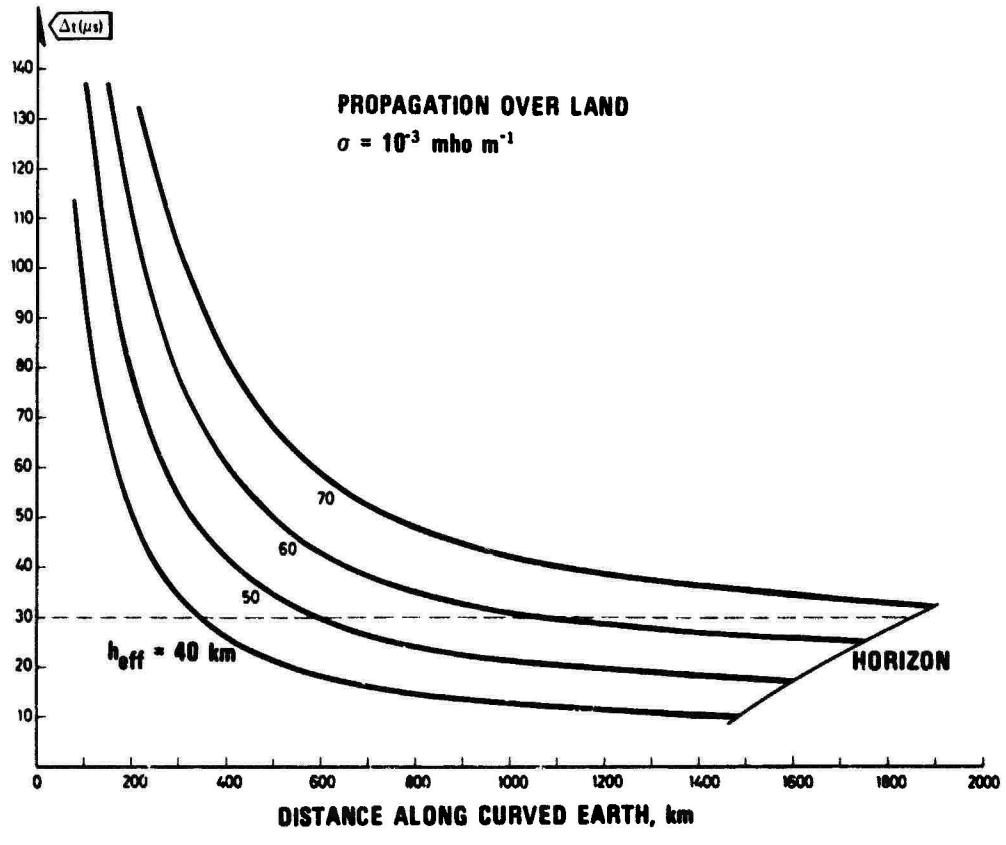


Figure 5. Difference in propagation time (in μs) between groundwave and skywave vs propagated distance along the Earth for various values of the effective reflection height. Propagation above poorly conducting land.

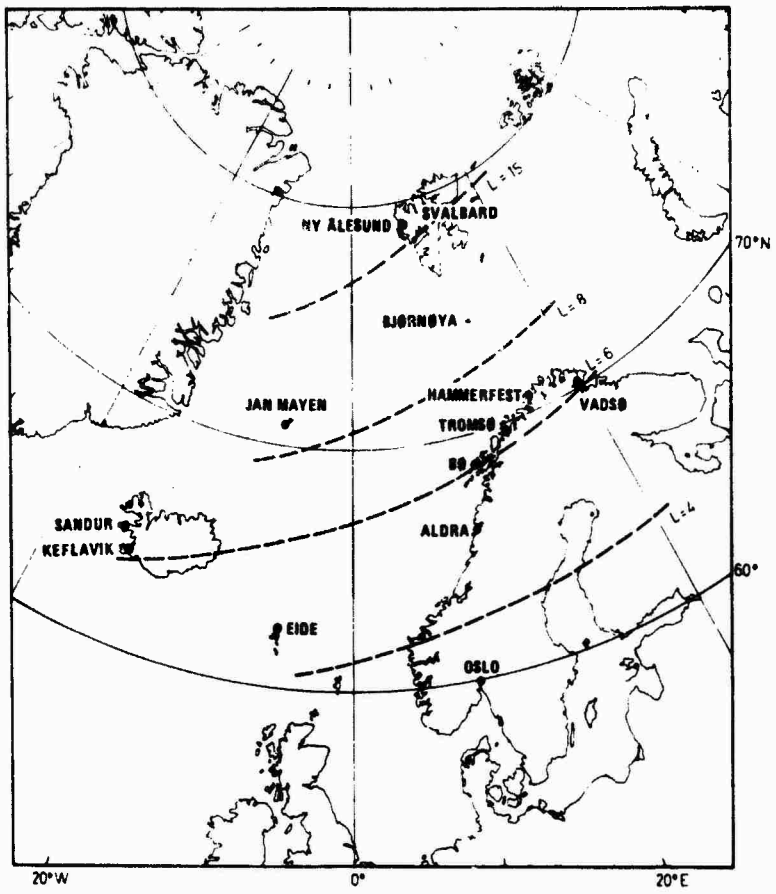


Figure 6. Map showing the positions of the LORAN C transmitters at Eide (Faraoe Islands), Sandur (Iceland), Bø (Norway) and Jan Mayen (Norway). The signals were received at Vadsø, North Norway.

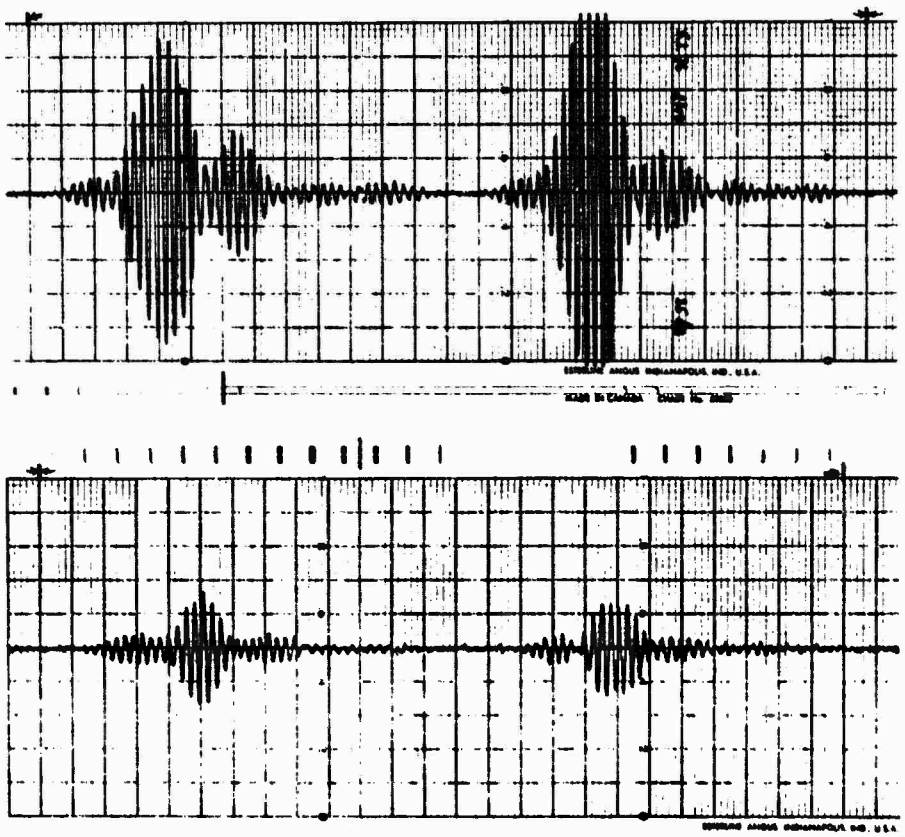


Figure 7. LORAN C pulses (integrated over 12 min using an AUSTRON 2000 receiver) from the Bø transmitter received in Vadsø on 5 Jan 76.

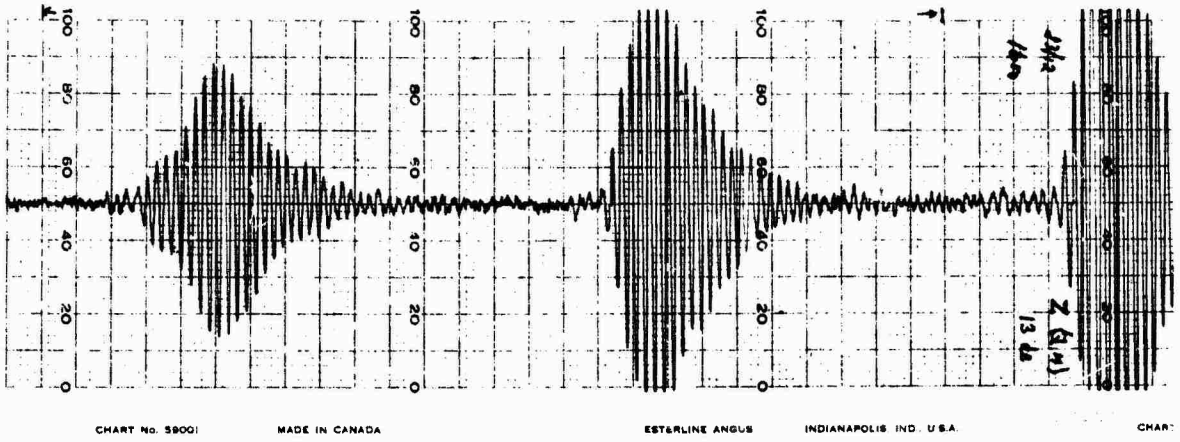


Figure 8. LORAN C pulses (integrated over 12 min) from the Jan Mayen transmitter received in Vadsø on 23 Dec 75.

PROPAGATION EFFECTS ON OMEGA

by

E. R. Swanson
Naval Electronics Laboratory Center
San Diego, CA 92152 USA

ABSTRACT

Broad attention is devoted to propagational effects on OMEGA. Primary emphasis is placed on experimental observations and empirical synthesis but with supporting reference to waveguide propagation theory. After a brief discussion of the history and background of OMEGA, a sketch of engineering and operational considerations is presented. Special propagational effects including Sudden Phase Anomalies (SPA's) associated with Sudden Ionospheric Disturbances (SID's) and Polar Cap Absorptions (PCA's) are defined. Normal propagational conditions are emphasized especially typical repeatability and typical diurnal behavior. Factors effecting the prediction of phase are summarized including characteristic velocity and phase variation with excitation of the waveguide mode, ground conductivity over a propagation path, variations with geomagnetic dip angle and path azimuth, effects of the aural zone and polar cap regions, and solar cycle variation of velocity. Temporal variations are also addressed including variation with solar zenith angle and anomalous changes at sunrise and sunset. The zenith angle effect is shown to explain seasonal variations as well as most diurnal change. General propagational considerations are shown to represent the received OMEGA phase over 95% of the time. A few percent of the time, however, phase is shown to be abnormally perturbed due to SPA's on sunlit paths or PCA's on arctic paths. Note is made of the importance of single mode propagation by the same dominant mode throughout the day. Coverage limitations for various individual station arising from complicated modal structure is noted.

1. OMEGA

Omega is a very low frequency (VLF) navigation system operating in the internationally allocated navigation band between 10 and 14 kHz. Eight stations, seven permanent, presently provide nearly global coverage. Global coverage from eight permanent stations is expected toward the end of the decade. In marine navigation, Omega normally is used as a single frequency hyperbolic system although several frequencies are available and the system can be used in the range-range mode.

The present stations are in essentially continuous 24-hour operation. At some sites this has been true for over ten years less incidental outages and major outages for facility improvement to the present 10 kW radiated power at the system base frequency of 10.2 kHz. Thousands of receivers have been constructed most being sold commercially. The U. S. Navy is installing Omega on all ships capable of independent operation except some. The U. S. Military Airlift Command is presently acquiring nearly one thousand receivers for some of its aircraft. Writing six years ago on the then existing four station developmental system, S. S. D. Jones noted: ". . . (Omega) is already the most widely deployed ground-based navigation aid in the world by a very substantial margin" (Jones, 1970).

1.1. History

Omega itself may be considered the oldest VLF navigation system. However, the system has evolved considerably from its earlier forms. In 1947 J. A. Pierce first proposed a hyperbolic navigation system based on phase difference techniques rather than time differences. In particular, a system operating in the vicinity of 50 kHz with a sine wave modulation of 200 Hz was suggested. An experimental system of this type was constructed by the Navy Electronics Laboratory and called Radux. Subsequently, in 1955, it was suggested that the Radux information from the LF Signal be combined with a separate VLF transmission near 10 kHz. This system was called Radux-Omega and initial 10.2 kHz transmissions were made in 1955. Later the LF transmissions were discontinued and ranges were expanded to a single frequency Omega system and, later, a multifrequency Omega system. Omega can thus trace a twenty-nine year evolutionary history and has included transmissions at the system base frequency for twenty years.

Modern transmissions using the stations in Norway, Trinidad, Hawaii, and Forestport, New York began in 1966. Previously, stations were used in a conventional slave-master configuration. Modern transmissions are derived from a bank of cesium frequency standards at each station and each station is controlled as a source of standard signals. The modern arrangement is especially practical for a global system in that the navigator can pair stations in any convenient way to obtain useful hyperbolic geometry and signals. The modern configuration has also proved more reliable than older arrangements.

1.2 Implementation

Practical implementation of Omega requires work in three broad areas: station design and construction; receiver design, manufacture and installation; and propagational theory development and implementation leading to either practical models for incorporation in automatic receivers or charts and tables for use manually. Work in all three areas has been proceeding concurrently for a number of years. However, it is the station construction schedule which has received the most attention. Rather apparently, crude navigation may be attempted with crude receivers and poor propagational knowledge; no navigation can be attempted without signals.

Anticipated stations sitings are shown in figure 1. Also shown is a station in Trinidad which was part of an earlier four station network and will continue to transmit at least through the end of calendar year 1976. All stations shown are presently constructed except the station in the South Pacific. The stations are given in Table 1 according to letter designation, location, antenna type and operating agency (modified from Tables I and II of Herbert and Nolan, 1975). The nominal radiated power of all stations is 10 kW at 10.2 kHz except for Trinidad which radiates less than 1 kW. At the present time there is no firm date for commencement of construction of the South Pacific station but past experience has indicated about thirty months is required from final sites selection to on-air (Herbert and Nolan, 1975).

Eventual system management has been described by Herbert and Nolan (1975). Omega stations not on U. S. soil will be operated by host nation agencies who will be responsible for maintaining the Omega signal without interruption and in phase with the world-wide Omega Navigation System. These agencies were listed in Table 1.

TABLE I
TRANSMITTING STATIONS

STATION	LOCATION	ANTENNA TYPE	OPERATING AGENCY
A	Bratland, Norway	Valley Span	Norwegian Telecommunications Administration
B	Liberia	Grounded Tower	Dept. of Commerce, Industry, and Transportation
C	Haiku, Hawaii	Valley Span	U. S. Coast Guard
D	La Moure, North Dakota, USA	Insulated Tower	U. S. Coast Guard
E	La Reunion Isl, Indian Ocean	Grounded Tower	French Navy
F	Golfo Nuevo, Argentina	Insulated Tower	Argentina Navy
G (Permanent)	South Pacific	(Grounded Tower proposed)	–
G (Temporary)	Trinidad	Valley Span	U. S. Coast Guard
H	Tsushima Island, Japan	Insulated Tower	Japanese Maritime Safety Agency

1.3 Format

Basic Omega signals consist of very low frequency 10.2 kHz continuous wave pulses transmitted sequentially from each station. Since the transmissions are time shared, a commutator is required to separate each station within the 10-second commutation pattern shown in Fig. 2. The commutation scheme is unambiguous and is synchronized in a known relationship to International Time as usually transmitted. However, the Omega commutation pattern is no longer simply related to regularly broadcast time pulses as was once the case when UT-2 was broadcast. The commutation pattern was originally defined to commence beginning at 0000 hours and repeat every 10 seconds. Present standard time transmissions employ "leap" seconds which are used typically once or twice per year. For Omega to follow leap seconds would require each user to jump his commutator whenever leap seconds were introduced to standard time. This is operationally inconvenient. Worse, it is potentially dangerous in the event some user does not get informed of the forthcoming time change. A one second change is sufficiently close to a one segment misalignment that it is possible for a receiver to appear to be tracking satisfactorily but actually be tracking on different pairs of signals. The inconvenience during set-up of having off-sets between Omega and International Time must be weighed against the inconvenience and incipient hazard of having occasional leap seconds.

A hyperbolic Omega receiver measures the phase of two or more Omega stations against a reference generated from an internal oscillator. The internal oscillator permits storage of the phase information so that the relative phases of the different stations can be intercompared. Readout is the phase difference in centicycles between selected stations and ordinarily is recorded continuously on strip chart recorders. Since the comparisons are all between signals of the same frequency, no internal ambiguity can arise within the receiver due to divider jumps. Further, a hyperbolic Omega receiver has no intrinsic specification for absolute radio frequency phase since shifts common to all stations will be removed in the hyperbolic differencing. The system does, however, have an inherent physical ambiguity. Since adjacent carrier cycles cannot be distinguished, the measured phase of each carrier is inherently ambiguous by one cycle. Hyperbolic phase differences on a baseline between two stations are ambiguous by the hyperbolic spacing of one-half wavelength.

Because of continuous operation, the navigational lane ambiguity problem inherent in a single frequency system generally is not troublesome to the navigator. However, additional frequencies are included in the full Omega format to permit reestablishment of lane should difficulties occur (Fig. 3). Additional frequencies included for lane identification are 13.6 and 11 1/3 kHz; they may be used independently in exactly the same manner as the 10.2 kHz transmissions. However, ambiguities in the 13.6 kHz lines-of-position will be coincident with those at 10.2 kHz only every 24 miles. Similarly, LOP's at 11 1/3 will only be coincident every 72 miles. Comparison of the coincidence of lines-of-position obtained at the various frequencies can thus serve to establish the proper lane of the 10.2 kHz carrier within the 72 mile ambiguity of the lowest beat frequency (Pierce *et al*, 1956; Swanson, 1955; Swanson and Hoppelley, 1962; Burgess, 1969; Burgess and Walker, 1970). Several practical aspects of lane resolution should be emphasized: 1) lane is usually established in-port and maintained by continuous tracking; 2) in the event of single station outage using a four station receiver, no ambiguity has occurred; 3) brief outages present no problem since lane is easily re-established by dead reckoning; 4) if a lane ambiguity arises, it is probably not on all LOP's therefore not all need be re-determined; 5) if not all lanes have been lost, lane can be reset by intersecting existing LOP's with only one external navigational LOP such as sun line, depth contour, radar-range, etc.; and 6) if coincidence of lanes is not reasonably "close," navigation directly on the difference frequency is possible, or alternatively, stations may be paired differently to present less accurate fixing within unambiguous, more highly divergent lines-of-position. The latter options are particularly important. Omega lane resolution need never be forced. Lanes may be resolved selectively. If an ambiguity of 8 miles should develop on an LOP, it is still possible to use the 3.4 kHz beat position derived from 13.6 and 10.2 kHz to an unambiguous position accuracy of about two miles. If, at some later time, coincidence improves, then the carrier lane identification can confidently be made (Swanson, 1971).

Figure 4 also shows additional transmissions on unique frequencies indicated by $f_1, f_2, \ldots f_8$. These signals are presently transmitted by some but not all stations. They are useful for vlf navigation as described in section 2 and, coincidentally, are very useful for frequency dissemination. Some stations presently transmit two unique frequencies separated by 250 Hz instead of the one shown. The 250 Hz Omega difference frequency has been used successfully to identify carrier cycles directly in specialized timing applications (Wilson *et al*, 1972; Chi and Wardrip, 1974; Kugel, 1975 [contains bibliography]). Recently, implementation of a fourth commutated frequency at 11.050 kHz has been suggested by Swanson *et al* (1975), who reviewed the evolution and potential of the Omega format. Although transmissions at 10.2, 11 1/3, and 13.6 kHz are firm, additional format details are somewhat open and may be decided at an international meeting in the Fall of 1976.

1.4 Practical Navigation

Omega is most primarily familiar as Stringer (1970) observed at a recent meeting of the British Institute of Navigation. "It satisfies the three R's -- reliability, redundancy and range." The existing operational stations individually maintained an average 99.7% of the scheduled availability (99.4% total time availability) during the last period for which statistics have been computed (Vannicola, 1976). In an evaluation of one receiver, maintainability could not be assessed since, in 20,000 hours of operation, the receiver did not fail. As previously noted, VLF signals propagate to great range so that Omega will provide redundancy. In many locations, *all* eight stations will be detectable although perhaps only five or six may be usable. A range limitation will often be due to a station interfering with itself by propagating around the world by both a short and long propagation path, e.g., repeatable signals from Trinidad are received in Hawaii at certain times by propagation 3/4 of the way around the world eastbound from Trinidad.

Manual Omega navigation requires: antenna, antenna coupler, receiver (preferably with strip chart recorders), skywave correction tables, and hyperbolic charts (or lattice tables). Equipment has been built by numerous manufacturers. Available charts and tables are listed in the catalog of the U. S. Oceanographic Office. Once the equipment is installed, it must be properly checked out and synchronized to the commutation pattern. Checkout procedures can vary between equipments of various manufacturers. Synchronization with the commutation pattern is obtained by one of three methods: 1) look at (or listen to) the Omega signals and move the commutator into alignment; 2) knowing the relation between Omega and International Time, start the commutator from a standard 10-second pulse such as might be obtained from WWVB; or 3) knowing the strongest signal in the area, align the commutator until the correct segment corresponds. In practice the third alternative is the most commonly used although, because of anisotropy in signal propagation, a guess of the strongest signal expected in an unfamiliar area can be hazardous. Once commutation is properly established, phase tracks should settle within a few minutes and a fix may be taken. Knowing the actual position and working backwards, correct lane counts can be annotated on strip chart recorders and/or lane counters set. Before sailing, it is desirable to prepare hyperbolic skywave corrections for the station pairs to be used on the voyage. If the scale of available Omega charts is not satisfactory, lattice tables should be used to plot Omega lanes on the local charts.

Underway, the fixing procedure is simply to read the various phase differences indicated by the receiver, add the respective skywave corrections and plot the resulting LOP's on the chart. The reduction takes about one minute. Of course some reasonableness checks should be made prior to taking the fix. Commutation synchronization can be checked. If the receiver is equipped with a "no signal" light or with field strength indication, the navigator should assure himself that adequate signals are available. Also, strip chart records should be scanned to verify continuous tracking and lane count. Ironically, the high reliability and automatic tracking of Omega place an unusual premium on self-discipline in routinely performing reasonableness checks. If Omega were not reliable, the navigator would appreciate the obvious need for performing elementary checks. However, since a hundred fixes may be obtained without difficulty, it is human nature to assume there will never be any difficulty.

There are automatic Omega receivers as well as manual. Automatic usually implies a receiver which can automatically synchronize itself as well as read-out directly in latitude and longitude. Some receivers exist which have an automatic synchronization capability but which read-out hyperbolic phase differences rather than latitude and longitude. Most automatic equipment includes some approximation to the propagation model used to compute global phase differences. Most, but not all, automatic equipment is intended for aircraft use. A pioneering equipment was the AN/ARN-99 Omega receiver (Smith, 1971) of which the AN/SRN-17 multi-frequency receiver intended for surface application and the AN/BRN-7 receiver intended for submerged application are derivative. Another pioneering automatic receiver was the Canadian Marconi Corporation CMA-719 (Mactaggart, 1971). Numerous other manufacturers have recently produced equipment. Accuracy specifications for automatic receivers have been addressed by Sakran (1974). More recently airborne specifications have been addressed by Special Committee 126 of the Radio Technical Commission for Aeronautics (RTCA). Evaluations of various receivers have been carried out under the U. S. Air Force 2041 program.

Over 53 manufacturers have constructed Omega receivers.

1.5 Special Uses

A number of specialized uses or forms of Omega have developed to exploit particular aspects of the system. Ordinarily, Omega accuracy will not be significantly influenced by poor signal to noise ratio. Usually signals will not be marginal and the primary accuracy limitation will be due to inherent day to day fluctuations in the transfer (or "mapping") function of the propagation medium itself (Swanson, 1970).

Since errors will be introduced by all segments of the long transmission paths to an area, it follows that two receivers in the same area will exhibit similar errors. This spatial correlation results in excellent rendezvous accuracy and also excellent accuracy for relative navigation such as might be needed within a task force or for traffic control. Differential or micro-Omega are systems designed to upgrade nominal system accuracy in local areas through use of spatial correlation (Swanson *et al*, 1974a; Swanson *et al*, 1974b; Nard, 1974; Hastings and Comstock, 1975). Differential Omega concepts have been used for oceanographic survey. OPLE and PLACE were designed by NASA for tracking vehicles or traffic control (Laughlin *et al*, 1965; Laughlin *et al*, 1967). They employ an on-board Omega transponder and satellite relay to a central processing site. Most recently work was conducted on the Global Rescue Alarm Net (GRAN) wherein it was envisaged that mariners and airmen could carry a small transponder which would telemeter the Omega format through a satellite to ground much as in OPLE and PLACE so that position of the unit needing rescue could be determined. Work has been described by Crawford and Rupp (1972). Although the GRAN program is not presently being pursued, it showed considerable technical promise.

In addition to the spatial correlation features exploited by Rendezvous and Differential Omega, exploitation is made of the dispersive correlation between variations on one frequency as closely selected to those on a second over the same propagation path. Dispersive correlation significantly reduces the errors expected using beat frequency navigation or lane resolution. A special application called "Composite Omega" has been suggested by J. A. Pierce to take optimum advantage of dispersive correlation (Swanson, 1965; Pierce, 1966; Swanson, 1966; Pajousek and Reder, 1973; Pierce, 1974; Brown and Van Allen, 1976).

Auto-correlation of propagation variations is also significant being 8 to 10 hours for normal propagation conditions at night. However, diurnal prediction errors correlate over shorter periods and may become the dominant consideration in practical applications. Measurements by Wright (1969) indicate a potential for determining velocity to 1/3 kt when the propagation paths are either all dark or all sunlit.

Special use of Omega has also been made to track weather balloons. Use has been made of the multiple frequency transmissions for geophysical exploration. Timing applications using narrowly spaced difference frequencies were mentioned in section 1.3. Timing using carriers and usual difference frequencies have been described by Swanson and Kugel (1972) while lead edge capability was further studied by Swanson and Adrian (1973). Accuracies of a few microseconds are possible using carrier phase measurements.

2. VLF NAVIGATION

VLF offers tremendous advantages for navigation: Coverage is excellent; reception is reliable. This is a characteristic of the frequency band and is not necessarily restricted to the particular navigational implementation we know as Omega (Litchford, 1971; Swanson and Robie, 1973). In addition to the generally favorable propagation characteristics, there is also a compatibility between bandwidth capabilities which can be practically engineered and its inherent bandwidth requirements as developed by Palmer (Swanson, 1974). Considering the inherent capabilities of the frequency band and the considerable variety which can be employed in designing navigation systems, it may be at first surprising that there is not a great proliferation of vlf navigation systems. Cost, combined with the attempt to make the Omega format itself flexible, have served to restrict other approaches. However, there are two techniques which should be noted.

Many vlf communications stations throughout the world are stabilized to standard frequency. In the past, transmissions have been continuous wave with on-off keying. Presently, many U. S. Naval transmitters are being converted to minimum shift keying. For navigational purposes the change constitutes a complication but not an insurmountable difficulty. With simple carrier transmissions from three or more stations, the various signals can be received; then shifted in frequency to some common frequency and then paired and phase differenced (Palmer, 1972). Alternatively, they can be used in a range-range mode against a precision oscillator. Such systems must be initialized. They cannot refine an approximate location as can Omega. In operation a navigation system using such signals is somewhat like an inertial system in that it is exactly correct at the time and place of initialization and then deteriorates with separation and elapsed time. However, the "error drift" is not unlimited as is true with inertial equipment. The error budget for such equipment has been discussed by Swanson and Dick (1975). A substantial number of receivers employing this principle have been manufactured. Flight results from one particular equipment have been described by Tymczyszyn (1973).

Beuker (1973) describes a vlf navigation system maintained by the USSR. The frequencies are 11.905, 12.649 and 14.881 kHz with stations located in the east, west, and center of the Soviet Union. Radiated power appears high: 50 to 100 kW.

Although this paper addresses primarily Omega propagation, much of the material covered is applicable to any vlf navigational technique. However, there are significant differences in emphasis between propagational effects at the vlf communications frequencies, usually above 20 kHz, and those in the 10–14 kHz navigation band.

3. PROPAGATION AND THE MEDIUM

The propagation characteristics which permit use of vlf at great range also introduce limitations on Omega. To a greater extent than at higher frequencies, the use of vlf assumes a reliance on nature to provide repeatable propagation. Repeatable propagation usually occurs but with some unwanted temporal and spatial complexity. Also, on occasion, irregular variations occur.

One convenient analytical model for vlf radio propagation is that of a concentric spherical waveguide formed between the earth and the ionosphere. If only one propagation mode is supported by the waveguide, then there will be a simple relationship between phase and distance at long distances from a transmitter. Phase at any given point will be related to distance to the transmitter, ground conditions especially ground conductivity, and characteristics of the lower D-region of the ionosphere which forms the upper boundary of the waveguide. As the ionosphere is a magneto-plasma, propagation is anisotropic so that phase is not only a function of distance from a transmitter but also dependent on geophysical path details.

The ionosphere undergoes regular predictable seasonal and diurnal changes as a function of the solar illumination. One practical problem is the prediction and removal of this unwanted temporal variation.

A limitation is the occasional unpredictable variation caused by unpredictable variations in the geophysical environment. The two most important anomalous variations are Sudden Phase Anomalies (SPA's) caused by Sudden Ionospheric Disturbances (SID's) and Polar Cap Absorptions (PCA's). These effects are related to solar flares. SPA's are caused by the X-ray flux impinging on the sunlit ionosphere thus anomalously changing the width of the waveguide. PCA's result from flares which also emit protons. The charged particles are guided by the earth's magnetic field to the polar regions where they cause an anomalous reduction in the width of the waveguide. Effects of SPA's and PCA's are discussed in section 4.2. Usual stable propagational characteristics will be discussed next.

3.1 General Characteristics

Characteristics of vlf propagation in general and Omega propagation in particular will now be developed by adopting specific propagation models as convenient. A subsequent section discusses the relationship of various propagation models.

The concept of a spherical waveguide leads immediately to several valid predictions. For example, energy will propagate around the world in all directions and may reinforce at the antipodal point from the transmitter. This phenomena has been observed and illustrates the extreme range obtained at VLF. Also, since the lower ionosphere or "D-region" is controlling at VLF, one would expect a severe attenuation when the wavelength becomes comparable with the height of the guide at about 70 to 90 Km, i.e., a frequency of about 4 kHz. A severe increase in attenuation is observed as expected. Another generality, very important to navigation systems, is that if one mode should be dominant, then phase and amplitude should vary regularly as a function of distance from the transmitter without fluctuations due to interference between various modes. This condition may occur at any distance but in practice it is especially true for frequencies near 10 kHz at distances greater than about 1,000 Km. Note that the regularity is occurring at *large* distances. This is just the opposite from the usual experience with, say, Loran A, where the groundwave propagation at short distances is regular and multihop skywave can cause irregularities at longer distances.

VLF propagation has been studied for many years. This is due not only to its practical use as the mainstay of fleet communications for a half century but also because of the extreme repeatability of measurements. Although diurnal variations occur, measurements over paths from 5,000 Km to 10,000 Km long typically show repeatability of about 1 dB in field strength while phase variations are measured in microseconds. Although detailed prediction of ULF fields is difficult, some notable success has been achieved (see figure 4)

The most accurate theoretical work is now being done using digital computers to solve the waveguide problem. However, expressions developed by J. R. Wait and others may provide some insight. For long paths where only one mode is dominant, the waveguide theory of the sharply bounded ionosphere yields

$$E = \frac{3 \times 10^5}{h} \left[\frac{P\lambda}{a \sin d/a}\right]^{1/2} e^{-\alpha' d} \Lambda \qquad (1)$$

where

E = Received field strength in μV/m
h = Height of the ionosphere
P = Radiated power in kilowatts
λ = Wavelength
a = Radius of the earth
d = Path length (great circle distance)
α' = Attenuation coefficient (Napiers/Mm); $\alpha_{dB/Mm} = 8.68\,\alpha'$
Λ = Excitation factor

Note that the field strength decreases exponentially with distance modified by a "focusing" term, sin d/a, to take into account the convergence of the spherical field in the far half of the world. The excitation factor, Λ, accounts for the transfer of energy into the propagating mode. Expressions similar to equation 1, or similar equations employing a superposition of contributions from various modes, result from a number of ionospheric models.

3.1.1 Theoretical Characteristics

Wait and Spies (1964) made extensive computations to determine the theoretical characteristics of VLF propagation under a variety of conditions. From equation 1 and the associated discussion, it is apparent that the important parameters are: the propagating mode, characteristic attenuation, characteristic velocity, and characteristic excitation. Theoretical assumptions include mode, n, the effective ionospheric height, h; the effective ionospheric gradient, β; the ground conductivity, σ_g; and in the case of anisotropic models, the path orientation with respect to the magnetic field. Other assumptions, which include an ionospheric conductivity parameter related to the effective electron collision frequency, and the specific definitions of various terms are beyond the present scope (cf. Wait, 1970). For present purposes it is sufficient to note that the assumptions may be adequately realistic to suggest behavior, although detailed calculations for specific propagation paths may differ significantly because of the sensitivity of many of the modal parameters to precise path characteristics.

Figure 5 shows computed attenuation rates as might apply over a seawater propagation path for various frequencies. The curve for h = 70 is approximately applicable during day while the curve for h = 90 Km approximates night.* Note that the minimum attenuation occurs around 16 kHz increasing slowly for higher frequencies but increasing more rapidly at lower frequencies because of the aforementioned cut-off when the waveguide height approaches one wavelength. While Figure 5 was for the first mode, Figure 6 shows the attenuation rate for the second mode. Note that at night at 10 kHz the attenuation rate of the second mode is 9 db/Mm compared with 1-1/2 dB/Mm for the first mode, but at 20 kHz it is only 3 dB/Mm compared with 1.7 dB/Mm. Figures 7 and 8 show the excitation factors for the two modes which indicate the relative ease with which propagation may be established. Excitation factors for night conditions at 10 and 20 kHz for the two modes are shown in the following table:

Excitation Factors (dB)

	Frequency	
Mode	10 kHz	20 kHz
1	-1.0	-10
2	+1.3	+2.5

In both examples shown it is easier to excite the second mode than the first. However, at 10 kHz the difference is only 2.3 dB while at 20 kHz it is 12.5 dB. These differences should be compared with differences in attenuation rates between modes for the respective frequencies. At 10 kHz the first mode will equal the second mode after 0.3 Mm while at 20 kHz equality will not occur for nearly 10 Mm. As a consequence, the interference pattern formed between the first and second modes is much less consequential at the lower VLF frequencies than in the upper portion of the frequency range. The spatial interference in loop antennas may be negligible at long distances at frequencies near 10 kHz but must be considered at almost any range for frequencies near 20 kHz propagating at night. (In practice, the situation is considerably more complicated due to radical variation of the propagation parameters with latitude and direction at night.)

The previous conclusions have been based on the isotropic model of Wait and Spies. Their results provide a useful starting point as an introduction to VLF propagation but it will now be worthwhile to make specific comparisons with more general theory and with observation. Pappert has developed a computer solution to the problem of propagation in the earth-ionosphere waveguide as noted in section 3.2.2. The formulation provides for quite general ionospheric electron density profiles, exponential variation of gyro frequency with height, magnetic dip angle, magnetic field strength, magnetic path orientation, and ground conductivity. Snyder and Pappert (1969) have performed numeric calculations for the specific propagation paths between San Diego and Hawaii. Results for both the isotropic case and the anisotropic case were obtained. Since their formulation is different from that of Wait, they have a somewhat different definition of height, i.e., h' instead of h, but the distinction is not significant here. It is important to note that they are computing for a different height but one which is also considered to be representative of propagation conditions at night. Further, their mode numbering system is necessarily arbitrary. Attenuation rates and excitation factors for seven modes are shown in figures 9 through 12 for the isotropic case and also for both propagation from Hawaii to San Diego and also San Diego to Hawaii. Note that the attenuation rates are generally less for propagation to the East. Table II gives the contributions to the mode sum from the first few modes for three separate cases: 1) The isotropic model of Wait and Spies with h = 90 Km; 2) Full wave calculations from Snyder and Pappert for isotropic propagation between San Diego and Hawaii with h' = 84 Km, and 3) Anisotropic calculations for propagation from San Diego to Hawaii. As mentioned previously, the isotropic nighttime results from Wait and Spies show equality between contributions from the first and second modes at 1/3 Km for 10.2 kHz but at 10 Mm for 20 kHz. Modal constants for the isotropic model with slightly lower effective ionospheric

*More exact estimates are given in DCA Rept. 960-TP-74-5 (1974) and DCA MEECN Rept. C650-TP-76-4 (1976).

TABLE II
COMPARISON OF RELATIVE AMPLITUDES OF VARIOUS MODES

Model	Frequency	Propagation Constants n	α	Re Λ	Relative Field Strength (dB) at 1/3 Mm	1 Mm	10 Mm
Wait & Spies	10 kHz	1	1.5 dB/Mm	-1 dB	-1.5	-2.5	-16
h = 90 Km		2	8.8	1.3	-1.6	-7.5	-86.7
	20	1	1.7	-10	-10.6	-11.7	-27
		2	3.0	2.5	1.5	-0.5	-27.5
Snyder &	10	1,2	1.7	0	-0.6	-1.7	-17
Pappert		3	12.0	1.8	-2.2	-10.2	-118.2
SD–Hawaii isotropic h′ = 84 Km	20	1,2	1.5	-6.5	-7.0	-8.0	-21.5
		3	3.7	2.5	1.3	-1.2	-34.5
Snyder &	10	2	1.2	-18.5	-18.9	-19.7	-30.5
Pappert		1	2.2	-2.0	-2.7	-4.2	-24.2
SD to		3	11.5	1.8	-2.0	-9.7	-113.7
Hawaii h′ = 84 Km	20	2	1.7	-11.0	-11.6	-12.7	-28.0
		1	2.3	-21.8	-22.6	-24.1	-44.8
		3	3.7	3.0	1.8	-0.7	-34.0

reflection height as calculated by Snyder and Pappert are generally similar to those of Wait and Spies except that the second mode is attenuated somewhat more rapidly. The anisotropic modal constants are quite different. Mode 2 shows the lowest attenuation rate but is so weakly excited as to be suppressed even at 10 Mm at 10.2 kHz. At 10.2 kHz, mode 3 is nearly equal to mode 1 at 1/3 Mm but is rapidly dominated by the first mode at greater distances. At 20 kHz, mode 1 is suppressed at all distances but there is a change in dominance between modes 2 and 3. Indeed, propagation near 20 kHz is even more complex than indicated by Table II since more than the first three modes are important. Figure 13, from Snyder and Pappert, shows the phasor mode sum as a function of distance as obtained by summing a partial set of modes and summing the first seven modes. Propagation near the upper end of the vlf band is thus quite complex. Since the variations in amplitude with distance shown in figure 13 also imply irregular variation of phase with distance, propagation prediction for navigation using a frequency near 20 kHz would be considerably more difficult than prediction for a frequency near 10 kHz. However, Figure 4 indicated the feasibility of such calculations.

Phase velocities for the first mode are shown in Figure 14. Note that the velocity is lower at night. In practice, navigation using phase measurements places very stringent demands on knowledge and predictability of phase velocity. For example, an error of one part on 10^4 in phase velocity would cause a position error of one-half mile at 5,000 miles. At this precision, curves such as those in Figure 14 are not especially useful as many secondary factors can contribute to change the velocity significantly. For example, Figure 15 from Wait and Spies (1965) shows the theoretical variation of phase velocity with ground conductivity. Note both the minor differences occurring between sea water and normal ground with a conductivity near 10 mmho/M and the significant dependence of the velocity when the conductivity is exceptionally poor. Theoretical variation of velocity with ground conductivity for 10.2 and 13.6 kHz is shown in section 3.3 for both day and night conditions.

A marked variation on attenuation rate by variations in ground conductivity also occurs. For example, at the Omega frequencies an attenuation rate of a few decibels per 1,000 Km might be considered typical for propagation over sea water or over typically conducting land. However, over ice with ground conductivity 0.01 millimho/meter the attenuation would be on the order of 20 dB/1,000 Km.

Perusal of figures 5 through 15 and of the hundreds of functions given in the references will provide a good qualitative and quantitative introduction to VLF propagation. However, the reader is again cautioned that significant variations can occur with latitude and direction especially at night (see section 4.4).

3.1.2 Experimental Observations

The most important changes are the diurnal variations of phase and amplitude. They are typically associated with ionospheric change related to variation of the solar zenith angle over the propagation path being studied. Two typical examples of diurnal variation of amplitude are shown in Figures 16 and 17. Note that the field strength tends to be constant at night and decreases following the transit of the sunrise line over the propagation path to a lower value just after the entire path becomes sunlight and finally reaches a nominal value during the day. The decrease just after sunrise is very common for frequencies near 10 kHz and is typically about 4 dB and lasts for about an hour. Although the details of this particular sunrise decrease are not well understood, they presumably are related to both the dynamics of ionospheric dissociation and recombination rates leading to daytime equilibrium and another phenomena especially important in the 20–30 kHz frequency range, viz: mode conversion caused by the sunrise line. Typical diurnal variation of phase over a long west to east path at 10.2 kHz is shown in Figure 18. A constant or "flat" night is observed shifting into a slow variation or "curvature" during the day. Diurnal phase variation at the Omega frequencies is discussed further in section 3.3. At higher frequencies within the VLF range phase tends to be somewhat less stable during the night but more stable in mid-day. Figure 19 shows that the diurnal variation can be considerably more complex than is usually observed at 10.2 kHz. The "steps" during transitions are probably due to modal conversion at the terminator. A theory for the step phenomena has been proposed by Crombie (1964 & 1966) and has had some success over long paths (Lynn, 1971 & 1973). However, if the results shown are close to those for a similar path treated by Pappert and Morfitt (1975), then a more complicated modal conversion model is necessary for this particular path.

As previously mentioned and shown in Figures 16 and 17, field strength is typically repeatable to a standard deviation of about 1 dB from day to day at 10.2 kHz. Repeatability of phase is, however, one of the most useful properties of VLF. Phase variations occur due to ordinary random variations in the ionosphere or, usually to a much lesser extent, to variations in ground conductivity. Occasional larger variations occur. Already mentioned were sudden phase anomalies (SPA's) and polar cap absorptions (PCA's). Their effects will be discussed in section 4.2. Typical stabilities appropriate for single propagation paths of various lengths from about 4 Mm to 10 Mm are given in Table I.

TABLE III
STANDARD DEVIATIONS OF VLF SIGNALS
(microseconds)

Period	Frequency 10.2 kHz	13.6 kHz
Day	3 μs	2 μs
Night	5	4
Transition	4	4

The higher Very Low Frequencies tend to be more stable during the center portion of the day but less stable at night. Typical navigational effects on Omega are discussed in section 4.1.

In all discussions of the repeatability of phase measurements it is very important to draw a distinction between day to day statistical scatter, or "noise," and electromagnetic noise. Repeatability is limited by the stability of the transfer or "mapping" function affecting the signal after it leaves the transmitting antenna and until it is received at the receiver (Swanson, 1970). This type of repeatability is effected by random variations in the ionosphere or ground and will not improve by increasing the electromagnetic signal-to-noise ratio. Normally, the signal-to-noise ratio does not significantly affect repeatability except over exceptionally long paths and/or with unusually weak transmitters.

Natural noise at VLF is primarily due to electromagnetic signals radiated from thunderstorm lightning. The total thunderstorm noise can be computed on a global basis from all "storm centers" (Maxwell and Stone, 1965). In addition, the noise is highly impulsive and proper receiver techniques may yield improvements on the order of 15 dB over what could be obtained if the noise were Gaussian (Swanson and Adrian, 1973). While the ambient noise is usually high, arctic areas are typically quiet and reception of weak signals is common. These characteristics, as well as the wide variation in signal strength which may occur depending on propagation path length, combine to give stringent specifications for VLF receivers.

A second area of low signal strength but adequate signal-to-noise ratio is underwater reception. VLF signals are unique in their ability to penetrate sea water to useful depths. At 10 kHz the attenuation rate is about 3 dB/meter which indicates reception at 50 feet 8 Mm removed from a 10 Kw station.

Propagation blackouts, of the type sometimes experienced at higher frequencies, are rare to the point of being virtually unknown at VLF.

3.2 Approaches to VLF Propagation Prediction

It will be convenient to identify various approaches to vlf propagation prediction and then elaborate and interrelate certain approaches. These methods can be categorized as:

1. Complete
2. Full Wave Waveguide
3. Hop
4. Traditional Mathematical
5. Parametric
6. Statistical

A brief background on each of the methods follows.

3.2.1 Complete

The "complete" approach is that theoretically optimum method which has never been implemented. In principle, it should be possible to solve field equations directly in terms of specified boundary conditions at the surface of the earth and in the ionospheric magnetoplasma. In practice, of course, full boundary conditions at the earth boundary would have to reflect existing geographical features in detail (especially continental shapes and ground conductivity variations), while specification of the ionospheric boundary would require detailed information on electron and ion densities as affected by geographical position and the sunrise-sunset variation. Even if a full, complete formulation were possible, it is questionable whether present computers are of sufficient size to support solution. Certainly, it is expected that the complete approach would be unnecessarily complex and uneconomical in practice.

3.2.2 Full Wave Waveguide

In simplest form, the full wave waveguide approach considers a spatially homogeneous earth-ionosphere wave guide wherein the lower boundary is specified by a specific ground conductivity and the upper boundary by an arbitrary electron density profile. Correct differential equations for propagation are specified and the solution is worked out through appropriate algorithms within a large digital computer. The approach differs from the traditional in which problems in mathematical physics are solved toward as nearly a closed form of solution as possible, and then computations, if any, are used simply to evaluate the solution. Many workers have used full wave solutions in conjunction with wave guide notions. Pappert, Gossard, and Rothmuller (1967), following Budden's formulation, were first to report the results of a program which fully allowed for earth curvature, ionospheric inhomogeneity, and anisotropy. Subsequent contributions on different aspects have been published by various investigators and a more general book was written by Galejs (1972).

3.2.3 Hop

The hop or ray approach to radio prediction has been traditional at lf and above for many years. Extension to the vlf range is complex in that the wavelength becomes comparable with the height of the wave guide. Extension of hop theory to the very low frequencies has been primarily due to the work of Pitteway (1970) and Berry and Herman (1971).

3.2.4 Traditional Mathematical

Considerable work along traditional lines has been accomplished to obtain closed form solutions for vlf propagation within an idealized spherical earth-ionosphere wave guide. Variables include ground conductivity, exponentially varying electron density, and exponentially varying collision frequency. Magnetic field effects are accounted for in certain cases. Major work along these lines was conducted in the early 1960's, particularly by Wait (1970). Numerical results were published by Wait and Spies (1964) and Watt (1967). See also Galejs (1972).

3.2.5 Parametric

Parametric approaches do not intrinsically seek to relate observed radio quantities such as phase or amplitude to inherent geophysical parameters such as electron density profiles. Instead, relevant propagation parameters such as attenuation, velocity, and excitation are assumed to be specified in terms of readily defined path characteristics such as orientation, latitude, ground conductivity, diurnal period, etc.

Two parametric programs are of special importance because of their present widespread application; both are historically early developments. The Omega skywave correction program first became operational in its present format in March 1964. The program is restricted to phase prediction at specified Omega frequencies, but predicts phase 24 hours per day at all seasons. Variations were initially incorporated with orientation, ground conductivity, latitude, sunspot number, and diurnal period. The second program of major interest is the RCA-NRL program developed by Schwartz. This program was based on earlier work at higher frequencies, and computes amplitude for various frequencies within the vlf communications band. The vlf program originally included variations with frequency, diurnal period, orientation, ground conductivity, zenith angle, and latitude. The first coverage predictions were published in 1965.

Both the Omega skywave correction program and the RCA-NRL program have been extensively modified and refined since inception.

3.2.6 Statistical

Generally, vlf fields tend to be either quite stable and simply explained – e.g., at long range at navigational frequencies near midday – or extremely complex – e.g., at short range at communications frequencies at night. Perhaps because of this, a moderately complex statistical model tends to be either unnecessary or inadequate. Accordingly, the author knows of no generally applicable statistical models for vlf propagation, although most models, especially the parametric ones, may be statistical in that input parameters are determined statistically.

One model of special application is the Kasper-Ell model for use with Omega data. This model has been extensively developed by Kasper (1970) after initial feasibility studies at the Naval Electronics Laboratory Center (NELC). The model is especially interesting in that it operates in conjunction with the Omega skywave correction model. The parametric skywave correction model is first applied to account for recognized spatial or diurnal variations, and then an error field is developed for actual observations. Spatial correlation is then evaluated and smoothing conducted to whatever extent may be statistically justified.

3.2.7 Comments and Prognosis

The complete approach may be recognized as the "correct" approach in a fundamental if not practical sense. Other approaches must be compared with the complete approach, particularly as to assumptions of spatial homogeneity.

The full wave mode, hop, and traditional approaches all seek to determine propagation parameters such as attenuation rate, velocity, and excitation based on input such as electron density profiles, ground conductivity, magnetic field, orientation, and/or reflection coefficients. In the mode solutions, the propagation parameters are deduced for the various modes supported by the wave guide. Evaluation of the true received field thus implies a summing of contributions from various modes. The hop or ray approach implies the summing of various rays. One mode may usually be expected to become dominant at long range, while one hop (or groundwave) may be expected to become dominant at short range; therefore, hop theory may offer some advantages in quicker convergence for shorter paths or for fixed transmitter-receiver separation, while the wave guide approaches offer advantages for longer paths. However, the advantages of hop theory are more marked at lf than vlf, and the method will not be considered further here.

The traditional mathematical approach to the simply bounded wave guide solution, although refined to considerable elegance and sophistication, is obviously more limited than the computer approach: exponential ionospheric electron distributions are used in the traditional method, while the more advanced full wave methods can compute modal parameters for arbitrary electron density profiles and also incorporate the effects of ions. Because of the complex interrelationships accounted for in the general full wave formulation, large variations in propagation parameters can occasionally occur due to minor path changes. In one extreme case noted by Snyder (1968), a variation in path azimuth of 2° caused a variation in the excitation factor of 10dB.

In application, the full wave waveguide approaches require various geophysical data as input including path azimuth, magnetic field, ground conductivity, and electron density profile. Output are the propagational parameters for the modes supported under the input conditions. When applied to practical paths, two complications arise: (1) the required geophysical input, especially the electron density profile, may not be adequately known, and (2) conditions will vary along propagation paths, so some type of path segmenting must be employed.

Extensive efforts to deduce electron density profiles have been conducted for many years. Most methods of deducing electron density profiles do not work well in the ionospheric D-region. The best data appear to be the result of vlf sounding. Vertical sounding was first conducted at vlf by NRL at Smithland, Arizona in 1953 (Knecht *et al.* 1962). Subsequently, more extensive installations have been built by Hildebrand at Thule, Hawaii, and in the Mojave Desert. Shellman (1970), using the full wave solution as a subroutine, has developed and refined electron density profile determination from sounding data to obtain confidence limits on the results. Profiles have also been deduced manually by comparing full wave predictions based on various assumed profiles with observed flight data as shown, for example, in a recent MEECN report (1976).

Variable propagation conditions along the path have been handled by Bickel using the WKB technique, which is valid provided the propagation parameter variation is slow. An application of full wave theory using an assumed electron density profile and the WKB

approximation is shown in figure 4 together with data later observed during flight measurements (Bickel *et al*, 1970). The agreement is excellent, although the calculations do not include allowance for ionospheric roughness. Gossard and Paulson (1964) have shown typical ionospheric fluctuations of 2 to 3 km at night. Ionospheric roughness is likely to cause increased relative attenuation of higher order modes. Recently theory has been extended to include mode conversion (Pappert and Shockey, 1974; Pappert and Shockey, 1975; Pappert and Morfitt, 1975).

Parametric approaches have considerable economic advantages. Also, since the bulk of such programs is ordinarily devoted to bookkeeping and summing of the various functions as applied to path segments, the programs are inherently compatible with the full wave solution. The full wave solution provides guidance on the separability of the variables and suggests appropriate functional forms for the variations; the parametric program incorporates these and performs calculations over long paths. Parametric approaches work best when the range of variables which must be considered is least and when the effects of geophysical variations on the propagation parameters can be separated. Although some sophistication can be economically incorporated into parametric programs to account for variations which are not readily separated, close coupling of propagation parameters is likely to necessitate use of a full wave solution.

In addition to economic advantages, parametric approaches offer a precision and sensitivity to output not readily obtained with a full wave program. To vary the result of a full wave program, the geophysical input must be changed – most likely the assumed electron density profile. Such changes of input may cause significant variations in phase and/or amplitude not only at the frequency of interest but at other frequencies as well. The difficulties in properly deducing profiles based on observations have already been mentioned. With a parametric approach, simpler methods of deducing the parameters are available (Swanson, 1971).

The statistical approach has the disadvantage of not incorporating any physical variations no matter how well understood. Confidence in the results derives from huge quantities of data and the associated measurement statistics. The principal disadvantages of the statistical approach are overcome in the Omega Force-Fit technology by applying the approach to the error field developed after application of the Omega skywave correction parametric model. Ideally, the parametric model would not produce an error field, and, hence, further corrections would not be needed. In practice, predictions based on any physical model will exhibit errors due both to intrinsic model limitations and inaccuracy of the required input such as ground conductivity. The statistical approach allows removal of prevailing errors, at specific sites where sufficient data have been collected, without necessitating basic model changes.

Present and future use of the various prediction approaches is difficult to assess. Clearly the "complete" approach is of only academic interest. Traditional mathematical methods will be frequently applied to practical problems but their importance should not be underestimated as it is often through such efforts that the fundamental nature of problems can be understood. Statistical adjustments to published Omega predictions are likely to remain in wide use as the principal means of removing known local biases. However, their importance may wane somewhat as automatic receivers come into widespread use and further improvements are made in global prediction theory. The global model for Omega phase prediction developed by the author and described further in section 3.3 is, in a sense, perhaps the most widely applied model. It is parametric and use is restricted to the prediction of Omega phase. It is used for production of voluminous propagation correction tables issued by the Hydrographic Center of the U. S. Defense Mapping Agency. The model, or simplifications, have also been incorporated in various automatic Omega receivers. Because of calculation efficiency and ease of adjusting parameters for best accuracy, the method or extensions of the basic method can be expected to be in widespread use into the future. However, increased use of full wave or hop calculations may be anticipated to assess coverage areas wherein Omega can be used reliably with the model. A parametric computation for field strength at the vlf communications frequencies is extensively employed by the U. S. Naval Research Laboratory (NRL). The remaining prediction techniques are full wave mode and hop developments.

There has been a significant proliferation of full wave mode and hop computer programs which can be used for determining Omega or vlf fields. Results of the various methods have been well cross-checked. Various full field approaches have been inter-compared while results from more complex programs have also been compared, e.g., that of Johler with that of Berry and also the NELC program developed by Pappert with that developed by Berry (Morfitt, 1972). Choice of program thus depends on the sophistication available and the need for complexity, cost, and familiarity of specific workers with the various approaches. In practice, use patterns suggest that the latter consideration may be most important. Most full wave work in Europe, especially in the U. K., seems derivative from work by Pitteway (see Smith and Pitteway, 1974). Some work in the U. S., particularly at lf and the communications band in the vlf range has used Johler's program using zonal harmonics. The U. S. Air Force is using the Wave Hop program by Berry for a number of calculations in the vlf and lf ranges. The wave hop has also been used for limited Omega computations to determine fields at relatively short ranges (< 1000 Km). The U. S. Navy and U. S. Coast Guard make extensive use of full wave mode calculations based on the work of Pappert but highly refined by numerous workers. One program of special note is the Integrated Prediction Program (IPP) developed by Snyder and Ferguson (Ferguson, 1972). This program determines geophysical parameters, finds the modes, and computes the mode sum over a long great circle path over the earth. The advantage of such a sophisticated program for practical applications is obvious. Current work is centered in two areas. Theoretical extensions to include mode conversion have already been noted. Since Pappert and Morfitt (1975) note that their most recent mode conversion model can be implemented with about the same ease as a WKB model, it seems likely that these techniques will receive wider practical application in the future. A second active area is that of insuring that a full set of modes is identified in the course of solution.

3.3 Phase Prediction

A paper presenting the theoretical formulation used for Omega phase prediction has been written by Swanson (1971). This work was expanded by Swanson and Brown (1972) into a document including not only the background and theoretical formulation of Omega phase prediction but also listing the FORTRAN program used in the production of Omega propagation prediction (or skywave correction) tables used with Omega. More recent program documentation has been written by Morris and Cha (1974). Considering the detailed documentation available, this paper will present but a superficial review. Principal attention will be directed at the factors effecting the spatial variation of phase. This section has been abstracted primarily from the aforementioned 1971 paper including some editorial revision done in preparing a more recent paper (Swanson, 1975).

3.3.1 Path Definition

The first problem is to define the propagation path. It is mathematically easiest to consider a line path of no lateral extent. However, Fermat's principle of least time indicates that deviations from the straight optical path will occur if propagational velocities are markedly different to one side of the optical path. Such asymmetry across the path could occur, for example, if the path paralleled a coastline or if the sunrise or sunset terminator were parallel to the propagation path. Calculations have been performed under idealized geometry assuming a nearby terminator, and the path deviations were found to be negligible.

Figure 20 shows the ray hop geometry for propagation from a transmitter to a receiver. Rays launched near the tangent plane suffer the fewest reflections and, hence, contribute maximum field strength at the receiver. Rays launched near the vertical contribute almost nothing at the receiver. The illustration suggests that ionospheric conditions directly over the transmitter or receiver are of little importance in determining the total field at great distance, and, hence, some regions of excitation (and de-excitation) should be treated separately. For ionospheric heights near 80 km, the geometry indicates that the excitation regions should extend inward from the ends of the path by somewhat less than 9° of arc on the earth's surface.

The wave guide approach does not yield quite as simple a visualization of the relative influence of the path ends. However, calculations have been conducted to determine the spatial extent of the first oblique Fresnel zone assuming various propagational parameters (Bradford, 1971). The midpoint of the two edges was typically 8° removed from the transmitter, while central zone (secular ray) calculations were typically grouped near 5°.

The spatial extent of the excitation region is best determined experimentally (Swanson and Bradford, 1971). A typical feature of a vlf phase recording is a marked variation at the beginning of sunrise. The time of occurrence of the sunrise onset was first determined for certain predominantly north-south paths near the equinox when path illumination would be occurring simultaneously. It was found that sunrise began when the zenith angle reached 98°. Using the deduced zenith angle to define the commencement of sunrise, ephemeral data were used to compute the time of the beginning of sunrise at the eastern end of various predominantly east-west paths. As expected, the computed times occurred significantly earlier than indicated by corresponding observations. Knowing path orientation and latitude, the time discrepancies were converted into the distance along path at which the effect could be assumed to take place. The effective path shortening is 6.1° at each end, which is in accord with both theoretical treatments.

3.3.2 Segmenting

Having deduced that the propagation path on a spherical earth can be approximated as a segment of the great circle between transmitter and receiver but shortened by regions of excitation and de-excitation at the ends, it is now necessary to develop a treatment for the central or midpath portion. Assuming a single mode model and slow variations of the propagation parameters, a simple path averaging technique, where phase shift is added from one path segment to another, will be appropriate. Phase prediction over a long path then becomes the addition of excitation phase variations at the ends plus a summation of incremental phase shifts over each path segment in the central portion:

$$\varphi = \varphi_{\text{excitation}} + \varphi_{\text{de-excitation}} + \sum_{\text{path}} \delta\varphi_i \tag{2}$$

Phase shifts on excitation and de-excitation are developed based on respective local geophysical conditions at the path ends, whereas each incremental phase shift in the midpath is a function only of geophysical conditions prevailing over the path segment. The phase shift expected over an incremental midpath segment is a complicated function of various path and geophysical parameters $X_1, X_2, \ldots,$ such as path orientation, magnetic field orientation and magnitude, ground conductivity, etc.:

$$\delta\varphi = \delta\varphi(X_1, X_2, X_3 \ldots) \tag{3}$$

An alternate form of practical convenience is

$$\delta\varphi = \delta\varphi[X_1, f(X_1, X_2), X_3, g(X_3, X_4), \ldots] \tag{4}$$

where various geophysical functional relationships have been specified, i.e., certain terms have been arbitrarily coupled. Since one of the component dependencies of the arbitrarily coupled dependencies remains as an independent variable, equation 4 can retain the same complete generality of equation 3 if f and g are single-valued functions of their respective arguments. If forms of the functional relationships are chosen so that cross products in the generalized Taylor expansion drop, Equation 4 may be expanded to

$$\delta\varphi = \delta\varphi_0 + \left.\frac{\partial\delta\varphi}{\partial X_1}\right|_0 X_1 + \frac{\partial\delta\varphi}{\partial f} f + \ldots$$

which is a linear equation of the form

$$\delta\varphi = \delta\varphi_0 + (K')^T X'$$

where

$$K' = \left.\begin{pmatrix} \frac{\partial\delta\varphi}{\partial X_1} \\ \frac{\partial\delta f}{\partial f} \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}\right|_0 \quad \text{and} \quad X' = \begin{pmatrix} X_1 \\ f \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}$$

Substitution into equation 2 yields the phase over any given path

$$\varphi = \varphi_{\text{excitation}} + \varphi_{\text{de-excitation}} + (K')^T \Sigma X_i' \tag{5}$$

or

$$\varphi = K^T X \tag{6}$$

where K and X represent K′ and Σ X′ redimensioned to incorporate representations for excitation and de-excitation. Phase differences may be computed by applying equation 5 to the component path segments and subtracting. K may be found by regression analysis of observed phase differences.

3.3.3 Spatial and Temporal Partitioning

Obviously, temporal and diurnal variations can be incorporated directly into the formulation presented. In practice, however, it is more convenient to treat diurnal variation separately. Addressing a selection of data, all obtained when the associated propagation paths were dark, allows determination of the spatial variation at night without simultaneously addressing a wide range of diurnal conditions. Spatial variation during the day may be similarly determined after adjusting data to correspond to normal illumination over all component propagation paths simultaneously. Spatial predictions can thus be developed for extreme illumination conditions. Phase predictions for conditions of intermediate illumination are obtained by proportioning the day and night predictions based on a diurnal function and the illumination conditions that have developed on each path segment. Spatial and diurnal variation are thus derived independently.

3.3.4 Spatial Variation

The spatial theory for phase developed by the author, and widely used for global Omega propagation predictions, is a parametric theory. The art in the development is choosing a suitable form for the parametric forms, as used in equation (4), so that the cross products drop in the general Taylor expansion, leading to equation (6) which is the matrix linearization used to compute phase shift over each path segment. Detailed parametric forms and associated coefficients, as determined by both theoretical considerations and regression analysis of Omega data. are given in the references, particularly Swanson, 1971.

Velocity variation with ground conductivity is addressed by discrete coefficients representing velocity variation from sea water for each conductivity range used to represent a discrete conductivity grade. In both daytime and evening, the velocity shows a rather slow monitonic decrease, with decreasing ground conductivity, down to a conductivity of 10^{-4} mho/m. Further decreasing ground conductivity from 10^{-4} to 10^{-5} mho/m means a rapid increase in velocity is predicted. Three parameters are used to represent phase shift upon excitation or de-excitation of earth-ionosphere waveguide. One term represents a nominal phase shift on excitation, which is always expected, while a second represents additional phase shift expected at various bearings. The second term is also sensitive to omagnetic dip angle. A final excitation term allows for additional phase shift when the excitation occurs over regions of extremely low ground conductivity. There is a substantial variation of velocity with geomagnetic dip angle. Additionally, five parameters are used to represent a complex variation of velocity with both azimuth and geomagnetic dip angle. Auroral terms represent an abrupt increase in velocity for propagation under the auroral zone (beginning just above 60° absolute geomagnetic latitude and increasing to a maximum variation from nominal at about 66°). Anomalous velocity variation is also found to occur throughout the polar cap. Velocity variation with sunspot number is included in regression results. As previously noted, prediction coefficients for day and night conditions are developed separately.

3.3.5 Diurnal Variation

A key element of the phase prediction theory used with Omega is the notation of a Diurnal Function (figure 21) (Swanson and Bradford, 1971). This is defined to operate within the range 0 to 1 where 0 corresponds to idealized, normally illuminated day conditions, and 1 represents stable night conditions. It may be thought of as an ionospheric forcing function defined by the solar zenith angle. Given sufficient time for the ionosphere to stabilize, the phase delay across any path segment, at any given time, can be determined from the diurnal function as follows: First, the solar zenith angle, cosine of the solar zenith angle, and diurnal function are determined for the path segment. Secondly, the phase delays expected during normally illuminated day or at night are computed from spatial theory. Finally, the intermediate delay is computed by weighting illumination conditions by the diurnal function. Delay over a long path can be computed by summing incremental delays over the various path segments.

In practice, there are dissociation and recombination rates associated with the ionosphere and, hence, the ionosphere possesses an effective time constant at any given height. As the solar zenith angle is usually changing, stable illumination conditions rarely occur. More generally, the ionosphere approximates a system, with an effective time constant being forced by the diurnal function. The effective time constant ranges from 1/8th hour during the day to 1-1/8th hours at night. During much of the 24-hour day, when the diurnal function is slowly varying, the historical influence in determining effective ionospheric phase height is slight, and the diurnal function alone is a good predictor of relative, temporal phase variation.

At sunset, and especially at sunrise, anomalous phase changes occur. The abrupt phase shift at sunrise occurs with the first incidence of weak light which has already passed through the lower layers of the atmosphere. Hence we speculatively envision a process such as photo-detachment. Anomalous transitional shifts are computed with the aid of "dump schedules."

The diurnal function explains not only diurnal behavior near mid-day, but also seasonal variation. However, there is a slight relationship between the slope of the diurnal function at the zenith and the phase of the solar cycle as represented by the sunspot number.

3.3.6 Prediction Accuracy

Spatial prediction biases of long-term averages from global prediction theory were evaluated by Swanson and Kugel (1972b), and found to be typically 3.2 cec, both day and night. At the time of the evaluation the existing data base heavily reflected observations in North America, Hawaii, and Western Europe, based on transmissions from Hawaii; Forestport, New York; Panama; Trinidad, the United Kingdom, and Norway. New data, not yet compiled and analyzed, will provide a better sampling of predictions over a wider range of geophysical conditions. Some, at least temporary, decrease in predictional accuracy should be expected for conditions substantially different from those used in determining prediction coefficients. One region of exceptionally large prediction errors is in prediction of the phase of the North Dakota signal in Europe and the Mediterranean (Swanson, 1976). Prediction errors of about 18 cec, or five standard deviations, are prevalent. It has been speculated that the biases may be due to propagational effects of a stratified, irregular ground in central Canada.

Predictional accuracy during transitions is difficult to assess. Whereas conditions are relatively stable during night or near mid-day so long-term averages can confidently be developed and readily associated with given illumination conditions, no similar situation exists during transitions.

One method of assessing predictional accuracy during transitions is to consider the rms variation of individual hourly phase differences from those predicted. This technique then reflects not only predictional biases, but also day to day scatter. One such

analysis was conducted by Calvo and Bortz (1974) who noted a typical 24-hour rms discrepancy from predictions based on global theory of 12.5 centicycles (one mile on a baseline). (They also evaluated line-of-position rms discrepancy after application of the statistical force fit adjustments, and found an error on the order of 0.6 nmi.) As typical repeatability of the phase of a signal over a long path was given in table II, and also determined by Swanson and Kugel (1972a) as on the order of 3-4 centicycles, one would expect an rms discrepancy on a typical hyperbolic line-of-position to be the root-sum-of-squares combination of predictability over both paths forming the hyperbolic pair combined with repeatability over the same paths, or about 7 cec either day or night. Predictions during transitions are not expected to be as poor as would be required to explain the difference in expectation in the rms computed by Calvo and Bortz over the 24-hour day. However, in either case, general rms accuracy is expected to be in the range 7 to 12.5 cec. Prediction biasis are clearly less.

4. NAVIGATIONAL PERFORMANCE

For practical purposes it is convenient to divide consideration of performance into consideration of typical median navigational accuracy and blunders.

Nominal repeatability is meant to include statistical variations through perhaps the 95th percentile, and is well represented by the standard deviation or circular error probable. It is directly applicable to typical operation and nominal wander about intended track. Blunders may be induced by the equipment, the operator, or sudden anomalous propagation conditions. Blunders relate primarily to navigational safety. They are best expressed as the probability of some specified large error, as might be associated with navigational disaster.

4.1 Median Results

Typical performance depends on navigational geometry, but fundamentally on signal properties, including measurability and predictability, as well as repeatability. Signal-to-noise ratios are typically adequate to allow accurate, timely measurement. Many equipments have adequate accuracy. Typical predictability was discussed in section 3.3.6 whereas signal repeatabilities were given in table II.

Owing to the rather recent advent of signals from eight stations on a worldwide basis, a modern evaluation of Omega accuracy on a global basis has not been done. As previously noted, there is a known serious bias in the predictability of North Dakota signals in Europe and the Mediterranean. Other biases may also be expected. If these biases can be removed, as was done with predictional errors in the earlier developmental system, then an accuracy assessment of the earlier system geometry may be indicative of eventual accuracy. An evaluation to determine potential operational accuracy based on fixed site monitoring was conducted in 1969 (Swanson, 1969). Observations were systematically sampled at various hours and various seasons at different sites so as to estimate year-round navigational accuracy. Results from sampling observations at San Diego, California and Rome, New York over a 2-year period are shown in figure 22. Figure 23 shows similar data for May 1969 from San Diego, New York, Miama, Bermuda, and Wales, Alaska. Results from both samplings indicate a circular error probable of one mile. This is the single most useful estimate of Omega accuracy, that is, that applicable over a wide area 24-hours per day.

Accuracy may be better or worse than nominal for various special applications. Clearly for any specified local area or particular period, predictive accuracy should be at least no worse than nominal. Removal of prediction bias through extensive monitoring and "force fit" was already noted as useful especially for high or critical traffic areas.

4.2 Disturbed Conditions

Since the human and instrumental blunder modes can be addressed through design changes or changes in procedures, the blunders of fundamental concern are those associated with propagation. Experimental verification of highly usual deviations is extremely difficult without a causative model, as there is high likelihood that extremely anomalous data variations are due to errors in the reduction rather than operation, e.g., recording of 38 instead of 83. Some care must also be given to the relation between predictability of various events and blunders. Events which are not now predictable may become predictable in the future. Even if predictable, blunders will occur if predictive methods are not implemented. An example is that of the effects of eclipse on vlf. This effect has been observed by various workers; subsequently predicted by Noonkester and Sailors (1971). Although the occurence of eclipse is highly predictable, no effort is made to incorporate such knowledge into practical prediction, and thus eclipses although rare, are an occasional source of significant error.

Because of the aforementioned difficulty of confidently insuring against anomalous measurement or data processing errors, blunders are best addressed by discussing their geophysical causation. Nominal phase changes due to ordinary random geophysical variations are not associated with particular geophysical fluctuations, but are statistically noted in determining typical navigational performance. Occasional severe variations are addressed through geophysical causation. By intent, this paper addresses primarily nominal propagation conditions, and is weak in the geophysical morphology of various events capable of producing significant errors. A companion paper by Larsen (1976) addresses morphology of errors in detail. However, two effects, SIDs and PCAs, are of sufficient importance that they must be noted here. Larsen (1976) presents the morphology and fundamental data base for determining the effects of SIDs and PCAs. The fundamental data cited by Larsen is valid for determining SID activity over long sunlit paths and PCA activity over long transpolar and transauroral paths near maximum solar activity. It has been speculatively extrapolated by Swanson (1974) to more typical conditions (figure 24). The renormalization of the function representing SIDs includes extrapolation to typical conditions rather than near-noon, 24-hour operation and allowance for nominal hyperbolic cancellation of phase advances. PCA renormalization is similar except that reduction over the 24-hour day is not applicable. PCA renormalization to nominal solar activity is speculative since there is an insufficient data base near solar minimum.

Figure 24 shows that about 2 percent of the time the nominal scatter expected from a normal statistical distribution will be exceeded due to the effect of SIDs and PCAs. The difference is especially important when assessing the probability of some unusual event such as a navigational error greater than 5 miles. There is virtually no probability of an error greater than 5 miles from the normal distribution (an excursion of over 10 standard deviations). However, figure 24 shows that the probability of an error that large, or greater, due to an SID is 0.02 percent, and from a PCA, 0.2 percent. That is, a PCA could induce an error of that magnitude on one out of 500 occasions.

4.3 Anomalous Diurnal Phase Change

Several cases may be defined when two or more waveguide propagation modes are present: I - a highly dominant mode with slight contribution from a second mode; II - two nearly equal modes with one remaining dominant; and III - a temporal or spatial change in

modal dominance. For illustration, only two modes will be considered. In Case I with the competing mode very small, the ideal propagation condition of a single dominant mode is closely approximated. There will be a slight nonlinearity in the variation of phase with distance wherein the error is quasi-sinusoidal, with a wavelength determined by the difference in the phase velocities of the two modes. Case II with the competing mode almost equal yields a phase error as a function of range which is basically similar to that developed in Case I, only larger. That is, there is an irregularity in phase as a function of distance analogous to the Hollingsworth pattern observed in amplitude as a function of distance. The maximum phase deviation results from the resultant signal being pulled in quadrature from the dominant mode, and is thus 90° or 25 cec, corresponding in hyperbolic geometry to perhaps 2 nmi anomalous position variation. Gallenberger has computed anomalous phase variation for Cases I and II (Swanson, 1971b). The effect can usually be considered an incidental, although sometimes moderately large, source of error, but not a blunder mode. However, the effect may be of importance in lane resolution, and has been studied extensively by Burgess (1969) for this reason. Dicke (1970) has also made extensive measurements of Hollingsworth patterns to determine propagation parameters. It is Case III when the modal dominance changes that special care is needed. Diurnal changes of mode dominance may occur on numerous paths at the vlf communications frequencies, and can also occur at the Omega frequencies in the case of certain westbound equatorial or trans-equatorial paths. Signals which may undergo changes in modal dominance are best considered unusable, but can be used under special circumstances (Swanson, 1975a).

When modal dominance changes between day and night, it is self-evident that at some time the two modes must be of equal amplitude. If the signals are in phase at equality, an arithmatic addition of field strength would occur and there would be no anomalous phase shift; if out of phase, the signals would cancel. Generally, some intermediate condition will prevail. Depending, in detail, on the relative phasing at equality, the spatial or diurnal transition may be normal or abnormal. In particular, a sunrise which might normally be associated with a one cycle phase change may actually show no phase change, or a two-cycle variation. The navigational effect of this cycle slippage is an anomalous position displacement into an adjacent "8-mile" navigational lane without loss of phase tracking.

Cycle slippage has been discussed by Swanson (1972) using the phasor sum from field contributions by two modes. Diurnal variation of phase and amplitude are represented by the phasor loci shown in figure 25.

Figure 25a corresponds to a first mode over a long path which experiences a one-cycle phase variation during a transition while the amplitude varies slightly. Figure 25b shows the corresponding second mode varying by two cycles while the amplitude would vary from very low during the day (but exaggerated in the illustration) to substantial at night. The reference cited shows the phasor sum for realistic parameters (figure 26). Two types of transitions are shown, depending on the relative phase between the two modes during the day. For a relative phasing of three radians, the transition tends to ramp, largely as expected, from a single mode alone, but with a marked minor anomaly in mid-transition, and some phase pulling from one cycle at night. For a relative phasing of two radians, however, there is an anomalous cycle slip where the transition gains an extra cycle. As can be seen from the accompanying phasor loci, whether or not the cycle slip occurs depends upon whether or not the "loop" in the composite locus encompasses the origin. This, in turn, is critically dependent on the absolute phasing between the two propagation modes during the day (or equivalently at night in this model). The absolute phasing over long paths is, however, variable as a result of nominal ionospheric changes. Thus, cycle slippage may or may not be a repeatable feature from day to day.

Although a severe fade will often accompany cycle slippage, the danger is that adequate signal may remain for continuous tracking. In this case a large positional error could occur without the user being aware. The best safeguard is respect of coverage restrictions to areas of single mode dominance.

4.4 Coverage Limitations

The importance of coverage considerations to avoid cycle slippage was mentioned in the preceding section. In addition, if some significant contribution is being made by other than the dominant mode, it is necessary to determine that the anomalous phase variation with distance be sufficiently small for the intended use. For example, carrier perturbations amounting to a mile or so can be accepted in normal track navigation, but not for lane resolution. Significant limits are imposed by self-interference wherein a signal propagated around the earth on the long-path will interfere with that obtained over the short path. The aforementioned limitations all apply even though associated signals may of substantial amplitude and apparent phase stability. Another coverage restriction is that the signals present adequate signal-to-noise ratio for accurate, timely measurement.

Omega coverage is only recently beginning to receive the attention the subject warrants. Although the full wave theory necessary for assessment became available in the mid-1960s, the first scientifically documented full wave work was that of Sailors (1970) in 1970. Most work has been within the last two years. Swanson and Dick (1975b) published a limited assessment for North America and the North Atlantic. More recently, Swanson (1975c) examined coverage in India in considerable detail by computing both long and short path signals from all Omega transmitters, and thus contributing to the available inventory of computed Omega radials. The most substantial effort to determine Omega coverage was, however, that of Bortz, Gupta, Scull and Morris (1974). They employed a modification of the Naval Electronic Laboratory Center (NELC)-developed Integrated Prediction Program (IPP), described by Ferguson (1972), for the full wave computations. Supporting radials are given in a report by Gupta (1974). The work of Bortz, et al included estimation of prevailing noise conditions, but is conservative in that no allowance was made for effective improvement in signal-to-noise ratio through nonlinear processing techniques, while an improvement of 15 dB may be expected. However, the results are comprehensive and are reproduced as figures 27 and 28. Note that signal coverage is given wothout addressing the influence of navigational geometry.

5. CONCLUSIONS

Omega propagation has been reviewed. Primary attention has been devoted to a summary of the Omega system and development of nominal propagation conditions based primarily on waveguide notions. Mechanisms for progatationally induced navigational blunders have been noted. However, detailed discussion of the morphology of propagational events has been avoided in this paper. A companion paper by Larsen describes these events in detail.

BIBLIOGRAPHIC COMMENT

Many review papers are developed through combining and revising past works. This is no exception. Some comment on foundation may be useful. This paper has a substantial resemblance to Technical Note 2101 (Swanson and Brown, 1972) but without the detailed attention to phase prediction. This is less due to direct abstraction from the note than to common origins. Both are based heavily on updates of lectures given at the University of Michigan in 1967 and 1968, particularly the 1969 revision. Also extensive use was made of a

paper on phase prediction (Swanson, 1971b) and a paper on Omega (Swanson, 1971c). An additional reference, not cited explicitly, is Burgess and Jones, 1975, which is a review of LF and VLF propagation, and can profitably be read in conjunction with this work. No attempt was made to make the references exhaustive. Rather, recent work by active authors has been cited, as well as certain fundamental works. Finally, apology is given for extensive reference to work by Swanson. However, as Jr. J. R. Wait noted in an aside in a review paper of his at a meeting in Spaatind, I am most familiar with my own work.

REFERENCES

Berry, L.A. and Herman, J.E., 1971, A Wave Hop Propagation Program for an Anisotropic Ionosphere, OT/ITS Res. Rept. 11, April 1971.

Bickel, J.E., Ferguson, J.A. and Stanley, G.V., 1970, "Experimental Observations of Magnetic Field Effects on VLF Propagation at Night," Radio Sci., v.5., pp 1127-1136, August-September 1970.

Bortz, J.E., Sr., Gupta, R.R., Scull, D.C., and Morris, P.B., 1974, "Omega-Signal Coverage Prediction," Proc. 2nd Omega Symposium, 4-7 November 1974, Washington, D.C., pp. 1-8

Bradford, W.R., 1971, Fresnel Zone Theory of VLF Coupling into the Earth-Ionosphere Waveguide, NELC TN 1896, 22 July 1971.(*)

Brown, R.G. and Van Allen, R.L., 1976 "Three Frequency Difference Omega," paper presented at the Air Navigation meeting of the U.S. Inst. of Navigation, Warminster, Pa., April 17-28, 1976.

Burgess. b., 1969, "Propagation effects and lane-ambiguity resolution in Omega," Proc Instn Elect Engrs, v 116, 8, pp 1297-1303.

Burgess, B., and Jones, T.B., 1975, "The Propagation of l.f. and v.l.f. Radio Waves with Reference to Some Systems Applications," Radio & Elect. Engr., v. 45, n. 1/2, January/February 1975, pp 47-61.

Burgess, B. and Walker, D., 1970, "Effects in Omega from Propagation Variations," J. Inst. of Navigation (Br.), 23, 1, pp 49-59.

Calvo, A.B. and Bortz, J.E., Sr., 1974, "Evaluating the Accuracy of Omega Predicted Propagation Corrections," NAVIGATION, v. 21., 2, Summer 1974, pp 102-112.

Chi, A.R. and Wardrip, S.C., 1974, "Recent Field Test Results Using OMEGA Transmissions for Clock Synchronization," Proc. 6th Annual Precise Time and Time Interval (PTTI) Planning Meeting, 3-5 December 1974, NASA/GSFC Report X-814-75-117, pp. 187-197.

Crawford, W.R., and Rupp, W.E., Jr., 1972, "Locating Downed Aircraft by GRAN (Global Rescue Alarm Net)," Navigation (US), 19, 4, Winter 1972-3, pp 311-316.

Crombie, D.D., 1964, "Periodic Fading of VLF Signals Received Over Long Paths During Sunrise and Sunset," J. Res. Nat. Bur. Stand. Sect. D. 68D(1), pp. 27-34.

Crombie, D.D., 1966, "Further Observations of Sunrise and Sunset Fading at Very-Low-Frequency Signals," Radio Sci., 1(1), pp. 47-51.

Ferguson, J.A., 1972, An Improved Integrated Prediction Program, Naval Electronics Laboratory Center Technical Note 2231, 11 December 1972.(*)

Galejs, J., 1972, Terrestrial Propagation of Long Electromagnetic Waves. (Pergamon Press), 362 pp.

Gupta, R.R., 1974, Mode Interference Studies: Vol. I, Transmitter-Excited Interference, The Analytic Sciences Corporation, TR-343-10, 29 November 1974.

Hastings, C.E. and Comstock, A.L., 1975, "The Optimum Implementation of Differential Omega," Navigation (U.S.), 22, 1, pp 76-80.

Herbert, N.F., and Nolan, T.P., 1975, "The Current Status of the OMEGA Navigation System," Paper presented at the OMEGA Symposium, Tokyo, June 1975 (Abridgement published in Navigation [Japanese], Oct. 1975 n. 47, pp 55-56 [in Japanese]).

Johler, J.R., 1970, "Spherical Wave Theory for MF, LF, and VLF Propagation," Radio Sci., v 5, pp 1429-1443, December 1970.

Jones, S.S.D., 1970, "VLF Techniques for Navigation," Journal of the (British) Institute of Navigation, v. 23, n. 1, January 1970. p. 23.

Kasper, J.F. Jr., 1970, "A Skywave Correction Adjustment Procedure for Improved Omega Accuracy," Proc. of the ION Nat'l Marine Meeting, U.S. Coast Guard Academy, New London, Conn., 12-14 October 1970, v. II, pp. 17-30

Kugel, C.P., 1975, Propagation Effects on the Operation of the Omega Dual-frequency Timing Receiver, NELC TN 3046, 23 September 1975.(*)

Larsen, T.R., 1976, these proceedings.

Laughlin, C., Hilton, G., Hollenbaugh, R. and Lavigne, R., 1965, Meteorological Experiment Using the Omega System for Position Location, NASA GSFC Rept. X-713-65-416, October 1965.

Laughlin, C.R., et al, 1967, PLACE Experiment Description, NASA GSFC Rept. X-133-67-577, November 1967.

Litchford, G., 1971, "Making General Aviation Safer and More Effective through Universal Electronic Design," Astro. and Aero., January 1971, p. 36.

Lynn, K.J.W., 1971, "Frequency Dependence of vlf Modal Interference Effects Observed on East-West Propagation Paths," JATP, 33 (No. 6), pp. 951-958.

Lynn, K.J.W., 1973, "VLF Mode Conversion Observed at Middle Latitudes," JATP, 35(3), pp 439-452.

Mactaggart, D., 1971, "Design and Performance of the CMA 719 Computerized Airborne Omega Receiver," Proc. 1st Omega Symposium, 9-11 November 1971, Washington, D.C., pp 63-74.

Maxwell, E.L. and Stone, D.L., 1965, 10 kc/s Atmospheric Noise Prediction, DECO Rept. 54-F2.

MITRE, 1976, Determination of Effective Ionospheric Electron Density Profiles for VLF/LF Propagation, MITRE Technical Report Defense Communications Agency, 1 January 1976, C650-TP-76-4.

Morfitt, D.G., 1972, Numerical Comparison of the Mode and Wave-hop Theories of Radio Wave Propagation at VLF and LF, Naval Electronics Laboratory Center Technical Report 2118, 26 August 1972.

Morris, P.B., and Cha, M.Y., 1974 Omega Propagation Corrections: Background and Computational Algorithm, Dept. of Transportation, U.S. Coast Guard, Omega Navigation System Operations Detail, Rept. ONSOD-01-74. December 1974. 226 p.

Nard, G., 1974, "State of Experimentations and Program of Development of Differential Omega in France," Proc. 2nd Omega Symposium, 5-7 November 1974, Washington, D.C., pp. 96-108.

Palmer, W., 1972, Technical Considerations Regarding the GNS System, Winslow Palmer P.E., 114 Montrose Dr., Ft. Myres, Florida 33901.

Papousek, W. and Reder, F.H., 1973, "A Modified Composite Wave Technique for OMEGA," NAVIGATION, 20, 2. pp 171-177.

Pappert, R.A., Gossard, E.E., and Rothmuller, I.J., 1967, "A Numerical Investigation of Classical Approximations Used in VLF Propagation," Radio Sci., v. 2, pp. 387-400, April 1967.

Pappert, R.A., and Morfitt, D.G., 1975, "Theoretical and Experimental Sunrise Mode Conversion Results at VLF," Radio Sci., v. 10 n. 5, May 1975, pp 537-546.

Paulson, M.R., Gossard, E.E., and Moler, W.F., 1962, "The Nature and Scale Size of Irregularities in the D-Region of the Ionosphere as Observed on a Near Vertical Incidence VLF Sounder," ch. 6 in Propagation of Radio Waves at Frequencies below 300 kc/s, W.T. Blackband, ed., NATO/AGARD Avionics Panel, Ionospheric Research Committee, 1962.

Pierce, J.A., 1968, The Use of Composite Frequencies at Very Low Radio Frequencies, Division of Engr. and Applied Physics, Harvard University, Technical Rept. 552, February 1968.

Pierce, J.A., 1974, Omega: Facts, Hopes, and Dreams, Division of Engr. and Applied Physics, Harvard University, Technical Rept. No. 652, June 1974. 154 pp.

Pierce, J.A., Palmer, W., Watt, A.D., and Woodward, R.H., 1966, Omega: A World-wide Navigational System – System Specifications and Implementation. Published by Pickard and Burns Electronics for the Omega Implementation Committee. (Pickard and Burns Pub. No. 886B). 2nd ed., May 1966.

Sailors, D.B., 1970, An Assessment of Omega Signals at NASA Tracking Stations, Naval Electronics Laboratory Center Technical Note 1777, 4 December 1970.(*)

Sakran, F.C., Jr., 1974, "Accuracy Specifications for Automatic Omega Navigators," Proc. 2nd Omega Symposium, 5-7 November 1974., Washington, D.C., pp 9-20.

Shellman, C.H., 1970, "Electron Density Distributions in the Lower Ionosphere with Associated Error Limits Derived from VLF and LF Sounder Data," Radio Sci., v. 5, pp 1127-1136, August-September 1970.

Smith, E.J., 1971, "AN/ARN 99 Method of Removing Omega Lane Ambiguity," Proc. 1st Omega Symposium, 9-11 November 1971, Washington, D.C., pp 75-82.

Smith, G.H., and Pitteway, M.L.V., 1974, "FORTRAN Program for Obtaining Wavefields of Penetrating, Nonpenetrating, and Whistler modes of Radio Waves in the Ionosphere," ELF-VLF Radio Wave Propagation (Being the Proc. of the NATO Adv. Study Inst. held at Spaatind, Norway, Apr 17-27, 1974), J.A. Holtet, ed., (1974: D. Reidel, Dordrecht, Holland), pp 69-86.

Snyder, F.P., 1968, Effect of Magnetic Dip-Angle and Azimuth-Angle Variations on Mode Numbering at VLF, Naval Electronics Laboratory Center Report 1587, 7 October 1968.

Snyder, F.P., and Pappert, R.A., 1969, "A Parametric Study of VLF Modes Below Anisotropic Ionospheres," Radio Sci., New Series, v. 4, p 213-226, March 1969.

Stringer, F.S., 1970, "A Hybrid System for World-Wide Navigation," Journal of the (British) Institute of Navigation, v. 23, n. 1 p. 26, January 1970.

Swanson, E.R., 1965, Omega Lane Resolution: Phase Measurements at Two Very Low Frequencies, Naval Electronics Laboratory Center Report 1305, 5 August 1965.

Swanson, E.R., 1969, Accuracy of Omega Navigation System Using 1969 Skywave Corrections, 31 December 1969 (AD704504).

Swanson, E.R., 1970, "Time Dissemination Effects Caused by Instabilities in the Medium," Phase and Frequency Instabilities in Electromagnetic Wave Propagation, K. Davies, ld. (Proc. AGARD/EPC Symposium, Ankara, 9-12 October 1967), pp 181-198.

Swanson, E.R., 1971a, "Application of Omega to Aircraft Navigation and Traffic Control," Journal of the (British) Institute of Navigation, v. 24, n. 1, January 1971, pp 125-128.

Swanson, E.R., 1971b, "VLF Phase Prediction," VLF-Propagation, G. Bjøntegaard, ed., (Proc. from the VLF-Symposium, Sandefjord, Norway, 27-29 October 1971). Norwegian Institute of Cosmic Physics Report 7201, January 1972.

Swanson, E.R. 1971c, "Omega," NAVIGATION, v. 18, 2, pp. 168-175, Summer 1971.

Swanson, E.R., 1972, "Use of Propagation Corrections for VLF Timing," Proc. 4th Ann. DOD Precise Time Interval (PTTI) Strat. Plan. Mtg. 14-16 November 1972, pp 310-323.

Swanson, E.R., 1974, "ELF-VLF Applications in Navigation and Communications," ELV-VLF Radio Wave Propagation, J.A. Holtet, ED" (Proc. of the NATO Advanced Study Institute held at Spaatind, Norway, April 17-27, 1974). (D. Reidel: Dordrecht, Holland), pp 371-384.

Swanson, E.R., 1975, VLF Navigation Monitor: Purpose and Function, Naval Electronics Laboratory Center Technical Report 1955, 30 May 1975, 15 pp. (Also designated as Dept. of Transportation Report FA74WAI-425-4). (AD A013 651).

Swanson, E.R., 1975b, "OMEGA Propagation," Paper presented at the Omega Symposium, Tokyo, 12-13 June 1975. Abridgement published in: (Japanese) Navigation, October 1975, n. 47, pp 57-63 (in Japanese).

Swanson, E.R., 1975c, Omega Coverage in India: A Case Study, Naval Electronics Laboratory Center Technical Report 1974, November 1975 (AD A 022 035).

Swanson, E.R., 1975, Omega Prediction Errors in the Mediterranean, Naval Electronics Laboratory Center Technical Note in press(*).

Swanson, E.R., and Adrian, D.J., 1973, Omega Envelope Capability for Lane Resolution and Timing, Naval Electronic Laboratory Center Technical Report 1901, 20 November 1973 (AD 774 891).

Swanson, E.R., Adrian, D.J., and Lester, P.H., 1974, "Differential Omega Navigation for the U.S. Coastal Confluence Region," NAVIGATION, v. 21, n. 3, Fall 1974, pp 264-271.

Swanson, E.R. and Bradford, W.R., 1971, Diurnal Phase Variation at 10.2 kHz, Naval Electronics Laboratory Center Technical Report 1781, 11 August 1971 (AD 737 212).

Swanson, E.R., Britt, T.F., Wilson, J.J., and Hanselman, J.C., 1975, Omega Format Optimization, Naval Electronics Laboratory Center Technical Report 1966, 8 October 1975, 28 pp. (AD A019 364).

Swanson, E.R., and Brown, R.P., 1972, Omega Propagation Prediction Primer, Naval Electronics Laboratory Technical Note 2101, 3 August 1972(*).

Swanson, E.R., and Dick, M.J., VLF Airborne Navigation Requirements, Naval Electronics Laboratory Center Technical Report 1941, 4 February 1975, 22 p. (Also designated as Dept. of Transportation Report FA74-WAI-425-1). (AD A012 581).

Swanson, E.R. and Dick, M.J., 1975b, Propagational Assessment of VLF Navigation signals in North America and the North Atlantic, Naval Electronics Laboratory Center Technical Report 1944, 4 February 1975. (Also designated as Dept. of Transportation Report FA74-WAI-425-2). (AD A012581).

Swanson, E.R. and Hepperley, 1969, Composite Omega, Naval Electronics Laboratory Center Report 1657, 23 October 1969. (AD 863 791).

Swanson, E.R., and Kugel, C.P., 1972a, "VLF Timing: Conventional and Modern Techniques Including Omega," Proc. IELE, v. 60, n. 5., May 1972. pp. 540-551.

Swanson, E.R. and Kugel, C.P., 1972b, Omega VLF Timing, Naval Electronics Laboratory Center Technical Report 1740 (Rev.), 29 June 1972. (AD 743529).

Swanson, E.R., and Robie, F.C., VLF Navigation, Naval Electronics Laboratory Center Technical Document 233, 1 January 1973. (AD 761498).

Tymczyszyn, W.R., 1975, "Inflight Worldwide VLF Experience Using Global Navigation GNS-500," Proc. 7th Ann. Precise Time and Time Interval (PTTI) Applications and Planning Meeting, 2-4 December 1975. NASA Publ. No. X-814-76-45. pp 521-548.

Vannicola, V., 1976 (Personal communication, 15 June 1976).

Wait, J.R., 1970, "Electromagnetic Waves in Stratified Media," v. 3 of International Series of Monographs on Electromagnetic Waves. 2nd ed., (1970: Pergamon Press).

Wait, J.R. and Spies, K.P., 1964, Characteristics of the Earth-Ionosphere Waveguide for VLF Radio Waves, National Bureau of Standards Technical Note 300, December 1964.

Wait, J.R., and Spies, K.P., 1965, "Influence of Finite Ground Conductivity on the Propagation of VLF Radio Waves," Radio Sci J. Res. NFS, 68D, n. 10, October 1965.

Watt, A.D., 1967, VLF Radio Engineering, v. 14 of International Series of Monographs in Electromagnetic Waves, (1967:Pergamon Press).

Wilson, J.J., Britt, J.E., and Chi, A.R., "OMEGA Timing Receiver, Design and System Test," Proc. 4th Ann. Precise Time and Time Interval (PTTI) Planning Meeting, 14-16 November 1972, NASA/GFSC Publ. No. X-814-73-72, pp 345-361.

Note: Proceedings of the first and second Omega sympsia are available from: The Institute of Navigation, Suite 832, 815 15th St., N.W., Washington, D.C. 20005.

*Naval Electronics Laboratory Center Technical Notes and Memoranda are informal documents intended primarily for use within the Center.

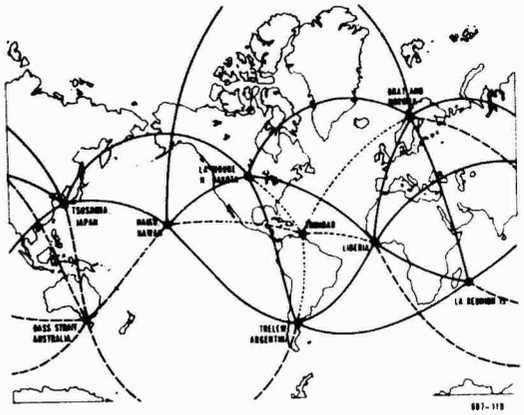


Figure 1. Omega station locations.

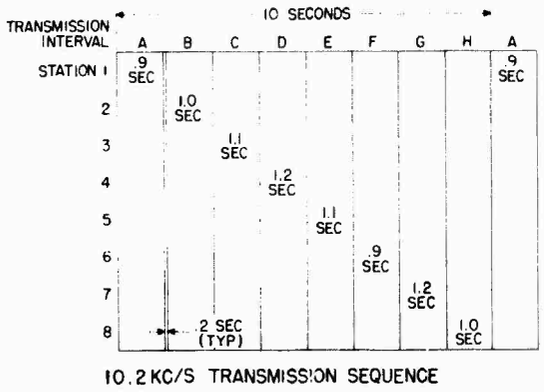


Figure 2. Simplified 10 2 kHz transmission format.

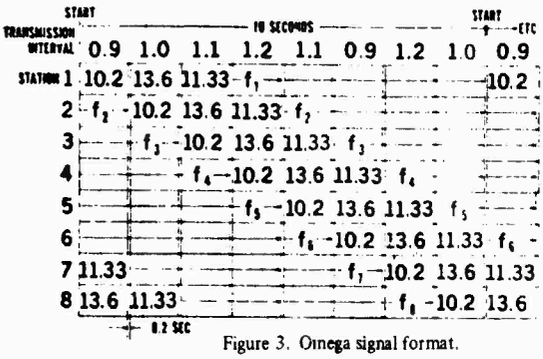


Figure 3. Omega signal format.

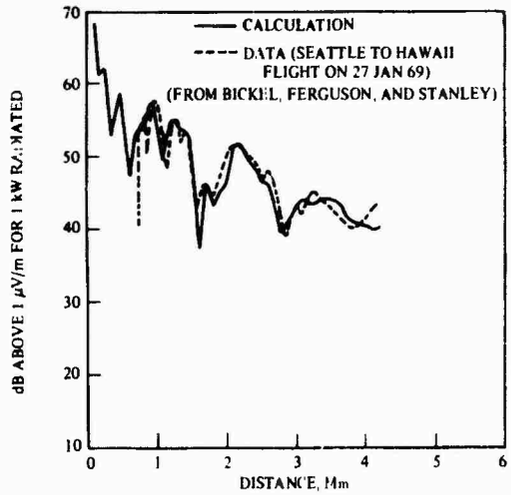


Figure 4. Comparison of predicted and observed amplitude at night, 23.4 kHz (from Bickel, Ferguson, and Stanley, 1970).

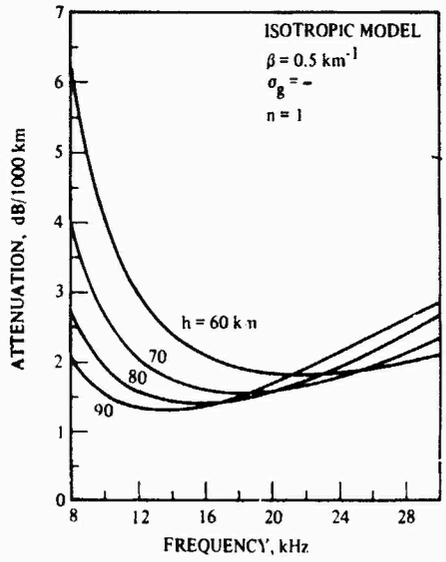


Figure 5. Theoretical attenuation rates for first mode (from Wait and Spies, 1964).

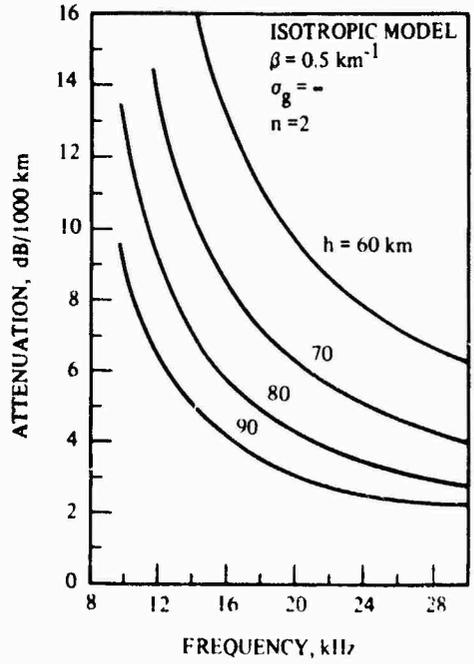


Figure 6. Theoretical attenuation rates for second mode (from Wait and Spies, 1964).

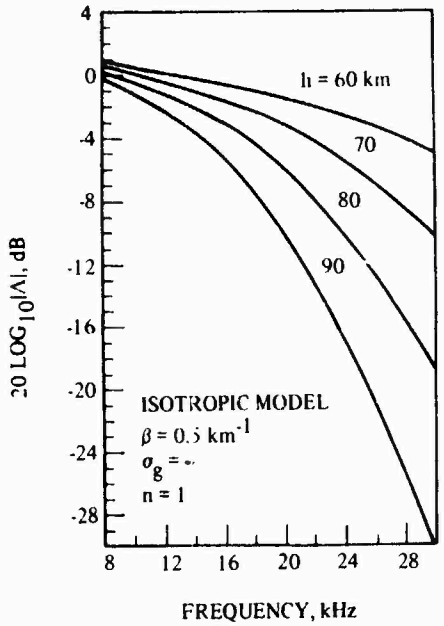


Figure 7. Theoretical excitation factors for first mode (from Wait and Spies, 1964.

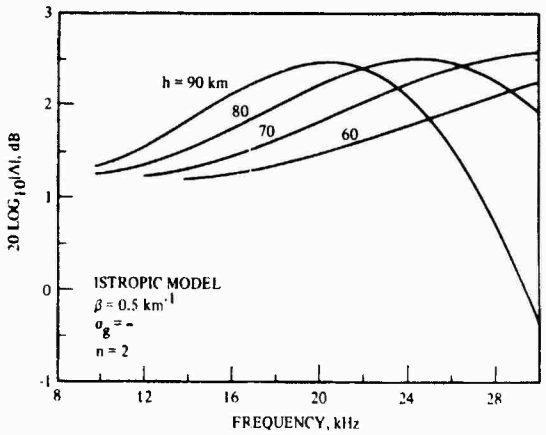


Figure 8. Theoretical excitation factors for second mode (from Wait and Spies, 1964).

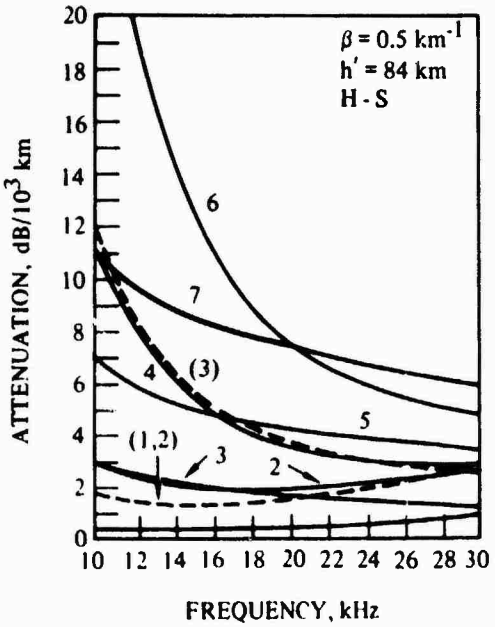


Figure 9. Attenuation versus frequency. Hawaii-to-San Diego path, solid curves. Isotropic ionosphere, dashed curves (from Snyder and Pappert, 1969).

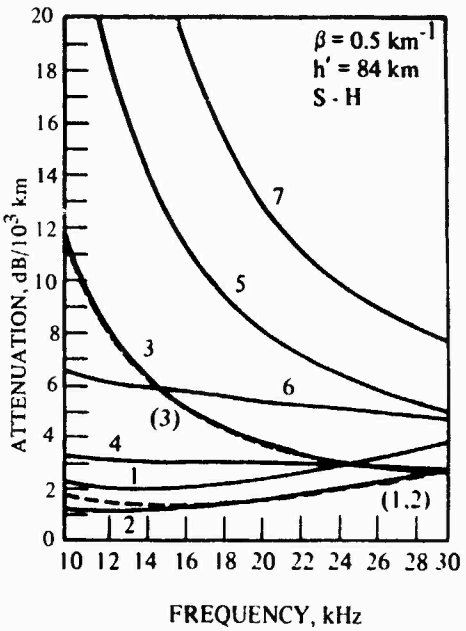


Figure 10. Attenuation versus frequency. San Diego-to-Hawaii path, solid curves. Isotropic ionosphere, dashed curves (from Snyder and Pappert, 1969).

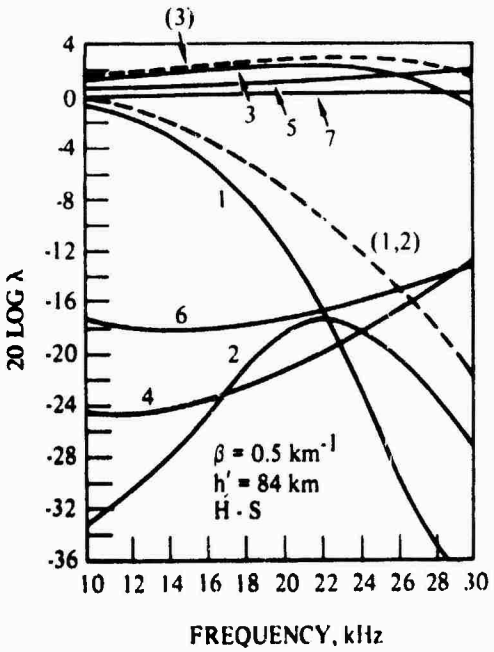


Figure 11. Magnitude of the excitation factor in dB versus frequency. Hawaii-to-San Diego path, solid curves. Isotropic ionosphere, dashed curves (from Snyder and Pattert, 1969).

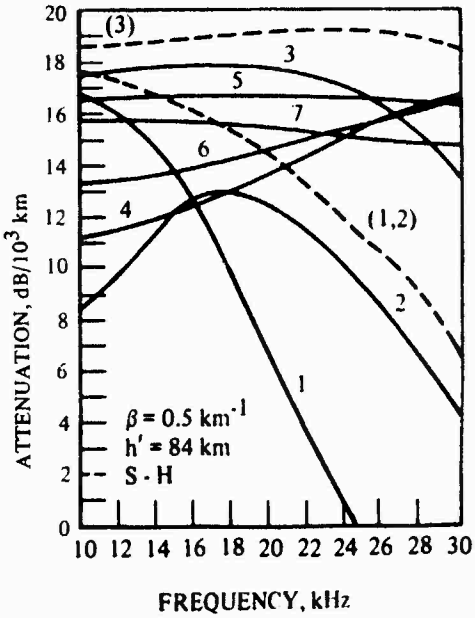


Figure 12. Magnitude of the excitation factor in dB versus frequency. San Diego-to-Hawaii, solid curves. Isotropic ionosphere, dashed curves (from Snyder and Pappert, 1969).

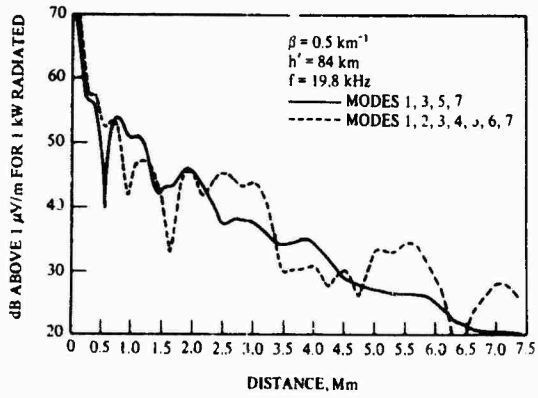


Figure 13. Amplitude of the vertical component of the electric field at the ground in dB above 1 μv/m for 1 kilowatt of radiated power versus distance for the SH path (from Snyder and Pappert, 1969).

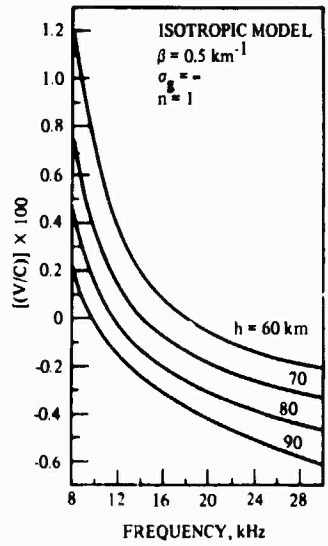


Figure 14. Theoretical phase velocities (from Wait and Spies, 1964).

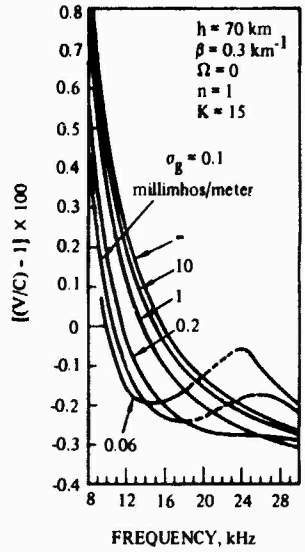


Figure 15. Phase velocity for a wide range of ground conductivities (from Wait and Spies, 1964).

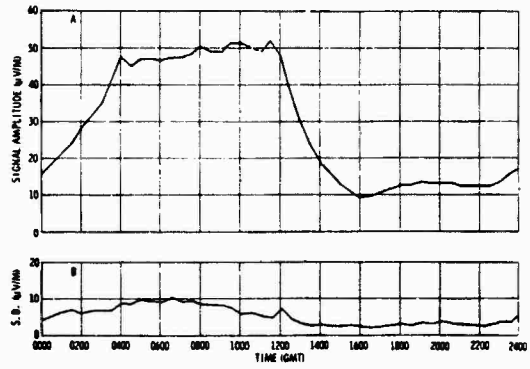


Figure 16. Plot of average field strength (A) and corresponding standard deviation (B) for 10.2 kHz signal recieved at Rome, N.Y. from Haiku, Hawaii, 24 October – 10 November 1962.

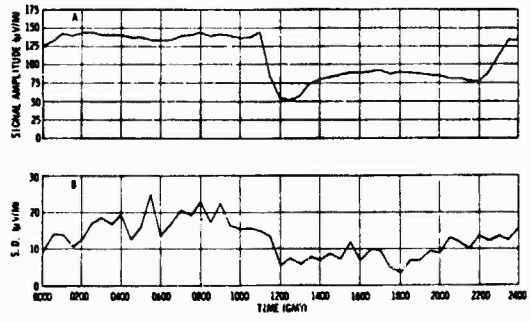


Figure 17. Plot of average field strength (A) and corresponding standard deviation (B) for 10.2 kHz signal received at Farfan, C.Z., from Forestport, N.Y., 27 November – 21 December 1962.

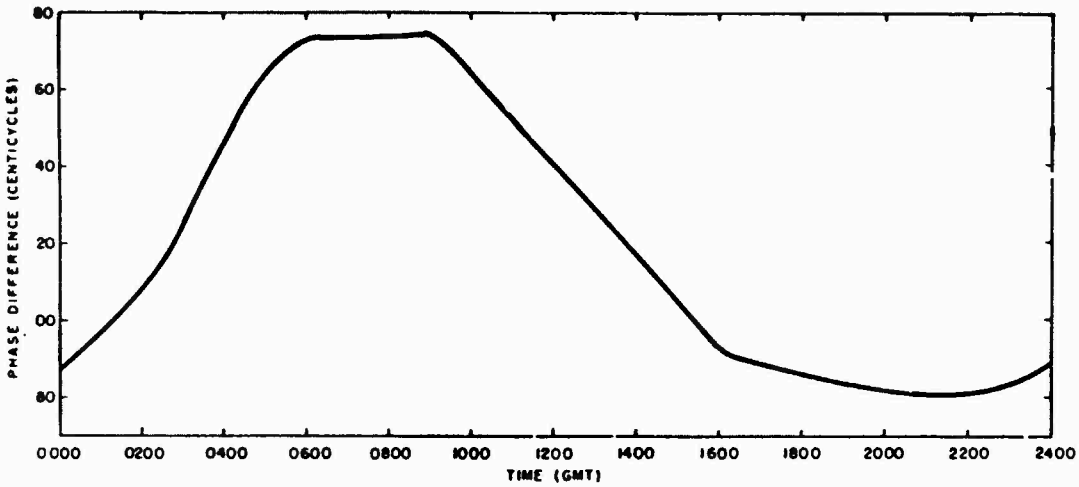


Figure 18. Average 10.2 kHz phase of Haiku, Hawaii received at Forestport, New York, 17–24 May 1966. Standard Deviation approximately 2 centicycles.

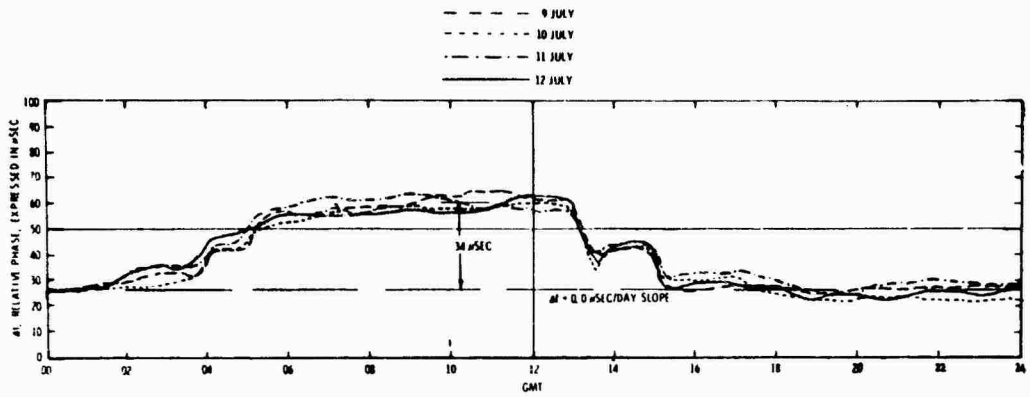


Figure 19. 19.8 kHz NPM (Hawaii) phase recorded in San Diego, 9–12 July 1965 (from Bickel).

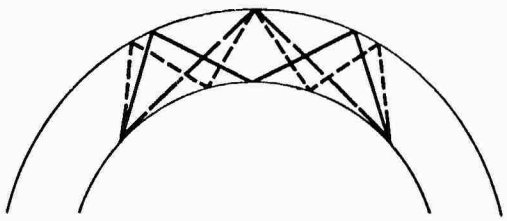
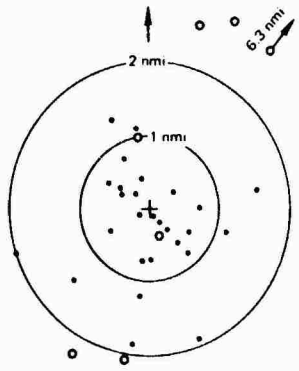


Figure 20. Ray hop propagation modes.



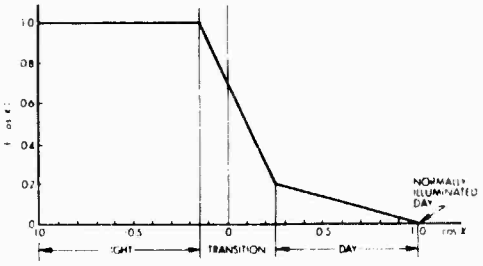


Figure 21. The diurnal function.

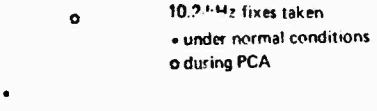


Figure 22. Polar plot of Omega fix errors. California and New York, two year sampling.

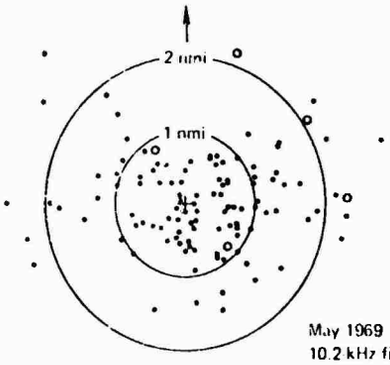


Figure 23. Polar plot of Omega fix errors. Five site sampling.

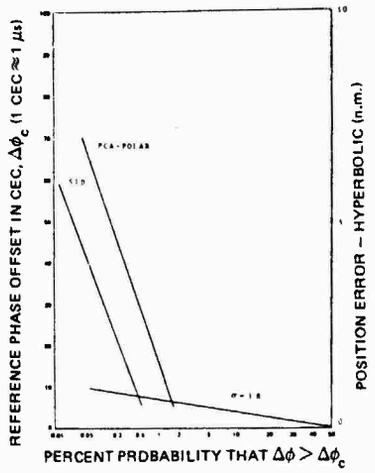
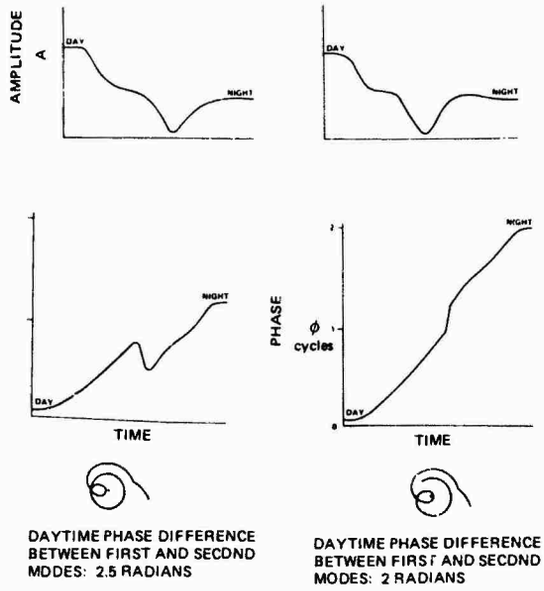


Figure 24. Typical probability of propagation induced blunders.

(a) FIRST MODE (≈ 1 CYCLE)

(b) SECOND MODE (≈ 2 CYCLES)

Figure 25. Vector loci for modal variation during transitions.



DAYTIME PHASE DIFFERENCE BETWEEN FIRST AND SECOND MODES: 2.5 RADIANS

DAYTIME PHASE DIFFERENCE BETWEEN FIRST AND SECOND MODES: 2 RADIANS

Figure 26. Sunset transition with change of mode dominance.

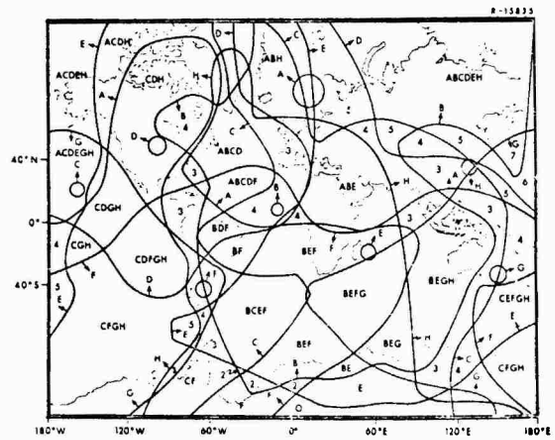


Figure 27. OMEGA 10.2 kHz signal coverage: Local noon. Fully implemented system.

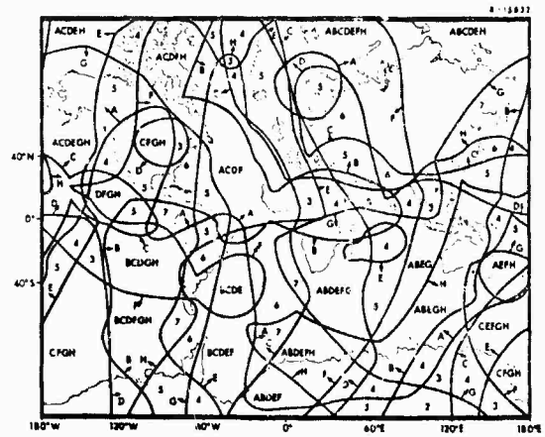


Figure 28. OMEGA 10.2 kHz signal coverage: Local midnight. Fully implemented system.

DIFFERENTIAL OMEGA : TESTS AND DEVELOPMENT IN FRANCE

D. ABADIE
SERVICE TECHNIQUE DES PHARES ET BALISES
Bonneuil/Marne
France

P. LAURENT
SERCEL
Nantes
France

SUMMARY

The operation of OMEGA radionavigation system worldwide coverage may be locally improved as far as accuracy and reliability are concerned, by taking into account corrections by a fixed shore station called differential OMEGA station.

This paper describes telemetering format used in France and gives some experimental results obtained at the differential OMEGA stations of Ile d'Yeu and Cap Couronne, on French coasts.

I. INTRODUCTION

Differential OMEGA is a navigation system which can improve the accuracy of standard OMEGA with a coverage of about 300 nautical miles around an OMEGA differential station located on land.

The principle of differential OMEGA is already described in other papers (BROGEN, K.W. and LUKEN, K.O.L., 1966 ; NARD, G., 1971 ; SWANSON, E.R. 1974) and is now well known : an OMEGA receiving station placed on land compares theoretical values of OMEGA coordinates with observed values and broadcasts observed corrections to sailors in a 100 to 300 nautical miles radius circle.

With the French coding system, the corrections are automatically used by the receiver, so that the sailor only reads coordinates on the receiver.

II. DESCRIPTION OF THE CODING FORMAT TESTED IN FRANCE

II.1. General

The format developed in France for transmission of differential OMEGA corrections is based on the following choices :

- phase corrections are transmitted for each single OMEGA station with no connection with corrections for other stations, in a sequence exactly superimposed to the transmission cycle of OMEGA stations. In other words, the phase correction for a given OMEGA station is transmitted in the real time where this station transmits,
- the phase reference is given by the local oscillator of the OMEGA receiver of the corrections transmitting station. The frequency stability of the local oscillator is the same as that of a standard receiver (i.e. around 10^{-7}),
- corrections are transmitted in an analog form,
- the information modulates in phase a 20 Hz sub-carrier which also modulates a carrier frequency in phase, with a modulation index smaller than 0.8,
- the carrier frequency is that of a maritime radiobeacon which transmits either on a continuous or sequential basis.

x

x x

To help understanding the reasons for these choices, some information will be given on the operation of correction transmission in paragraphs II.2 and II.3 hereafter.

II.2. Description of the transmission format

Figure 1 shows a block diagram of the operation of a differential OMEGA corrections transmitting station. This station is in three parts :

- a modified OMEGA receiver for the reception of fundamental OMEGA signals,
- a coder modulator which shifts the phases of the signals received from the OMEGA coordinates computed for the correction transmitting station, applies the resulting phases to a 20 Hz sub-carrier and modulates the radiobeacon master oscillator with this sub-carrier,
- the radiobeacon itselft.

The OMEGA receiver successively receives the signals from the OMEGA stations. Assume, for instance, that stations A, B, D and G are correctly received, E, F and H are more or less weakly received and C is not received(1).

After an amplification and reshaping of the 10.2 kHz signal before mixer M1, the signal has the form shown in figure 2 (signal 1).

The following notations will be used :

- d_S^A = distance between OMEGA station A and the correction transmitting station S
- d_M^A = distance between station A and ship M
- d_M^S = distance between station S and ship M
- F = OMEGA transmitting frequency (10.2 kHz)
- f = transmitting frequency of the radiobeacon associated with the correction transmitting station
- $C_S^A(t)$ = propagation speed of the resultant phase on A-S for time t
- $C_M^A(t)$ = propagation speed of the resultant phase on A-M for time t
- C_{SO}^A = theoretical phase propagation speed on A-S
- φ_i^A = phase of the signal from OMEGA station A during the i-th step of the signal processing
- φ_o = initial phase of OMEGA stations, the same for all stations
- φ_{oLi} = initial phase of the local oscillator i
- equivalent notations for signals from stations B, C etc...
- k_i^A, k_i^B etc... are integer numbers, positive, negative or null such that all phases be within $(0 - +2\pi)$.

The following conditions are considered as being met :

- phase propagation speeds $C_S^A(t)$, $C_S^B(t)$ etc... and $C_M^A(t)$, $C_M^B(t)$ etc... are constant during the duration of an OMEGA sequence (10 seconds),
- the frequency stability of local oscillators used in OMEGA receivers of station S and of ship M is good enough for the ratio between the OMEGA frequency and the local oscillators frequency to be considered constant during an OMEGA sequence.

Before mixer 1, the phase of the signal received from station A is :

$$\varphi_1^A = 2\pi F \frac{d_S^A}{C_S^A(t)} + \varphi_o + 2k_1^A \pi$$

Likewise :

$$\varphi_1^B = 2\pi F \frac{d_S^B}{C_S^B(t)} + \varphi_o + 2k_1^B \pi, \text{ etc ...}$$

(1) Such is the case in Western Mediterranean. This case is mentioned only for the sake of clarity ; the described format is adequate for any kind of reception of OMEGA stations around the earth.

After mixer M1, the sequential signal has the same form as that of signal 1 in figure 2 but its frequency is reduced to 1.2 kHz and the phase of the segments are :

$$\varphi_2^A = 2\pi F \frac{d_S^A}{C_S^A(t)} + \varphi_o + \varphi_{oLi} + 2k_2^A\pi$$

$$\varphi_2^B = 2\pi F \frac{d_S^B}{C_S^B(t)} + \varphi_o + \varphi_{oLi} + 2k_2^B\pi$$

The demultiplexer separates the signals from the different OMEGA stations, selects several of them, 5 for instance (it is the number chosen for the two differential OMEGA stations presently in service in France) stores their respective phases with 5 identical phase control φ_1 to φ_5 which time constant is around 100 seconds.

If the selected stations were A, B, D, E and G, then :

$$\varphi_1 = \varphi_2^A \qquad \varphi_2 = \varphi_2^B \qquad \varphi_3 = \varphi_2^D \qquad \varphi_4 = \varphi_2^E \qquad \varphi_5 = \varphi_2^G$$

Then phase shifters $\Delta\varphi_1 - \Delta\varphi_5$ substract from these 5 phases, phases such that :

$$\Delta\varphi_1 = 2\pi F\left(\frac{d_S^A}{C_{So}^A} - \frac{d_S^E}{C_{So}^E}\right) + 2k_1\pi$$

$$\Delta\varphi_2 = 2\pi F\left(\frac{d_S^B}{C_{So}^B} - \frac{d_S^E}{C_{So}^E}\right) + 2k_2\pi$$

$$\Delta\varphi_3 = 2\pi F\left(\frac{d_S^D}{C_{So}^D} - \frac{d_S^E}{C_{So}^E}\right) + 2k_3\pi$$

$$\Delta\varphi_4 = 0$$

$$\Delta\varphi_5 = 2\pi F\left(\frac{d_S^G}{C_{So}^G} - \frac{d_S^E}{C_{So}^E}\right) + 2k_4\pi$$

The terms $\Delta\varphi_1$, $\Delta\varphi_2$, $\Delta\varphi_3$, $\Delta\varphi_5$ actually are the fractionary part of the theoretical hyperbolic coordinates A - E, B - E, D-E, G-E, of station S as they can be obtained from the chart.

It should be noted here that :

- the values $\Delta\varphi_1 - \Delta\varphi_5$ result from theoretical calculations and are therefore fixed once when the station is commissioned but if the charting has been changed in the meantime,
- the choice of the reference OMEGA station (E in the chosen example) is arbitrary and makes no difference between this station and the seven others. As will be seen in the description of the corrections receiver, the system could even work, be this station off the air.

After the phase shifters, the phases on the storage channels are :

$$\varphi_3^A = 2\pi F\left(\frac{d_S^A}{C_S^A(t)} - \frac{d_S^A}{C_{So}^A} + \frac{d_S^E}{C_{So}^E}\right) + 2k_3^A\pi + \varphi_o + \varphi_{oLi}$$

$$\varphi_3^B = 2\pi F\left(\frac{d_S^B}{C_S^B(t)} - \frac{d_S^B}{C_{So}^B} + \frac{d_S^E}{C_{So}^E}\right) + 2k_3^B\pi + \varphi_o + \varphi_{oLi}, \text{ etc...}$$

The multiplexer located after the phase shifters, controlled in synchronism with the demultiplexer from the OMEGA format, sets the signals φ_3^A, φ_3^B etc... in a sequence according to a format identical to the OMEGA format ; the segments with no correction (C, F and H in the given example) are left void, and the 1.2 kHz signal has the shape of signal 2 in figure 2.

Mixer M2 brings the signal frequency to 20 Hz. Phases on segments A, B, D, E and G are then as follows :

$$\varphi_4^A = \varphi_3^A + \varphi_{OL2} + 2 k_4^A \pi$$

$$\varphi_4^B = \varphi_3^B + \varphi_{OL2} + 2 k_4^B \pi, \text{ etc...}$$

The shape of the signal remains that of signal 2 in figure 2. This sequential signal, which frequency is 20 Hz, phase modulates the radiobeacon carrier with a modulation index of 0.8. The mathematical value of the signal transmitted by the radiobeacon is, during transmission of segment A :

$$E \cos \left[2\pi f t + 0.8 \cos (40\pi t + \varphi_4^A) + \varphi \right]$$

then during segment B :

$$E \cos \left[2\pi f t + 0.8 \cos (40\pi t + \varphi_4^B) + \varphi \right], \text{ etc...}$$

II.3. Description of corrections reception

Block diagram 1 of figure 3 shows how corrections are received in the general situation.

The phases received by the OMEGA receiver on ship M are :

$$\varphi_M^A = 2\pi F \frac{d_M^A}{C_M^A(t)} + \varphi_0$$

$$\varphi_M^B = 2\pi F \frac{d_M^B}{C_M^B(t)} + \varphi_0, \text{ etc...}$$

At the same time the corrections receiver will receive the following signal on the radiobeacon frequency f during segment A :

$$E \cos \left[2\pi f t + 0.8 \cos (40\pi t + \varphi_4^A) + \varphi' \right]$$

where φ' is different of φ because of the time of propagation between station S and ship M.

Likewise on the following segments.

The discriminator which follows reception circuits gives a 20 Hz sequential signal which shape is similar to that of signal 2 in figure 2 and where φ' has disappeared.

The demultiplexer which follows the first phase discriminator, controlled by the OMEGA format, separates the segments and selects four of them.

The control of the demultiplexer of the OMEGA receiver and of that of the corrections receiver are coupled so that the same stations are selected. Suppose for instance the operator chose to display couples A-D and B-G. Segment A, D, B and G are sent to four 20 Hz detectors which give phases φ_4^A, φ_4^D, φ_4^B and φ_4^G as voltages. These phases are substracted from those stored in the demultiplexer channels of the OMEGA receiver.

Eventually, in the chosen example, the input of the first phasemeter of the OMEGA receiver receives 1.2 kHz signals which phases are :

$$\varphi_5^A = \varphi_M^A - \varphi_4^A + \varphi_{OL3} + 2 k_5^A \pi$$

$$\varphi_5^D = \varphi_M^D - \varphi_4^D + \varphi_{OL3} + 2 k_5^D \pi$$

where φ_{OL3} is the phase introduced by the local oscillator of the ship's OMEGA receiver.

The phasemeter then displays $(\varphi_5^A - \varphi_5^D)$, i.e :

$$\varphi_5^A - \varphi_5^D = \varphi_M^A - \varphi_M^D + \varphi_4^A - \varphi_4^D + 2(k_5^A - k_5^D)\pi.$$

φ_{OL3} has disappeared. Furthermore the substraction $\varphi_4^A - \varphi_4^D$ takes out terms $2\pi F \frac{d_S^E}{C_{So}^E}$, φ_0, φ_{OL1} and φ_{OL2}.

Then :

$$\varphi_S^A - \varphi_S^D = \left(\varphi_M^A - \varphi_M^D\right) + 2\pi F\left(\frac{d_S^A}{C_{S0}^A} - \frac{d_S^D}{C_{S0}^D}\right) + 2\pi F\left(\frac{d_S^D}{C_S^D(t)} - \frac{d_S^A}{C_S^A(t)}\right) + 2k'\pi$$

where the second and third terms are the differential OMEGA phase correction applied to the natural phase difference $\left(\varphi_M^A - \varphi_M^D\right)$.

II.4. Advantages of the chosen format

II.4.1. Transmission of corrections for each single OMEGA station

The very advantage to transmit phase corrections for each single OMEGA station is that it makes possible for the mariner to choose the station couples he wants to use.

The number of stations which can be received at a given location varies from 3 to 7. Practically reception trials performed by the French Navy in Western Mediterranean gave the following results :

Station	Name	Reception quality	Proportion of time
Aldra	A	5 on 5	24 h per day
Liberia	B	5 on 5	24 h per day
Hawaï	C	0 on 5	
North Dakota	D	1 to 3 on 5	24 h per day
Reunion	E	1 to 2 on 5	12 h per day
Argentine	F	2 to 3 on 5	18 h per day
Trinidad	G	4 on 5	24 h per day
Japan	H	1 to 4 on 5	18 h per day

This table shows that 7 of the 8 stations are received but that only 3 (A, B, G) are permanently received correctly, the reception of the others being bad during part of the time. Furthermore stations F and H are most often alternatively received.

The best choice of OMEGA stations in a given area then depends on the time, the season and therefore it is necessary to transmit phase corrections for the highest possible number of stations; furthermore the mariner is helped if he does not need to look in the sailing directions which stations are corrected in a given area and if his differential OMEGA receiver tells him automatically which segments are corrected and which are not.

II.4.2. Analog transmission of signals

The main advantage of **analog** transmission for the correction is that it is then possible to design corrections receivers which are easily fitted to standard OMEGA receivers : the principle is shown on figure 2 (block diagram 2). In this type of receiver, the phase modulated 20 Hz signal, obtained after one single phase demodulation of the carrier, is directly applied to the local oscillator of the OMEGA receiver which it modulates in phase. The storage and display circuits directly use the OMEGA signals which are automatically corrected with the local oscillator changes of **phase** which are proportional to the corrections to be applied to each segment.

II.4.3. Transmission by phase modulation of a sub-carrier which modulates the carrier in phase

With identical coding, the chosen modulation gives a greater range than a system which would use a phase modulation of the amplitude modulation as most of the radiated energy is in the carrier itself. It can be established that with the chosen modulation index (0.9), 98 % of the radiated energy is in the central component and the two first lateral components, the maximum frequency deviation being $\pm$16 Hz from the central frequency.

Furthermore, the transmission of the useful signal by phase modulation of a sub-carrier and not of the carrier itself, makes it possible to get rid of the influence of the length of the path of waves at frequency f between station S and ship M. Therefore the correction signals can be used either on the skywave or on the groundwave of the signal received from the corrections station.

II.4.4. Use of a radiobeacon signal as a carrier

The advantage of using a radiobeacon signal as a carrier is obvious because of the difficulty to get new frequencies for maritime radionavigation.

It should be fine to use only radiobeacons working continuously to transmit differential OMEGA corrections, both because of the S.I.D.'s which can reach 50 cecs in a few minutes and to simplify the design of corrections receivers as shown above in II.4.2.

But the correction transmission format can be adapted to a sequential transmission of the carrier by having a 3 second interval before the sequence of transmission of differential OMEGA. This interval makes it possible for the correction receiver to identify the transmitting stations. During this identification interval, the carrier is modulated in phase with a frequency of 13, 14, 15, 16, 17 or 18 Hz. The corrections receiver is designed in such a way that it can recognize the low frequency phase modulation of the carrier, the operator can then choose the nearest one in a group of radiobeacons which transmit the differential OMEGA corrections.

Furthermore, the corrections are stored in the receiver during the time interval which separates two consecutive transmissions of the same radiobeacon.

III. RESULTS OF EXPERIMENTATIONS

Now two OMEGA differential stations are operating in France : one at Ile d'Yeu, on the French Atlantic coast; the second at Cap Couronne, near Marseilles, on the Mediterranean coast. Figure 4 shows the locations of the stations and the expected ranges.

III.1. Experiments carried out from Ile d'Yeu

Results of these experiments were described in other reports (NARD. G., 1971, 1972) and only the main results will be given here.

III.1.1. Readings recorded at a fixed point

Distance from the station		68 % phase error (centicycles)		Average accuracy gain compared to standard OMEGA corrected with USNOO corrections tables
km	Nautical miles	day	night	
30	16	1,3	1,5	8.0
85	46	1,5	1,9	7.0
111	60	1,6	2,1	6.7
170	92	1,7	2,5	6.2
370	200	2,3	3,5	4.6
550	297	4,0	6,0	2.6

These results show that :

- the maximumrepeatability is better than 1.5 centicycle on every LOP during 68 % of the time. The correspondant repeatability in position is better than 0.3 nautical mile (530 m),
- the diminution of space correlation increases with distance but the improvement of accuracy given by differential OMEGA is perceptible at more than 300 nautical miles from the correction transmitting station.

Figure 5 shows the efficiency of differential OMEGA during an S.I.D., observed at fixed point.

III.1.2. Readings taken at sea

Comparisons have been made in 1972 on the Atlantic ocean between standard OMEGA (corrected by USNOO tables) and differential OMEGA, on a ship, with the help of a radiobeacon system (TORAN). Several thousands of positions were noted and gave the following results at distances from the corrections

transmitting station comprised between 110 and 160 nautical miles (200 and 300 km) :

- 68 % phase error on pair A-B : 2.8 cecs
- 68 % phase error on pair B-D : 2.15 cecs
- 68 % resulting position error : 0.72 nautical mile (1340 m) with the OMEGA stations in operation before 1974.

III.2. Results of experiments carried out from Cap Couronne station

Experiments are now carried out from Cap Couronne station. Some results are already available.

III.21. Repeatability

Distance from the station		Station bearing	Pairs recorded	Centicycles	
km	Nautical miles			Day	Night
111	60	286°	A - B	1	2.5
			B - D	2	3
168	91	254°	A - B	1	2
			B - D	2	3
			A - G	1	2.5

The number of fixed observation stations was too small for the effect of distance from the correction transmitting station to be evaluated but the results obtained do not much differ from results of III.1.1.

III.2.2. Accuracy

Differences observed between theoretical values and average readings during 48 h are the following:

- at 60 nautical miles from the corrections transmitting station :

  2 cecs on pair A-B
  1 cec on pair B-D

- at 90 nautical miles from the corrections transmitting station :

  2 cecs on pair A-B
  3 cecs on pair B-D
  2 cecs on pair A-G

III.2.3. Range as a function of radiated power

The differential OMEGA station of Cap Couronne is associated with a maritime radiobeacon, the range of which is less than 50 nautical miles at 50 microvolts/metre . Tests showed that differential OMEGA corrections were received at 91 nautical miles from the station, 24 hours per day and with a good signal to noise ratio, with only sea path. In Ajaccio, situated 160 nautical miles from the station and with a sea path the corrections are received 15 hours per day. It then can be said that the range of a differential OMEGA station associated with a maritime radiobeacon is more than twice the radiobeacon range at 50 microvolts/metre.

IV. CONCLUSION

Presently available results proved the efficiency of a differential OMEGA system and its utility for mariners. Therefore, several countries have asked international bodies as CCIR, IMCO and IALA to design an internationally agreed format as quickly as possible.

BIBLIOGRAPHY

- BRODGEN, J.W., and LUKEN, K.O.L., "differential OMEGA", N.R.L. Memorandum (1966);
- NARD, G. "Results of recent experiments with differential OMEGA", OMEGA symposium of the U.S. Institute of Navigation (1971) and the International Hydrographic Review (1973),
- NARD, G., "Results of experiments and developments of differential OMEGA in France", OMEGA Symposium of the U.S. Institute of Navigation (1974),
- SWANSON, E.R., "differential OMEGA Navigation for the U.S.coastal confluence region", N.E.L.C., (1974).

ORGANIGRAMME DE LA STATION DE TRANSMISSION DE CORRECTIONS

(OMEGA corrections transmitting station)

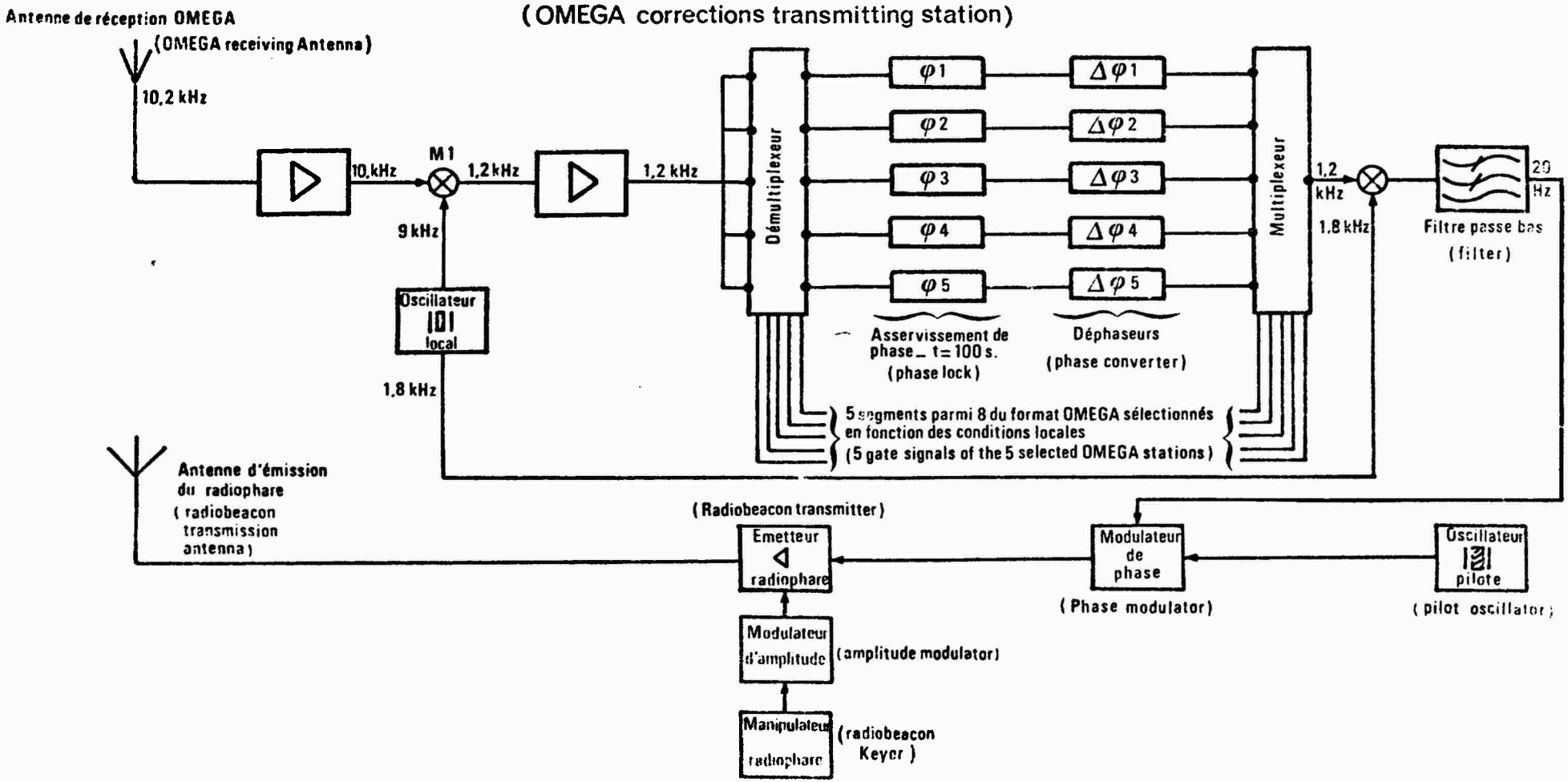
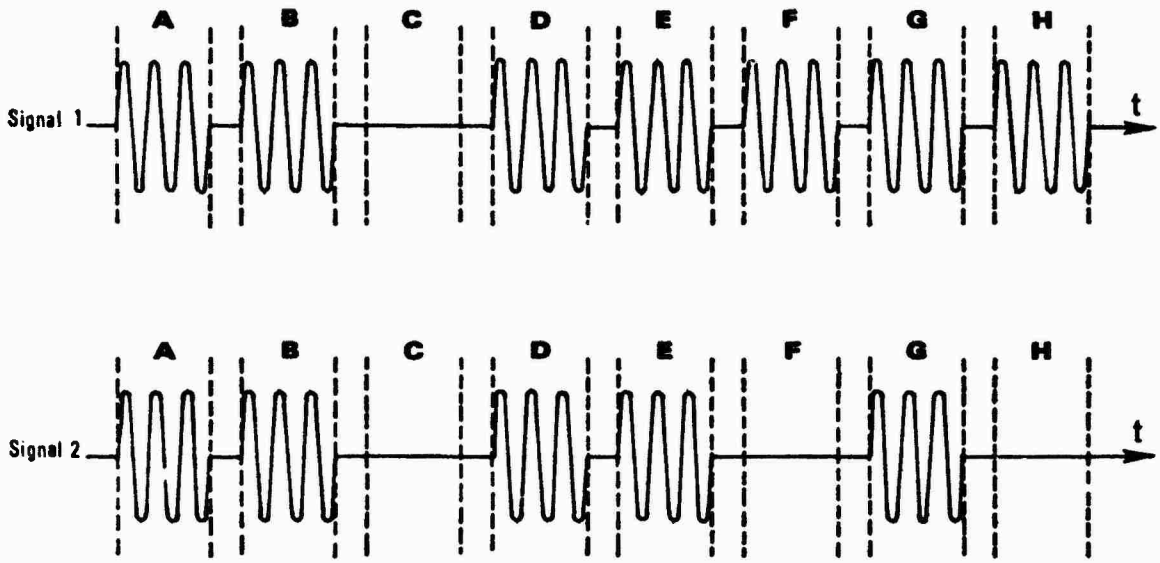


Figure 1



CHRONOGRAMME D'ÉMISSION D'UNE STATION OMEGA DIFFERENTIEL CONJUGUÉE A UN RADIOPHARE SÉQUENTIEL

(Signal sequence when differential OMEGA station coupled with sequential radiobeacon

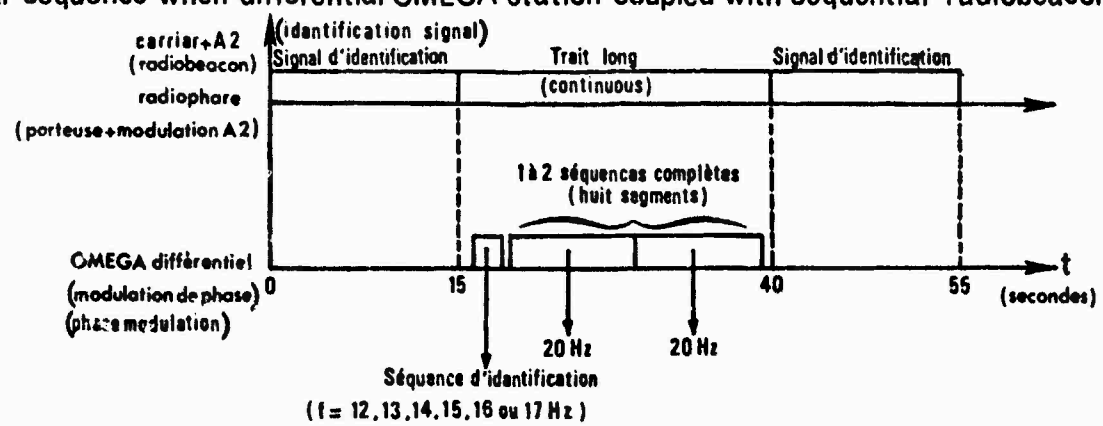


Figure 2

ORGANIGRAMME DU RECEPTEUR

OMEGA DIFFERENTIEL
(mobile receirer station)

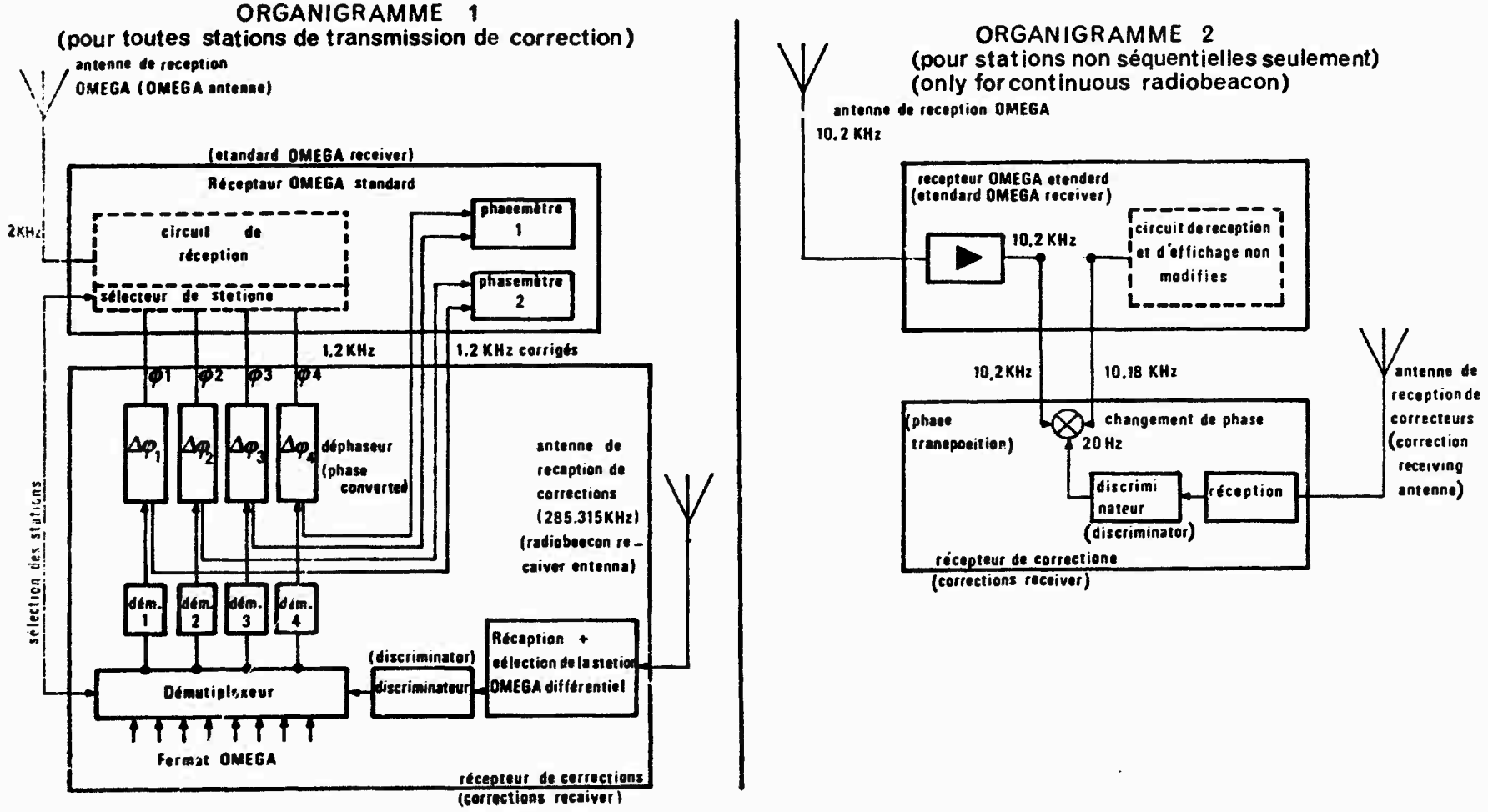


Figure 3

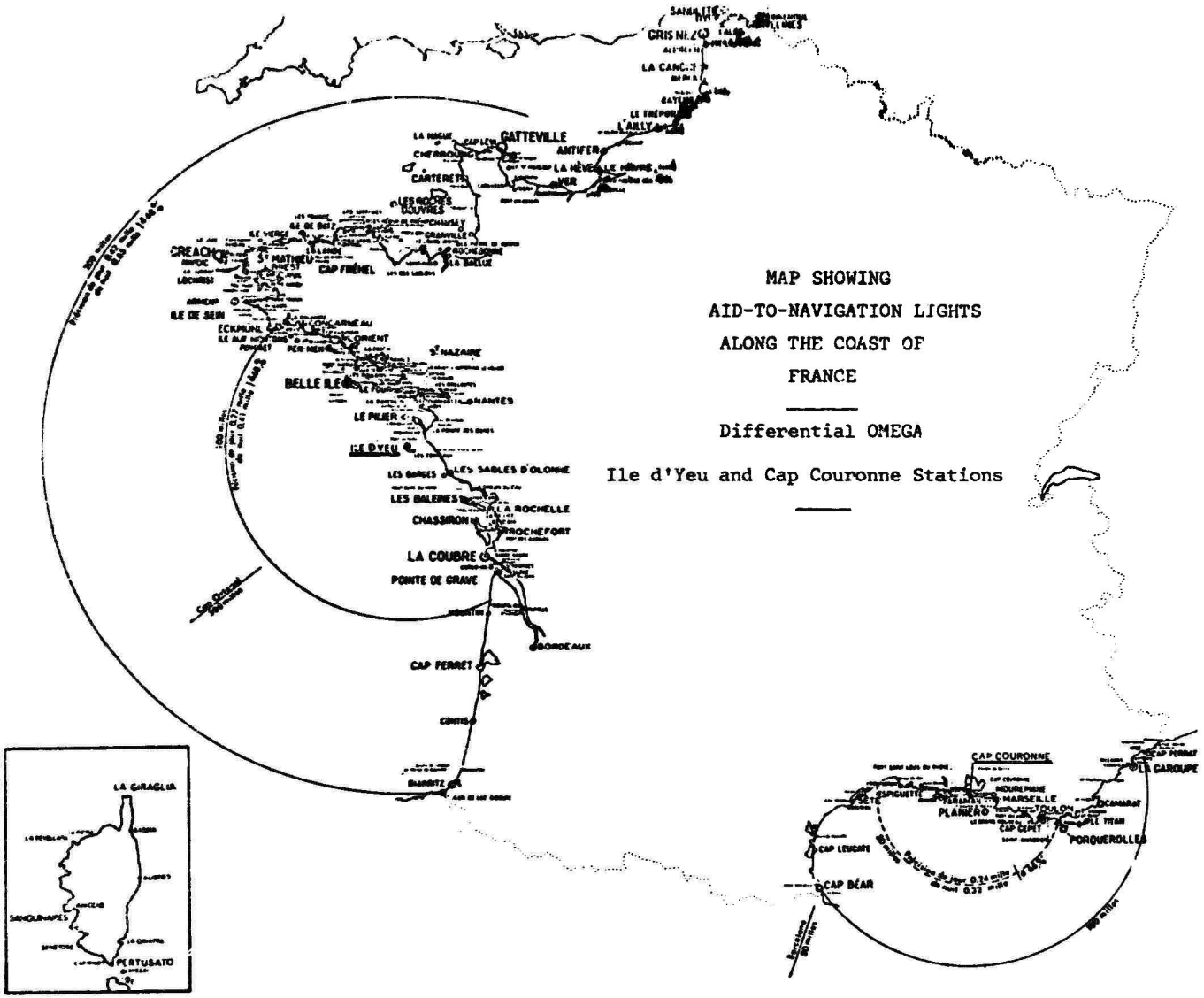
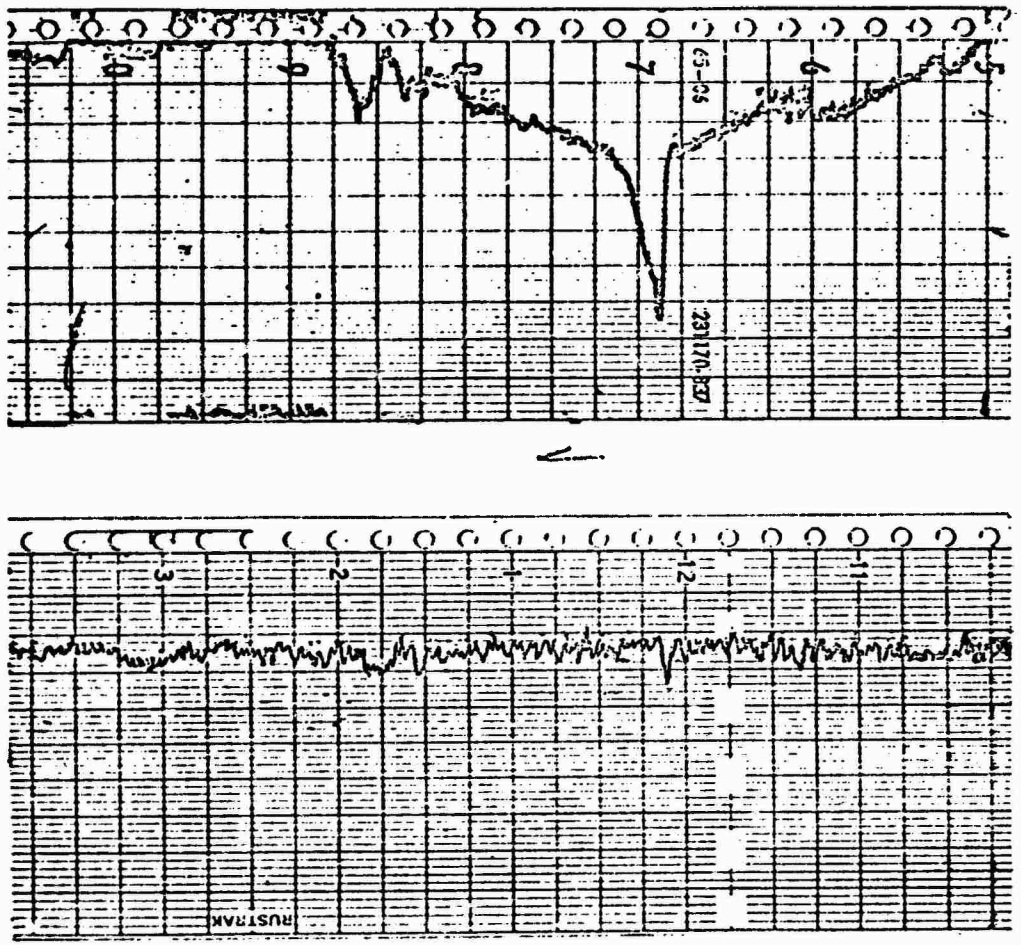

MAP SHOWING
AID-TO-NAVIGATION LIGHTS
ALONG THE COAST OF
FRANCE
Differential OMEGA
Ile d'Yeu and Cap Couronne Stations
GATTEVILLE
CARTERET
CREAC'H
ST MATHIEU
CAP FRÉHEL
ILE DE SEIN
BELLE ILE
LE PILIER
ILE D'YEU
LES SABLES D'OLONNE
LES BALEINES
LA ROCHELLE
ROCHEFORT
CHASSIRON
LA COUBRE
POINTE DE GRAVE
BORDEAUX
CAP FERRET
CAP COURONNE
MARSEILLE
PLANIER
CAP CANAILLE
PORQUEROLLES
CAP LEUCATE
CAP BÉAR
LA GIRAGLIA
PERTUSATO

Figure 4

A-D NATURAL AND DIFFERENTIAL

NANTES - ILE D'YEU
Distance 90 km

Effect of ionospheric disturbances

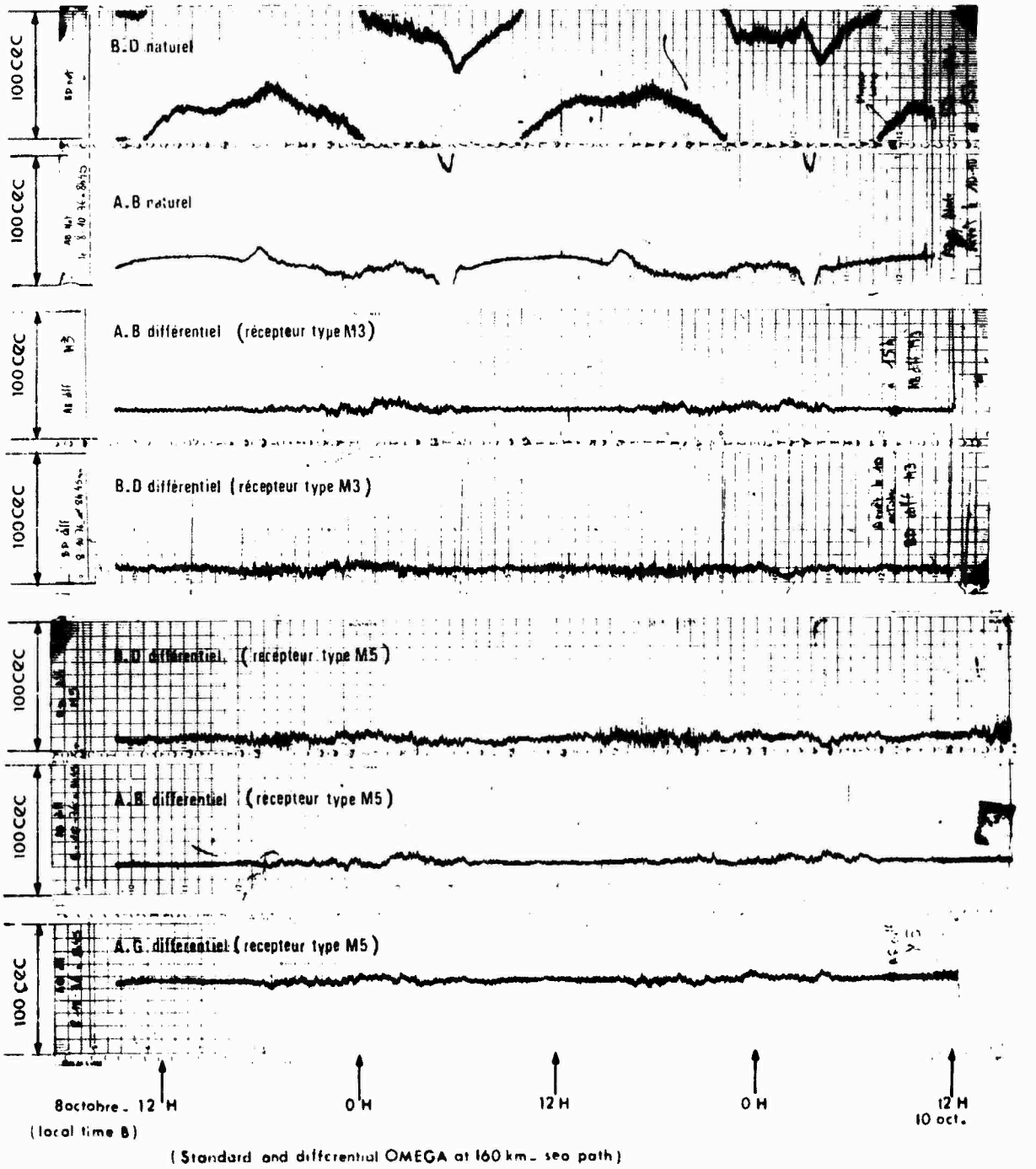


A-D normal

A-D Differential

Figure 5

ENREGISTREMENTS OMEGA NATUREL ET OMEGA DIFFERENTIEL AU CAP LEUCATE A 91 M (160 km) DE LA STATION DU CAP COURONNE. 8-10 octobre 1976



(Standard and differential OMEGA at 160 km - sea path)

Figure 6

DISCUSSION

T.R.Larsen, Norwegian Defence Research Establishment, Kjeller, Norway

These were interesting results. Are there plans for more stations in France?

Author's Reply

French Lighthouse Authority intends to build two new stations: one in the Mediterranean Sea, and one in the Atlantic (in Brittany). The two new differential OMEGA stations will work with sequential radio beacons. If an international format is quickly obtained, other stations will be built in the near future with this format.

A COMPARISON OF THE CALCULATED AND MEASURED DAYTIME PROPAGATION CHARACTERISTICS OF THE OMEGA TRINIDAD TRANSMISSIONS

G. Foley and T.B. Jones
University of Leicester, U.K.

and

B. Burgess
Royal Aircraft Establishment
Farnborough, U.K.

SUMMARY

The variation of the phase and amplitude of a VLF wave as a function of distance from the transmitter can be calculated by means of the waveguide mode analysis (Budden, 1961; Wait, 1962). This procedure evaluates the transverse electric and magnetic modes of propagation in the spherical waveguide formed by the earth as one wall and the ionosphere as the other. The conductivity of the earth between transmitter and receiver is known and for propagation over sea water, a very accurate value of the conductivity is available. Considerably greater uncertainty exists regarding the conductivity of the ionospheric wall since this is dependent on the electron density and collision frequency height distribution. The electron density profile will vary according to the time of day, season and latitude and will also be sensitive to a wide range of geophysical disturbances. The accuracy of the calculated propagation characteristics of the Omega signals therefore depends critically on the validity of the model ionosphere adopted.

During the summer of 1966 the phase and amplitude of the 10.2 and 13.6 kHz transmissions from Omega Trinidad were measured in an aircraft flying radially away from, and then towards, the transmitter. Data were obtained during day-time for propagation in three azimuthal directions E-W, W-E and S-N.

The variations of the phase and amplitude of both frequencies has been calculated for the three azimuthal directions, using a number of D-region models. The calculated variations are compared with the measured changes of phase and amplitude in order to assess the accuracy of the analysis procedure. Very close agreement is obtained between the measured and calculated values of both phase and amplitude when the ionospheric model employed is similar to that suggested by Deeks (1966a). Other D-region models yield calculated values which do not show the same degree of agreement with those measured. This investigation has shown that it is possible to calculate accurately the phase and amplitude of both the 10.2 and 13.6 kHz transmission at any distance from the low latitude Omega Trinidad transmitter for daytime conditions.

1. INTRODUCTION

A comparison has been made between measurements of the phase and amplitude of VLF waves as a function of distance from the transmitter and theoretical calculations of these parameters based on the waveguide mode theory. The measurements were recorded during aircraft flights near the Omega transmitter in Trinidad. The calculations were made with three separate electron density profiles in order to derive, by comparison between the calculations and measurements, information on the ionospheric D-region at this low latitude. In particular, an electron density profile was constructed from a mid-latitude profile by reducing the low-lying C-layer ionisation, thereby simulating the effect of the geomagnetic cut-off of the cosmic radiation.

2. EXPERIMENTAL RESULTS

The phase and amplitude variation of a VLF signal as a function of distance may be conveniently measured by means of a phase tracking receiver carried aboard an aircraft. A comprehensive set of flight measurements were made close to the Omega Trinidad transmitter by the Radio Department, Royal Aircraft Establishment, Farnborough, during August, 1966. Two day flights were made in which the aircraft flew radially away from the transmitter in a magnetic east and west direction respectively. A further flight from Bermuda to Barbados via Trinidad allowed propagation in a northerly (magnetic) direction to be studied (Burgess, 1970).

The field strengths at frequencies of 10.2 kHz and 13.6 kHz were recorded as a function of distance and in addition the differential phase delay between these two frequencies was also measured. The flight times were chosen so that the zenith angle of the sun at the receiver was never greater than 20°, and, thus, throughout the observation periods, conditions over the path corresponded to within 2 hours of the local noon.

3. PHASE OBSERVATIONS

Figures 1 to 3 show the measured time delay of the 10.2 kHz carrier relative to the 13.6 kHz carrier (i.e. $\tau_{10.2} - \tau_{13.6}$ sec) as a function of distance from the transmitter in the three azimuthal directions.

Figure 1 relates to two flights between Bermuda and Trinidad where propagation is approximately towards magnetic north. The general agreement between the measurements taken on the two flights is quite good. The data are considered to be representative of conditions within 2 hours of midday over the path. The increases in the relative phase delay (or time delay) at approximately 200-300 km and 650-850 km from the transmitter are due to the destructive interference between the first two propagating modes at 10.2 kHz.

Similarly, the decreases in phase delay at distances of [illegible] interference at 13.2 kHz. It may be shown that the sense of these phase delay variations are such that the magnitude of the dominant (first) mode exceeds that of the next largest (third) mode at each frequency.

The above remarks also apply to the phase delay variations measured on flights in directions approximately magnetic east and west of the transmitter; Figures 2 and 3, respectively. In comparing these results, it should be noted that the magnitude of the phase delay variations are greatest for propagation towards magnetic east and least for propagation towards magnetic west. This is in agreement with waveguide mode theory which predicts that VLF propagation is generally non-reciprocal with respect to the earth's magnetic field.

4. AMPLITUDE OBSERVATIONS

For each of the aircraft flights detailed in section 2, the changes in signal strengths of the 10.2 kHz and 13.6 kHz signals have been recorded as a function of distance from the transmitter (that is for propagation in the magnetically north, east and west directions). The 13.6 kHz data for the flight between Burmuda and Trinidad (S-N propagation) is reproduced in Figure 4. A well-defined minimum occurs at a distance of 300-400 km due to the destructive interference between the dominant quasi-TM modes. At 800 km, the signal strength again decreases as the second position for destructive interference is approached. The results for propagation at 13.6 kHz in the W-E and E-W directions are reproduced in Figures 5 and 6, respectively, and show the same general features as for S-N propagation. It should be noted that for the W-E case, the decrease in signal level at the first and second minima are considerably greater than for the two other directions of propagation.

Similar results have been obtained for the 10.2 kHz transmission and the signal strength variations for this frequency are shown in Figures 7 to 9. Destructive interference between dominant modes produces minima at distances of approximately 200-300 km and 800-1000 km from the transmitter. For E-W propagation, the minima are quite small and the second minimum at 900 km was only observed on the return flight. An anomalous decrease in signal level, extending for a range of about 50 km, was noted on the W-E path at a distance of about 350 km on the outward path and at 420 km on the return flight. It was not clear why this sudden drop in signal occurred or why the position is different on the outward and return parts of the flight.

Little change occurs in either the dip angle or magnetic field intensity over the E-W and W-E flight paths, but much larger changes in these parameters are present in the flight from Barbados to Trinidad (S-N propagation). Thus, any inaccuracies that might arise from field variations will be accentuated in the N-S results.

5. COMPARISON WITH CALCULATED RESULTS

The variation of phase and amplitude as a function of distance from the transmitter has been calculated for frequencies of 10.2 kHz and 13.6 kHz by means of the MODE computer program.

The ionospheric models adopted for this study were those of Deeks (1966a) and Smith (1968), since these exhibit quite different features. The Deeks' profile represents summer noon sunspot maximum conditions over southern England. Above about 70 km, however, the profile corresponds to measurements of the electron density at low latitudes reported by Thrane (1966) and Van Biel (1966). A main feature of the Deeks' model is a pronounced maximum at 63 km, the so-called C-layer, thought to arise from ionisation by cosmic rays. The present experiments were conducted at low latitudes where the C-layer maximum might be considerably reduced due to the geomagnetic cut-off of the ionising cosmic radiation. In order to study the effect of reducing the low-lying ionisation, a model has been constructed which is basically the Deeks' profile with the C-layer removed. This is marked as profile III in Figure 10, while the Deeks' profile is marked I. The Smith profile, marked II in Figure 10, represents conditions for a zenith angle of 15.5° in the southern hemisphere.

The effective collisional frequency profile used in the calculation was derived from the CIRA mean atmosphere model by the methods suggested by Deeks (1966b).

The calculations, using MODE, were made for the same propagation directions as those of the experimental observations. The magnetic dip and gyrofrequency were assumed constant and equal to the values at the path mid-point. Although no allowance was made for the variation of the magnetic field along the path, this is not expected to be a major source of error. Bickel et al. (1970) have shown that the assumption of mid-point value of the field for the whole path produces little error except when propagation is in a southernly direction. In this study, no measurements were made for this direction of propagation. The values of the geomagnetic field parameter, and other relevant constants, are given in Table 1.

The aircraft flight took place at a height of approximately 30,000 ft. Due account was taken of this situation by the use of the height gain factor of each propagating mode. The height gain was calculated of each mode using the theory of Pappert et al. (1967) and preliminary calculations showed the calculated field strength variations to be appreciably different at zero and 30,000 ft.

Superimposed on Figures 1 to 3 are the calculated phase delay of the 10.2 kHz carrier relative to the 13.6 kHz carrier as a function of distance from the transmitter in the three azimuthal directions using each of the three profiles in Figure 10. The total field strength was assumed to be a sum of five waveguide modes, of which two (modes 2 and 4) are quasi-TE modes and relatively unexcited. It can be seen that the peak-to-trough variation in the relative time delay is reproduced well by all three profiles except in the case of propagation towards magnetic west. In this latter case, the Deeks' model (profile II) underestimates the variations while, in the other two cases, this variation is overestimated. The positions of the maxima and minima of the phase delay as predicted by the models are, in general, about

50-100 km in error. This difference, however, is about the same as that between the results obtained on the outward and inward flights. The greatest variation in phase delay of the carrier frequencies, and the greatest discrepancy between the measured and predicted results, are seen for propagation in the direction of magnetic west. This may be attributed in part to the greater effective reflection height and ionospheric penetration, for propagation in this direction (Snyder and Pappert, 1969).

The fine structure variations in the phase delay measurements are not, in general, reproduced by the calculated results. This is because such fine detail is produced by the higher order modes. However, it should be noted that the most obvious feature, the cusp-like minima at 400 km from the transmitter for westerly propagation, is predicted for the modified Deeks' profile (II).

Superimposed on Figures 4 to 6 are the calculated signal strength of the 13.6 kHz carrier as a function of distance from the source transmitter for propagation in magnetic north, east and west directions, respectively. Consistent with the time delay results, the peak-to-trough variations in the signal strength is predicted with acceptable accuracy by the mode theory. The minima in the field strength at about 400-500 km, resulting from interference between the two dominant modes, is calculated to be about 50-100 km further from the transmitter than those observed experimentally. The use of the modified Deeks' profile (II) is justified by these results in that the depth of these minima are predicted with greater accuracy by this model. The depth determined using the basic Deeks' model (I) is much less than was measured during the flights, especially in the case of propagation towards magnetic east.

Similar comments may be made for the predicted field strength calculations of the 10.2 kHz carrier as shown in Figures 7 to 9. The anomalous decrease in signal strength at about 400 km for easterly propagation is not reproduced for any of the three ionospheric electron density models.

6. CONCLUSIONS

The measured amplitude and phase delay variations at distance of less than 1,000 km from the transmitter may be reproduced and interpreted by waveguide mode theory. At greater distances, the least attenuated mode is likely to dominate and therefore the interpretation of measured results is easier. The fine structure in the measured variations is not, in general, reproduced by the calculated results since such fine detail is produced by the higher order modes.

The greatest discrepancy between measured and calculated variations is obtained for propagation towards magnetic west. This has been attributed to the greater degree of ionospheric penetration expected for this direction of propagation. The daytime calculated results compare well with the measured variations.

Three distinct profiles have been used in the study of the daytime results. The comparisons do not favour any one of the three profiles used in the study, although there is some support for the use of a profile with a reduced C-layer electron density maximum at these latitudes.

7. REFERENCES

BICKEL, J.E., J.A. FERGUSON and G.V. STANLEY, 1970, Radio Science 5, 19.

BUDDEN, K.G., 1961, Radio Waves in the Ionosphere, Cambridge University Press.

DEEKS, D.G., 1966a, Proc. Roy. Soc. 29, 413.

DEEKS, D.G., 1966b, J. Atmosph. Terr. Phys. 28, 839.

PAPPERT, R.A., E.E. GOSSARD and I.J. ROTHMULLER, 1967, Radio Science 2, 387.

SMITH, R.A., 1968, "Ground Based Radio Studies of the Lower Ionosphere", Proc. DRTE Conf., Ottawa, Canada, p.335.

SNYDER, F.P. and R.A. PAPPERT, 1968, Radio Science 4, 213.

THRANE, E.V. and W.R. PIGGOTT, 1966, J. Atmosph. Terr. Phys. 28, 721.

VAN BIEL, H.A., 1966, "Ground Based Studies of the Lower Ionosphere", Proc. DRTE Conf., Ottawa, Canada, p.216.

WAIT, J.R., 1962, "Electromagnetic Waves in Stratified Media", Pergamon Press.

BURGESS, B., 1970, Proc. IEE 117, 51.

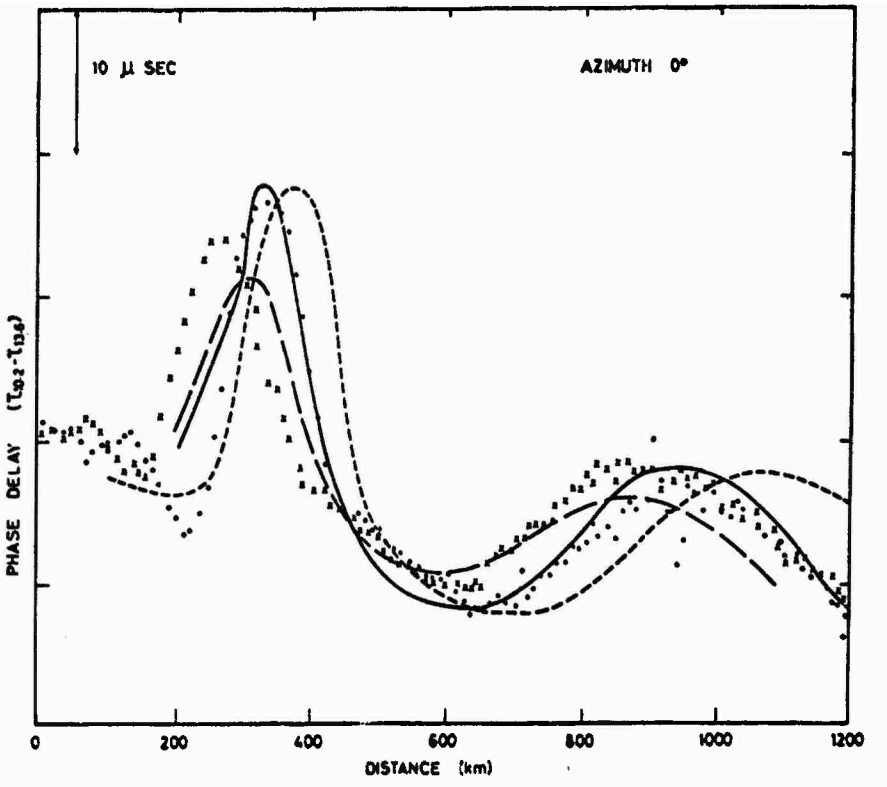


Fig.1 Comparison of measured and calculated phase delays between the 10.2 and 13.6 kHz transmissions from Omega Trinidad. S–N propagation. Daytime (9 and 23 August 1966)

Key:

Measured values	Flight towards transmitter	...
	Flight away from transmitter	xxx
Calculated values	Smith profile	——
	Deeks profile	– –
	Modified Deeks profile	- - - -

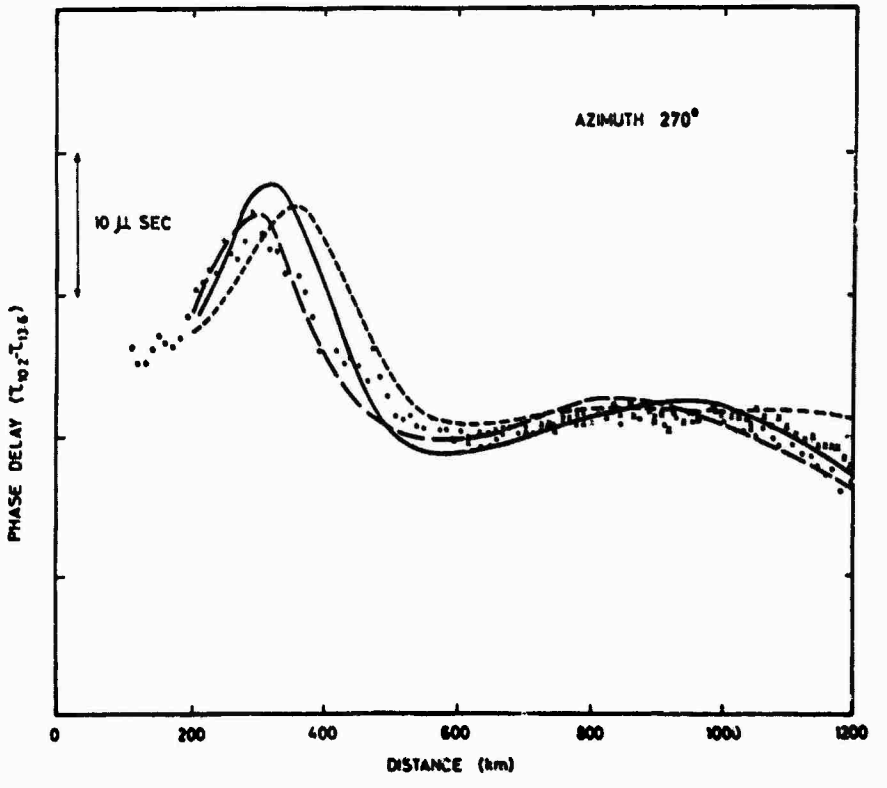


Fig.2 Comparison of measured and calculated phase delays between the 10.2 and 13.6 kHz transmission from Omega Trinidad. W–E propagation. Daytime (15 August 1966)

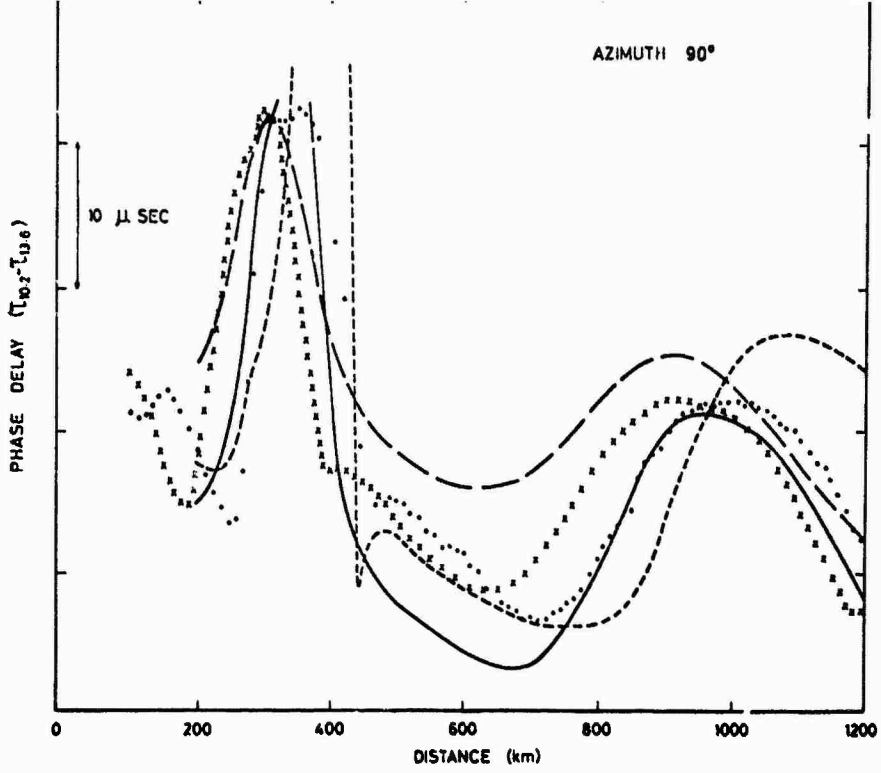


Fig.3 Comparison of measured and calculated phase delay between the 10.2 and 13.6 kHz transmission from Omega Trinidad. E–W propagation. Daytime (16 August 1966)

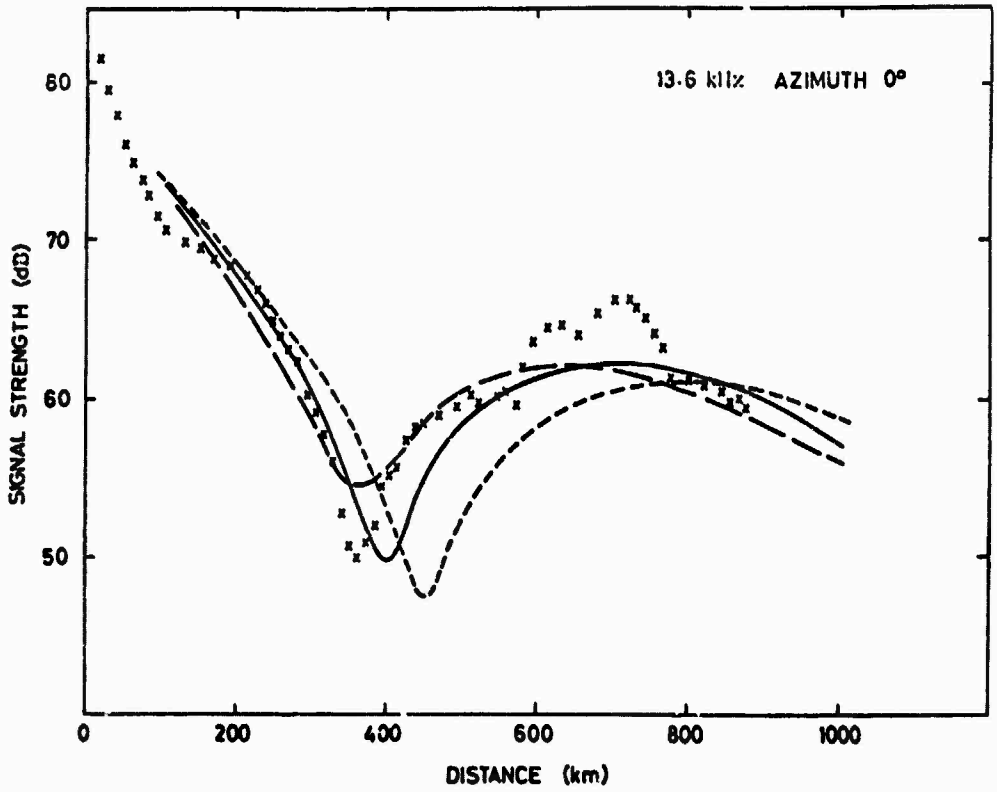


Fig.4 Comparison of the measured and calculated amplitude variations of the 13.6 kHz transmission from Omega Trinidad. S–N propagation. Daytime (9 and 23 August 1966)

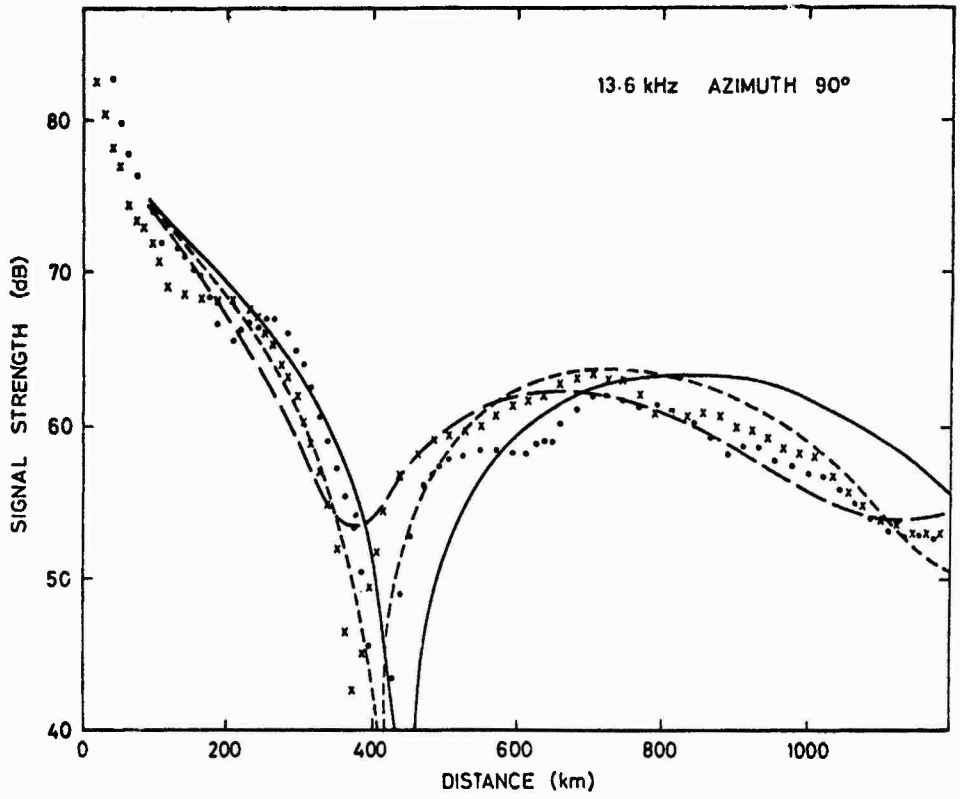


Fig.5 Comparison of the measured and calculated amplitude variations of the 13.6 kHz transmission from Omega Trinidad. W–E propagation. Daytime (15 August 1966)

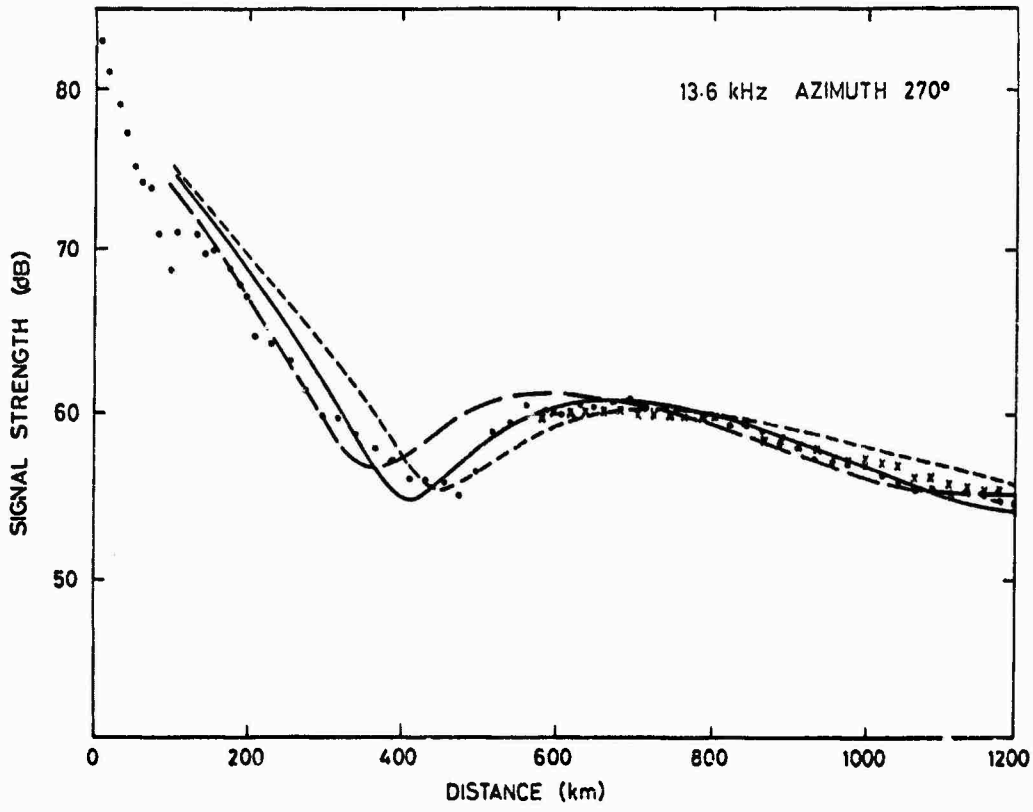


Fig.6 Comparison of the measured and calculated amplitude variations of the 13.6 kHz transmission from Omega Trinidad. E–W propagation. Daytime (16 August 1966)

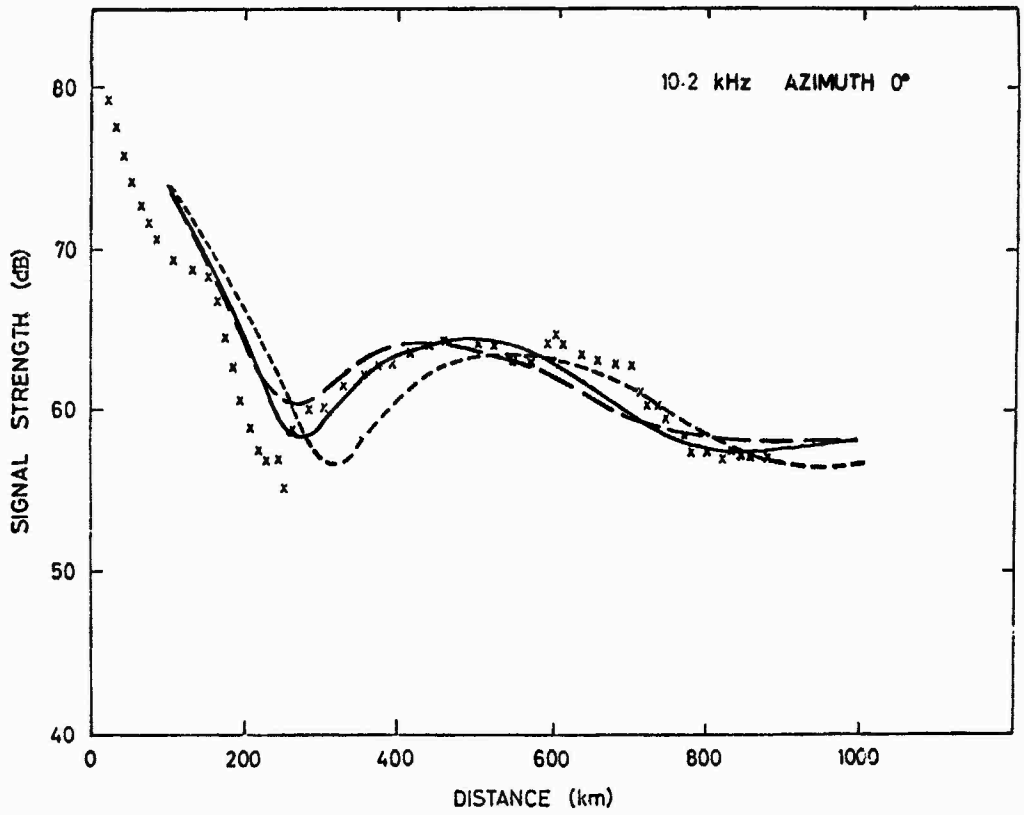


Fig.7 Comparison of the measured and calculated amplitude variations of the 10.2 kHz transmission from Omega Trinidad. S–N propagation. Daytime (9 and 23 August 1966)

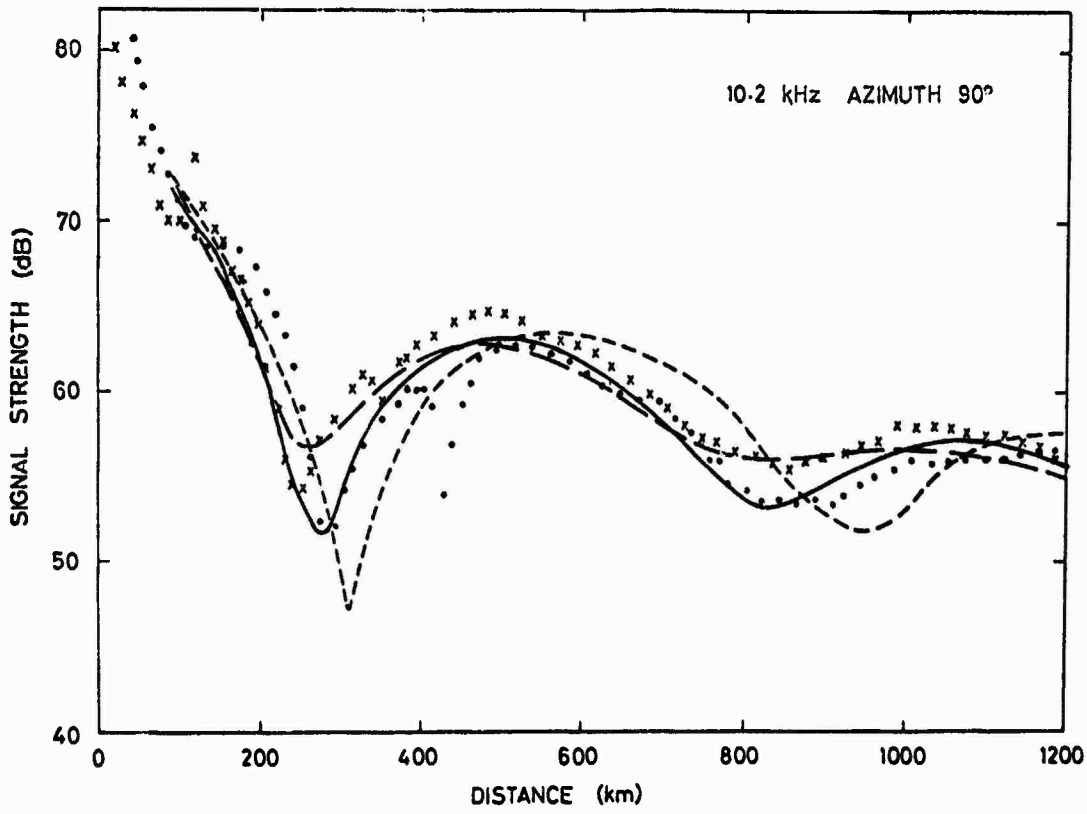


Fig.8 Comparison of the measured and calculated amplitude variations of the 10.2 kHz transmission from Omega Trinidad. W–E propagation. Daytime (15 August 1966)

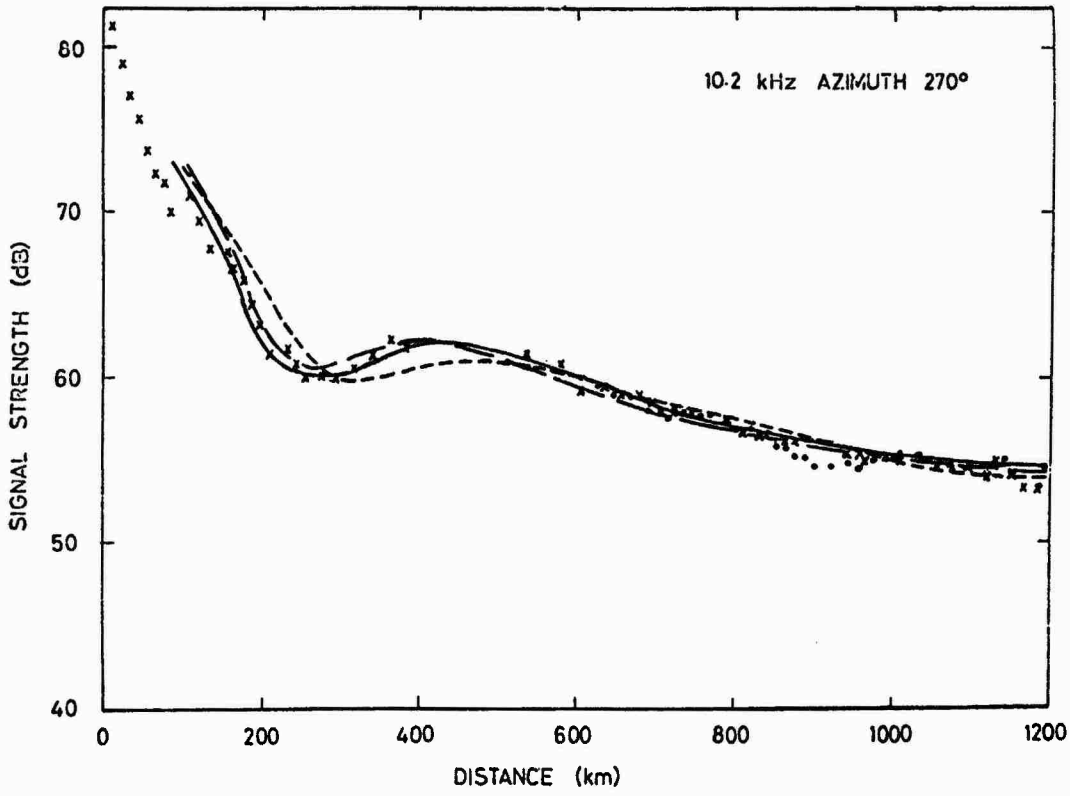


Fig.9 Comparison of the measured and calculated amplitude variations of the 10.2 kHz transmission from Omega Trinidad. E–W propagation. Daytime (16 August 1966)

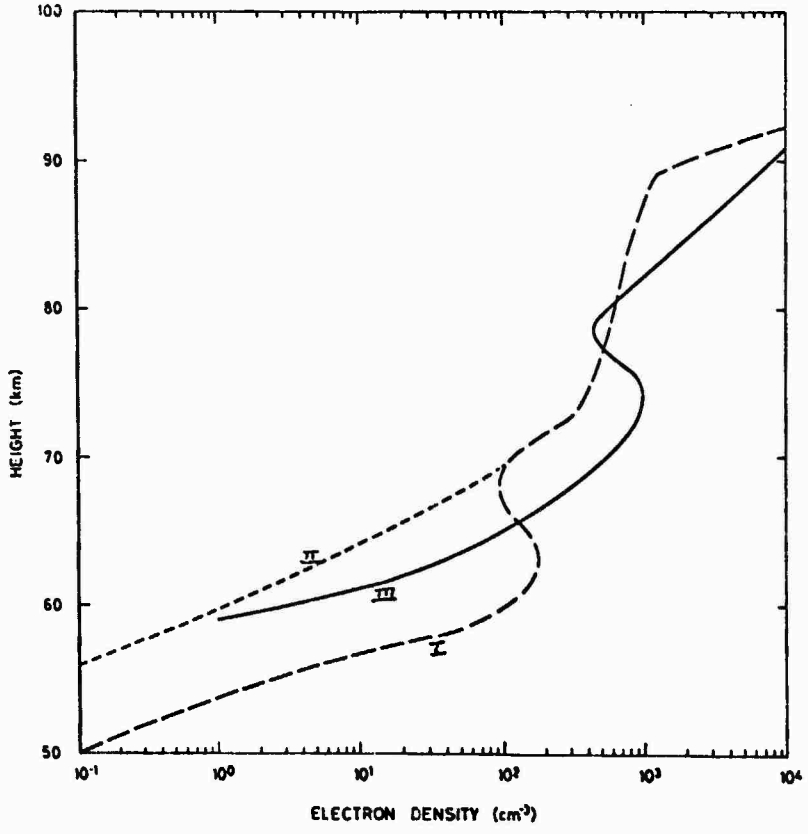


Fig.10 Electron density profiles used in these calculations

Key:

Profile I – Summer noon, sunspot maximum conditions (Deeks 1966)

Profile II – Profile derived from I by decreasing C-layer electron density

Profile III – Daytime over NSW Australia (Smith et al. 1968)

DISCUSSION

T.R.Larsen, Norwegian Defence Research Establishment, Kjeller, Norway

The "revised" Decks N(h) profile seemed to predict the right amplitude for the phase for the Trinidad measurements. Can the offset be explained if you make a slight height change of the whole N(h) profile?

Author's Reply

We have not yet repeated the calculation for the profiles displaced in height. It is quite likely that such a displacement will affect the agreement between the experimental and calculated values, but until we have done the calculations, I cannot give any indication of the magnitude of the changes involved.

(Added reply to Author's by B.Burgess) The field strength measurements taken by mounting VLF transmissions are normally smoothly varying; the perturbations on the field strength measurements given in the paper we believe are a real phenomena. A possible explanation could be in the fact that the propagation paths monitored on the flights in the West Indies, passed over a number of small islands, which could act as small perturbations in the lower wall of the earth-ionosphere waveguide; these physical perturbations excited evanescent modes of propagation.

T.R.Larsen

Can the fine structure, that was apparent on some of the phase records, be interpreted as time variations due to ionospheric changes?

Author's Reply

It seems unlikely, because the experiments were carried out at times when the ionosphere was undisturbed, and geophysical activity was low. There may be some transient phenomena which disturb the conductivity of the ionosphere but it is difficult to envisage what these might be. There is a feasibility that changes in ground conductivity (e.g., small islands) may have an influence on the observations.

OMEGA ACCURACY IN POLAR REGIONS DURING IONOSPHERIC DISTURBANCES

T.R. Larsen
Norwegian Defence Research Establishment
N-2007 Kjeller, Norway

SUMMARY

The paper presents some results of a study on OMEGA propagation variability at high latitudes. More than 300000 hourly phase difference values obtained between 1968 - 1970 for various OMEGA pairs and several observation sites have been analyzed. The statistics for signal variability at 10.2 kHz show typicallly an r.m.s. variability of 4 - 9 centicycles (cec) for the hourly standard deviations grouped into 24-hour periods on a monthly basis. Highest variability is obtained during winter months.

Phase difference observations during solar X-ray flares have been analyzed. In 1968 such events were occurring during less than 1% of the total observation time; mean duration of each event was ≈55 min with a standard deviation of ≈30 min. Mean phase offset (at 10.2 kHz) was 23 (cec) with a standard deviation of 16 cec. These values refer to the OMEGA pair A-B received at Oslo (60°N); similar results were obtained for A-C received at Ny Ålesund, Svalbard (79°N).

The effects of 13 solar proton events have been analyzed for the pair A-C received at Ny Ålesund. The mean value of the phase offset (at 10.2 kHz) was 38 cec with a standard deviation of 11 cec. During the event which started on 2 November 1969 offset in excess of 60 cec was recorded during the initial hours of the event.

1. INTRODUCTION

1.1 Background

Radio wave propagation in the Arctic is hampered by several types of ionospheric disturbances unique for these areas, such as auroral and solar proton events (SPE). In addition, effects from stratospheric warmings, energetic electron precipitation and solar X-ray (SXR) flares may influence the radiowave propagation characteristics.

For VLF navigation systems the phase of the radio wave is often a critical parameter, whose stability and predictability contributes to the navigational accuracy of the system. The present study presents some results of three years of monitoring 10.2 kHz OMEGA signals at several recording sites.

Two such stations were operated by the Norwegian Defence Research Establishment between 1968 and 1970 near Oslo (≈60°N) and at Hammerfest (≈71°N). The Hammerfest station was moved to Tromsø (≈70°N) in late 1969, cf Figure 1 which shows the geographical positions of sites referred to in this paper. Furthermore, The Norwegian Institute of Cosmic Physics has operated a receiving station at Ny Ålesund on Svalbard (≈79°N). These data as well as data from recordings at Keflavik, Iceland (≈64°N) and Hestmona (very close to Aldra), Norway (≈67°N) have been incorporated in this study. (The data from Ny Ålesund has been supplied by Prof A Egeland, whereas the latter set was made available through Mr E Swanson, Naval Electronics Laboratory Center, San Diego.)

Table 1 lists the various recording sites and gives their distances (in km) from the OMEGA transmitters.

1.2. Observational arrangements

At each recording site the phase difference at 10.2 kHz of selected pairs of OMEGA transmitters was monitored continuously on strip-chart paper. At Hestmona some recordings were also made at 13.6 kHz. The phase difference values were read off from the charts at each hour and these data formed the basis for a subsequent statistical evaluation.

1.3. The data set

The data reccrding during times when solar proton evenst were occurring, have been excluded from the statistical study. These data, however, will be discussed separately.

1.4. Scope of this study

In chapter 2 the results of an analysis of the OMEGA phase variability are presented. More than 300000 hourly phase difference values are included in this study. In chapter 3 predicted OMEGA phase difference values will be compared with the measured ones. Chapter 4 is devoted to effects from solar X-ray flares on OMEGA propagation and the influence of solar proton evenst are likewise treated in chapter 5. Some conclusions are presented in a final chapter 6.

2. OMEGA PHASE DIFFERENCE VARIABILITY

2.1. Monthly statistics on phase difference variability

The phase difference values obtained each hour for selected OMEGA station pairs have been averaged over a monthly basis for each hour of the day. The corresponding standard deviations have been calculated. The r.m.s. variation (in centicycles, cec) of these 24 monthly standard deviation values has been evaluated. Such results are tabulated in Table 2 which shows the data for the OMEGA pairs NORWAY minus TRINIDAD (A-B) and NORWAY minus HAWAII (A-C) at 10.2 kHz for the months of 1968 at five recording sites. Tables 3 and 4 show similar data for the years 1969 and 1970.

The individual entries in these tables are, disregarding some exceptions, computed from approximately 650 - 700 hourly phase values. The tables shows that generally there is rather good agreement between the results obtained at the various sites, although the r.m.s.-values for the pair A-C exhibit larger variations. There is a tendency for lower phase variability during summer months.

Most r.m.s. values range between 4 and 9 cec. Approximately, a measurement accuracy with a standard deviation of 10 cec on each of two lines of positions will translate into a positional fix accuracy of ≈1.8 km (cep) under favourable systems geometry.

Table 5 gives similar data for the OMEGA pair NORWAY minus FORESTPORT (A-D) at the Oslo and Keflavik sites for the period 1968-70. The Forestport transmitter operated at low output signal levels, but reception was possible at the Oslo and Keflavik sites.

2.2. Monthly statistics on phase difference variability for data grouped into periods of daytime, nighttime and the full 24-hours

The standard deviations (in cec) of the signal variability against the monthly mean value were computed for each hour of the day. In Table 6 the r.m.s. variation of these standard deviations are shown when grouped into periods during which complete daytime or complete nighttime existed along the propagation path. The last entry gives the results for the complete 24-hour period. These data refer to OMEGA NORWAY minus TRINIDAD (A-B) at 10.2 kHz in 1968, whereas Table 7 gives similar data for OMEGA NORWAY minus HAWAII (A-C).

The relatively large r.m.s. values for daytime (cf Table 6) during fall and late winter should be treated with some caution. Due to short days fewer hourly values were included when the r.m.s. values were computed. Low transmitter output power may also have resulted in erratic phase variations when reception was marginal.

2.3. Monthly statistics on phase difference variability as a function of frequency

Most of the recordings were made at 10.2 kHz but at Hestmona, Norway data were also collected at 13.6 kHz. The OMEGA pairs that were monitored were TRINIDAD minus NORWAY (B-A) and HAWAII minus NORWAY (C-A). As Table 1 shows, the receiving site at Hestmona is very close to the Aldra transmitter (18 km away). The signal variability is therefore essentially due to the variations on the path from Hawaii to Hestmona which crosses the center of the polar cap and traverses the auroral zone twice.

Table 8 shows the data for 13.6 kHz during the years 1967 - 1970, and Table 9 gives a comparison of results for 10.2 kHz and 13.6 kHz for the same period. No significant difference is detected in the phase variability between 10.2 and 13.6 kHz. The data show a tendency for higher r.m.s. values during the March/April and September/October months. Similar findings are reported by Martin (1972).

3. OMEGA PHASE DIFFERENCE PREDICTABILITY

3.1. Monthly statistics on OMEGA predictability

The received phase difference values have been adjusted by the "skywave corrections" according to the skywave correction model of 1969 (Swanson, private communication). The r.m.s. variation between these adjusted values and the predicted values for the specific sites have been computed for each hour of the day on a monthly basis. Data obtained during periods of solar proton events (SPEs) have been excluded. The r.m.s. variation of these hourly r.m.s. values have been evaluated for each month. The results from four receiving sites for the OMEGA pair NORWAY minus HAWAII at 10.2 kHz are presented in Table 10 for the years 1968-70.

Table 11 shows similar results for OMEGA NORWAY minus TRINIDAD during 1969.

Discussion of these data is beyond the scope of this paper. It should, however, be noted that a better prediction model exists today in as much as the correction model of 1969 has been updated.

4. EFFECTS FROM SOLAR X-RAY FLARE EVENTS ON OMEGA PHASE DIFFERENCE VALUES

4.1. Results for solar X-ray (SXR) induced effects in 1968

During solar X-ray flares the lower ionosphere on the daylit hemisphere of the Earth experiences abnormal ionization at VLF reflecting heights. The characteristics of such events will not be discussed at length here, but a sample recording for a flare on 8 July 1968 is shown in Figure 2, in which the A-C phase difference at 10.2 kHz recorded at Ny Ålesund is reproduced. Typical for such SXR events on VLF is the rapid increase to maximum phase offset, followed by a much slower recovery (time is running towards left in Figure 2). The example shown resulted in ≈60 cec offset and total duration was ≈175 min. This event was the largest (in phase offset and duration) detected during 1968 at Ny Ålesund for the A-C pair. The X-ray flux in the 2 - 12 Å band increased by a factor of 150 during peak disturbance (Solar Geophysical Data, IER-FB-293, Jan 1969).

The corresponding apparent movement of the receiving site is shown on the map for a few selected times during the disturbance, cf Figure 3. Data for A-B and A-C were used and the corresponding lane numbers are given.

In Table 12 are shown the mean duration and mean phase difference offset for SXR induced effects on OMEGA pairs A-B (received at Oslo) and A-C (received at Ny Ålesund) during 1968. The corresponding standard deviations are also given. Only clearly identifyable SXR induced effects (offset > 5 cec) were included. Total duration of disturbed periods was ~1% of total observation time in 1968.

Similar results using a much larger database (9 years) were obtained by Swanson (private communication, 1976).

In Figure 4 is shown the phase difference offset vs number of flares in 1968 for the same two OMEGA pairs. The really large flare effects (>50 cec) are occurring for about 10% of the flares.

5. EFFECTS FROM SOLAR PROTON EVENTS ON OMEGA PHASE DIFFERENCE VALUES

5.1. Results for solar proton events during 1968-70

Some typical effects on VLF phase and amplitude resulting from the ionization by energetic solar protons of the lower strata of the ionosphere at high latitudes (>55-60° geomagnetic) are illustrated in Figure 5. The corresponding propagation paths are shown in Figure 6. The records in Figure 5 show that such events may last for several days, up to 10 - 14 as a maximum. The shaded area in the three phase records represent the offset from average variation; as can be noted the effects started a few hours after the solar flare, which in addition to hard X-rays also produced energetic protons responsible for the prolonged effect.

Figure 5 also shows the amplitude of the transmitters monitored. Typically, the amplitude is not changed significantly at OMEGA frequencies during SPE events, except for paths crossing certain areas, e g the Greenland icecap. Thus for NPG to Tromsø (upper inset) reductions (>10 dB) are recorded.

Figure 7 exhibits the phase deviation recorded during an SPE in February/March 1969 on selected OMEGA pairs at two observing sites. Phase offset in excess of 70 cec were observed for A-C recorded at Oslo. The horizontal lines give the approximate boundaries for the diurnal phase variations during undisturbed conditions. Negligible effects are present on the A-B pair recorded at Oslo. This SPE event therefore did not produce significant ionospheric effects below ≈60° geomagnetic; note, however, that the path from Forestport to Oslo was affected.

Figure 8 gives a comparison with derived phase difference offset for A-C recorded at Hammerfest and the solar proton fluxes (>10 MeV) for this event. A weak, but probably consistant tendency to slightly lower phase offsets at night is noticeable. This is due to the D-region chemistry which reduces the number of free electrons during nighttime (given constant production rate of ionpairs). The proton flux data are taken from Solar Geophysical Data (301/303 - Part II, Sept and Nov, 1969).

Position accuracies are degraded during SPE events. As an example, Figure 9 shows the apparent movement of the receiving site at Ny Ålesund during the initial hours of the SPE starting on 2 November 1969. This SPE was the largest during the recording period from 1968-70. Position errors up to 7-8 km are apparent.

Table 13 summarizes the maximum phase offsets during SPEs for A-C recorded at Ny Ålesund during 1968-70. (The data do not cover all events during this period due to equipment failure etc.) A few of the periods listed cover the duration of several successive SPE disturbances. The median offset value, excluding the event in January 1969, for the remaining 13 events listed in Table 13 equals 38 cec with a standard deviation of 11 cec.

As to the total duration of disturbances due to SPE effects no accurate analysis of these OMEGA data has been made. As an upper limit, however, we may take the time during which protons >10 MeV were present in the Earth's environment. Protons above these energies will penetrate to levels below 63 - 65 km. From the data published in Solar Geophysical Data (issues in 1968-69) it is estimated that fluxes >10 MeV were present above background during ≈1500 hours in 1968, corresponding to about 17% of total time in that year. This is probably a too high estimate of duration of SPE activity on OMEGA. Westerlund et al (1968) derived a value for the minimum detectable proton flux using VLF waves as a diagnostic tool of 0.5 protons/cm^{2}s sr above 25 MeV. From the tables in Solar Geophysical Data protons >30 MeV with larger fluxes than 0.9 protons/cm^{2}s sr were present during approximately 550 hours in 1968. This corresponds to ≈6% of total time in 1968, during which the solar sunspot activity maximized.

This value is in general agreement with Martin (1972) who found that SPE events disturbed OMEGA phase measurements during ≈5% of the time for OMEGA paths to Wales, Alaska in 1969/70.

6. CONCLUSION

The phase measurements reported in this paper refers to signal transmissions from OMEGA stations in 1968-70. During these times the radiated power was not that foreseen in the implemented OMEGA system. Unfavourable signal to noise ratios may therefore at times have caused erratic phase variation in addition to possible propagation effects.

The monthly values of OMEGA phase variability are less than 10 cec for most months between 1968-70. Typical values are 4-9 cec, with lowest values during summer months. Approximately, a measurement accuracy with a standard deviation of 10 cec on each of two lines of positions will translate into a positional fix accuracy of 1.8 km (rms) under favourable systems geometry.

In 1968 solar X-ray (SXR) flare effects on OMEGA had a mean duration of individual events of ±55 min with a standard deviation of ±30 min. One SXR event disturbed the OMEGA phase measurement for 175 min. Mean phase deviation was found to be ±23 cec with a standard deviation of ±16 and 11 cec, for two OMEGA pairs.

The maximum phase deviations observed during 1968-70 on A-C at Ny Ålesund during 13 solar proton events amounted to 36 cec as mean for the maximum offset with a standard deviation of 11 cec. Maximum observed value was 60 cec.

7. ACKNOWLEDGEMENT

The author is grateful to Prof A Egeland, the Norwegian Institute of Cosmic Physics, and Mr E R Swanson, Naval Electronics Laboratory Center, for having supplied OMEGA phase difference data. The OMEGA phase measurements made by the Norwegian Defence Research Establishment were carried out under contract with Norwegian Defence Communications Administration/US Department of the Navy.

REFERENCES

Martin, J.N., 1972, "OMEGA phase variations during PCA events", Naval Electronics Laboratory Center, TR-1835, San Diego, USA.

Solar Geophysical Data, ESSA Research Laboratories, US Department of Commerce (various issues in 1968-1969).

Westerlund, S., F.H. Reder and C. Åbom, 1969, "Effects of polar cap absorption events on VLF transmissions", J Atmosph Terr Phys, 17, 1329-1374.

Table 1 List of OMEGA recording stations used during this study. Distance between recording stations and the OMEGA transmitters are given (in km).

Recording sites		OMEGA STATIONS A Aldra	B Trinidad	C Haiku	D New York
Oslo	60°N	730	8022	10933	5768
Hestmona	67°N	18	8233	10223	5889
Keflavik	64°N	1671	6653	9783	3983
Tromsø	70°N	434	8530	9901	5706
Hammerfest	71°N	639	8715	9791	5826
Ny Ålesund	79°N	1396	8493	8853	5239

Table 2 MONTHLY STATISTICS ON OMEGA SIGNAL VARIABILITY AT 10.2 kHz FOR SEVERAL RECORDING SITES in 1968.

The standard deviation of the observed phase difference values was computed for each hour of the day on a monthly basis. (Data recorded during solar proton events were excluded.)

The table gives the r.m.s. variation (in cec) of these hourly standard deviations grouped into a 24-hour period for OMEGA pairs NORWAY minus TRINIDAD (A–B) and NORWAY minus HAWAII (A–C)

OMEGA PAIR
A-B (10.2 kHz)

Sites/Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
	1968											
Oslo	5.65	6.17	4.28	4.91	4.82	2.22	5.82	3.88	4.12	6.07	8.03	4.88
Hestmona	6.24	8.21	6.85	5.19	5.13	5.79	4.28	4.94	5.09	9.49	8.02	6.35
Keflavik	-	-	-	5.05	7.08	6.19	-	7.62	-	7.08	8.64	5.00
Hammerfest	5.13	7.53	6.95	5.57	5.01	4.16	4.25	4.60	5.32	8.60	6.99	6.92
Ny Ålesund	5.71	7.19	6.79	6.32	5.60	6.05	7.33	6.50	6.15	5.70	7.56	6.79

OMEGA PAIR
A-C (10.2 kHz)

Oslo	9.32	8.70	8.76	7.97	7.29	5.70	7.97	6.98	8.08	8.47	9.07	7.28
Hestmona	9.99	8.39	8.52	11.47	4.58	4.68	4.73	5.17	6.71	7.13	9.80	6.76
Keflavik	No recordings											
Hammerfest	6.82	6.38	8.12	8.87	5.11	4.87	5.35	6.55	7.56	7.82	7.67	7.88
Ny Ålesund	6.34	5.84	6.54	8.26	3.95	3.67	4.11	5.01	6.47	5.40	5.69	5.73

Table 3 MONTHLY STATISTICS ON OMEGA SIGNAL VARIABILITY AT 10.2 kHz FOR SEVERAL RECORDING SITES IN 1969.

The standard deviation of the observed phase difference values was computed for each hour of the day on a monthly basis. (Data recorded during solar proton events were excluded.)

The table gives the r.m.s. variation (in cec) of these hourly standard deviations grouped into a 24-hour period for OMEGA pairs NORWAY minus TRINIDAD (A–B) and NORWAY minus HAWAII (A–C).

OMEGA PAIR A-B (10.2 kHz) 1969												
Sites/Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Oslo	4.74	5.05	8.69	5.02	5.01	-	3.06	3.73	4.71	-	-	-
Hestmona	5.79	7.16	8.62	6.14	7.00	5.73	4.03	5.05	5.28	9.17	7.91	6.44
Keflavik	5.00	5.28	6.94	4.54	4.71	3.89	2.86	4.36	4.01	5.29	7.38	5.56
Hammerfest	6.30	6.76	7.49	6.96	5.90	5.61	4.32	4.52	5.82	-	-	-
Ny Alesund	6.75	6.54	6.90	4.59	5.34	6.20	6.52	6.52	-	-	-	-
OMEGA PAIR A-C (10.2 kHz)												
Oslo	7.63	6.27	11.29	8.47	8.50	-	5.76	6.02	7.19	-	-	-
Hestmona	5.36	6.20	12.97	12.11	6.47	5.75	3.40	3.73	6.31	12.11	7.96	5.81
Keflavik	No recordings											
Hammerfest	7.47	7.32	10.83	5.74	5.95	5.74	3.67	-	10.42	-	-	-
Ny Alssund	6.21	7.20	10.40	8.41	4.74	6.79	2.69	4.03	-	-	-	-

Table 4 MONTHLY STATISTICS ON OMEGA SIGNAL VARIABILITY AT 10.2 kHz FOR SEVERAL RECORDING SITES IN 1970.

The standard deviation of the observed phase difference values was computed for each hour of the day on a monthly basis. (Data recorded during solar proton events were excluded.)

The table gives the r.m.s. variations (in cec) of these hourly standard deviations grouped into a 24-hour period for OMEGA pairs NORWAY minus TRINIDAD (A–B) and NORWAY minus HAWAII (A–C).

OMEGA PAIR A-B (10.2 kHs) 1970												
Sites/Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Oslo	5.53	6.20	-	-	4.36	5.54	3.90	1.45	4.82	5.72	5.40	5.50
Hestmona	6.59	7.21	8.54	7.86	5.30	5.10	4.45	-	4.97	7.27	7.80	7.79
Keflavik	6.47	7.81	6.67	7.41	-	-	-	-	-	-	-	-
Tromsø	11.60	8.95	-	-	4.91	6.47	5.39	2.49	6.44	9.93	8.35	6.93
Ny Alssund	-	11.94	13.07	13.89	13.63	11.79	9.71	6.50	6.33	6.56	7.16	7.48
OMEGA PAIR A-C (10.2 kHs)												
Oslo	7.89	6.79	-	-	8.91	7.78	7.94	2.78	10.19	12.80	7.76	7.17
Hestmona	7.56	7.94	10.34	9.14	4.54	4.85	5.46	-	9.81	11.90	6.79	7.17
Keflavik	No recordings											
Tromsø	14.27	11.16	-	-	6.94	9.20	10.21	3.27	12.03	12.55	7.78	10.58
Ny Alesund	-	6.65	18.17	10.37	6.26	5.65	7.00	2.54	5.52	9.13	7.12	8.09

Table 5 MONTHLY STATISTICS ON OMEGA SIGNAL VARIABILITY AT 10.2 kHz RECORDED AT OSLO, NORWAY AND KEFLAVIK, ICELAND DURING 1968–1970.

The standard deviation of the observed phase difference values was computed for each hour of the day on a monthly basis. (Data recorded during solar proton events were excluded.)

The table gives the r.m.s. variation (in cec) of these hourly standard deviations grouped into a 24-hour period for the OMEGA pairs NORWAY minus FORESTPORT (A–D).

OMEGA PAIR
A-D (10.2 kHz)

OSLO SITE

Year/Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1968	6.85	8.42	8.40	7.08	7.18	5.51	7.37	7.03	7.69	8.60	10.29	8.48
1969	8.63	7.85	7.53	7.22	7.52	-	5.40	5.76	7.43	-	-	-
1970	7.21	7.87	-	-	6.89	6.74	7.16	2.24	7.34	8.92	9.82	7.96

OMEGA PAIR
A-D (10.2 kHz)

KEFLAVIK SITE

1968	-	-	-	-	-	-	-	-	-	-	4.51	6.38
1969	6.46	6.07	5.85	4.88	4.71	4.02	3.17	-	-	6.30	6.84	5.59
1970	5.99	5.80	7.62	4.99	-	-	-	-	-	-	-	-

Table 6 MONTHLY STATISTICS ON OMEGA SIGNAL VARIABILITY AT 10.2 kHz FOR SEVERAL RECORDING SITES IN 1968.

The standard deviation of the observed phase difference values are computed for each hour of the day on a monthly basis. (Data recorded during solar proton events were excluded.)

The table gives the r.m.s. variation (in cec) of these hourly standard deviations grouped into daytime, nighttime and 24-hour periods for the OMEGA pair NORWAY minus TRINIDAD (A–B).

OMEGA PAIR
A-B (10.2 kHz)

Daytime, 1968

Sites/Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Oslo	-	4.97	2.70	3.44	3.86	2.31	5.59	2.64	2.99	6.01	7.46	-
Heetmona	-	9.16	7.36	4.09	3.30	4.88	3.03	2.90	4.05	10.35	9.35	-
Keflevik	-	-	-	4.79	6.84	6.69	-	7.18	-	8.90	-	-
Hammerfest	-	8.45	7.03	3.67	3.40	3.60	3.37	4.08	3.88	9.81	-	-
Ny Alesund	-	-	-	6.48	5.91	6.71	8.21	7.39	7.48	-	-	-

Nighttime, 1968

Oslo	5.48	5.32	5.56	4.68	-	-	-	4.16	5.95	5.56	5.15	4.57
Heetmona	5.68	6.31	5.07	4.55	-	-	-	6.20	5.62	7.61	6.96	5.71
Keflevik	-	-	-	4.79	-	-		-	-	4.97	7.18	4.81
Hammerfest	4.07	5.08	5.62	4.94	-	-	-	-	4.91	5.18	5.86	5.33
Ny Alesund	4.45	5.83	4.61	-	-	-	-	-	4.45	3.42	5.98	6.36

24-houre, 1968

Oslo	5.65	6.17	4.28	4.91	4.82	2.22	5.82	3.88	4.12	6.07	8.03	4.88
Heetmona	6.24	8.21	6.85	5.19	5.13	5.79	4.28	4.94	5.09	9.49	8.02	6.35
Keflevik	-	-	-	5.05	7.08	6.19	-	7.62	-	7.08	8.64	5.00
Hammerfest	5.13	7.53	6.95	5.57	5.01	4.16	4.25	4.60	5.32	8.60	6.99	6.92
Ny Alesund	5.71	7.19	6.79	6.32	5.60	6.05	7.33	6.50	6.15	5.70	7.56	6.79

Table 7 MONTHLY STATISTICS ON OMEGA SIGNAL VARIABILITY AT 10.2 kHz FOR SEVERAL RECORDING SITES IN 1968.

The standard deviation of the observed phase difference values was computed for each hour of the day on a monthly basis. (Data recorded during solar proton events were exluded.)

The table gives the r.m.s. variation (in cec) of these hourly standard deviations grouped into daytime, nighttime and 24-hour periods for the OMEGA pair NORWAY minus HAWAII (A–C).

OMEGA PAIR
A-C (10.2 kHz)

Daytime, 1968

Sites/Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Oslo	-	-	-	-	5.57	-	7.03	-	-	-	-	-
Hestmona	-	-	-	10.12	3.82	4.36	4.57	4.69	-	-	-	-
Hammerfest	-	-	-	9.46	4.94	4.68	5.19	6.58	-	-	-	-
Ny Alesund	-	-	-	-	3.43	3.64	3.69	4.64	-	-	-	-

Nighttime, 1968

Oslo	7.99	-	-	-	-	-	-	-	-	-	6.86	6.37
Hestmona	9.01	7.00	-	-	-	-	-	-	-	-	6.31	5.52
Hammerfest	7.95	-	-	-	-	-	-	-	-	-	8.00	7.47
Ny Alesund	6.42	6.96	-	-	-	-	-	-	-	-	4.61	5.91

24-hours, 1968

Oslo	9.32	8.70	8.76	7.97	7.29	5.70	7.97	6.98	8.08	8.47	9.07	7.28
Hestmona	9.99	8.39	8.52	11.47	4.58	4.68	4.73	5.17	6.71	7.13	9.30	6.76
Hammerfest	6.82	6.38	8.12	8.87	5.11	4.87	5.35	6.55	7.56	7.82	7.67	7.88
Ny Alesund	6.34	5.84	6.54	8.26	3.95	3.67	4.11	5.01	6.47	5.40	5.69	5.73

Table 8 MONTHLY STATISTICS ON OMEGA SIGNAL VARIABILITY AT 13.6 kHz RECORDED AT HESTMONA, NORWAY DURING 1967–1970.

The standard deviation of the observed phase difference values was computed for each hour of the day on a monthly basis. (Data recorded during solar proton events were excluded.)

The table gives the r.m.s. variation (in cec) of these hourly standard deviations grouped into a 24-hour period for the OMEGA pairs TRINIDAD minus NORWAY (B–A) and HAWAII minus NORWAY (C–A).

OMEGA PAIR
B-A (13.6 kHz)

Year/Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1967	-	-	-	-	2.77	3.80	4.34	4.96	5.92	4.52	5.60	6.58
1968	5.86	7.45	6.92	4.45	3.96	4.64	4.35	4.48	5.60	7.90	7.02	5.54
1969	4.97	6.02	7.54	5.19	5.67	4.90	3.48	4.99	4.35	7.90	6.58	6.06
1970	6.44	6.55	7.27	6.72	4.74	3.87	3.58	-	4.47	6.05	6.94	6.78

OMEGA PAIR
C-A (13.6 kHz)

1967	-	-	-	-	2.90	5.71	4.22	4.22	7.14	12.20	6.10	8.37
1968	8.72	7.20	8.87	10.08	3.22	4.25	4.41	4.30	5.41	6.66	8.18	6.76
1969	4.78	5.16	11.78	10.04	5.53	4.50	2.69	3.36	4.97	11.69	7.23	5.63
1970	7.02	7.80	8.86	7.51	-	-	-	-	-	-	-	-

Table 9 MONTHLY STATISTICS ON OMEGA SIGNAL VARIABILITY AT 10.2 kHz and 13.6 kHz RECORDED AT HESTMONA, NORWAY DURING 1967–1970.

The standard deviation of the observed phase difference values was computed for each hour of the day on a monthly basis. (Data recorded during solar proton events were excluded.)

The table gives the r.m.s. variation (in cec) of these hourly standard deviations grouped into a 24-hour period for the OMEGA pair HAWAII minus NORWAY (C–A).

OMEGA PAIR

C-A 13.6 kHz

Year/Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1967	-	-	-	-	2.90	5.71	4.22	4.22	7.14	12.20	6.10	8.37
1968	8.72	7.20	8.87	10.08	3.22	4.25	4.41	4.30	5.41	6.66	8.18	6.76
1969	4.78	5.16	11.78	10.04	5.53	4.50	2.69	3.36	4.97	11.69	7.23	5.63
1970	7.02	7.80	8.86	7.51	-	-	-	-	-	-	-	-

10.2 kHz

1967	-	-	-	-	4.15	7.71	4.66	4.97	7.63	13.56	8.46	7.96
1968	9.99	8.39	8.52	11.47	4.58	4.68	4.73	5.17	6.71	7.13	9.80	6.76
1969	5.36	6.20	12.97	12.11	6.47	5.75	3.40	3.73	6.31	12.11	7.96	5.81
1970	7.56	7.94	10.34	9.14	4.54	4.85	5.46	-	9.81	11.90	6.79	7.17

Table 10 MONTHLY STATISTICS ON OMEGA PREDICTABILITY AT 10.2 kHz FOR SEVERAL RECORDING SITES DURING 1968–1970.

The r.m.s. variation of the observed phase difference values from the predicted values for OMEGA station pairs were computed for each hour of the day on a monthly basis. (Data recorded during solar proton events were excluded and the skywave correction model of 1969 was used.)

The table gives the r.m.s. variation (in cec) of these hourly r.m.s. values grouped into a 24-hour period for the OMEGA pair NORWAY minus HAWAII (A–C).

OMEGA PAIR: A-C (10.2 kHz) 1968

Sites/Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Oslo	20.33	20.21	11.14	11.27	8.82	7.91	10.31	8.49	12.35	12.24	12.99	11.03
Hestmona	14.00	14.16	14.96	17.36	8.10	9.82	13.18	10.39	7.17	12.39	8.83	7.78
Hammerfest	19.75	19.01	9.10	9.16	7.44	9.35	7.70	10.23	13.52	12.22	16.41	15.33
Ny Alesund	10.68	11.32	10.70	11.13	4.65	5.20	6.84	7.55	9.08	10.17	10.12	10.82

1969

Oslo	13.69	9.37	11.96	13.02	10.08	-	18.47	14.37	10.01	-	-	-
Hestmona	6.13	10.79	16.11	20.33	11.29	16.50	24.24	18.19	9.03	19.02	9.15	7.51
Hammerfest	15.99	10.79	10.62	11.58	7.13	7.34	9.15	13.38	-	-	-	-
Ny Alesund	12.49	9.18	11.36	17.46	6.02	9.20	13.96	9.59	-	-	-	-

1970

Oslo	8.16	8.81	-	-	9.79	10.35	12.18	11.18	12.37	14.05	10.37	11.49
Hestmona	10.38	10.25	23.52	15.36	10.35	12.41	16.21	-	12.57	19.12	7.83	7.93
Tromsø	15.90	13.69	-	-	8.54	9.87	11.38	6.77	12.12	16.55	12.56	16.22
Ny Alesund	-	9.75	20.30	11.15	7.85	6.82	10.61	13.19	14.54	11.10	9.01	10.41

Table 11 MONTHLY STATISTICS ON OMEGA PREDICTABILITY AT 10.2 kHz FOR SEVERAL RECORDING SITES DURING 1969.

The r.m.s. variation of the observed phase difference values and the predicted values for OMEGA station pairs were computed for each hour of the day on a monthly basis. (Data recorded during solar proton events were excluded and the skywave correction model of 1969 was used.)

The table gives the r.m.s. variation (in cec) of these hourly r.m.s. values grouped into a 24-hour period for the OMEGA pair NORWAY minus TRINIDAD (A–B).

OMEGA PAIR: A-B (10.2 kHz)

1969

Sites/Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Oslo	7.38	7.62	10.44	7.51	7.19	-	8.83	7.44	7.93	-	-	-
Hestmona	6.96	14.95	12.27	13.38	17.23	10.90	5.75	8.56	11.15	12.50	10.58	8.09
Keflavik	8.71	11.33	8.90	6.44	12.28	9.35	7.82	6.51	6.63	9.67	8.76	6.76
Hammerfest	12.50	14.12	14.42	18.47	16.78	12.87	7.72	8.61	11.60	-	-	-
Ny Ålesund	14.17	14.09	16.23	10.97	11.23	14.65	19.92	19.11	-	-	-	-

Table 12 Solar X-ray flare effects upon OMEGA propagation in 1968. The duration (in min) and maximum phase offset (in cec) have been evaluated for all SXR events evident in the OMEGA NORWAY minus TRINIDAD (recorded at Oslo) and OMEGA NORWAY minus HAWAII (recorded at Ny Ålesund) data. The table gives the mean duration and offset as well as the corresponding standard deviations.

SXR-EFFECTS ON OMEGA PROPAGATION AT 10.2 kHz IN 1968

	A–B OSLO	A–C NY ÅLESUND
MEAN DURATION (min)	56	53
ST DEV (min)	30	31
MEAN PHASE OFFSET (cec)	23	22
ST DEV (cec)	16	11

PRESENCE OF SXR-EFFECTS IN 1968:

< 1%

OF TOTAL OBSERVATION TIME

Table 13 Estimated maximum phase deviation during solar proton events for OMEGA PAIR A–C (10.2 kHz) recorded at Ny Ålesund, Svalbard.

Solar proton events (dates)	Phase deviation (in cec)
26 Apr - 28 Apr '68	25
9 Jun - 12 Jun "	35
6 Jul - 30 Jul "	30
26 Sep - 9 Oct "	45
18 Nov - 21 Nov "	40
3 Dec - 9 Dec "	30
24 Jan - 27 Jan '69	>10 *
25 Feb - 5 Mar "	50
21 Mar - 23 Mar "	35
30 Mar - 3 Apr "	45
11 Apr - 17 Apr "	40
13 May - 15 May "	20
2 Nov - 10 Nov "	60
5 Nov - 8 Nov '70	30

* No data for 24 and 25 January 1969

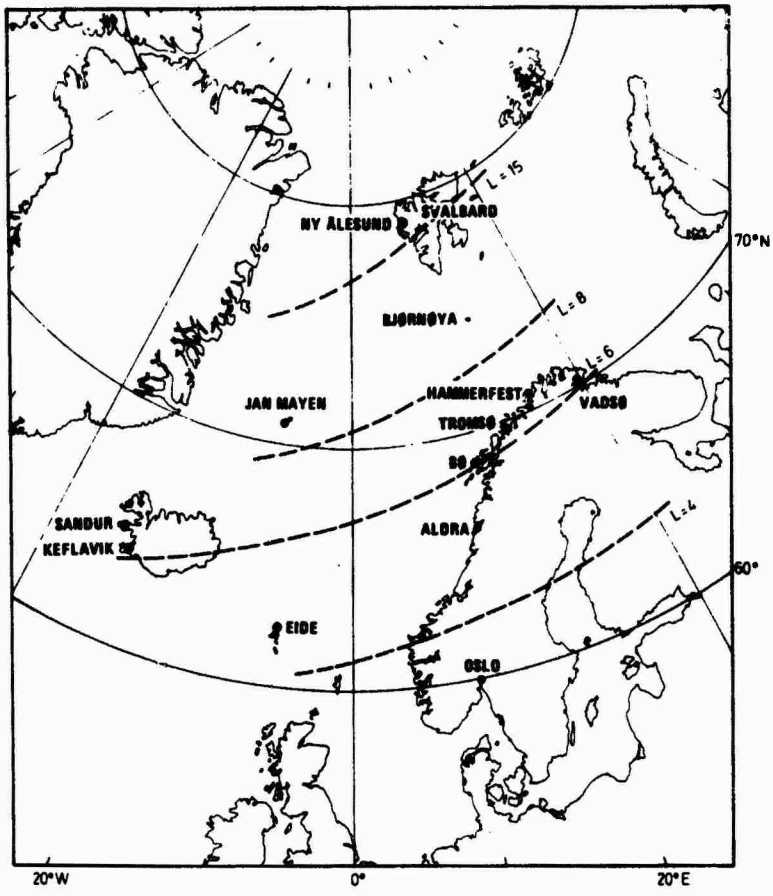


Figure 1. Map showing the positions of OMEGA recording sites at Oslo, Hestmona (Aldra), Keflavik, Tromsø, Hammerfest and Ny Ålesund. Lines of constant L-value are drawn for representative values.

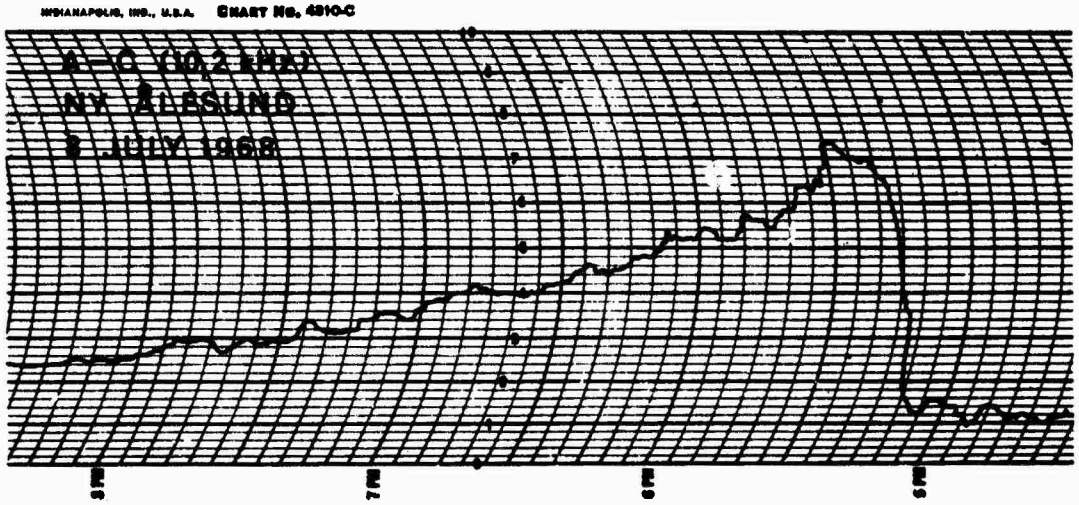


Figure 2. Phase difference recording of OMEGA NORWAY minus HAWAII (10.2 kHz) made at Ny Ålesund on 8 July 1968. This SXR flare disturbed the phase values for ≈175 min; maximum phase offset ≈60 cec.

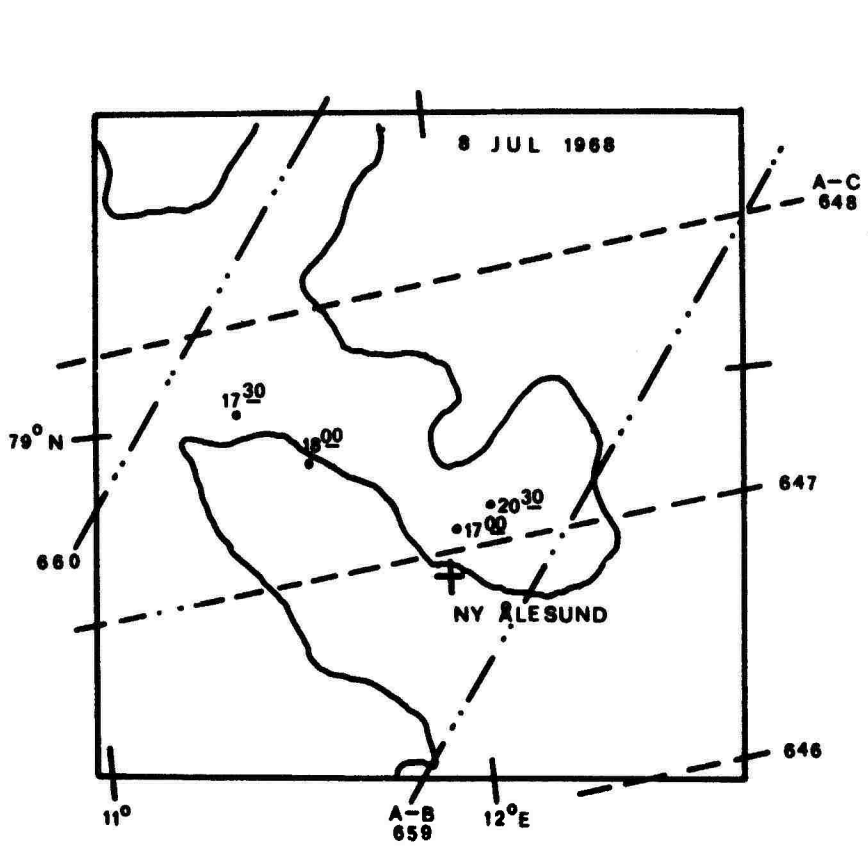


Figure 3. OMEGA position fixes at Ny Ålesund during the SXR flare on 8 July 1968.

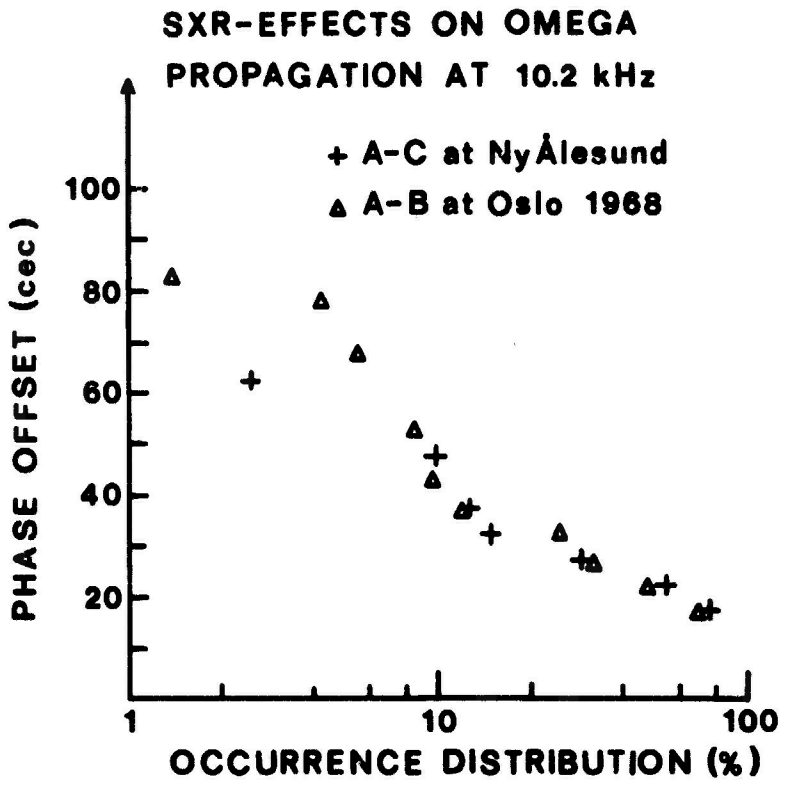


Figure 4. OMEGA phase difference offset vs frequency of occurrence of SXR-effects in 1968 for two station pairs.

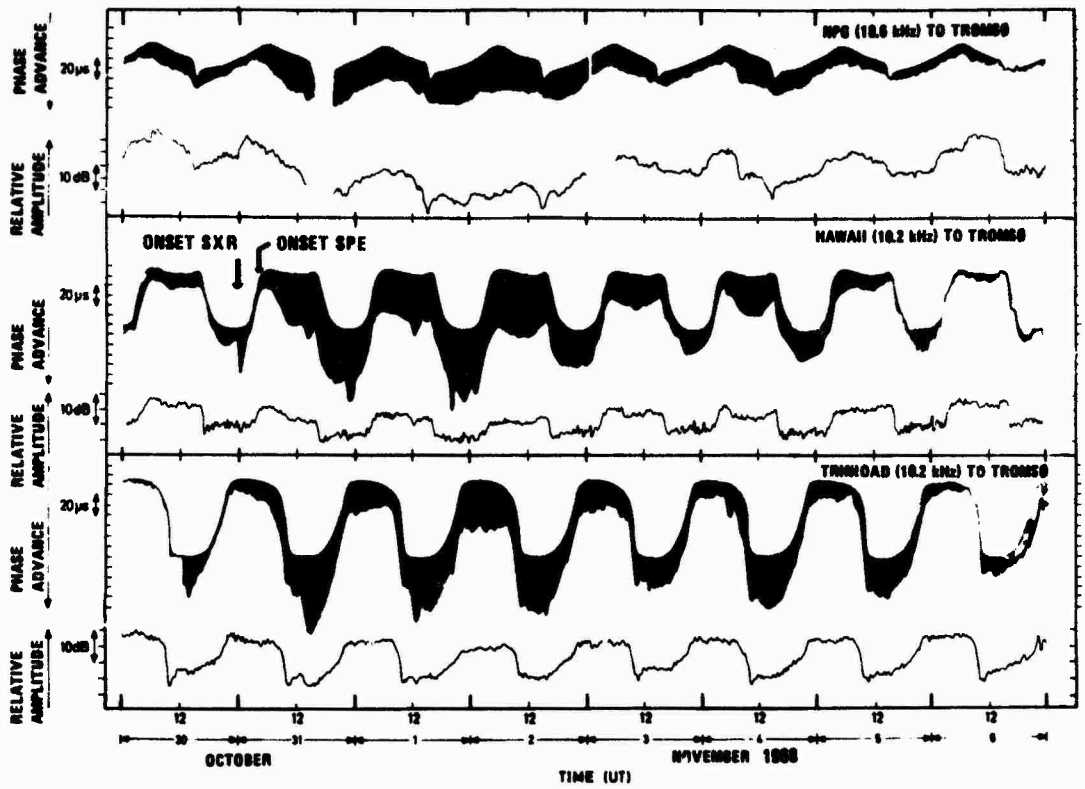


Figure 5. VLF recordings at Tromsø during a solar proton event starting on 31 October 1968. Phase and amplitude curves are given for three VLF paths (cf map in Figure 6). The shaded area represents the deviations from average phase variation.

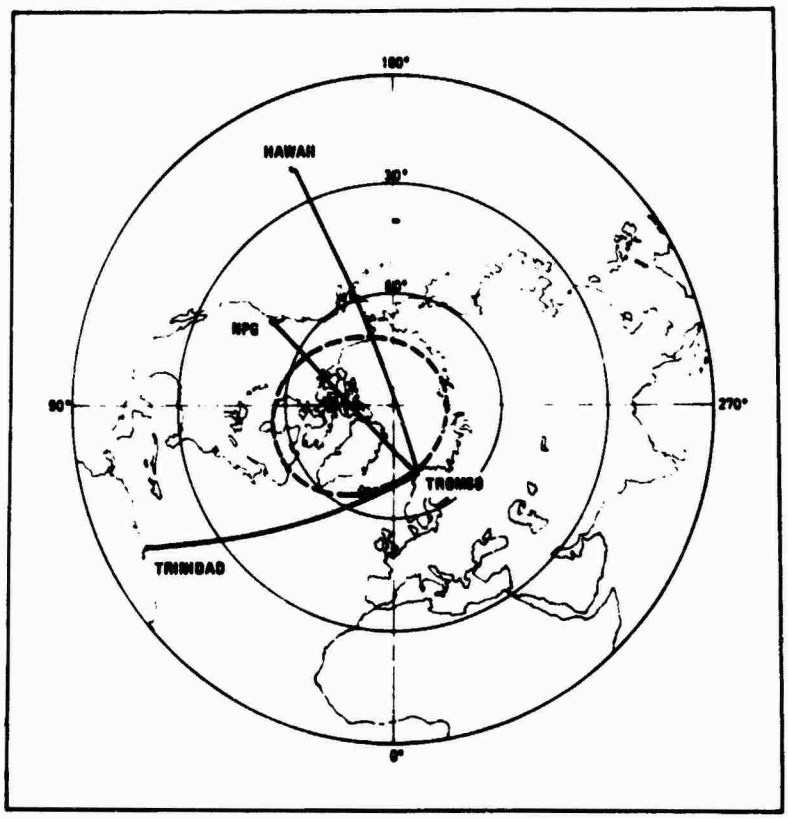


Figure 6. Map showing the propagation paths from NPG (Seattle, USA) and the OMEGA transmitters Trinidad and Hawaii to Tromsø. The position of the auroral zone is indicated by the dashed curve.

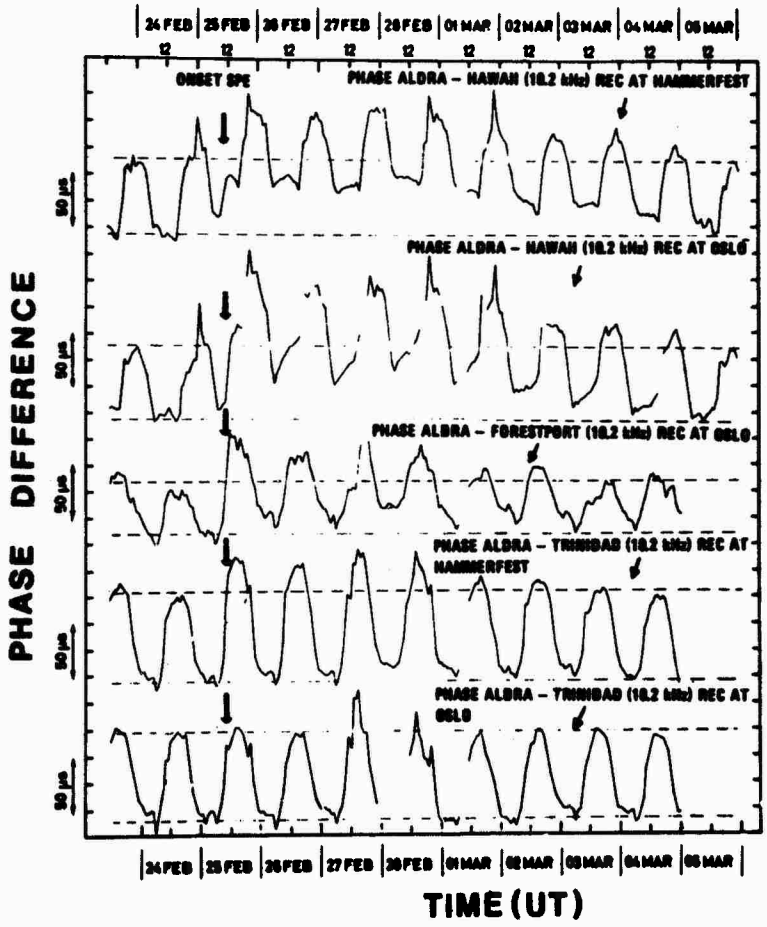


Figure 7. Phase difference variations for selected OMEGA pairs recorded at the Oslo and Hammerfest sites during a solar proton event in February/March 1969. The pairs of dashed, horizontal lines mark the boundaries for normal diurnal phase variations for the different OMEGA station combinations.

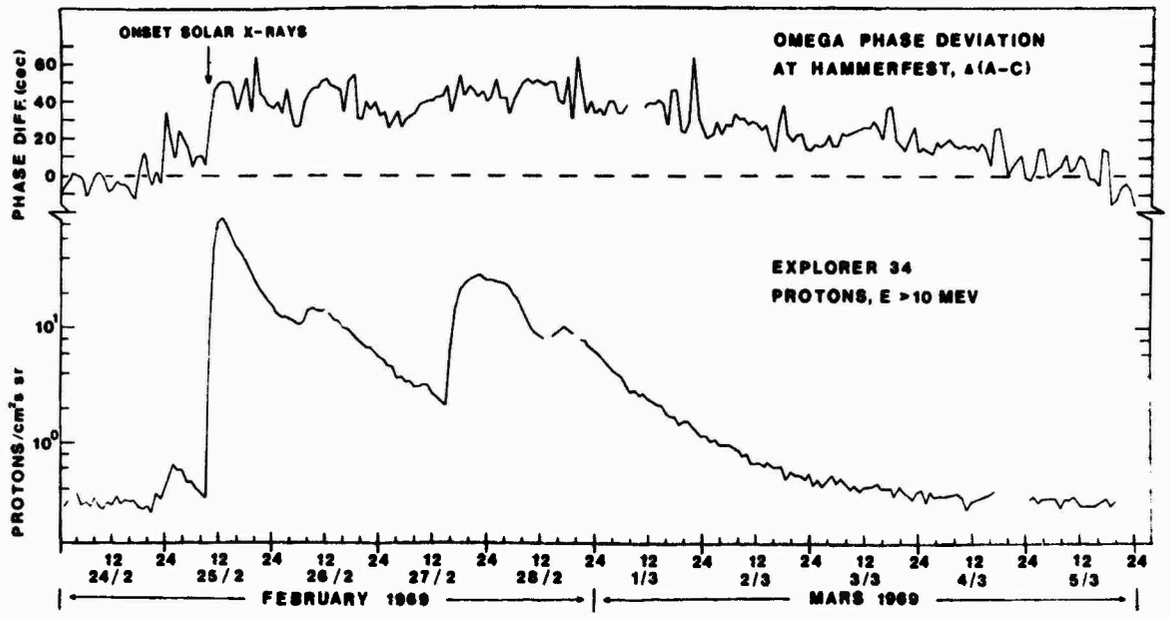


Figure 8. OMEGA phase difference offset for NORWAY minus HAWAII (10.2 kHz) received in Hammerfest during SPE in February/March 1969 (top) compared with fluxes of energetic solar protons (>10 MeV).

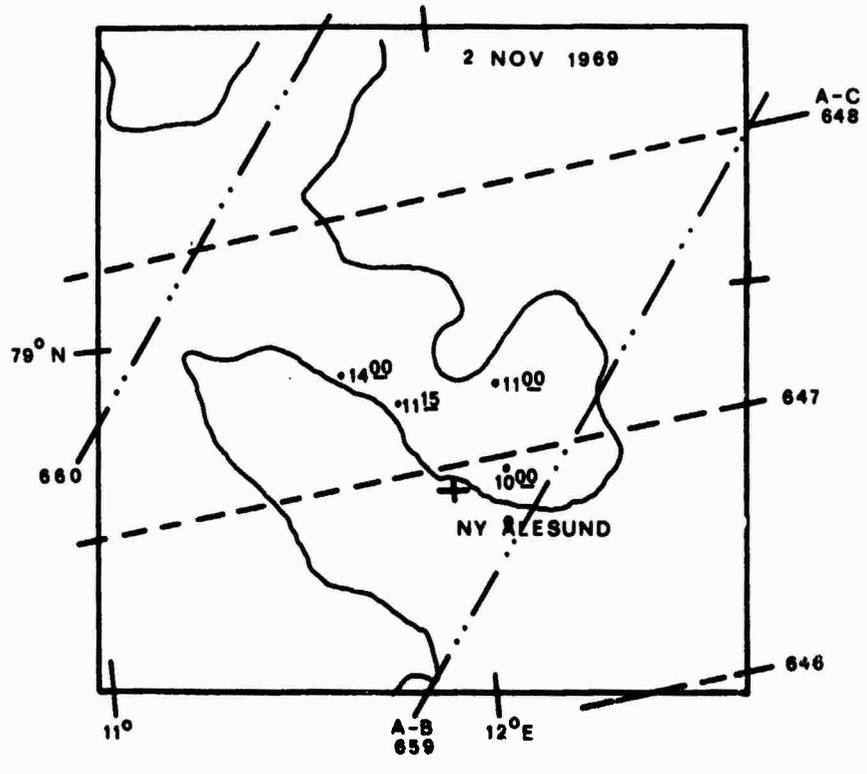


Figure 9. OMEGA position fixes at Ny Ålesund during the SPE on 2 November 1969.

DISCUSSION

E.R.Swanson, Naval Electronics Laboratory Center, San Diego, CA 92152, USA

The reference to me as a source of data is appreciated. However, I would like to clarify its meaning. I will admit to being associated with the monitoring program and take the blame for any errors in computation or editing, however, I did not personally make the measurements. 300,000 hours of data did not measure themselves. Several Norwegians were responsible. H.P.Rise was in charge at Hestmona, while Dr Larsen was personally responsible for the monitoring at Oslo and Hammerfest (Tromsø).

SHORT RANGE NAVIGATION REQUIREMENTS FOR TRANSPORT SYSTEMS

Louis W. Roberts
George G. Haroules
U. S. Department of Transportation
Transportation Systems Center
Cambridge, Massachusetts 02142
United States of America

SUMMARY

The Short Range Navigation Requirements for Trsnsport Syatems are presented. A brief discussion of the overell errors affecting the accuracy of radio navigations syatema is discussed in terms of their impact on the requirement. Several scensrios typical of transport problems with their inherent accuracy requirements are also presented to emphasize the propsgation problems that need to be satiafied as a function of frequency. A recently discovered principle of interferometry is submitted as a solution to solve the transport requiremenra.

1. INTRODUCTION

The evaluation of methods for the generation, rediation and reception of radio waves has made practical their use for the purpose of navigation and position determination. Early techniques produced results of limited accuracy. However with advancea in radio techniques and technology have come improvements in the accuracy end reliability and extensions to the applicability of radio methods of distance measurement and position fixing.

The measurement of distance by radio had its origin in the original experiments of Heinrich Hertz when he demonstrated that electromagnetic waves could be concentrated into beams by means of parabolic mirrors. The first prectical applications of radio to diatance measurement employed the principle of direction finding. In fact, the early history of poaition finding end distance measurement by radio is the history of direction finding. From the direction finding measurement, the line of position of the receiver from the transmitter can be determined and drawn on a map. The intersection of two lines of position from two known separated transmitter locations gives a position fix. The distance of the receiver from each of the transmitters then can be determined by calculation or by scaling from a map.

At the present time, radio techniques are used in such diverse requirements aa marine and air navigation and guidance, aircraft terrein cleerance and landing, colliaion avoidance of vehiclea, aurveying and geodesy, in meteorology for measurement of air mass movements, storm location and hurricane tracking, radio astronomy, and the precise determination of the location of people, places, things end events. Such a surveillence system of land vehicles is emerging as a new service called Automatic Vehicle Monitoring.

The feasibility, range and accuracy of redio distance measurement depends ultimately on the propagation characteristics of the redio waves which form the besis of the measurement.

Since in a radio measurement, distance is determined by e meesurement of time, it is necessary to know the velocity end the peth of the wavea in order to determine the geometric distence. The propagation velocity end the path both ere determined by the propagation cherecteristics of the medium. It is possible in generel because of the complex structure of the etmosphere for radio waves to arrive at receiving location vie more than one peth. In discussing propagation cheracteristics, it is convenient to divide them into the categories of ground wave, tropospheric end ionospheric charecteristics. To these must be added the cheracteristics of noise end interference, since the emplitude of the signel which is required for a useful meesurement depends on the noise level in the receiver. Those propegation errors which limit the performance in terms of the accurecy of a radio/navigation system may be clessified es the propagation velocity error, the propegation peth error end the propegetion multipeth error. The velocity error erises from the use of an incorrect value of velocity to convert the time meesurement into distance. The peth error is usually ceused from the effects of atmospheric refraction and dispersion. The multipeth error arises as the result of two redio wevea of varying phese difference being veried in phase with respect to the reference weve. An example would be the contamination of ground weves by sky weves or noise. Another example is shown in Figures 1 end 2 in which energy arrives at the receiver via different paths.

1.2. Clesses of Radio Nevigation and Positioning Systems.

Conventional redio systems used for distance measurement and position fixing can be classified into three cetegories es follows; (1) Phase systems; (2) Pulse systems; end (3) Pulse end phesed systems. Distence is determined correspondingly in terms of peth phese, pulse deley or a combinetion of the two. Thus, e system which provides a position fix may be effected either by the meesurement of two beerings, one distance end one beering; or two distences. The last method cen be further subdivided since either the distence may be measured directly or the sum or difference of two distences may be measured. Radio systems may be charecterized by the types of signel used es well es by the principel application.

Radio systems used for an angle only measurement do not perform a distance measurement but can be used to obtain a position fix. Such radio systems determine only a direction. If transmitter or receiver position is known, then this direction leads to a line of position. To obtain a position fix with such a system required that at least two intersecting lines of position be measured. Such a fix requires a known beseline. The two types of angle only systems most commonly used are redio direction finding end redio beacons.

Teble I presents e classificetion of systems under the categories of type of distence meesurement, type of position fix, type of signel trensmitted end primary applicetion.

1.3. Nature of Accuracy of Radio Navigation Systems.

The overall error of a radionavigation measuring system may be divided into the following parts: (1) the propagation error; (2) the instrumentation error; and (3) the observation error.

The first type of error is external to the instrumentation of the navigation system while the last two types are internal. For position fixing systems, these errors may combine in different ways for different parts of the operating area. The propagation error for example, may be a function of direction, distance and station configuration.

The division of internal errors into categories of instrumental and observational is not necessarily based on the function performed. For example, in some systems a given function is performed by a human observer while in other systems the same function is performed by a part of the electronic equipment. In the former case an error in performing the function would be classed as an observational error while in the latter case, it would be classified as an instrumental error.

Errors in radio navigation systems are also classed as systematic errors or random errors. A systematic error is one which obeys a definite law with respect to some cause; usually not known. An example is the use of an incorrect value of velocity of propagation. Systematic errors are constant or of the same sign from one observation to the next. Random errors, on the other hand result from chance variations in equipment or observation and may be positive or negative with equal likelihood. Hence, the mean of a large number of observations will reduce the effect of random errors. The mean will not eliminate systematic errors however.

System errors eventually can be detected and allowed for, even though the cause may not be known. Random errors on the other hand impose a limit on the reliability of the measurement.

1.4. Errors in Position Fixing.

The error of a position fix depends on the errors in the two position lines whose intersection determines the fix. The size and shape of the region of uncertainty surrounding a fix is different for different systems. Hence, the shape of the error contours varies with the type of system. Radio Navigation Systems fall into the following categories:

(1) Angle only or azimuth systems; (2) Rho-theta systems; (3) Range systems; and (4) Hyperbolic systems.

Systems using multilateration techniques or interferometric fall into the above categories or their combination.

Table II shows the comparative characteristics of some of the more recent electronic positioning systems classified basically as hyperbolic systems.

In general, the radial error of a position fix will include both system and random errors. By calibration system errors can be eliminated.

If the radial error is small compared to the distance from a station, the position line in the vicinity of the fix point may be considered to be parallel. The situation then may be represented as in Figure 3.

In this figure, P represents the true position while F represents the fix position as determined by the intersection of the position lines L_1 and L_2. The parallelogram FF^1pp^1 defines the distance error through its major diagonal $FP=D_E$.

Lines drawn through P parallel to L_1 and L_2 then define the errors of the position lines for this particular measurement which are E_1 and E_2. The angle of intersection of the position line is denoted by θ. An elementary derivation gives

$$(1) \qquad D_E^2 = \frac{E_1^2 + E_2^2 + 2E_1E_2 \cos\theta}{\sin^2\theta}$$

This applies for a single measurement. For a large number of measurements n the mean square distance error will be

$$(2) \qquad d^2 = \frac{1}{n} \Sigma D_E^2 = \frac{1}{n \sin\theta} [\Sigma E_1^2 + \Sigma E_2^2 + 2 \cos\theta \Sigma E_1E_2]$$

$$(3) \qquad = \frac{1}{\sin^2\theta} (\sigma_1^2 + \sigma_2^2 + 2R_{12} \sigma_1\sigma_2 \cos\theta)$$

where σ_1 = the standard deviation of the error of position line L_1

σ_2 = the corresponding quantity for L_2

R_{12} = the correlation coefficient between the two errors

1.5. Economy of Spectrum Use.

Radio Frequencies are allocated in accordance with international treaty agreements expressed in the Radio Regulations of the International Telecommunications Union. Table III is a list of Radionavigation aids and their associated frequency bands that are used in Transport Systems.

1.6. Common Denominator Transport System Requirements.

The common denominator requirements of short range navigation for transport systems namely high angular precision, precise range and high data rate are the performance requirements that are imposed on a vehicle control system. In fact the requirements for navigation, position determination and surveillance systems are functionally similar, frequently identical and often not dependent on a specific vehicle type. This results as a consequence of the fact that all transport systems consist of terminals, vehicle ways interconnecting the terminals, vehicles which move on the vehicle ways and a control system to assure the efficient operation of the system.

Efficiency includes a broad spectrum of considerations. One important consideration is navigation of vehicles within the defined boundaries of the vehicle ways. Another is the avoidance of collisions with other vehicles using a common vehicle way. Such functional requirements are common to all transport vehicle systems. Though the specific nature of the control system for each vehicle system is influenced by restraints peculiar to each system, the parts of a transport vehicle control system are functionally quite similar. In transport systems the prime emphasis is on air and marine requirements. Aircraft and marine vehicles require a continuous availability of position information exclusive of the time of day, the visibility restrictions or adverse weather conditions. The safety of marine and air vehicles is predicated upon the availability and reliability of a stated or measured position as well as the prudent use of available navigation information. However, there is an emerging requirement in supplying land users with navigational aids which have similar operational features to those used by marine and air. Land users also have the prime objective of taking the most efficient path from origin to destination with the avoidance of hazards and the prevention of confliction between vehicles.

Table IV is a broad summary of navigation requirements by operational area categories for air, marine and land. Those propagation anomalies are also identified in terms of the particular system being used.

Tables V, VI, VII are the individual vehicle requirements for Air Navigation, Marine Navigation and Land Navigation. The navigation system requirements to satisfy the users needs are also identified.

Though the sensing of vehicle position in a transport system is a necessary input function of a control system, the required measurement accuracy is intimately interrelated with time as a consequence of vehicle mobility and velocity. The velocity vector (speed and direction) of a vehicle can be obtained in delayed time from a quantized picture of the vehicle position with time. A time picture of the velocity vector includes knowledge of changes in vehicle speed and direction that is the acceleration vector.

Therefore, it is clear that the precision to which vehicle velocity is known is determined by both the accuracy of position measurements and the time between position measurements on position data renewal rate.

It is from the above that the relatively simple and obvious conclusion that the knowledge of vehicle position is equally as important to the effective control system of transport vehicles as is position data renewal rates.

Short range navigation requirements for transport systems require inherent precision pointing and ranging capabilities in order to satisfy a broad variety of navigation applications.

1.7. Vehicle Position Measurement

The determination of vehicle position is just one functional part of a vehicle control system. It is, however, a very important part. The common denominator nature of this function in three specific modes of operation can be described as surveillance, guidance, and collision avoidance.

Surveillance: Vehicle surveillance refers to the measurement of the positions and motions of vehicles in a common fleet. Position measurements are performed from one or more fixed sites. A further requirement of surveillance is the identification of each vehicle with its position and motion.

One example of the surveillance function as used in this context is an air traffic control radar beacon capability. In the skin track modes, the radar measures aircraft positions and motions relative to the fixed location of the radar. The radar, however, does not provide aircraft identification. This is accomplished through simultaneous use of an air traffic control beacon system.

The function of vehicle surveillance is to provide a vehicle fleet controller with the necessary information concerning vehicle positions and identifications needed to space and sequence the vehicle along the vehicle ways in an orderly, safe and efficient manner.

Three specific surveillance mode applications of ground transport system requirements are identified and will be discussed to emphasize the problem. They are: Vessel traffic surveillance; Airport ground traffic surveillance; Automatic vehicle monitoring.

In the vessel traffic surveillance system application, the positions and motions of all tall ships in transit through a Harbor Lane are made available to a Vassal Traffic Controller. The system is fully automated. The introduction of this capability aliminates the need for marina radio raporting and significantly reducas radar surveillance and raporting requirements. An automated system of this typa would qualify as the prime system for shipping traffic surveillance with radar as a redundant support system. Marina radar (voice communication) would ba usad only to identify vassals of large hull size detected by radar, which are not equipped with the discrete address beacon system. For a proposed application of the beacon to vessel traffic surveillance, the system configuration and anticipated performance is as follows:

Requirements and Assumptions

A current objective is to provide a Category II landing capability at Logan Airport, for which the usable length of Runway 4R will be greater than the present 7,500 feet. An important requirement is introduction of a reliable means for detecting and reporting the presence of tall ships when located in or in the vicinity of the 4R approach zone. It would appear that the implementation of a responsive Vessel Traffic System is needed to solve this requirement.

Figure IV is a simplified functional block diagram depicting the integration of a Vessel Traffic System with the Air Traffic Control System at the Logan Airport. Detection and reporting of tall ships is predicated on voice communication by marine radio, radar surveillance and a ship-installed radio beacon system. A competitive solution to beacon system requirements is desired.

Beacon Configuration

Each tall vessel would be equipped with a beacon transmitter. All transmitters would operate on the same frequency. Each transmitter would include a discrete address subsystem similar to the Bell Boy page system, shown in Figure V. The receiving array would be located on Governor's Island, as shown in Figure VI. This array would provide full coverage of the Harbor Lane in two contiguous 90° sectors.

Beacons would be addressed by role call via the discrete address system which without modification can accommodate 100 ships every 20 seconds.

Performance

Ship positions are measured in azimuth angle and range, similar to a radar scope presentation. The azimuth angle accuracy is 0.05° independent of range. The range accuracy is range dependent and varies from a minimum of 9m at the intersection of the centerline of the Harbor Lane with the centerline of Runway 4R to a maximum of 43m in the EAST DECISION ZONE and 82M in the WEST DECISION ZONE.

Airport Ground Traffic Surveillance

The application of radar and beacon to airport ground traffic surveillance differs in the nature of the functional requirement, but is identical to the vessel traffic surveillance configuration in terms of functional parts. Shipping traffic surveillance in the Boston Harbor Lane is performed for the purpose of alerting landing aircraft of tall ships penetrating or about to penetrate the approach zone clearance limits to Runway 4R. Airport ground traffic surveillance is needed for the efficient control of aircraft ground traffic movement, under the conditions of poor visibility. At present, only the positions and motions of ground traffic are available to the Ground Traffic Controller through observation of the ASDE Radar display. Lack of aircraft identification seriously impedes the release of prompt and effective ground traffic control advisories. During inclement weather, the present Airport Surface Detection Equipment (ASDE) is limited in performance at its maximum range because of its operational frequency 24GHz and severe clutter.

Requirements and Assumption

The important parameters of an effective Ground Traffic Management System include the identification, position and movement of each aircraft. For other than visual identification, a radio signal must be transmitted by each aircraft. This can be accomplished by either a discrete address beacon system operating on a roll-call basis or interrogation of ground space in discrete segments. Either approach assumes availability of an aircraft installed beacon. The two techniques differ only in the method of interrogation.

System Configuration

Each aircraft is equipped with a beacon transmitter. All transmitters operate on the same frequency. Each transmitter includes a discrete address subsystem. The ground transmitter and receiving array is located on Governor's Island as shown in Figure VII. This array provides full coverage of the airport in two contiguous 90° sectors.

Performance

Aircraft positions are measured in azimuth angle and range, similar to a radar scope presentation. The azimuth angle accuracy is 0.05° independent of range. The range accuracy is range dependent and varies from 25 feet at a range of 500 feet to 40 feet at a range of 18,000 feet.

Automatic Vehicle Monitoring

Automatic Vehicle Monitoring (AVM) refers specifically to that class of electronic means used for ascertaining at any time at a central control point the location and status of individual vehicles in a fleet dispersed in an urban area. Typical fleets include buses, police cars, ambulances, etc. Vehicle surveillance in this case allows prompt redirection and scheduling of fleet vehicles to meet unanticipated variations in the need for fleet functions.

Requirements and Assumptions

An AVM system is an electronic means for ascertaining, at any time at a central control point, the location and status of individual vehicles in a fleet dispersed in an urban area. Desirable features include:

Coverage area typically 10 to 15 miles in diameter;
A minimum of 500 vehicles;
A position update rate of at least once every 10 seconds;
Vehicle position resolution of at least 50 feet;
Vehicle identification and position to be relayed to the central display point without vehicle operator assistance (automatic);
Availability of a voice communication link with each vehicle;
Availability of a visual and aural alert signal and data link (printout) for each vehicle;

Vehicle Configuration

Each vehicle would be equipped with a discrete address receiver modified to operate at a 50Hz call rate and equipped with data printout capability. Each vehicle would also be equipped with a Land Mobile transceiver, modified to transmit a 10 millisecond signal, on command from its receiver, (See Figure VIII)

Three receiving terminals (R_1,R_2,R_3/R) would be disposed at the points of an equilateral triangle approximately 10 miles on each side. (The exact triangular form and distances are not critical.) The monitoring station would be located at one of the three receiving stations. The receiver in each case would be a modified Land Mobile Band receiver, tuned to the frequency of the transmitted signal. Vehicles would be interrogated in roll-call sequence, via the discrete address system at a 50Hz rate. This system is capable of transmitting messages within the address format. This allows the capability for data printout at each vehicle as well as initiation of visual and aural signals, to alert the operator to shift to a voice channel for direct communication.

Performance

Based on a nominal separation distance between receiving terminals of 10 miles, and the proper sideband separation frequency of 20Hz, a 10 millisecond transmission period and a 0.1° phase measurement accuracy, the vehicle position/resolution capability of the system should be typically 20 feet in both the x and y map coordinates of the urban area.

Guidance

Guidance refers to those modes of operation in which each vehicle in a fleet is equipped with a receiver capable of detecting and processing signals to derive navigation input parameter values, such as the direction and distance to fixed points in a local framework of reference. The guidance function implies determination of the identification and position of one or more fixed reference points in a local reference framework. A magnetic compass is perhaps the simplest example of a guidance signal receiver and display system.

Two specific guidance control system requirements of a transport system which are identified are channel navigation and railroad grade crossing control.

Navigation of cargo shipping traffic through the narrow confines of the St. Lawrence Seaway during the winter months has been selected as the specific example of channel navigation. The system configuration and anticipated performance are as follows:

Requirements and Assumptions

All-weather navigation of cargo shipping through narrow channelways requires improvement in present methods for generating radio guidance signals. Typical examples of this need are the several narrow channel sections located along the 108 mile extent of the St. Lawrence Seaway from Lake Ontario to the 45° parallel. 48.4 miles consist of six separate sections which are relatively narrow (300 feet or less). The two longest narrow sections are each 15 miles in length. The narrowest is the Wiley Dondero Ship Channel near Massena, New York, shown in Figure IX.

A recent study of navigation guidance requirements for this part of the Seaway led to the conclusion that a position accuracy of ±5 feet cross channel and ±100 feet down channel was required. The cross channel requirement negates the efficacy of techniques such as LORAN-C or radar, unless the latter is supported by installation of a large number of land-based reflectors at precisely known locations.

Configuration

Each ship would be equipped with a 4-channel receiver operating from a single omni-directional receiving antenna. Two transmitting stations would be required for the Wiley Dondero Ship Channel; one located at the Eisenhower Lock and a second approximately 5 to 6 miles north or south of the channel near the midpoint of the narrow section. Both stations would transmit in a two element IFI Guidance Mode. The transmitter at the Eisenhower Lock would illuminate front and back 15° sectors, as shown in Figure X. Separation of these transmitting elements would be a nominal 1,500 feet. The transmitter at Station B would illuminate a single 120° sector covering the entire length of the narrow channel section. The nominal separation of transmitting elements at this site would be 250 feet.

Performance

The cross-channel position of the vessel is provided by the signal transmitted from Station A, which allows a measurement accuracy of 5 feet at a distance of 7½ miles. The down channel position of the ship is obtained by processing the signal transmitted from Station B. The down channel position accuracy varies with the position of the ship along the channel and reaches a maximum of 35 feet at either extremity of the narrow section.

Railroad Crossing

The objective at a railroad grade crossing is to prevent collisions between trains and automobiles. Within the context of this subject the associated technology does not usually fall in the category of collision avoidance. Collision avoidance modes, as used here, refer to avoidance of collisions between vehicles in a common fleet, and not between vehicles in two different fleets. The technology associated with collision avoidance of the latter type (typical of grade crossings) falls in the category of guidance.

Most frequently the right of way at the intersection is given to one fleet of vehicles, while vehicles in the other fleet assume responsibility to avoid collisions at the intersection. Vehicles charged with the responsibility to avoid collisions at an intersection must of course be informed of the presence or anticipated presence at the intersection. This process involves some form of guidance signal transmission. Simple examples of conventional guidance signals at railroad crossings are flashing lights and barriers. Visual reception of these guidance signals alerts the motor vehicle driver to bring his vehicle to a stop until the flashing lights are extinguished and the barrier raised. Accidents at grade crossings persist, since all grade crossings are not equipped with these visual guidance signals, as a consequence of the mathematics of capital expenditure requirements and maintenance costs.

In the grade crossing application of a guidance system, the motor vehicle receives a guidance signal from a transmitter located at the grade crossing, which, when appropriately coupled to the vehicle's speed control and braking system, automatically brings the motor vehicle to a safe stop, prior to the intersection. The stop control signal is deactivated after the train passes the intersection. Further discussion of the concept and the anticipated performance follows the initial discussion of requirements.

Requirements and Assumptions

Reduction of collisions and safeguarding of personnel and property at railroad grade crossings has been the subject of several investigations. A fundamental requirement is some means of communicating with motor vehicles as they approach a grade crossing. The conventional method of flashing lights, barriers, etc. provides a form of visual communication which, though effective, would require a substantial investment in installation, maintenance and operation, if introduced at all grade crossings. Some form of radio communication with motor vehicles when approaching a grade crossing would appear to be more reliable and cost effective. Communication directly with the operator of a motor vehicle introduces consideration of human factors, particularly operator reaction to the form of the alert signal. A reasonable alternative is direct radio communication with the speed control and braking mechanism of the motor vehicle. The efficacy of this form of solution would, of course, require that all vehicles be equipped with the radio control system. This would require a gradual transition over a period of several years. A logical starting point would be introduction of these systems in fleet vehicles, such as school buses, public buses, trucks and ultimately automobiles.

Configuration

Each motor vehicle in a designated fleet would be equipped with a receiver, similar to the three element receiving bar shown in Figure 8. A transmitter would be located at each grade crossing, as shown in Figure 11. Each transmitter would include a discrete address subsystem. Each train would also be equipped with a discrete address terminal manually programmed in the sequence of the address codes, corresponding to the grade crossings in their order of occurrence. The beacon address would be transmitted from the train, prior to reaching an intersection. Fleet vehicles equipped with the receivers would immediately derive a range and range rate measurement relative to the intersection. These parameter values would be introduced into the speed control and braking system in a manner that would gradually bring the vehicle to a full stop at a pre-programmed distance from the intersection.

The intersection beacon would be turned off by a signal transmitted from the train after passing the intersection.

Performance

The range resolution would be 26 feet at 500 feet, 7 feet at 250 feet and 1 foot at 100 feet. The data rate would be 20Hz. Velocity and deceleration parameters could be obtained directly from the vehicle speedometer, or by the first time derivative of the remaining distance to the intersection.

Collision Avoidance

As indicated in the prior section, collision avoidance modes of application, within the context of this report, refer to methods for preventing collisions between vehicles in a common fleet. Rear-end collisions are among the most frequent type. Approximately one-third of all collisions are rear-end collisions, resulting in a "societal cost" estimated to be measured in billions of dollars annually.

Two specific applications of a distance measurement capability to rear-end collision avoidance of vehicles in a common fleet have been identified. These include Personal Rapid Transit and rail type vehicles, as well as automobiles.

It is apparent that an effective rear-end collision avoidance system for automobiles requires that all vehicles be equipped with a cooperative control system. This capability could be gradually introduced into the fleet via new vehicles and the system become effective over a period of several years. Candidates for early introduction without awaiting infusion through new fleet stock would be school buses and trucks.

Requirement and Assumptions

The parameters which enter into maintaining a safe distance between vehicles are the relative distance, the closing rate (relative velocity) and the ground velocity of the vehicle to be controlled. The vehicle installed speedometer, of course, provides a continuous measure of ground speed. What is needed is a device for measuring the distance and relative velocity between vehicles.

A device which provides a precise range measurement capability at a high data rate is the essential requirement, since relative velocity can be obtained directly as a first time derivative of the measured distance between vehicles.

System Configuration

A field test configuration of a collision avoidance system is shown in Figure 12. The transmitter with a single radiating element is top-mounted on one automobile and the receiver with a three-element antenna is top-mounted on a second vehicle. In this test configuration, the automobile equipped with the transmitter is the "lead" vehicle and that equipped with the receiver is the "following" vehicle. The measurement of relative range and range rate between vehicles is displayed in the vehicle equipped with the receiver.

Performance

In the proposed configuration, shown in Figure 8, the receiver "bar" would be a nominal 2 meters in length to provide a sector angle coverage of ±22.5° in the forward direction. At a range of 300 meters (1,000 feet) the range resolution is 9.5 meters (31 feet). The data rate is 20Hz.

The range resolution improves as the square in reduction of relative distance between the vehicles i.e., at a distance of 100 meters (300 feet), the range resolution is 1.06 meters (3,5 feet),

Comment

Collision avoidance effectiveness of any radio technique of this type is of course dependent of equipping all automobiles with the necessary transmitters and receivers. It is not likely that this will occur in the near future; however, it is appropriate to investigate the problems and their technology now as an aid to future decisions.

Spacing Control - Personal Rapid Transit and Rail

Requirement and Assumptions

Future high performance personalized rapid transit (HPPRT) systems are predicated on the assumption that individual vehicles will carry four (4) to six (6) passengers. This allows the vehicles to be small in size and weight, leading to a concomitant small size and weight for the supporting guideway (an important cost consideration).

To meet high capacity requirements with a large number of low capacity vehicles, it is necessary that the vehicles move at nominal speeds of 20 to 40 mph between terminals at fraction of a second "headways". A one-half second headway at a speed of 20 mph is approximately 15 feet, and 30 feet at 40 mph. Though these "tailgating" distances and speeds can be accommodated by an automobile driver, the objective for an HPPRT system is to control the separation distance automatically, without operator assistance.

This requirement is markedly similar to the relative range and range rate requirements associated with automobile rear-end collision avoidance.

System Configuration

A single transmitting antenna element would be top-mounted at the rear of each vehicle, and a three-element receiving "bar" would be top-mounted at the front of each vehicle, as shown in Figure 13. The signal received at each vehicle would be introduced in the speed control loop for that vehicle.

Performance

For a receiving "bar" two meters in length, the range resolution at a separation distance of 15 feet (½ second headway at 20 mph) would be 0.25 inch. The data rate would be 100Hz. At a vehicle spacing of 30 feet, the relative range resolution would be slightly less than one inch at a data rate of 100Hz.

Intermediate Frequency Interferometry

The recently discovered principles of Intermediate Frequency Interferometry eliminate many of the restraints associated with conventional interferometry and allow opportunity to apply the inherent precision pointing and ranging capabilities of interferometry to a broad variety of navigation problems. There are no special components required for an IFI navigation sensor. In fact, component requirements are well within the demonstrated capabilities of present technology, at the level of commercially available units.

An engineering design study of the applicability of IFI techniques to the navigation requirements of transportation systems reveals the competitive nature of this technique when assessed with conventional techniques in areas of application requiring precision pointing and ranging over large azimuth angles at relatively short ranges.

Intermediate frequency interferometry techniques differ from carrier frequency interferometric techniques in the method of transmitting these related frequencies (sideband and carrier) from radiating elements located at either end of a common baseline. In carrier frequency interferometry, one antenna radiates a carrier signal, while the other radiates a sideband signal. In intermediate frequency interferometry, each antenna simultaneously transmits both a carrier and a sideband signal. The carrier frequencies are displaced to allow separate identification of the signals transmitted by each antenna. The modulation frequency, or difference frequency between the sideband and the carrier, however, is the same for each carrier and its corresponding sideband signal.

In the simplest configuration, transmitting antenna A and B are located at either end of a baseline. The nominal spacing between the antennas is one wavelength at the modulation frequency. At the receiver, the carrier and sideband signals transmitted by antenna A are heterodyned and filtered to derive the modulation frequency. In a similar manner, the carrier and sideband signals transmitted by antenna B are heterodyned at the receiver to derive the modulation frequency. Though these two modulation frequencies derived by the heterodyning of their corresponding sideband and carrier signals are at the same frequency, their phase difference is directly related to the path length difference between the receiver and the antennas A and B located at either end of the transmitting baseline. By the method of transmission and signal processing, however, the path length difference is measured in fractional wavelengths at the modulation frequency, rather than at the carrier frequency as in carrier frequency interferometry. As a result, the transmitting antennas can be configured as completely separate apertures, physically located at either end of the baseline, a significantly less complex configuration than required by carrier frequency interferometry methods, wherein the total extent of the baseline is frequently occupied by a reflecting surface illuminated by a complex multi-element feed structure.

A functional block diagram of a simple two element intermediate frequency interferometric transmitter is shown in Figure 14. The signal outputs of carrier frequency generators f_{c1}, and f_{c2}, are not necessarily phase related. Their frequency separation is determined by the ease with which they can be separately filtered and identified at the receiver. Each carrier frequency is modulated by a signal derived from a common modulation frequency generator operating at a frequency f_m. The output of each modulator is filtered to derive the original carrier frequency and one sideband.

For the purpose of the discussion which follows, assume that the upper sideband has been selected in each case. The signal transmitted from antenna A will consist of carrier f_{c1} and upper sideband $f_{c1} + f_m$. The signal transmitted from antenna B is carrier frequency f_{c2} and upper sideband frequency $f_{c2} + f_m$.

For the purposes of explanation and with reference again to Figure 13, we will assume that the geometric relationship between the receiver location and the transmitting baseline is such that the distance from antenna A to the receiver is R, and from antenna B to the receiver is R + ℓ. On arrival at the receiver the sideband signal S_1 transmitted from antenna A will have the form

$$S_1 = \cos\left[\left(\omega_{c_1} + \omega_m\right) t + \frac{2\pi R}{\lambda_{s_1}}\right] \qquad (1)$$

where:

$\omega_{c_1} = 2\pi f_{c_1}$

$\omega_m = 2\pi f_m$

λ_{s_1} = wavelength of sideband $(f_1 + f_m)$

On arrival at the receiver the carrier signal C_1 transmitted from antenna A will have the form:

$$C_1 = \cos\left[\omega_{c_1} t + \frac{2\pi R}{\lambda_{c_1}}\right] \qquad (2)$$

where:

λ_{c_1} = wavelength of carrier, f_{c_1}

Similarly, the sideband signal S_2 transmitted from antenna B will have the following frequency and phase relationship on arrival at the receiver:

$$S_2 = \cos\left[\left(\omega_{c_2} + \omega_m\right) t + \frac{2\pi R}{\lambda_{s_2}} + \frac{2\pi \ell}{\lambda_{s_2}}\right] \qquad (3)$$

where:

$\omega_{c_2} = 2\pi f_{c_2}$

λ_{s_2} = wavelength of sideband, $f_{c_2} + f_m$

and the carrier signal C_2 the following frequency and phase relationships:

$$C_2 = \cos\left[\omega_{c_2} t + \frac{2\pi R}{\lambda_{c_2}} + \frac{2\pi \ell}{\lambda_{c_2}}\right] \qquad (4)$$

where:

λ_{c_2} = wavelength of carrier, f_{c_2}

The carrier and sideband signals transmitted by each antenna are separately heterodyned at the receiver to extract their difference or modulation frequency in each case.

The heterodyning of sideband S_1 with carrier C_1 produces a modulation frequency and phase relationship M_1 in the form:

$$M_1 = \cos\left[\omega_m t + 2\pi R\left(\frac{1}{\lambda_{s_1}} - \frac{1}{\lambda_{c_1}}\right)\right] \qquad (5)$$

In a similar manner the output frequency and phase relationship obtained by heterodyning sideband S_2 with its carrier C_2 provides an output modulation frequency and phase relationship, M_2 in the form:

$$M_2 = \cos\left[\omega_m t + 2\pi R\left(\frac{1}{\lambda_{s_2}} - \frac{1}{\lambda_{c_2}}\right) + 2\pi \ell\left(\frac{1}{\lambda_{s_2}} - \frac{1}{\lambda_{c_2}}\right)\right] \qquad (6)$$

Since the modulation frequency or difference frequency between each carrier and its corresponding sideband is the same, and bears the same phase relationship to its carrier at time of transmission:

$$\frac{1}{\lambda_{s_1}} - \frac{1}{\lambda_{c_1}} \quad \frac{1}{\lambda_m} = \frac{1}{\lambda_{s_2}} - \frac{1}{\lambda_{c_2}} \qquad (7)$$

where:

λ_m = wavelength of the modulation frequency (f_m).

Introducing equation 7 in equations 5 and 6, it is apparent that the phase difference ϕ between the modulation frequency M_1 and the modulation frequency M_2 is:

$$\phi = \frac{2\pi \ell}{\lambda_m} \qquad (8)$$

Since the bearing angle θ of the receiver relative to the normal to the transmitting baseline is given by:

$$\theta = \sin^{-1}\left[\frac{\ell}{d}\right] \qquad (9)$$

where:
d is the baseline length, then

$$\theta = \sin^{-1}\left[\frac{\lambda_m \phi}{2\pi d}\right] \qquad (10)$$

The important feature to note is the marked similarity in the form of equations 6 and 7 in connection with the discussions of carrier frequency interferometry and equations 8, 9, and 10, respectively, as developed above in connection with intermediate frequency interferometry. In both cases, the path length difference ℓ between either end of the transmitting baseline and the receiver is obtained by a phase measurement which involves a determination of the path length difference ℓ in wavelengths. In the case of carrier frequency interferometry ℓ is measured in fractional wavelengths at the carrier frequency. However, in intermediate frequency interferometry, the same path length difference ℓ is measured in fractional wavelengths at the modulation frequency. Consequently, through the use of intermediate frequency interferometric techniques, measurement of the bearing angle is independent of the carrier frequency of operation and depends only on the modulation frequency. This eases many of the design restraints experienced by methods based on carrier frequency interferometry.

It is of further interest to note the use of intermediate frequency interferometric principles in applications wherein the bearing angle of a transmitting point source is measured by a two-element receiving array. Though this application can be accommodated by carrier frequency interferometric techniques as a consequence of the spatial separation of the transmitted signal at the receiving site, carrier frequency interferometric techniques suffer the same receiving element spacing and angular ambiguity restraints experienced in the mode in which the paired antennas radiate signals to a remote receiver. Methods for overcoming these restraints in order to provide improved angle sector coverage have the disadvantage that the antenna systems employed are complex and generally require critical proportioning of circuit constants be established and maintained between the elements of the receiving aperture if satisfactory operation is to be realized.

The application of intermediate frequency interferometric principles to this receiving mode differs from carrier frequency interferometric techniques in the method of signal transmission. In carrier frequency interferometry the transmitting antenna radiates a single carrier frequency. In intermediate frequency interferometry the transmitting antenna simultaneously radiates both a carrier and a sideband signal. The receiving system in its simplest configuration for a single angle coordinate consists of two receiving antennas, A and B, located at either end of a common baseline. Typical baseline separation distances are of the order of one wavelength at the modulation frequency.

The transmitter simultaneously radiates a carrier frequency f_c and a phase related sideband frequency f_c and f_m. The paths over which the carrier and sideband signals travel to antennas A and B are determined by the angle θ of the transmitter relative to the normal to the baseline of the receiving antennas A and B. If we assume that the transmitted signal travels over distance R from the transmitting antenna to receiving antenna A, and over distance R+ to receiving antenna B, then on arrival at receiving antenna A, the carrier signal C_A will have the frequency and phase relationship:

$$C_A = \left[\cos \omega_c t + \frac{2\pi R}{\lambda_c}\right] \tag{11}$$

where:

$\omega_c = 2\pi f_c$

f_c = carrier frequency

λ_c = carrier wavelength

The sideband signal S_A on arrival at antenna A will have the form:

$$S_A = \cos \left[(\omega_c + \omega_m)t + \frac{2\pi R}{\lambda_s}\right] \tag{12}$$

where:

$\omega_m = 2\pi f_m$

f_m = modulation frequency

λ_s = sideband wavelength

In a similar manner the transmitted carrier and sideband signals on arrival at antenna B will have the frequency and phase relationships shown in equations 13 and 14 respectively.

$$C_B = \cos\left[\omega_c t + \frac{2\pi R}{\lambda_c} + \frac{2\pi \ell}{\lambda_c}\right] \tag{13}$$

$$S_B = \cos\left[(\omega_c + \omega_m)t + \frac{2\pi R}{\lambda_s} + \frac{2\pi \ell}{\lambda_s}\right] \tag{14}$$

The carrier and sideband signals received by each antenna, A and B, are separately heterodyned to extract their difference frequency. This difference frequency is of course the modulation frequency. Heterodyning of the carrier and sideband signals received at antenna A produces a modulation frequency M_A with a frequency and phase relationship given by:

$$M_A = \cos\left[\omega_m t + 2\pi R\left(\frac{1}{\lambda_s} - \frac{1}{\lambda_c}\right)\right] \tag{15}$$

In a similar manner the heterodyning of the carrier and sideband signals C_B and S_B received at antenna B provide a difference frequency or modulation frequency component with the frequency and phase relationship:

$$M_B = \cos\left[\omega_m t + 2\pi R\left(\frac{1}{\lambda_s} - \frac{1}{\lambda_c}\right) + 2\pi \ell\left(\frac{1}{\lambda_s} - \frac{1}{\lambda_c}\right)\right] \tag{16}$$

It is important to recall at this point the wavelength relationship:

$$\frac{1}{\lambda_s} - \frac{1}{\lambda_c} = \frac{1}{\lambda_m} \tag{17}$$

where:

λ_m = modulation frequency wavelength

Introducing this relationship in equations 15 and 16 it is apparent that the phase difference between the modulation frequency M_A and the modulation frequency M_B is:

$$\phi = \frac{2\pi \ell}{\lambda_m} \tag{18}$$

If the separation distance between the two antennas A and B is d, then the bearing angle θ to the transmitter relative to the normal to the baseline between the two antennas A and B is given by:

$$\theta = \sin^{-1}\left[\frac{\ell}{d}\right] \tag{19}$$

and hence:

$$\theta = \sin^{-1}\left[\frac{\phi \lambda_m}{2\pi d}\right] \tag{20}$$

If carrier frequency interferometric methods were used, a single carrier frequency would be transmitted to receiving antennas A and B, and in this case the angle θ would be given by:

$$\theta = \sin^{-1}\left[\frac{\phi \lambda_c}{2\pi d}\right] \tag{21}$$

where:

λ_c = carrier frequency wavelength

since the phase difference measurement corresponding to the path length difference ℓ would be made at the carrier frequency.

Here again we reach the important conclusion that through the use of intermediate frequency interferometric techniques, measurement of the bearing angle is independent of the carrier frequency of operation and depends only on the modulation frequency. This allows a much greater physical separation between the two antennas A and B, allowing completely separate antenna apertures for A and B, thereby leading to the several advantages previously described in connection with that mode of operation wherein the paired antennas simultaneously radiate signals to a remote receiver.

TABLE I. CHARACTERISTICS OF RADIO NAVIGATION SYSTEMS

Type of Distance Measurement	Type of Position Fix	Type of Signal Transmitted	Principal Application
(a) None	(a) Two Bearings	(a) C.W.	(a) Navigation
(b) Path Phase	(b) Distance and One Bearing	(b) Pulse	(b) Surveying (1) Hydrographic (2) Land
(c) Pulse Delay	(c) Two distances (1) Separately (2) Difference		
(a) Beacons	(a) Beacons D.F.	(a) Beacons Decca Lorac Navarho FM Radar Radio Web/Multilateration RADU X RANA Raydist Tellurometer	(a) Beacons Cytac Decca Navarho Radar Radio Web RADU X
(b) Decca Lorac Navarho FM Radar Omega	(b) Navarho Radar	(b) Beacons Cytac EPI Loran Radar Shoren Tacan ATCRBS DABS Synchro-DABS	(b) EPI Elec.Position Ind. Lorac RANA Raydist Shoran/Hiran Tellurometer
(c) EPI Loran Radar Shoran/Hiran	(c) EPI Raydist Shoran/Hiran		
(d) Cytac VORTAC VOR-DME TACAN	(d) Cytac Decca Lorac Loran Radio Web Radar Rana Raydist	(a) C-W (b) Pulse	(a) Navigation (b) Surveying

TABLE II. CHARACTERISTICS OF ELECTRONIC POSITIONING SYSTEMS

SYSTEM	LORAN-A	DECCA	LORAN-C	SATELLITE	OMEGA	MICRO OMEGA	RAYDIST PASSIVE	RAYDIST RANGE-RANGE
Principle	Pulse	Phase	Pulse & Phase	Phase Doppler	Phase	Phase	Phase	Phase
Freq. Band	1900kHz	100kHz	100kHz	150MHz & 400MHz	10.2kHz	10.2kHz & 1750kHz	1600-4000kHz	1600-4000kHz
Range (N.Miles)	600	250	1200	World-Wide (Hourly)	World-Wide'72	200	350 Day 175 Night	200 Day 150 Night
Time for Fix	1-2 Min.	Instant-aneously	30 sec. Indexing	10 Min. (Hourly)	Instant-aneously	Instant-aneously	Instant-aneously	Instant-aneously
Sensitivity	500'	50'	50'	60'	400'	400'	1-1/2'	1-1/2'
Accuracy	1-2NM	1/4NM	1/4NM	0.1NM	1-2NM	Better than 0.1NM	10'	10'
Freedom from Dilution Problems	Medium	Medium	Good	Good	Best	Best	Fair	Good
Probable Max. Terrain Distortion	40' X Dilution	800' X Dilution	800'X Dilution	Velocity & Orbital	10,000 X Dilution	10'K X Dilution	40'	40'
Skywave Error	Negligible	Up to 1NM (Subtle & not easily forecast)	Negligible (if careful)	Negligible	10NM (but generally predictable to 10%)	Negligible	Negligible (Unless obvious)	Negligible (Unless obvious)

TABLE II. CHARACTERISTICS OF ELECTRONIC POSITIONING SYSTEMS (CONTINUED)

System	VOR	DME	TACAN	ILS	MLS	Radar
Principle	Phase	Pulse Range	Omni-Directional Beacon Phase & Pulse	Phase	Time-Ref. Beam Scan Rate	Pulse
Freq. Band	112-118MHz	960-1215MHz	960-1215MHz	108-111.9MHz 329-335MHz	5000-5250MHz	1000-2200MHz
Range	200NM	200NM	200NM	20NM	30NM	60-200NM
Time for Fix	Instant-eously	Instant-eously	Instant-eously	Instant-eously	Instant-eously	Instant-eously
Sensitivity						
Accuracy	±5.0°	±1.0Mile	±0.5° ±1.0 Mile	±.075°	.01°	0.25°
Freedom from Dilution Problems	Good	Good	Good	Good	Good	Good
Probable Maximum Terrain Distortion	Course Bends due to Multipath	Multipath	Multipath	Multipath	Multipath	Multipath Clutter
Skywave Error	Negligible	Negligible	Negligible	Negligible	Negligible	Negligible

TABLE III. FREQUENCY SPECTRUM OF RADIO NAVIGATION AIDS

Radio Navigation Aid	Frequency
Omega	10-14 KHz
Decca	80-120 KHz
Loran C	90-110 KHz
Non-Directional Beacons	
a) Aeronatical	200-285 KHz 325-405 KHz
b) Marine	285-325 KHz
Loran A	1700-2000 KHz
ILS	
a) Marker Beacons	75 MHz
b) Localizer	108-112 MHz
c) Glide Slope	328-335 MHz
VOR	112-118 MHz
DME	960-1215 MHz
TACAN	960-1215 MHz
MLS	5000-5250 MHz

TABLE IV. SUMMARY OF NAVIGATION REQUIREMENTS BY OPERATIONAL AREA

Navigation	Area	Requirement	System	Propagation Anomalies
Air Operations	Oceanic	Surveillance Navigation	Satellite VHF Communication LORAN-Inertial Nav.	Scintillation and Fading Scintillation and Fading Scintillation & Fading, Sky Wave Clutter
	Domestic Enroute and Terminal	Surveillance Navigation	Radar, Communications VOR-DME	Clutter, Multipath
	Approach and Landing	Surveillance Navigation Guidance	Radar VOR-DME ILS-MLS	Multipath Multipath
Marine Operations	Open Sea	Navigation	Loran	
	Coastal and Confluence Zone	Navigation	Loran	
	Inland and Restricted Waters	Navigation	Loran	Sky Wave Multipath
Land Operations	Automatic Vehicle Monitoring	Navigation Position Location	Loran Line of Sight Transmission by Radio	Blockage
	Dispatch	Communications	VHF Communications	RFI

TABLE V. VEHICLE REQUIREMENTS FOR AIR TRANSPORT SYSTEMS

Vehicle Requirements	System Requirements
Determination of flight direction	Improve efficiency of Air Traffic Control System
Determination of flight velocity	Accommodate future growth
Determination of altitude	Provide reliable and economic operation by the user
Determination of present geographic referenced position	Complete volumetric coverage
Landing manoeuvres under poor visual conditions	Transmissions should be 100% reliable and available
Avoid collision in air or ground mode	Transmissions should be free from ambiguities
	Transmissions and systems should provide maximum protection from blunders
	System should be accessible to all users
	System should be compstible with all other systems in use
	System should be 99.9% accurate in its operational areas

TABLE VI. VEHICLE REQUIREMENTS FOR MARINE TRANSPORT SYSTEMS

Vehicle Requirements	System Requirements
Determination of vessel direction	Precise position fixing plus procedure and knowledge
Determinstion of vessel velocity	Augmentation of visual aids
Determination of vessel position	Proximity warning about hazards or vessels
Determination of distance to go	Addition of communication with navigation
All weather position fixing	Continuously available and highly reliable complete coverage
Reduction of vessel collision during transit or docking	Accuracy comparable to visual navigation aids

TABLE VII. VEHICLE REQUIREMENTS LAND TRANSPORT SYSTEMS

Vehicle Requirements	System Requirements
Automatic position and location	Accuracy repeatable to within 50 ft.
Automatic maintaining of vehicle fleet size	Economic benefits are difficult to value in monetary terms
Tracking of an individual vehicle or person associated with vehicle	Increase operational efficiency
Location and position of highway accidents	Convenience of acquisition to system is difficult to measure
Coordination of ground vehicle with air and land vehicles	
Position location of disaster	
Repeatable accuracy of geographic coordinate references	

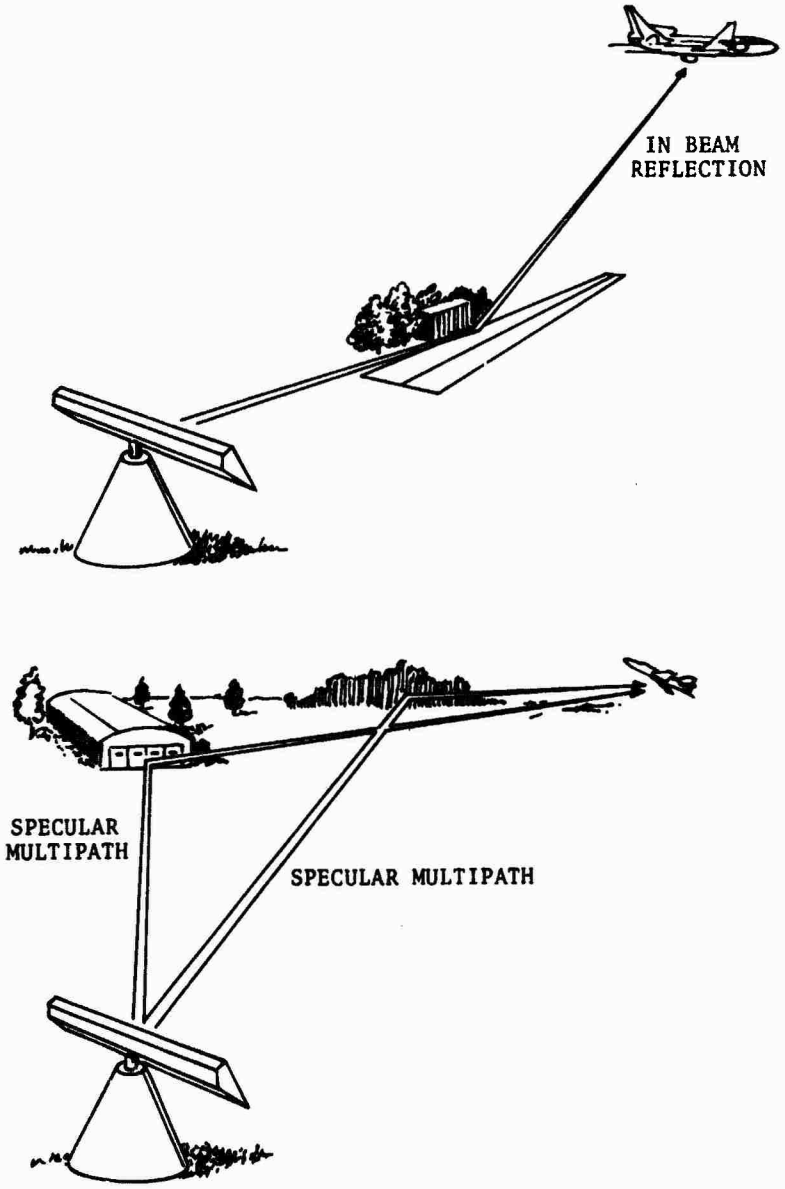


FIGURE 1. GEOMETRY OF IN BEAM AND SPECULAR MULTIPATH

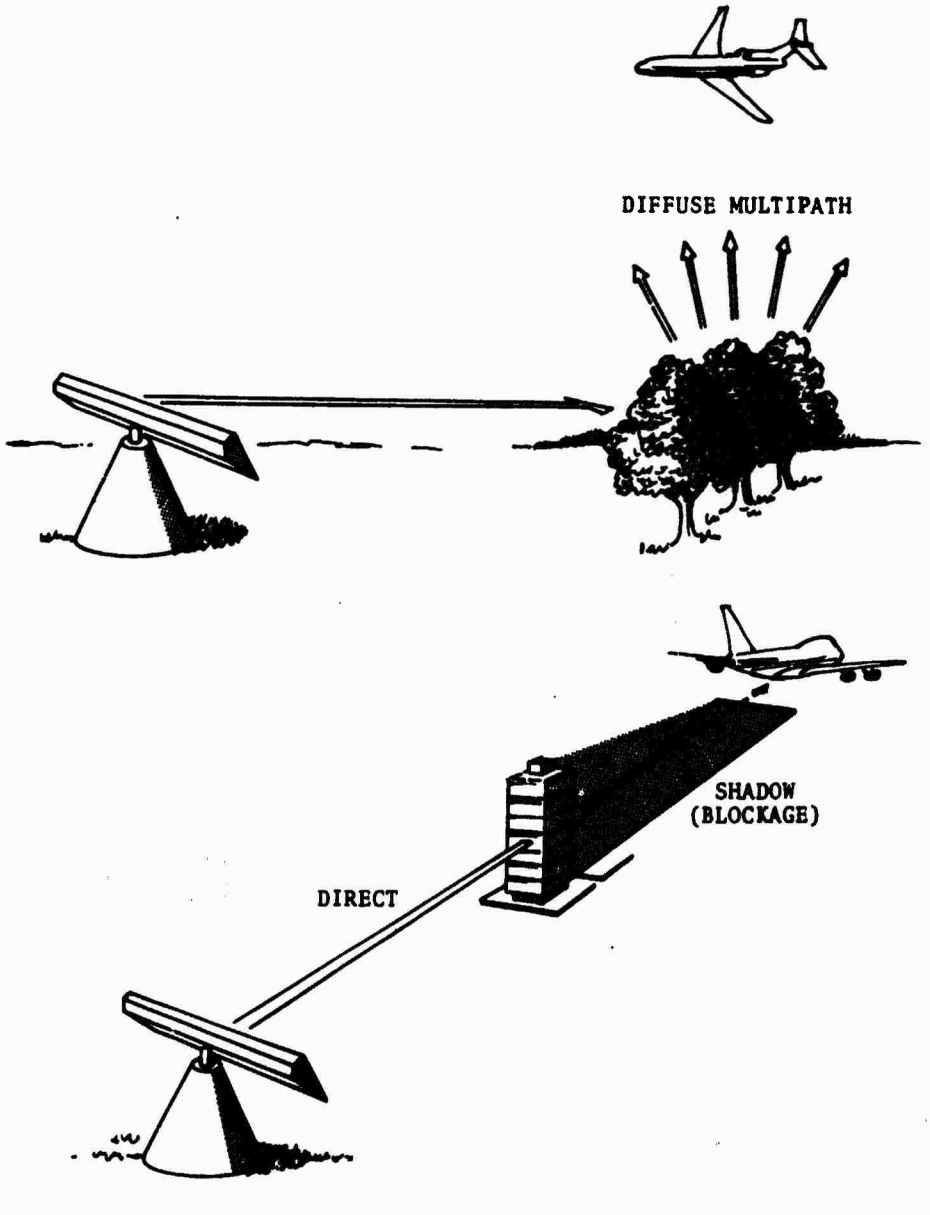


FIGURE 2. GEOMETRY OF DIFFUSE MULTIPATH AND RADIATION BLOCKAGE

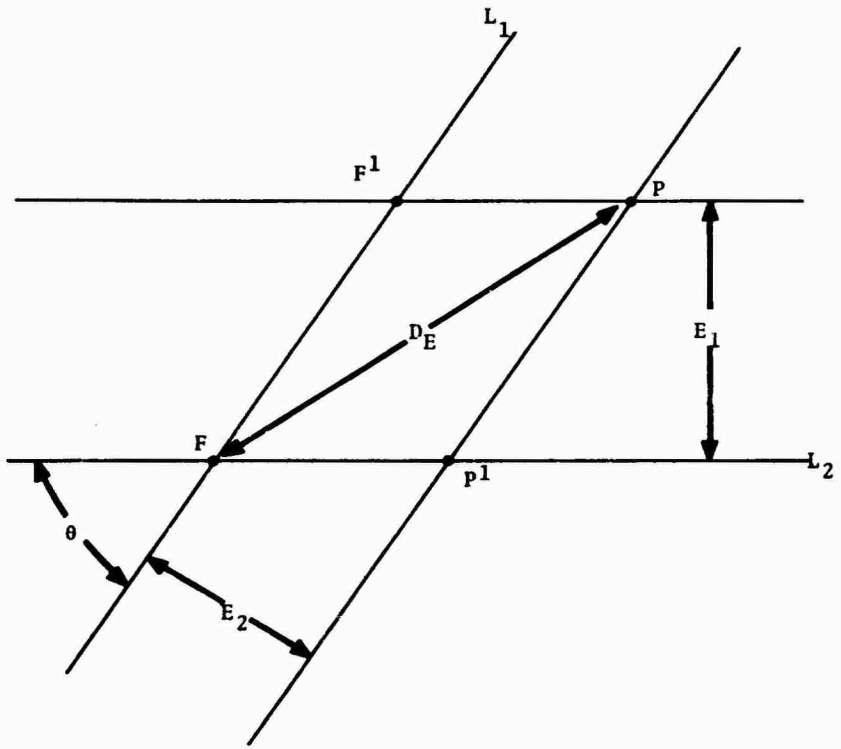


FIGURE 3. ERROR PARALLELOGRAM

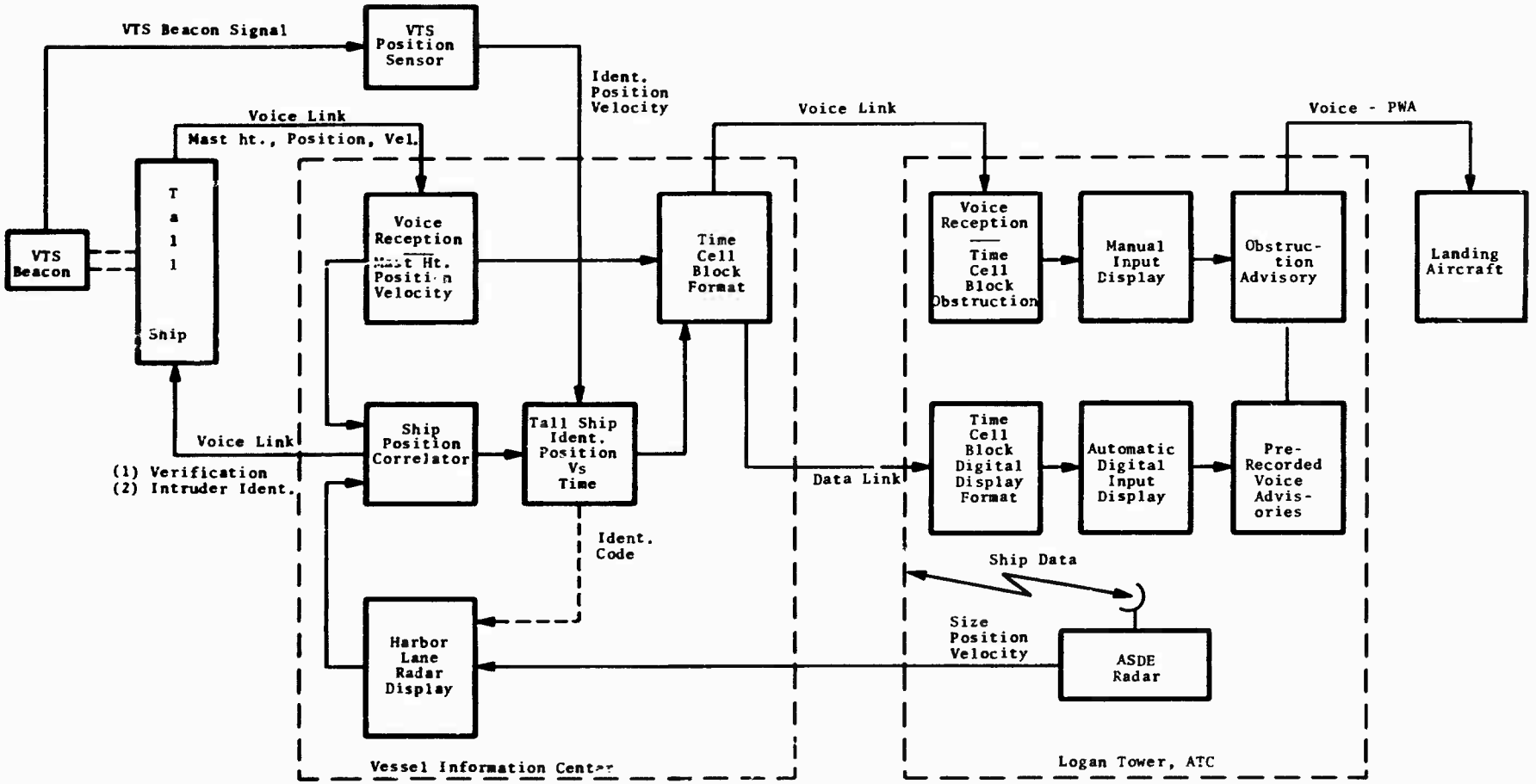


FIGURE 4. INTEGRATED VESSEL TRAFFIC SURVEILLANCE SYSTEM

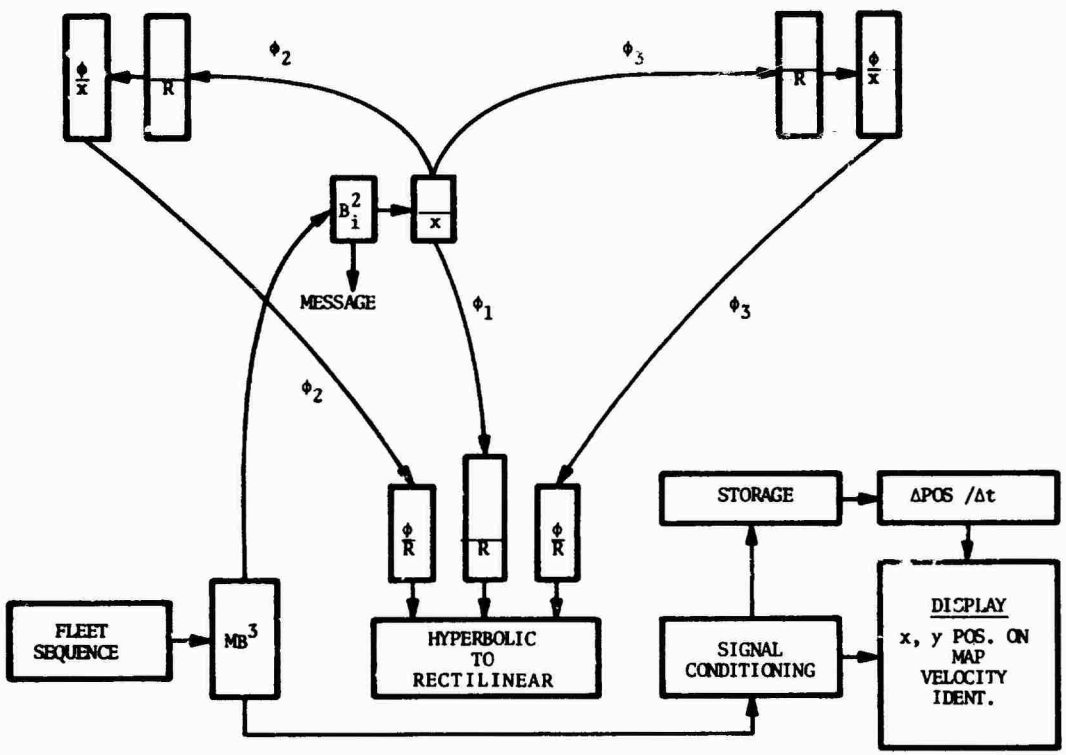


FIGURE 5. FUNCTIONAL BLOCK DIAGRAM OF A DISCRETE ADDRESS SURVEILLANCE SYSTEM

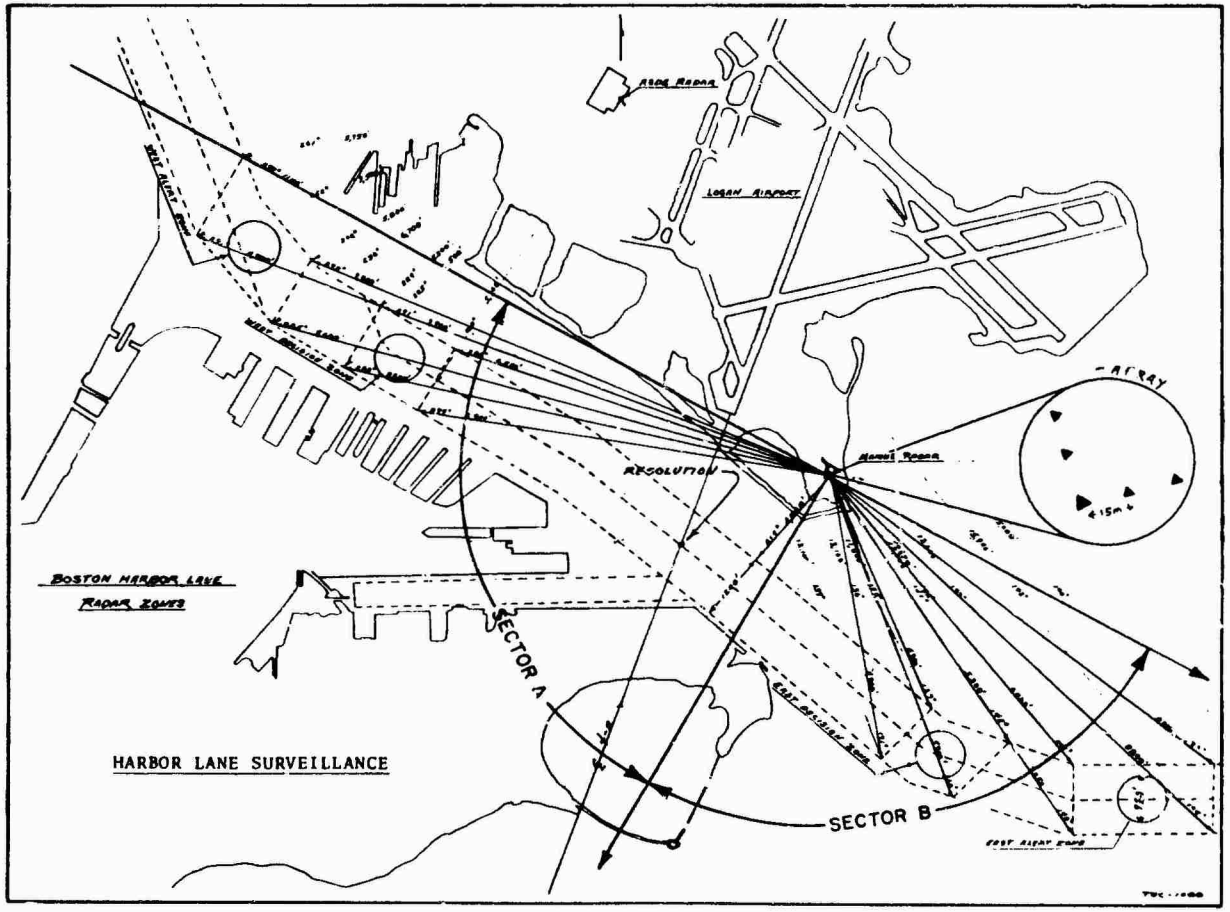


FIGURE 6. GEOMETRY OF A HARBOR SURVEILLANCE SYSTEM

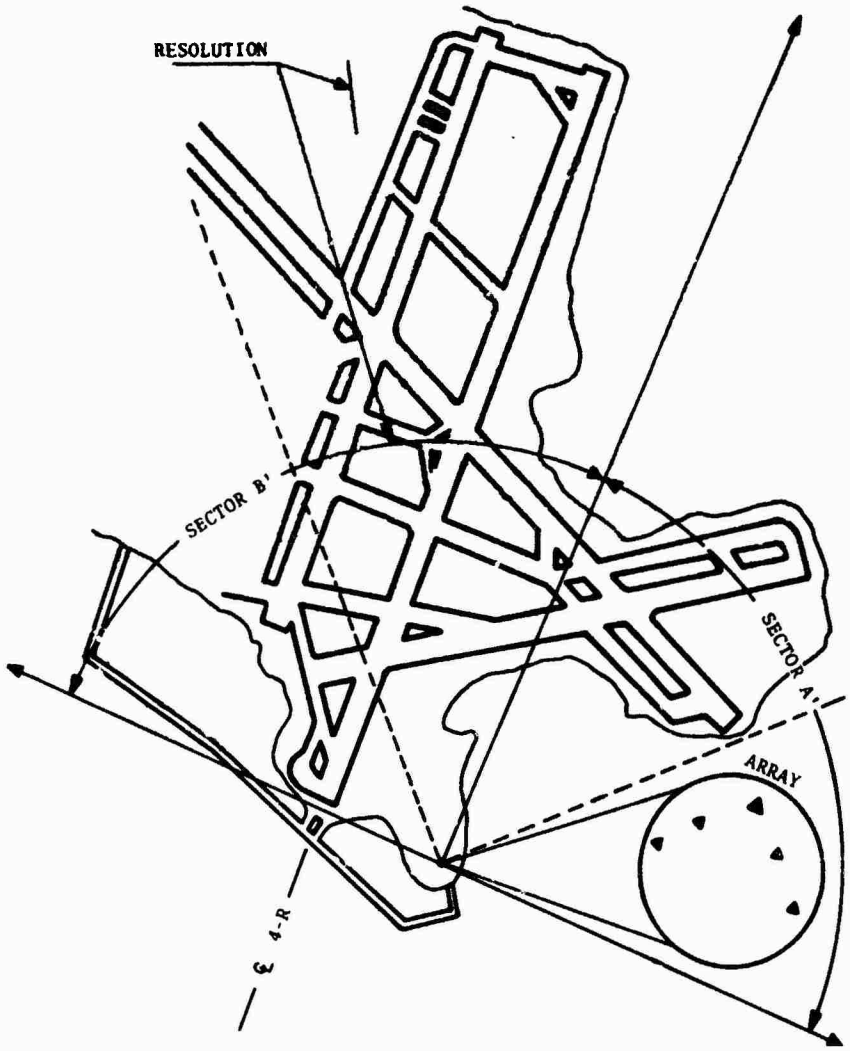


FIGURE 7. GEOMETRY OF AN AIRPORT GROUND TRAFFIC SURVEILLANCE SYSTEM

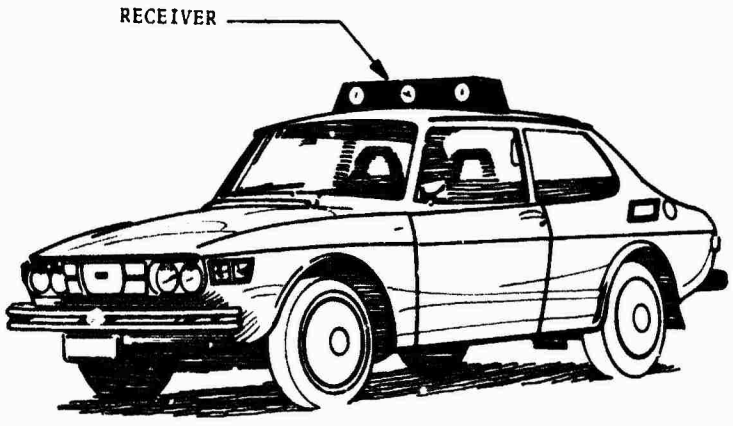


FIGURE 8. AUTOMOBILE RECEIVER

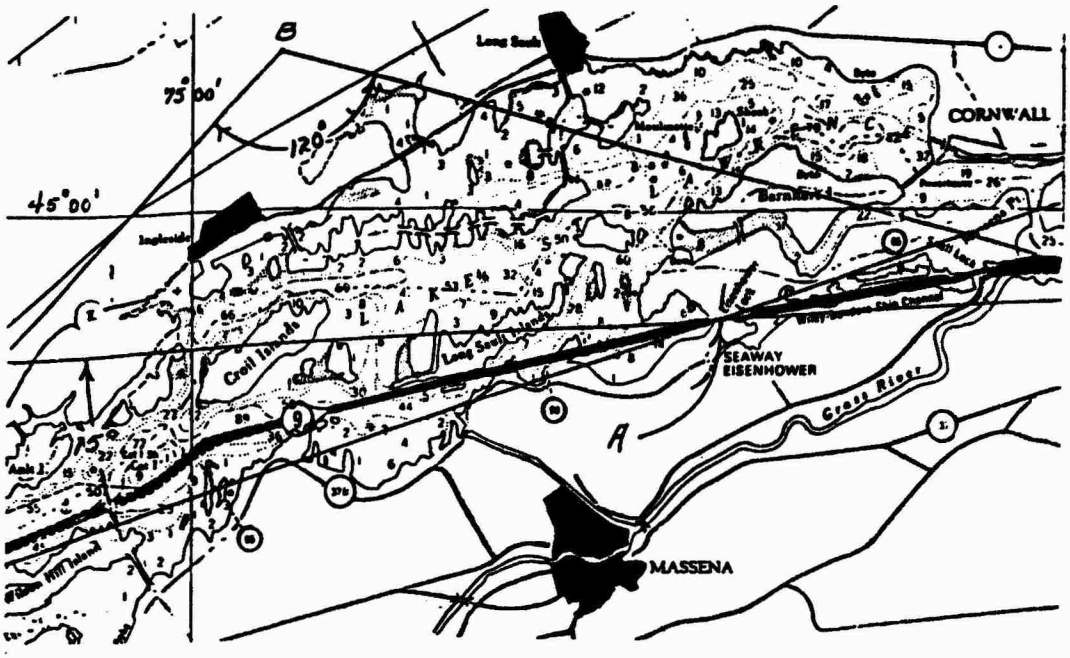


FIGURE 9. NAVIGATION CHARTS OF A TYPICAL NARROW MARINE CHANNEL

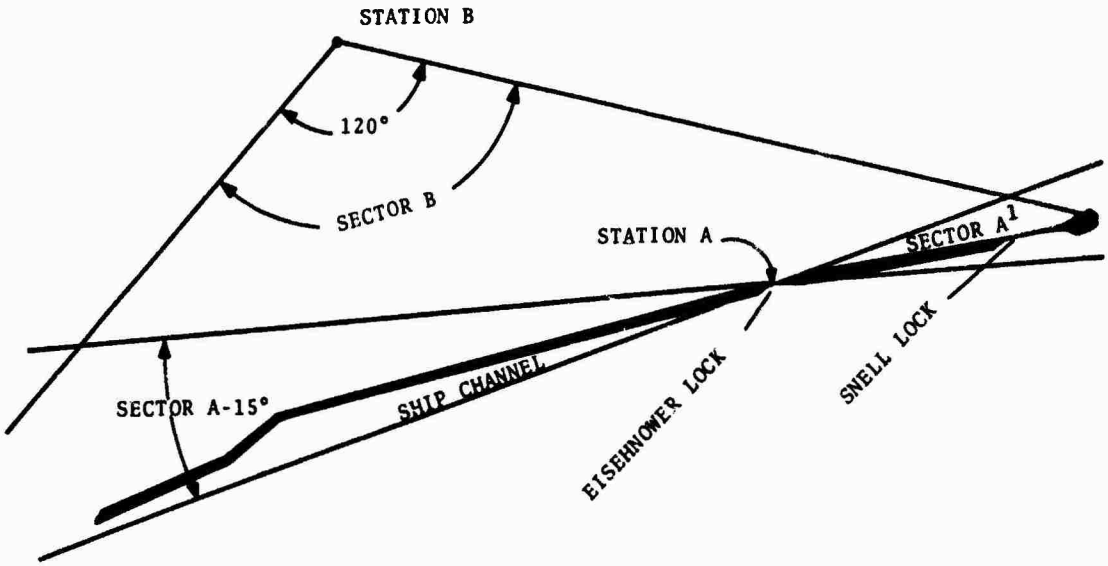


FIGURE 10. NAVIGATION SYSTEM CONFIGURATION GEOMETRY

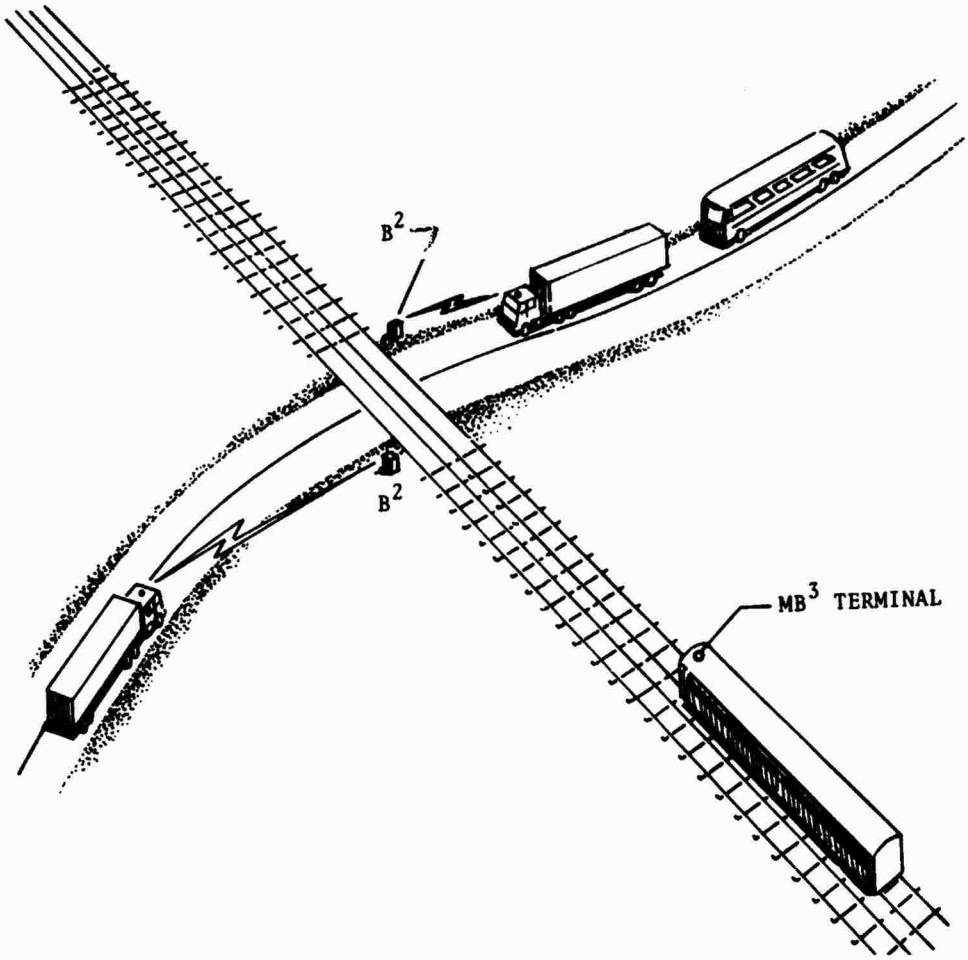


FIGURE 11. GRADE CROSSING CONTROL SYSTEM SCENARIO

FIGURE 12. AUTOMOBILE TRANSMITTER

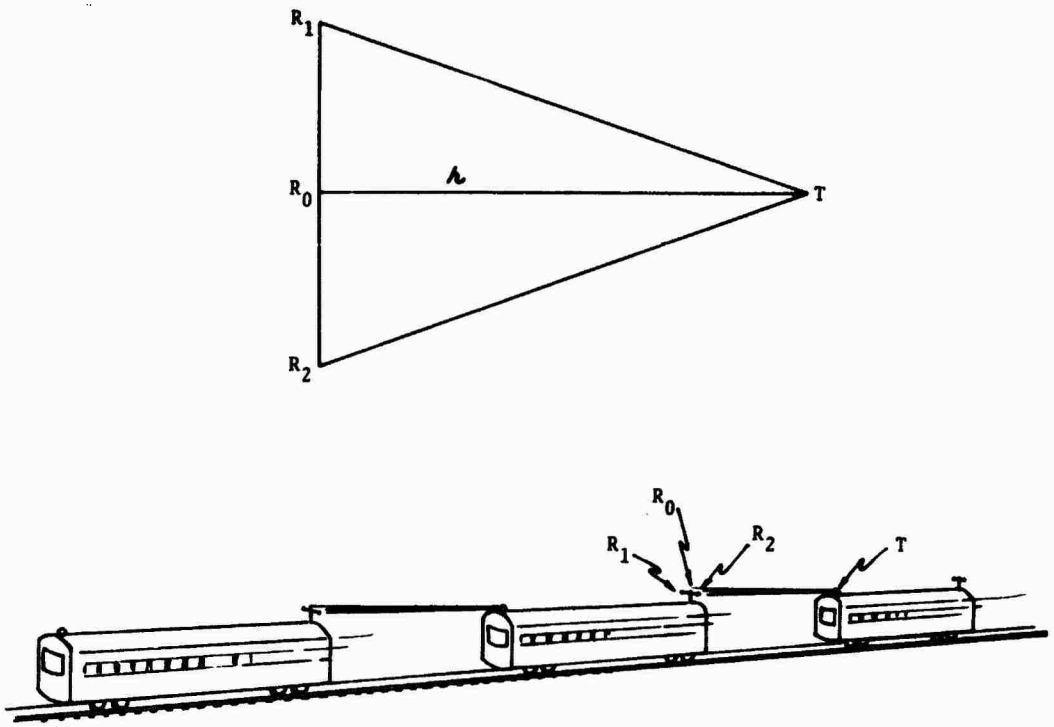
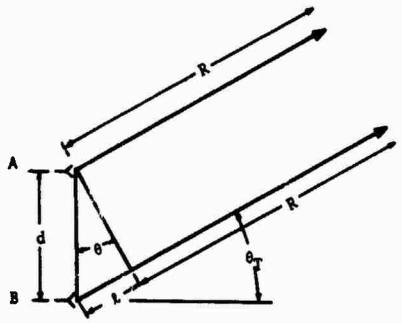


FIGURE 13. SPACING CONTROL GEOMETRY FOR PRT AND RAIL VEHICLES



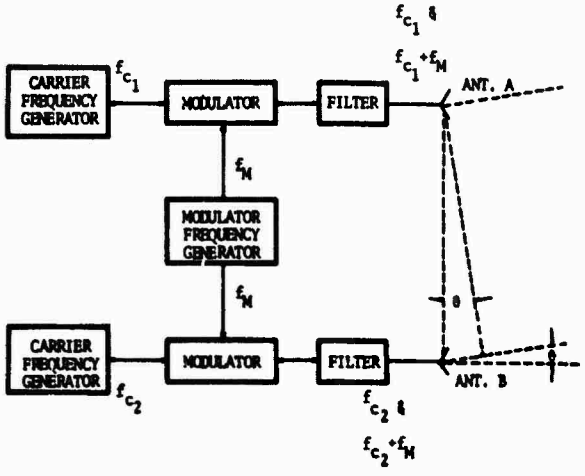


FIGURE 14. TWO ELEMENT INTERMEDIATE FREQUENCY INTERFEROMETRIC TRANSMITTER

APPLETON, E.V. and BARNETT, M.A.F., 1925, "Direct Evidence of Downward Refraction of Electromagnetic Waves" Proc. Roy. Soc., V. 109.

BREIT G and TUVE, M.A., 1926, "A Test of the Existence of the Conducting Layer", Phys. Rev., V. 28, pp. 554-575.

HEISING, R.A., 1928, "Experiments and Observations Concerning the Ionized Regions of the Atmosphere", Proc. IRE, V. 16, pp. 75-99.

BIRGE, R.T., 1941, "The General Physical Constants", London, Rept. on Prog. in Phys., V. 3, The Physical Soc., pp. 90-134.

DORSEY, N.E., 1944, "The Velocity of Light"., Am. Phil. Soc., New Ser., V. 34, Part I., Philadelphia.

U.S.A. Nat. Committee, U.R.S.I. 12th Gen. Assembly Rpt. p. 23, NAS-NRC Pub. 581, Washington, DC.

FROOME, K.D., "Precision Determination of the Velocity of Electromagnetic Waves", 1958, Nature, V. 181, p.258.

SMITH,Jr., E.K. and WEINTRAUB, S. 1953, "The Constants in the Equation for Atmospheric Refractive Index at Radio Frequencies", Proc. IRE, V. 41, pp. 1035-1037.

JOHLER, J.R., KELLAR, W.J. and WALTERS, L.C., 1956, "Phase of the Low Radiofrequency Ground Wave", NBS Circular 573, National Bureau of Standards, Washington, DC.

ADAMS, P.R. and COLIN, R.I., 1946, "Frequency, Power and Modulation for a Long-Range Radio Navigation System", Elec. Comm., V. 23, pp.144-158.

DEAN, W.N., 1957, "A Precision Multi-Purpose Radio Navigation System: Part II-Propagation Considerations" IRE Nat. Conv. Rec., V. 5, Pt. 8, pp. 86-93.

MILLINGTON, G., 1949, "Ground-Wave Propagation Over an Inhomogeneous Smooth Earth", Proc. IEE, V. 96, Pt. III, p. 53-64.

MILLINGTON, G., 1950, "Ground-Wave Propagation Over an Inhomogeneous Smooth Earth, Pt. 2, Experimental Evidence and Practical Implications", Proc. IEE, V. 97, Pt. III, pp. 209-217.

PRESSEY, B.G., ASHWELL, G.E., and FOWLER, C.S., 1956, "Change of Phase with Distance of a Low-Frequency Ground Wave Propagated Across a Coast-Line," Proc. IEE, V. 103, Pt. B, pp. 527-534.

GILLHAM, R.L., COATES, A., KITCHEN, F. TROW, G.H., LIPMAN, D.D., KEELING, H. and JENKINS, E., 1954, "An Evaluation of the Decca Navigation System," Ministry of Transport and Civil Aviation, CAP 124.

Anon., 1956, "Radio Aids to Maritime Navigation and Hydrography", Spec. Publ. No. 39, Internat. Hydrographic Bur., Monaco.

CROSSLAND, E.J., 1952, "Lorac-What it is and how it Works", Philco Techrep Div. Bull.

LEIFSON, G., 1953, "Hyperbolic Position Systems for Hydrographic Surveys," Internat. Hydrographic Rev., V. XXX, No. 1.

ROBERTSON, A.J.L., "A Mathematical Analysis of the Lorac Instrumentation", Lorac, Radio Surveying and Navigation System, Seismograph Service Corp.

KOEPPEL, B.W., 1956, "Final Report of Field Tests of the application of a Phase Comparison Radio Location System to Distance Measurement", Contr. No. DA-44-009 Eng-2696, Lorac Service Cor.,

LACOMBE, H., 1956, "The Rana Radionavigator-Characteristics, Testing and Operation of a Radionavigational Chain", Internat. Hydrographic Rev, pp. 22-48.

MOLLENDER, A.B., PHILLIPS, B.E., 1951, "Effect of Skywave Interference and Atmospheric Noise on Raydist Precision", Parsons-Aerojet Co. AF Tech. Rpt. No. 6., ASTIA Doc. No. R6253.

BENSON, J.M. and HAYWARD, V.M., 1952, "Final Engineering Report on Downrange Survey", Raydist Navigation Corp., Rpt. No. RN-4.

ERICKSON, E.E., 1952, "An Evaluation of Raydist", U. S. Navy Mine Countermeasures Station, Rpt. No. 41, ASTIA Doc. AD-9202.

VERNON, B., and REEDY, P. H., 1953, "Field Evaluation of the Three-Dimensional Raydist Systems," Parsons-Aerojet AF Tech. Rpt. No. 39, AD-34 219.

ROBINSON, F.W., and BLUST, F.A., "Final Report of Tests of Type "E" Raydist for Hydrographic Charting Over Fresh Water," U. S. Army Corps of Engineers, U.S. Lake Survey, AD-14 935.

CARSWELL, I., 1955, "A Study of Some Inherent Errors in the Three-Dimensional Raydist System," Rpt. No. 50, SRI Proj. 1197, Stanford Research Institute, AD-68 580.

Anon., "Tellurometer", prepared by Electronics Lab., U.S. Dept of Commerce, Coast and Geodetic Survey., Washington, DC.

McCALL, J.S., 1957, "Distance Measurement with the Geodimeter and Tellurometer", Proc. Am. Soc. Civ. Eng., J. Surveying and Mapping Div., Paper No. 1445.

ARMISTEAD, J.G., 1957, "Combined Engineering and Service Test Report-Tellurometer", U.S. Army Engineer Res. and Dev. Labs.

Tellurometer, Handbook, Operation and Field Maintenance, Mod. M/RAI, Tellurometer Ltd., April 1958.

CLEAVER, R.F., 1947, "Note on a Short-Range Radio Position-Finding System Using Modulated Continuous Waves", J. IEE, V. 94, Pt. 3A., pp. 984-989.

CLARK, C.I., COLIN, R.I., DISHAL, M., GORDY, I., and ROGOGG, M., 1954,"Navarho Long-Range Radio Navigation System," Elec. Comm., V. 31, pp. 155-166, and IRE Conv. Rec., V. 2, Pt. 5.

GOLDBERG, O.D., "Report on Navarho Evaluation", Rpt. No. RADC-TR-57-135, Pickard and Burns, Inc.,

HAUER, E.W., "Flight Evaluation of Airborne Navarho Equipment", WADC Tech. Note 57-345, AD-142031.

BARKER, R.A., 1957, "Survey of Various Navigational Aids Developed in Great Britain and France", AFCRC Tech. Rpt. No. TR-57-161, Stavid Engineering, Inc., AD-117014.

Stavid Engineering, Inc., 1957, "Radio Web Navigation System", Technical Description, Stavid Case No. Q-6-359.

WOLFF, I., LUCK, D.G.C., 1948, "Principles of Frequency-Modulated Radar", RCA Rev., V. 9, ; pp. 352-362, June 1948; pp. 531-555, 1948.

LUCK, D.G.C., 1949, "Frequency Modulated Radar", McGraw-Hill, New York.

SHARPE, B.A., 1947, "Low-Reading Absolute Altimeters," J. IEE, V. 94, Pt. 3A, pp. 1001-1011.

WIMBERLY, F.T. and LANE, Jr., J.F., 1954, "The AN/APN-22 Radio Altimeter, " IRE Trans. on Aero and Nav. Elec., V. ANE-1, No. 2, pp. 8-14.

CLARKE, A.R., 1880, "Geodesy", Chap. V., Oxford Univ. Press.

HELMERT, F.R., 1880, "Die Mathematishen und Physikalischen Theorie die Hoheren Geodasie," Liepzig, V. I., Chap. 5.

HAYFORD. J.F., 1909, "The Figure of the Earth and Isostasy", U.S. Coast and Geodetic Survey., p. 88.

ANDOYER, M.H., 1932, "Formule Donnant La Longuer de la Geodesique," Bu.. Geodesique, V. 34, pp. 77-ff.

LAMBERT, W. D., 1942, "The Distance Between Two Widely Separated Points on the Surface of the Earth," J. Wash. Acad. Sci., V. 32, p. 125.

U. S. Navy Hydrographic Office, "Loran ATechnical Report No. 3", 1945.

SODANO, E. M., 1958, "A Rigorous Non-Iterative Procedure for Rapid Inverse Solution of Very Long Geodesics," Bull. Geodesique, No. 47/48, pp. 13-25.

Aero. Chart and Information Center, 1957, "Geodetic Distances and Azimuth Computations for Lines Over 500 Miles," ACIC Tech. Rpt. No. 80.

WONG, M.S., 1951, "Ray-Tracing Picture of Radio-Wave Propagation in Arbitrary", AF Tech. Rpt. No. 6631, Wright Air Development Center.

GRANT, A.S.G., 1952, "The Application of Ray-Tracing Methods to the Reduction of Shoran to Map Distances", Trans. AGU, V. 33, pp. 645-649.

GOLDSON, W.L., 1953, "Aproposed Technique for Reduction of Shoran Data to Geodetic Distances," Trans. AGR, V. 34, p. 673-679 .

KROLL, C.W., 1949, "A Rigorous Method for Computing Geodetic Distance from Shoran Observations", Trans. AGU, V. 30, pp. 1-4.

BEAN, B.R., and CAHOON, B.A., 1957, "The Use of Surface Weather Observations to Predict the Total Atmospheric Bending of Radio Rays at Small Elevation Angles," NBS Rpt. 5082.

BEAN, B.R., and HORN, J.D., 1957, "On the World Wide Variations of Radio Refractivity as Derived from Standard Weather Observations", NBS Rpt. 5083.

BEAN, B.R., and CAHOON, B.A., 1957, "A Note on the Climatic Variation of Absolute Humidity", Bull. Am. Meteorol. Soc., V. 38, pp. 395-398.

BEAN, B.R. and HORN, J.D., 1958, "On the Climatology of the Surface Values of Radio Refractivity of the Earth's Atmosphere", NBS Rpt. 5559.

KALMUS, H.P., CACHERIS, J.C. and DROPKIN, H.A., 1954, "Nonquantized Frequency-Modulated Altimeter", IRE Trans. Aero and Nav. Elec., Vol. ANE-1, No. 2, pp. 15-19.

PIERCE, J.A., McKENZIE, A.A., and WOODWARD, R.H., 1948, "Loran", MIT Radiation Lab. Ser., V. 4,McGraw-Hill, New York.

WILLIAMS, Jr., R.B., 1953, "Continuous-Indicating Loran Navigator", Electronics, V. 26, pp. 166-169.

SITTERLY, B.W., "Methods of Computation of Loran Tables and Charts", Ref. 59, Chap. 6, pp. 170-194.

SEELEY, S.W., 1946, "Shoran Precision Radar", Elec. Eng., V. 65, pp. 232-240.

Air Photographic and Charting Service, 1953, "Basic Shoran Theory", 1st Computations Tech. Squadron. Orlando AFB, FL.

Air Photographic and Charting Service, 1954, "Hiran Surveying Handbook", 1st Computations Tech. Squadron. Orlando AFB, FL.

COLIN, R.I., and DODINGTON, S.H., 1956, "Principles of Tacan," Elec. Comm., V. 33, p. 11-25.

DeFAYMOREAU, E., 1956, "Experimental Determination of Tacan Bearing and Distance Accuracy," Elec. Comm. V. 33, pp. 67-73.

LATIMER, Jr., D.T., 1956, "Azimuth Errors of the TACAN System", IRE Trans. on Aero and Nav. Elec., V. ANE-3. pp. 150-156.

KAYTON, M. and FRIED, W.R., (ed.), 1970, "Avionics Navigation Systems", Wiley, New York.

U. S. Navy Hydrographic Office, 1963, "Air Navigation", H.O. Pub. No. 216, 717 pp., U.S. Gov't Printing Off., Washington, DC

REYNOLDS, L.G., 1953, "An Examination of Some Site and Transmission-Path Errors of the Decca Navigator System When Used Over Land," Proc. IEE, V. 100, Pt. III, pp. 29-35.

BATEMAN, R., FLORMAN, E.F., and TAIT, A., 1950, "A Source of Error in Radio Phase Measuring Systems," Proc. IRE, V. 38, pp. 612-614.

HUFFORD, G.A., 1950, "An Analysis of Some Anomalous Properties of Equiphase Contours," Proc. IRE, V. 38, pp. 614-618.

WILLIAMS, Caradoc, 1951, "Low-Frequency Radio-Wave Propagation by the Ionosphere with Particular Reference to Long-Distance Navigation", Proc. IEE, V. 98, Pt. III, pp. 81-99.

GRENNAN, D.G., and PHILLIPS, M. Lindeman, 1957, "Phase and Amplitude Variability in Medium-Frequency Ionospheric Transmission", MIT Lincoln Laboratory, Tech. Rpt. No. 93,

CRICHLOW, W.Q., HERBSTREIT, J.W., JOHNSON, E.M., NORTON, K.A., and SMITH, C.E., 1946, "The Range Reliability and Accuracy of a Low Frequency Loran System", Rpt. ORS-P-23, Appendix II, Operational Res. Staff, OCSigo,

CRICHLOW, W.Q., 1946, "The Comparative Accuracy of Various Existing and Proposed Radio Navigation Systems", NBS Rpt. CRPL-4-1.

TROW, C.H., and JESSELL, A.H., 1948, "The Presentation of the Fixing Accuracy of Navigation Systems", J. Inst. Nav., V. 1, pp. 313-337.

DOHERTY, R.H., 1957, "Pulse Sky Wave Phenomena Observed at 100kc", NBS Rpt. 6CB103.

PAWSEY, J.L. and BRACEWELL, R.N., 1955, "Radio Astronomy", Oxford Univ. Press.

DRAKE, F.D., and EWEN, H.I., 1958, "A Broad-Band Microwave Source Comparison Radiometer for Advanced Research in Radio Astronomy", Proc. IRE, V. 46, pp. 53-60.

STRUM, P.D., 1958, "Considerations in High-Sensitivity Microwave Radiometry", Proc. IRE, V. 46, pp. 43-53.

GIBSON, J.E., 1958, "Lunar Thermal Radiation at 35 kMc", Proc. IRE, V. 46, pp. 280-286.

MAYER, C.H., McCULLOUGH, T.P. and SLOANAKER, R.M., 1958, "Measurements of Planetary Radiation at Centimeter Wavelengths", Proc. IRE, V. 46, pp. 260-266.

MOSES, H.C., 1972, "Air Traffic Control Radar Systems Definition Report," FAA Rpt. No. FAA-EM-72-1, DOT/ Federal Aviation Administration, Washington, DC.

EMERSON, R.C., 1954, "Some Pulsed Doppler MTI and AMTI Techniques", Rpt. No. R-274, RAND Corp., Santa Monica, CA, AD-65881, .

BARLOW, E.J., 1949, "Doppler Radar", Proc. IRE, V. 37, pp. 340-355.

GRISETTI, R.S., SANTA, M.M. and KIRKPATRICK, G.M., 1955, "Effect of Internal Fluctuations and Scanning on Clutter Attenuation in MTI Radar", IRE Trans. Aero. Nav. Elec., V. ANE-2, pp. 37-41.

SHRADER, W.W., 1967, "MTI Radar", Tech. Mono. TMO-67/003, Selenia S.p.A., Rome, Italy.

WARD, H.R., and SHRADER, W.W., 1968, "MTI Performance Degradation Caused by Limiting", 1968 IEEE EASCON Rec., pp. 168-174, also IEEE Trans. Aerosp. Electron. Syst. (Abstracts), V. AES-4, pp. 913-923.

WARD, H.R., 1969, "The Effect of Bandpass Limiting on Noise with a Gaussian Spectrum", Proc. IEEE, V. 51, pp. 2089-2090.

GRASSO, G., 1968, "Improvement Factor of a Non-Linear MTI in Point Clutter", IEEE Trans. Aerosp. Electron. Syst., V. AES-4, pp. 640-644.

GRASSO, G. and GUARGUAGLINI, P.F., 1969, "Clutter Residues of a Coherent MTI Radar Receiver," IEEE Trans. Aerosp. Electron Syst., v. AES-5, pp. 195-204.

BARTON, D.K., and SHRADER, W.W., 1969, "Interclutter Visibility in MTI Systems", in EASCON '69, Tech. Conv. Rec., pp. 294-297.

SHRADER, W.W., 1960, "Antenna Considerations for Surveillance Radar Systems," in Proc. 7th Annual East Coast Conf. Aeronautical and Navigational Elec.

SHRADER, W.W., 1961, "Results of Antenna Pattern Considerations", in Proc. 8th Annual East Coast Conf. Aeronautical and Navigational Elec.,

PATRICK, A.M., 1961, "Primary Radar in Air Traffic Control", Interavia, V. 16, pp. 851-853.

MULHOLLAND, E.B., and SODEN, F.R., 1967, "Australia's ATC Radar Network", Interavia, V. 22, pp. 511-513.

WINTER, C.F., 1969, "Dual Vertical Beam Properties of Doubly-Curved Reflectors", presented at the 19th Annual USAF Antenna Research and Dev. Symp., Univ. Illinois, Robert Allerton Park, Monticello, IL.

EMSLIE, A.G., and McCONNELL, R.A., 1947, "Moving Target Indication", in "Radar Systems Engineering" Radiation Lab. Series, L.N.Ridenour, (ed.), McGraw-Hill, New York, No. 1, Ch. 16.

SHAPIRO, L.D., 1968, "Time Synchronization from Loran-C", IEEE Spectrum, V. 5, pp. 46-55.

POTTS, C.E., and WIEDER, B., 1972, "Precise Time and Frequency Dissemination via the Loran-C System", Proc. IEEE, V. 60, pp. 530-539.

WIEDER, B., 1971, "Use of Loran-C Over Land", in Proc. 25th Annual Symp. on Frequency Control.

SWANSON, E.R., and KUGEL, C.P., 1972, "VLF Timing: Conventional and Modern Techniques Including Omega", Proc. IEEE, V. 60, pp. 540-551.

PALMER, W., 1970, "The OMEGA Navigation System as a Source of Frequency and Time," in Proc. 24th Annual Symp. Frequency Control,

ALLEN, D.W., BLAIR, B.E., DAVIS, D.D. and MACHLEN, H.E., 1971, "Precision and Accuracy of Remote Synchronization via Portable Clocks, LORAN-C, and Network Television Broadcasts", in Proc. 25th Annu. Symp. Frequency Control, pp. 195-208.

NASA, "Study of Methods for Synchronizing Remotely-Located Clocks," Rpt. No. NASA-CR-738, National Aeronautics and Space Administration, Washington, DC

HANSON, D.W., and HAMILTON, W.F., 1971, "Clock Synchronization from Satellite Tracking", IEEE Trans. Aerosp. Electron. Syst., V. AES-7, pp. 895-899.

DAVIS, D.D., BLAIR, B.E., and BARNABA, J.F., 1971, "Long-Term Continental U.S. Timing Systems Via Television Networks," IEEE Spectrum. V. 8, pp. 41-52.

BAUGH, R.A., and CUTLER, L.S., 1970, "Precision Frequency Sources", Microwave J., V. 13, pp. 43-56.

GERBER, E.A., and SYKES, R.A., 1966, "State of the Art-Quartz Crystal Units and Oscillators", Proc. IEEE, Vol. 54, pp. 103-116.

McCOUBREY, A.O., 1966, "A Survey of Atomic Frequency Standards", Proc. IEEE, V. 54, pp. 116-135.

KARTASCHOFF, P. and BARNES, J.A., 1972, "A Survey of Atomic Frequency Generation" Proc. IEEE, V. 60, pp. 493-501.

SMITH, W.L., 1963, "Quartz Crystal Controlled Oscillators", Bell Telephone Lab., Rep. 25335-H.

DUCKETT, P.C., PEDUTO, R.J. and CHIZAK, G.U., 1970, "Temperature Compensated Crystal Oscillators", in Proc. 24th Annu. Symp. Frequency Control.

FRERKING, M.E., 1972, "Ruggedized Rubidium Frequency Standard", Proc. IEEE (Lett.), V. 60, pp.628-629.

HYATT, R. et a., 1971, "Performance of Newly Developed Cesium Beam Tubes and Standards", in Proc. 25th Annu. Symp. Frequency Control,. pp. 313-242.

KUPERSMITH, F. THORNBURG, C. and HO, J., 1970, "A New Primary Cesium Beam Frequency Standard", in Proc. 24th Annu. Symp. Frequency Control, pp. 308-314.

SYDNOR, R., CALDWEKK, J.J., and ROSE, B.E., 1966, "Frequency Stability Requirements for Communications and Tracking Systems", Proc. IEE, V. 54, pp. 231-236.

BLACKMAN, R.B., 1964, "Methods of Orbital Refinement", Bell Syst. Tech. J., V. 43, pp. 885-909.

ZABORSZKY, J., 1970, "A Practical Non-Diverging Filter", AIAA J., V. 8, pp. 1127-1133.

SORENSON, H.W, and SACKS, J.E., 1971, "Recursive Fading Memory Filtering", Inform. Sci., V. 3, pp. 101-119.

SORENSON, H.W., 1966, "Kalman Filtering Techniques", in "Advances in Control Systems", V. 3., C. T. Leondes (ed) Academic, New York. p. 219.

EWING, C.W. and MITCHELL, M.M., 1970, "Introduction to Geodesy", American Elsevier, New York.

RUPPERT, G., BERNSTEIN, R., AND BOWIN, C., 1969, "Precise Location of a Ship at Sea Utilizing VLF Transmissions", Navigation (J. Inst. Navigation). V. 16, pp. 111-128.

BESSON, J., 1970, "Comparison of National Time Standards by Simple Ovefflight", IEEE Trans. Instrumen. Meas., V. IM-19, pp. 227-232.

THOMPSON,Jr., M.C., WOOD, L.E., SMITH, D. and GRANT, W.B., "Phase Stabilization of Widely Separated Oscillators", 1968, IEEE Trans. Antennas Propagat., V. AP-16, pp. 683-688.

JESPERSEN, J.L., BLAIR, B.E., and GATTERER, L.G., 1972, "Characterization and Concepts of Time-Frequency Dissemination", Proc. IEEE, V. 60, pp. 502-521.

SORENSON, H.W., 1970, "Least-Squares Estimation: from Gauss to Kalman", IEEE Spectrum, V. 58, pp. 63-68.

BLAIR, B.E. (ed.), 1974, "Time and Frequency: Theory and Fundamentals", NBS Mono. 140, National Bureau of Standards, Washington, DC.

EVANS, J.E., BURCHSTED, R.B., CAPON, J., ORR, R.S., SHNIDMAN, D.A., and SUSSMAN, S.M., 1976, "MLS Multipath Studies, Vol. II, Application of Multipath Model to Key MLS Performance Issues", Rpt. No. FAA-RD-76-3, II, MIT/Lincoln Laboratory, Lexington, MA.

EVANS, J.E., BURCHSTED, R., CAPON, J., ORR, R.S., SHNIDMAN, D.A., AND SUSSMAN, S.M., 1976., "MLS Multipath Studies, Vol. I, Mathematical Models and Validation", Rpt. No. FAA-RD-76-3-, 1, MIT/Lincoln Laboratory, Lexington, MA.

CAPON, J., 1976, "Multipath Parameter Computations for the MLS Simulation Computer Program", Rpt. No. FAA-RD-76-55, MIT/Lincoln Laboratory, Lexington, MA.

DISCUSSION

M.Goutelard

Les systèmes que vous mentionnez concernant la protection des passages à niveau et des Trains en général dans les systèmes de traffic à vitesse élevée fait état d'émetteurs placés sur un véhicule et de récepteurs placés au sol ou sur d'autres véhicules. Cette solution n'a pas été retenue en France. Comment palliez-vous à une panne de l'emetteur et quelles sont les précautions prises pour éviter les collisions?

Author's Reply

The system proposed for the protection of fleet vehicles like transport and school buses has not been tried. It is offered as a possible solution to prevent grade crossing collisions. It assumes that there is no institutional barrier to controlling the speed of highway vehicles as the train approaches a crossing – which of course is not true. In the US this procedure is viewed as an invasion of civil rights and therefore will not happen for a very long time. In the meantime we should continue to explore technological approaches in order to assure the implementation of the most cost effective and socially desirable system possible. Existing systems in the USA still use gates, flashing lights or simple signs for those crossings that are not by-passed.

LES BESOINS EN SYSTEMES DE NAVIGATION A GRANDE ET COURTE PORTEE

POUR LES NAVIRES CIVILS ET MILITAIRES

A.M. Poubertou
Service Hydrographique et Océanographique de la Marine
3, avenue Octave Gréard - Paris
France

RESUME

Il existe actuellement en service ou en projet un grand nombre de systèmes de radiolocalisation à la mer, faisant souvent double emploi. L'analyse détaillée des besoins fait apparaître cependant une telle diversité qu'une solution unique de compromis est impossible. Si on pouvait négliger le facteur coût, un système très sophistiqué, tel que le projet NAVSTAR américain satisferait sans doute le plupart des besoins. Si on prend en compte le facteur coût, on aboutit à plusieurs familles de solutions de compromis, en fonction de la portée et de la précision requises, évidemment antagonistes. La présente étude essaie de définir ces familles, et de montrer que même à l'intérieur de chacune d'elles, le poids de certains besoins spécifiques, et les préoccupations nationales, rendent peu probable la conception d'un système unique. Les systèmes existants sont ensuite classés et analysés du point de vue de leur adaptation aux besoins reconnus, pour les différentes catégories d'utilisateurs civils et militaires, en s'attachant à souligner les perfectionnements qui seraient souhaitables à court terme.

1. INTRODUCTION

C'est un lieu commun d'observer que, depuis leur apparition il y a quelques décennies, les systèmes de radionavigation à l'usage des navires et/ou des aéronefs, et plus généralement les systèmes de radiolocalisation à l'usage de véhicules techniques plus spécialisés, se sont multipliés de façon anarchique et largement redondante. Je n'ai pas l'intention d'analyser les causes de cette prolifération, dont les principales - innovations techniques, particularismes nationaux, concurrence entre firmes, etc... - sont d'ailleurs évidentes. Je me contente de souligner, après bien d'autres auteurs, le grave inconvénient qui en résulte pour l'utilisateur, parfois acculé à des choix difficiles, ou à une accumulation de matériel coûteux, et les difficultés qui résultent de l'encombrement du spectre radioélectrique. Ces dernières doivent rester toujours présentes à l'esprit ,à la veille de la prochaine conférence de l'U.I.T. prévue pour 1979.

Beaucoup de ces systèmes, de conception déjà ancienne, ont été développés et mis en place sans que probablement une analyse très soignée sur le plan technique du besoin à satisfaire ait été effectuée. Ils représentaient ce que les laboratoires et l'industrie de l'époque pouvaient offrir ; leurs performances, quels que soient les critères envisagés, restaient très loin de l'idéal, et l'utilisateur ne pouvaient que s'en contenter. L'appellation encore courante, chez les usagers maritimes, d'"aides radioélectriques à la navigation", plutôt que de "systèmes de radionavigation", traduit bien ce caractère de méthode d'appoint, et non de procédé autonome, qui leur était reconnu par les navigateurs. Les systèmes plus récents ont été généralement développés dans le but exprès de satisfaire certains besoins, mais souvent défini étroitement : couverture d'une certaine région, satisfaction d'un certain type de navigation, ou d'une certaine catégorie d'usagers, etc... Économiquement, pour employer le langage des gestionnaires industriels, leur promotion correspondait souvent à l'occupation d'un "créneau", parfois à l'aide d'un matériel spécifiquement conçu pour un usage différent, et non révisé dans son principe. Il convient aujourd'hui d'examiner si une réponse appropriée plus globale, voire une réponse à l'ensemble de tous les besoins, ne serait pas possible sur une base plus rationnelle, avec une économie de moyens, ou même un système unique de compromis. C'est la question à laquelle je vais essayer de répondre dans les lignes qui suivent, en précisant une fois pour toutes que je me place du seul point de vue de l'usager maritime, à l'exclusion de tous les autres utilisateurs possibles de la radiolocalisation (aériens ou terrestres) et sans me dissimuler que ce point de vue réduit dans une certaine mesure la portée de mes conclusions.

2. ETUDE DES CRITERES

Pour évaluer le besoin en termes de performances des systèmes destinés à le satisfaire, il faut faire choix d'un certain nombres de critères. S'il s'agit de comparer entre eux plusieurs systèmes, ou de faire une analyse partielle assez fine, ces critères, qui sont en très grand nombre, doivent être soigneusement détaillés. Mais au niveau global où cette étude se place, les grandes solutions ne peuvent être que de compromis, il faut au contraire rester assez simple. C'est pourquoi, dans un premier temps, après une analyse assez détaillée, je m'efforcerai de définir un choix de critères simples et en petit nombre qui sera utilisé dans la suite. Ce choix n'est évidemment pas

le seul possible, ni peut-être le meilleur, mais il fournit une base de comparaison commode et menable.

L'ensemble des critères sera divisé en trois groupes :

- critères de caractère technique et opérationnel ;
- critères de caractère économique ;
- critères spécifiques du besoin militaire.

2.1. Critères techniques et opérationnels

2.1.1. Couverture du système

C'est-à-dire étendue de la zone géographique dans laquelle le système est capable de fournir un positionnement. Cette notion est en réalité assez complexe pour diverses raisons :

- la couverture d'un système est souvent fonction des autres caractéristiques, et notamment, pour les systèmes à base terrestre, de la précision de positionnement, soit pour des raisons radioélectriques, soit pour des raisons géométriques, soit normalement pour une combinaison des deux. Certains systèmes admettant un fonctionnement en mode dégradé (LORAN C sur onde de ciel, par exemple), qui permet une extension notable de la couverture, avec une précision évidemment beaucoup plus faible.

- la couverture d'un système peut être variable dans le temps, comme c'est le cas pour beaucoup de systèmes hyperboliques à moyenne portée soumis à l'effet de nuit (perturbation par les ondes ionosphériques).

- enfin, bien entendu, il faut distinguer la couverture "potentielle" d'un système et sa couverture "réelle" telle qu'elle est définie, à une époque donnée, par les implantations existantes. On pourrait définir la couverture potentielle comme la région qui serait couverte par l'ensemble maximal (et optimal) de stations qui pourrait être effectivement installées, compte tenu, notamment, des contraintes d'encombrement du spectre.

Nous retiendrons, pour les besoins de cette étude, un paramètre simplifié unique de couvertures (pour les systèmes à couverture non mondiale). Il décrira la zone dans laquelle le système considéré est capable de fournir en tout temps un positionnement de précision voisine de sa "précision typique annoncée" telle qu'elle est définie ci-après. Etant donné que le besoin maritime est, de façon très générale, d'autant plus exigeant que l'on est proche des côtes, cette zone sera définie par la distance entre la côte et sa limite extérieure vers le large : C'est bien là ce que les marins entendent par la "portée" d'un système.

2.1.2. Précision de positionnement

C'est également une notion complexe :

- pour la plupart des systèmes, elle est variable dans l'espace, comme nous l'avons déjà noté. Pour un système donné, elle peut s'établir à plusieurs niveaux, s'il existe des modes de fonctionnement dégradés, ou des modes d'utilisation distincts comme pour NAVSTAR. Elle peut varier dans le temps, par exemple par effet de nuit comme pour OMEGA. Elle est enfin fonction d'autres caractéristiques du système, telles que ses performances dynamiques relativement à la vitesse du véhicule, bien que ce dernier aspect soit tout à fait secondaire dans le cas de la navigation maritime.
- la notion de précision est complexe en elle-même, du fait qu'elle doit nécessairement s'exprimer en termes de probabilités. Le paramètre "écart circulaire probable" (ECP) ou toute autre analogue, en donne une bonne mesure dans le cas où la loi de dispersion des écarts ne s'écarte pas trop grossièrement du modèle gaussien, et c'est de cette façon que la précision est couramment définie. (Observons dès à présent que l'on ne serait pas dans ce cas si l'on décidait par exemple de considérer une loi de probabilité globale, tenant compte d'un pourcentage d'erreur sur le comote des chenaux, dans un système comportant une ambiguïté : Il devient alors impossible de caractériser par un paramètre unique tel que ECP le comportement de l'erreur. Ce point est évoqué plus loin).
- il convient également, du moins dans le cas des systèmes prétendant à des performances élevées, de distinguer les diverses notions recouvertes par le mot "précision" : précision relative (repeatability) ou fidélité, caractérisant l'aptitude du système à donner des indications identiques dans le temps (mais éventuellement biaisées) en une position donnée, et précision absolue (predictability). La première est fonction de la précision instrumentale (sensitivity) ou sensibilité, et du niveau d'erreurs aléatoires, la seconde inclut également les erreurs systématiques de toute nature. L'une et l'autre sont distinctes (mais évidemment liées) de la précision différentielle, qui caractérise l'aptitude du système à positionner simultanément deux utilisateurs l'un par rapport à l'autre. Cette dernière notion n'est pas indifférente et peut conduire à des modes d'exploitation systématiques, comme dans le cas de l'OMEGA différentiel.

Pour une analyse simple, nous pouvons nous contenter de décrire la précision d'un système par un paramètre unique caractérisant l'écart maximal de positionnement. Ceci suppose que nous ne tenons pas compte des erreurs grossières éventuelles liées par exemple au levé d'ambiguïté. D'autre part il convient de choisir un écart maximal correspondant à un haut niveau de probabilité, typiquement 95 % pour les systèmes à faible précision et 98 % ou plus pour les systèmes à grande ou très grande précision. Des valeurs inférieures, si elles sont significatives pour des techniciens (et correspondent d'ailleurs à une parfaite connaissance de la répartition des écarts si la loi de probabilité est connue) ne le sont pas pour l'usager courant qui estime pouvoir se fier de façon quasi absolue à la précision indiquée. Nous ne préciserons pas si cette précision est absolue ou relative : pour des systè-

mes à implantation fixe, ou bien en effet le positionnement absolu est de peu d'intérêt, l'essentiel étant de pouvoir retrouver une position déjà occupée ou une route déjà parcourue, ou bien il peut être atteint à partir du positionnement relatif par un jeu de corrections convenable, à condition que l'on dispose d'un champ de points d'étalonnage approprié. Enfin, pour tenir compte de la variabilité dans l'espace et dans le temps, nous conviendrons que la valeur indiquée correspond à la précision moyenne minimale assurée en tout temps dans la zone de couverture, dans des conditions normales d'utilisation, c'est-à-dire qui ne s'écartent pas exagérément des conditions habituellement rencontrées. Comme on le voit le paramètre ainsi défini correspond à la précision réellement obtenue avec un très haut niveau de fiabilité. Nous l'appelerons "précision annoncée" du système.

2.1.3. Contenu informationnel

Je regroupe sous ce titre un certain nombre de caractéristiques secondaires précisant la nature de l'information de positionnement fournie par un système. Il s'agit pour l'essentiel de la réponse à quatre questions :

- Le système fournit-il une position seule, ou comporte-t-il une information complémentaire, notamment la vitesse absolue du mobile sur le fond ? fournit-il une position à deux ou trois dimensions (pour mémoire, puisque cette information est sans intérêt pour le navigateur maritime, bien qu'elle soit susceptible, peut-être, de fournir un élément de contrôle de la qualité du résultat) ?
- La réponse du système est-elle continue (positionnement permanent), discontinue suivant un rythme plus ou moins lent (positionnement récurrent), discontinue sur interrogation du mobile ? La réponse du système est-elle en temps réel ou en temps différé ?
- Le positionnement fourni comporte-t-il une ambiguïté ? Ce point est particulièrement important pour deux raisons : la plupart des systèmes hyperboliques comportent une ambiguïté résiduelle qui nécessite, pour leur mise en œuvre, une connaissance approchée de la position par un moyen extérieur au système, avec une précision qui est parfois assez grande. D'autre part les dispositifs de "levé d'ambiguïté" propres au système, lorsqu'ils existent, outre qu'ils laissent subsister l'ambiguïté résiduelle mentionnée plus haut, sont susceptibles d'introduire une erreur de une ou plusieurs unités dans l'identification des chenaux des divers ordres. Nous avons négligé cette source d'erreur dans la définition de la précision, admettant ainsi que le levé d'ambiguïté doit fonctionner avec une sécurité de 100 %.
- L'information fournie par le système est-elle ou non redondante ? Il est clair qu'un système fournissant plus de deux lieux de position indépendants procure de ce fait un moyen de vérification de la position obtenue de nature à accroître considérablement la fiabilité et la précision.

De ces quatre critères, je ne retiendrai que le second, qui est important du fait que certaines catégories d'usagers ne s'accomodent pas par exemple d'un positionnement récurrent. Le caractère de redondance, au niveau de notre analyse, n'a pas à être examiné à part, mais seulement à intervenir dans la note de fiabilité générale définie plus loin. Il est évident enfin que le système idéal, dans tous les cas, ne devrait comporter aucune ambiguïté résiduelle. Cependant un certain niveau d'ambiguïté est la plupart du temps tolérable en pratique, en fonction notamment de la précision demandée : par exemple un système conçu pour positionner à 10 m près s'accomodera normalement d'un inconvénient d'une ambiguïté de plusieurs milles, les usagers exigeant un tel niveau de performance devant posséder par ailleurs les moyens de se placer sans erreur à cette précision. Autrement dit je considère qu'à chaque catégorie de besoin défini en précision est associé un certain niveau d'ambiguïté acceptable qu'il est inutile de préciser davantage.

2.1.4. Particularités diverses

Je classe ici pour mémoire un certain nombre de caractéristiques secondaires dont voici quelques unes :

- le système est-il passif (l'installation de bord ne comporte que des récepteurs) ou actif (le mobile doit émettre) ?
- dans le cas des systèmes actifs, quel est le "niveau de saturabilité" du système (c'est-à-dire combien admet-il d'interrogateurs mobiles en fonctionnement simultané) ?
- quelle est la nature de la donnée brute fournie par le système (azimuts, distances, différences de distances, etc...) ? Quel est le niveau de traitement requis (y compris les calculs de correction), du moins si ce traitement n'est pas systématiquement intégré au système, pour obtenir une donnée directement exploitable par l'usager ?
- quelles sont les diverses servitudes d'exploitation ? par exemple facilité de mise en œuvre et flexibilité d'emploi du matériel de bord, et le cas échéant du matériel des stations fixes (pour les systèmes "mobiles" à usage spécialisé) ; robustesse générale, facilité de réglage, d'entretien et de dépannage (disponibilité et qualité du service après vente ou du service de location-entretien), etc...

Je ne retiendrai pas le détail de ces critères, mais seulement le dernier, sous forme d'une "note de qualité" globale attribuée au système, avec trois niveaux :

- Excellente : matériel sûr et très robuste devant pouvoir être réglé et utilisé sans précaution particulière par du personnel sans compétence spéciale.
- Bonne : matériel de bonne robustesse, dont la mise en œuvre peut exiger certaines précautions et l'entretien le concours de personnel compétent.
- Courante : matériel sûr sous réserve de certaines contraintes d'environnement, dont la mise en œuvre correcte et l'entretien exigent des compétences particulières.

2.1.5. Disponibilité opérationnelle

Les caractéristiques regroupées sous ce titre sont d'une importance capitale, puisqu'elles décrivent l'environnement de mise en œuvre d'un système, indépendamment de ses qualités intrinsèques:

- quel est le niveau de fiabilité du signal reçu par le mobile, et notamment ce signal est-il correctement protégé en tout temps contre les bruits atmosphériques, les fluctuations et perturbations ionosphériques ou troposphériques, contre les interférences dues à l'encombrement du spectre radioélectrique ou à son propre fonctionnement (réflexions, parcours multiples, etc) ? Ce critère est évidemment à examiner en liaison avec d'autres (ambiguïté, caractère permanent ou fréquence de récurrence, redondance de l'information, etc.). Inversement, dans le cas des systèmes mobiles, les perturbations causées par le système ne risquent-elles pas d'en limiter l'utilisation à certaines périodes ou à certaines heures, en fonction de la réglementation ?

- la permanence du fonctionnement est-elle convenablement assurée, compte tenu, là encore, des éventuelles redondances ? Un réseau d'information concernant les dates d'interruption (pour entretien) et les dates de remise en service (après entretien ou incident) fonctionne-t-il de façon satisfaisante ?

- le Système a-t-il une durée de vie garantie raisonnable ? La compatibilité des matériels anciens est-elle assurée en cas de changement aux spécifications ?

Pour les besoins de notre analyse, je ne retiendrai pas ce troisième critère, non parce qu'il est secondaire, mais au contraire parce qu'il est capital. Tout système qui n'offrirait pas une garantie de durée suffisante serait impropre à satisfaire un besoin, quel qu'il soit, et devrait être rejeté. Les autres critères sont rassemblés en une "note de fiabilité" globale, à trois niveaux :

- Excellente : système offrant un très haut niveau de fiabilité en toutes circonstances.

- Bonne : système normalement très fiable, mais sujet à des défauts passagers, ou à certaines contraintes dans certaines zones.

- Courante : système de fiabilité parfois incertaine, et/ou soumis à des contraintes d'utilisation bien définies.

2.2. Critères économiques

Le coût d'un système, du point de vue de l'utilisateur, revêt deux aspects :

2.2.1. Coût du matériel

Il s'agit là du prix d'achat, d'amortissement et d'entretien du matériel de bord (et de la totalité du matériel dans le cas d'un système mobile), ou du coût de sa location. Un même système, s'il admet plusieurs modes d'exploitation (normal, dégradé, minimal) peut évidemment comporter plusieurs matériels de coûts différents.

2.2.2. Redevances d'utilisation

Jusqu'à ce jour, tous les systèmes en exploitation (autres que les systèmes mobiles, bien entendu) sont d'accès libre pour l'ensemble des utilisateurs, les frais d'installation et de fonctionnement des infrastructures étant à la charge des gouvernements ou de certaines autorités locales telles que les ports. Il n'est nullement garanti que cette situation se perpétuera, et il est même peut-être souhaitable que le financement de certains systèmes soit assuré, à l'avenir, par les usagers eux-mêmes, au moins en partie. Cette contribution prendrait alors la forme d'une redevance d'utilisation à acquitter sous une forme ou sous une autre par le détenteur d'un équipement de bord.

Nous retiendrons un critère unique global de coût, intégrant les deux composantes ci-dessus, et traduisant le coût total du positionnement au niveau de l'utilisateur, tel qu'il apparaîtrait par exemple dans le compte d'exploitation d'un navire (et comprenant donc le prix de revient de toutes les installations annexes devant éventuellement compléter le matériel de radiolocalisation proprement dit). Trois niveaux seulement sont retenus pour cette analyse :

- Modéré : Coût ne dépassant pas de façon sensible celui des systèmes les moins coûteux actuellement exploités.

- Moyen : Coût pouvant dépasser assez largement celui des systèmes actuels les moins coûteux.

- Peu Sensible : Coût acceptable relativement élevé ou très élevé.

2.3. Critères spécifiques militaires

En plus de tous les critères précédemment énumérés, et qui caractérisent évidemment aussi bien les besoins des navires militaires que ceux des navires civils, l'utilisation du positionnement radioélectrique pour les opérations militaires sur mer impose un certain nombre d'éléments de choix supplémentaires. Pour une analyse sommaire, on peut les ramener à trois critères principaux :

2.3.1. Disponibilité du système en cas de conflit

Ceci recouvre la garantie du fonctionnement de l'infrastructure (protection des stations

contre les coups de l'ennemi) et la résistance du système au brouillage. Un élément intéressant est la résistance au brouillage "intelligent", capable d'introduire dans le positionnement fourni un biais non décelable par l'utilisateur.

2.3.2. Discrétion de l'utilisation

Celle-ci peut être nécessaire (ce qui exclut alors l'emploi de systèmes actifs), mais ce n'est pas le cas systématiquement pour toutes les utilisations militaires.

2.3.3. Garantie contre l'utilisation par l'adversaire

3. REMARQUES GENERALES

A l'aide des critères simplifiés précédemment décrits, et de l'information à la disposition de l'auteur, il est alors possible de dresser le tableau ci-joint, qui décrit à l'aide de ces critères un certain nombre de besoins-types correspondant aux différentes catégories d'utilisateurs. Avant de passer à l'examen de ce tableau, un certain nombre de remarques sont à faire.

3.1. Tout d'abord, les sources utilisées sont essentiellement françaises. Au delà des enquêtes mentionnées en références, et d'un certain nombre d'études partielles non citées, l'auteur a utilisé sa propre expérience et les entretiens qu'il a eu au cours des années récentes avec de nombreux représentants des diverses catégories d'usagers. Il est improbable que des conclusions significativement différentes soient obtenues à partir de sources provenant d'autres pays que la France. L'analyse des articles de revue et de certaines communications présentées à des congrès de navigation ne fait en tout cas apparaître aucune divergence notable.

3.2. Les valeurs chiffrées indiquées dans certaines colonnes du tableau tiennent parfois compte d'une divergence assez large entre les différentes réponses obtenues au cours des enquêtes. Cela tient parfois aux différences intrinsèques entre les besoins de catégories d'usagers différents regroupés sur une même ligne, c'est-à-dire au caractère assez synthétique du tableau : il n'était pas question de détailler, au niveau de cette étude, des catégories aussi spécialisées, et de faible effectif, que les vedettes de patrouille des Douanes, les navires balisseurs des Services Maritimes, les bateaux de sauvetage, etc..... Mais cette incertitude est également liée à la formulation même de leur besoin par les différents usagers interrogés, qui n'est pas toujours aussi claire qu'il serait souhaitable, quelles que soient les précautions prises. Par exemple, il n'est pas toujours facile de savoir si la précision de positionnement indiquée correspond à la précision idéale souhaitée, à la précision considérée, au moins provisoirement, comme satisfaisante, ou seulement à une précision minimale requise, en dessous de laquelle un système de positionnement n'apporterait pas d'aide appréciable. Seul, bien souvent, le contact direct avec les usagers permet d'obtenir, sur ce point des réponses nettes.

3.3. Un autre élément qui peut être source d'ambiguïté tient à l'attitude des usagers vis-à-vis de leur propre besoin. On peut les classer en deux groupes principaux de ce point de vue : d'une part ceux qui, par vocation et par habitude intellectuelle, ont les moyens et la volonté de se livrer à une analyse objective de leur besoin et savent le définir en termes clairs, indépendamment des systèmes existants dont ils peuvent avoir l'expérience ; d'autre part ceux qui n'ont pu ou su faire cette analyse et se montrent incapable de définir leur besoin sinon en se référant aux performances des systèmes qu'ils connaissent. Si l'on interroge sans précaution, par exemple, une population de pêcheurs habitués à fréquenter une zone où deux systèmes différents coexistent en concurrence, il y a des chances que l'on obtienne une définition du système idéal qui se ramène à citer, critère par critère, les performances du meilleur de ces deux systèmes connus des usagers. Ces derniers, en outre, ont une grande difficulté à concevoir en termes de probabilités, comme il convient, la notion de précision. Les valeurs qu'ils indiquent correspondent, dans leur esprit, à l'erreur maximale de la position fournie par le système avec une quasi-certitude. C'est la raison pour laquelle j'ai adopté, pour la définition de ce critère, une indication d'écart maximal probable correspondant à une probabilité très élevée. Avec une définition différente, les valeurs indiquées par les usagers de cette catégorie n'auraient pas eu de signification.

On peut considérer que la première catégorie d'usagers comprend les "Techniciens" des opérations à la mer (navires de recherche, de prospection, de levé, câbliers, grande pêche "scientifique", etc.) et les militaires. La seconde catégorie comprend la plus grande partie des navigateurs proprement dits, marins du commerce, de la pêche et de la plaisance, et c'est bien entendu de beaucoup la plus nombreuse.

3.4. L'interprétation du tableau est je pense assez claire. Les treize lignes horizontales définissent autant de catégories d'utilisateurs, ou plus exactement d'utilisation, du positionnement radioélectrique, correspondant, comme je l'ai déjà indiqué plus haut, à un certain niveau de regroupement, pour les plus spécialisées. Les deux premières colonnes donnent la valeur des paramètres de couverture et de précision : il s'agit en principe des valeurs optimales jugées souhaitables par les utilisateurs, c'est-à-dire que pour chaque besoin une amélioration de ces performances est jugée sans intérêt opérationnel (mais bien entendu, dans bien des cas, des performances moindres seraient considérées comme déjà intéressantes). La troisième colonne indique si le positionnement doit être continu, ou si un positionnement récurrent est jugé suffisant. Les deux suivantes donnent pour chacun des deux critères de qualité et de disponibilité le niveau minimal requis pour que le système considéré soit considéré comme satisfaisant (étant bien entendu que le système idéal devrait presque toujours être noté "excellent" !). Il en est de même pour la sixième colonne, l'indication de coût donnée correspondant à la dépense maximale que l'utilisateur est disposé à consentir pour obtenir les performances décrites dans les colonnes précédentes.

3.5. Une conclusion se présente dès le premier examen d'ensemble de ce tableau : c'est l'extrême diversité du besoin, quel que soit le critère considéré, mais particulièrement en ce qui concerne couverture et précision. La satisfaction globale de ces besoins sur la base d'un système unique conduirait à un jeu de spécifications si contraignant que cette solution paraît hors d'atteinte. Et cependant, à condition de laisser de côté le critère de coût pour ne considérer que les critères techniques et opérationnels, il est clair qu'une telle solution n'est pas hors de portée. Le système GPS des forces armées américaines, actuellement en cours de développement satisfait en principe, dans sa version "protégée" de grande précision, la quasi-totalité des besoins exprimés, et pour beaucoup d'entre eux bien au-delà du minimum nécessaire. Il satisfait également les besoins militaires, puisqu'il a été conçu expressément dans ce but. Mais en pratique, il n'en est pas de même. En effet, d'une part, si le système GPS est finalement implémenté, la version de grande précision ne sera pas disponible pour les usagers civils. D'autre part, et pour autant que je sois informé, les conditions d'accès des usagers civils à la version "non protégée" de moindre précision, ne sont pas définies à l'heure actuelle. Il n'est pas certain que cet accès soit possible à tous. En tout état de cause la précision annoncée (100 m) est insuffisante pour couvrir certains besoins particuliers, assez spécialisés il est vrai. Mais surtout, si l'on prend en compte le critère économique, il semble bien que l'on doive aboutir à des matériels suffisamment sophistiqués pour que leur coût soit prohibitif aux yeux de certaines catégories d'usagers, même en admettant que l'accès au GPS ne comporte le paiement d'aucune redevance. Dans ces conditions, il faut constater que la solution unique satisfaisant en bloc l'ensemble des besoins n'est pas envisageable en pratique pour l'avenir prévisible.

L'examen plus détaillé du tableau permet de dégager trois catégories de besoins, qui correspondent :

- à un positionnement de faible précision mais de très grande couverture pour les besoins de la navigation océanique : typiquement, couverture mondiale et précision de l'ordre de 3 milles. Cependant les besoins actuels ou prévisibles de certaines catégories d'utilisateurs spécialisés sont beaucoup plus exigeants sur la précision.
- un positionnement de bonne précision jusqu'à une certaine distance des côtes, pour les besoins de l'atterrissage et de la navigation côtière courante, de la pêche, et des activités d'exploitation du plateau continental. Cependant les normes de précision nécessaires sont assez différentes d'une catégorie d'utilisateurs à l'autre, bien que les plus nombreuses soient relativement peu exigeantes, et la portée typique demandée, qui varie largement suivant la région géographique à couvrir, est mal définie.
- un positionnement de très bonne précision, pour la navigation dans les zones difficiles à forte densité de trafic, les zones dangereuses pour les navires à grand tirant d'eau, et le chenalage aux approches des ports importants. Ce besoin manifeste une tendance à croître en importance pour des raisons en partie liées au développement des mesures de réglementation du trafic.

4. ETUDE DU BESOIN

Nous allons examiner de plus près chacune de ces catégories d'utilisation, en fonction des classes de systèmes aptes à les satisfaire.

4.1. Systèmes à couverture mondiale ou étendue

Pour la plupart des usagers pratiquant la navigation océanique courante, le besoin d'un système de positionnement au large n'est pas ressenti de façon pressante. D'une part il est vrai que la conduite du navire n'impose pas une grande précision, et la plupart des commandants s'accomodent fort bien des procédés classiques de navigation, à l'estime recalée par observations astronomiques. D'autre part, pour des raisons psychologiques, on observe souvent une certaine réticence chez les marins envers les aides radioélectriques utilisés dans les zones où ils ne sont pas vraiment indispensables. Ils y voient, semble-t-il, une menace sur la liberté traditionnelle de la navigation en haute mer, et manifestent confusément la crainte de voir un jour le trafic maritime soumis à une réglementation contraignante sur le modèle du trafic aérien, bien que cette hypothèse ait bien peu de chances de se réaliser dans l'avenir prévisible. Le fait que ce besoin soit très faiblement ressenti explique le niveau élevé des exigences formulées par les usagers. Aucun système ne sera jugé satisfaisant s'il ne combine pas une disponibilité élevée et une excellente qualité, tout en restant d'un coût modéré.

Deux catégories de systèmes sont à l'heure actuelle susceptibles de satisfaire ce besoin, les systèmes hyperboliques à très basse fréquence (VLF) du type OMEGA, et les systèmes à satellites (NNSS, GPS ou tout autre). Les premiers, en l'état actuel de la technique d'utilisation, comportent une ambiguïté et la nécessité de corrections relativement complexes, exigeant une documentation abondante, mais fournissent un positionnement continu, et pour un coût normalement peu élevé. Un système tel que OMEGA, dont la précision correspond bien au niveau souhaité, sera pleinement satisfaisant, relativement à la formulation du besoin, si un système de calcul et d'application automatique des corrections peut lui être adjoint, et si un dispositif simple de contrôle du levé d'ambiguïté permet d'exploiter toute l'information du signal. Pour l'usager courant, qui ne dispose d'aucun système "intégré" impliquant l'emploi d'un ordinateur de bord capable de prendre ces fonctions en charge, les dispositifs correspondants devraient bien entendu faire partie du récepteur, tout en veillant à rester dans des limites de coût raisonnables.

Les systèmes à satellites du type NNSS ne fournissent qu'un positionnement recurrent, mais ce point n'est pas ressenti comme un inconvénient. Ils exigent de toute façon l'emploi d'un ordinateur de bord, ou d'un processeur spécialisé, qui peuvent éventuellement être utilisés pour intégrer les éléments de l'estime et fournir une position permanente. Du fait de leur coût plus élevé, et bien que des installations TRANSIT complètes soient désormais offertes à des prix très raisonnables, les systèmes à satellites n'ont pas actuellement la faveur des navigateurs. Le nombre de récepteurs de bord est ex-

trêmement faible dans la marine de commerce française.

Si le besoin courant est convenablement satisfait par les systèmes VLF, il n'en est pas de même en revanche pour celui des navires spécialisés (navires de recherche océanographique et géophysique, câbliers, navires de prospection des ressources sous-marins et bientôt navires d'exploitation en eau profonde, etc...), qui sont beaucoup plus exigeants sur le critère de précision. En contrepartie, le critère de coût est incomparablement moins critique pour ces usagers, qui sont également disposés à accepter des performances de qualité et de disponibilité moins sévères. Les navires spécialisés de cette catégorie ont tendance à s'équiper de systèmes intégrés capables d'élaborer automatiquement la position à partir des systèmes actuellement disponibles et d'équipements autonomes, tels que capteurs cap-vitesse ou inertiels. La gamme de précisions indiquée sur le tableau est très large, et correspond aux besoins des différentes opérations à la mer, extrapolés pour l'avenir prévisible. Les systèmes actuels n'en couvrent qu'une partie, et certaines opérations ne peuvent être exécutées par les seuls moyens du positionnement radioélectrique mais nécessitent le recours à des procédés auxiliaires tels que les balises acoustiques mouillées sur le fond. Il est évident que le véritable progrès serait ici l'accès au GPS, dans sa version protégée de haute précision, et ce besoin est je pense l'une des motivations du développement de ce système. D'autres systèmes conçus sur des spécifications différentes, tels que GEOLE proposés par Centre National français d'Etudes Spatiales (CNES) pourraient également apporter une solution au moins partielle. Il en est également ainsi pour le LORAN C (dans des régions géographiquement limitées, bien entendu) dans la mesure ou les techniques de traitement du signal et les méthodes de correction sont susceptibles d'améliorer encore la précision de ce système.

4.2. Systèmes à moyenne portée

Comme nous l'avons déjà constaté, la diversité du besoin est ici beaucoup plus grande. Une analyse simple permet de distinguer, pour le positionnement radioélectrique, les fonctions suivantes :

- atterrissage et navigation côtière, pour l'ensemble des navigateurs civils et militaires de toutes catégories. Le gain en précision, par rapport à la navigation océanique, provient évidemment des dangers de toute sorte que doit affronter le navigateur dans les petits fonds, et de l'augmentation parfois considérable de la densité du trafic. Les moyens traditionnels, positionnement optique sur amers terrestres, sur balisage flottant, sur balisage lumineux, emploi du radar et du sondeur, continuent d'être utilisés et sont souvent considérés comme satisfaisants. Les radiophares et les stations radiogonométriques sont toujours largement utilisés pour l'atterrissage. Mais de plus en plus, dans les zones à forte densité de trafic à tout le moins, la nécessité de systèmes de positionnement radioélectriques plus précis et plus sûrs est ressentie. La portée indiquée est relativement faible, ce qui indique bien que ce procédé est ressenti comme un substitut aux méthodes optiques ou radar, et la précision est modeste. Mais les performances de disponibilité et de qualité exigées sont élevées, ce qui est normal pour un positionnement destiné avant tout à garantir la sécurité du navire.
- navigation précise en zone sensible. Il existe cependant un certain nombre de régions peu éloignées des côtes où les normes de précision souhaitées sont plus strictes que pour la navigation côtière courante : Ce sont en particulier celles ou la densité du trafic est plus importante en même temps que l'environnement naturel est plus difficile : régime de la marée, des courants, conditions météorologiques, profondeurs faibles éventuellement encombrées d'obstructions diverses, présence d'installations off-shore. Il convient de mentionner ici le cas particulier des très grands navires, que leur tirant d'eau contraint souvent, en de telles zones, à emprunter des couloirs de navigation dont ils ne peuvent s'écarter sans danger. Nous appelerons ces régions "zones sensibles", l'exemple le plus caractéristique étant sans doute actuellement la Mer du Nord. La nature du besoin s'y caractérise par une exigence accrue en précision, ainsi qu'en fiabilité, avec en contrepartie l'acceptation d'un coût plus élevé, qui traduit l'intensité croissante avec laquelle le besoin est ressenti. Hors de vue de terre, il n'existe en effet aucune méthode classique autre que la navigation astronomique. En présence d'un balisage, ou pour se tenir à l'écart des structures off-shore, on peut évidemment utiliser le radar, mais il ne constitue un moyen de navigation proprement dit qu'aux approches des côtes. On observe d'ailleurs une tendance à le considérer désormais surtout comme une aide anticollision, le positionnement du navire devant être obtenu par d'autres moyens.
- besoins particuliers de la pêche. Le besoin tel qu'il est défini sur le tableau est formulé essentiellement par les chalutiers. La fonction à remplir est ici de permettre à l'usager de retrouver de façon précise les emplacements et les routes qu'il a déjà pratiqués et où son expérience lui indique que son activité sera fructueuse, et surtout de se tenir à l'écart des accidents de relief et obstruction de toutes sortes susceptibles de détériorer son matériel, dont il connait la position par sa propre expérience et celle de ses confrères. Ceci conduit à pratiquer une navigation souvent extrêmement précise, dans des régions qui s'étendent normalement à la totalité des plateaux continentaux, en attendant le développement du chalutage par grande profondeur. Le besoin exprimé est donc assez sévère sur les critères de couverture et de précision, et c'est semble-t-il en contrepartie qu'il manifeste une exigence moindre dans le domaine de la disponibilité. La qualité doit être excellente et le coût rester modéré, comme on pouvait s'y attendre de la part d'usagers ayant des moyens financiers limités, des bateaux souvent frustes et un personnel sans qualification particulière.
- exploitation off-shore du plateau continental. Ces usagers indiquent pour les critères de couverture et de précision des valeurs tout à fait comparables à celles des pêcheurs, mais leurs exigences quant aux trois autres critères sont inversées. Une excellente disponibilité est jugée nécessaire, tandis que les moyens financiers et la qualification du personnel disponible permettent d'accepter une qualité moindre et un coût plus élevé.

Il existe un assez grand nombre de systèmes actuellement exploités, expérimentés ou proposés, et capables de faire face à ce besoin, pour la plupart des systèmes à ondes entretenues ou à impulsions du type hyperbolique. En ce qui concerne la navigation côtière courante et l'atterrissage, un procédé tel que l'OMEGA Différentiel est extrêmement séduisant dans son principe : sa précision es-

comptée répond aux préoccuppations de la plupart des usagers. La répartition spatiale de cette précision peut être ajustée de façon logique et efficace en fonction de la répartition géographique du besoin par un choix judicieux des sites des stations de compensation au voisinage des zones sensibles. De plus l'équipement de bord reste d'un coût modéré et se présente comme un simple complément à l'équipement destiné à la navigation océanique. Il devient ainsi possible de satisfaire en bloc les besoins des navigateurs courants à grande et moyenne portée, sur la base d'un système unique, et au plan mondial, avec une grande économie de fréquence. Il est évidemment indispensable, pour atteindre ce but, que les émissions de compensation soient normalisées, faute de quoi l'OMEGA Différentiel perdrait l'essentiel de son intérêt. Je crois savoir qu'un processus est engagé pour atteindre cet objectif, et je formule le vœu qu'il aboutisse rapidement.

Les besoins plus spécifiques de la pêche au chalut mentionnés plus haut ne sont pas couverts, en précision et en portée par l'OMEGA Différentiel. Il en est de même pour d'autres types de navigation, tels que celle des navires à grand tirant d'eau dans les zones sensibles étendues. Il apparaît donc inévitable de recourir, au moins localement, à des systèmes plus performants, mais aussi sans doute plus coûteux. La panoplie existante est assez large (LORAN C, DECCA, TORAN, RANA P 17, pour ne citer que les procédés généralement connus des usagers français). Le problème n'est donc pas technique (hormis le point délicat à résoudre de l'encombrement du spectre), mais tient plutôt aux graves inconvénients de la prolifération. Les navigateurs, de toute évidence, n'admettront que des systèmes très normalisés à forte implantation mondiale, tels qu'ils soient assurés de pouvoir en bénéficier dans toutes les zones qu'ils peuvent avoir à fréquenter sans multiplier leurs équipements. Le cas de la pêche hauturière est peut être à considérer à part : Les zones fréquentées par un ensemble d'usagers donné sont suffisamment limitées et stables pour qu'il soit possible d'envisager des systèmes différents dans des régions différentes. A titre d'exemple, une région telle que la Mer du Nord pourrait être couverte en OMEGA Différentiel et en LORAN C ; une région telle que le Golfe de Gascogne et ses abords, en OMEGA Différentiel et un système spécifique orienté vers les besoins de la pêche, du fait que la navigation des navires à grand tirant d'eau n'y pose aucune contrainte particulière. Bien entendu cet exemple n'est donné que pour illustrer (peut être improprement) les remarques qui précédent, et ne prétend nullement définir une solution idéale !

On ne peut que souhaiter, sans se faire trop d'illusions, que soit recherchée par tous les moyens appropriés, une réduction de cette prolifération de systèmes différents. Parallèlement, on devra garder à l'esprit les contraintes très différentes des diverses catégories d'usagers en matière de coût. Il se trouve par exemple que ceux qui sont parmi les plus exigeants en matière de portée et de précision, ainsi qu'en matière de qualité, les pêcheurs hauturiers, demandent des prix modérés. Le prix de revient de la localisation est une part importante du bilan d'exploitation d'un petit ou moyen chalutier, et les avantages de la précision seront soigneusement pesés en regard de ce prix. La solution de compromis doit donc être soigneusement étudiée, et la solution idéale, pour l'avenir lointain, ne sera pas simple.

Les besoins des activités de recherche, prospection et exploitation se différencient nettement des autres. Les deux premières ont un caractère la plupart du temps très localisé et essentiellement transitoire. Si l'on tient compte du fait que les organismes intéressés, individuellement ou en coopération, disposent en général de moyens financiers importants, la solution logique réside dans l'utilisation de systèmes mobiles appropriés, dont l'infrastructure peut être déplacée en fonction des programmes de travail. De tels systèmes peuvent être la propriété de leurs utilisateurs ou utilisés éventuellement en location. Quant aux besoins permanents des activités d'exploitation, ils ne se différencient pas notablement de ceux des autres navigateurs fréquentant les mêmes zones du plateau continental, sinon peut-être par l'exigence d'une meilleure précision, pour certaines fonctions telles que la surveillance et l'entretien des pipes ou autres installations immergées non marquées en surface.

4.3. Systèmes à courte portée

Ces aides radioélectriques ont pour objet de permettre une navigation très précises dans un certain nombre de zones, en général très localisées, où les conditions sont particulièrement difficiles, du fait d'une forte concentration du trafic, et de la nécessité, au moins pour les plus grands navires, de se confiner à des chenaux de navigation ou aires de manœuvre souvent étroite. Ces zones sont essentiellement constituées par certains "passages obligés" (détroits de faible largeur, par exemple), certains dispositifs de navigation réglementés, qui vont devenir prochainement obligatoires pour tous les navires, dès l'entrée en vigueur des règles de la Convention de 1972, et bien entendu les chenaux d'accès, zones d'attente et de mouillage, etc, au voisinage des ports importants. Dans la très grande majorité des cas, ces zones difficiles se trouvent à proximité immédiate de terre, et elles sont fréquemment soumises à l'obligation de pilotage. Les méthodes traditionnelles de navigation optique ou au sondeur y sont normalement pratiquées. Cependant l'augmentation constante de la taille des navires, et les problèmes liés à leur manœuvrabilité et à leur tirant d'eau, la sensibilisation croissante des autorités aux risques de pollution entrainée par les accidents de mer, le souhait des navigateurs eux-même, rendent nécessaire que soient mis à la disposition des usagers des moyens plus précis, plus disponibles et plus fiables que ce qu'offre la tradition.

Une grande diversité de systèmes radioélectriques peut être envisagée. Les systèmes hyperboliques à moyenne portée permettent souvent de satisfaire convenablement ce type de besoin : utilisés à courte distance de leurs stations, ils procurent normalement une très bonne précision. Il est quelquefois possible d'aménager le réseau de stations en vue de couvrir convenablement une zone difficile, par exemple en faisant coïncider sensiblement l'axe d'un chenal important avec un lieu de position fourni par le système, et en s'arrangeant pour que l'angle de recoupement des lieux y soit optimal. Mais, étant donné la proximité des côtes, il est également possible de recourir à des systèmes fonctionnant en très haute fréquence, dont la portée est limitée à l'horizon. De tels systèmes sont capables des performances exigées en matière de fiabilité et qualité, et offrent une précision excellente, souvent au-delà du besoin.

L'écueil est ici comme précédemment la prolifération de systèmes différents suivant les pays et les régions, ce qui conduirait à accumuler sur les navires, en fonction des ports fréquentés, des installations coûteuses et disparates. Il est difficile d'imaginer un remède pratique réaliste à ce risque. Cependant, dans le cas particulier des zones de pilotage, il faut mentionner une solution simple et séduisante : sur la base d'un système de positionnement choisi, réaliser un matériel mobile suffisamment léger et compact pour que le pilote puisse embarquer et débarquer avec lui. Le navire n'a plus dès lors à acquérir aucun matériel, ni à se familiariser avec sa mise en oeuvre, tout en disposant du positionnement suivant son besoin. Cette solution a été retenue pour la couverture des chenaux d'accès au nouveau terminal pétrolier du HAVRE, au Cap d'ANTIFER. Je souligne que dans le cas du chenalage et de la manoeuvre, le contenu informationnel du système, dont je n'ai pas retenu le détail dans la définition de nos critères généraux, n'est pas sans importance : au contraire de la plupart des autres applications, une indication en temps réel sur la vitesse absolue du mobile, en plus de l'information normale de position, est ici d'un haut intérêt. Ce critère peut permettre de guider un choix, dans la mesure où peu de systèmes classiques sont conçus pour fournir normalement cette indication.

Les autres applications du positionnement à courte portée, en dehors de la couverture spécifique des zones difficile, ont toutes un caractère technique : certains besoins militaires, qui seront sommairement examinés plus loin, et ceux des opérations de levé hydrographique ou de travaux publics maritimes. La satisfaction de ce besoin relève soit des systèmes fixes examinés ci-dessus dans le cas des services portuaires (entretien du balisage, entretien des chenaux d'accès, etc.) soit de systèmes mobiles dont tous les Services Hydrographiques ont aujourd'hui une pratique courante.

4.4. Besoin militaire

J'écarte délibérement certains problèmes très particuliers tels que par exemple la navigation des sous-marins nucléaires, ou la mise en oeuvre de l'aviation embarquée contre des objectifs terrestres. Ces cas spéciaux mis à part, le besoin militaire ne diffère pas substantiellement du besoin civil et on peut y reconnaître en gros les mêmes catégories. Les normes indiquées pour les différents critères sont cependant en général différentes, en précision et en couverture notamment. La répartition des zones d'intérêt militaire ne coïncide pas entièrement avec les zones d'intérêt pour la navigation commerciale, bien que naturellement la protection de ces dernières, s'agissant de navigation alliée, ou leur attaque, pour celles de l'ennemi, leur confère ipso facto une importance stratégique. Par ailleurs le critère économique doit être dans ce cas considéré comme peu sensible, les armées de mer ayant pour vocation de se donner les moyens de faire face à leurs missions dans les conditions les plus efficaces. Le besoin militaire est par nature gouvernemental, ce qui signifie que le coût à considérer dans ce cas est le prix global du système, y compris son infrastructure, le "coût de positionnement" au niveau de l'utilisateur individuel ayant peu de signification. A ce propos, je souligne simplement, sans y insister, la différence d'éclairage du problème, suivant qu'on l'examine dans un cadre multinational, au sein d'une alliance telle que l'OTAN, ou que l'on envisage une solution capable d'assurer une autonomie nationale, pour des pays de la taille des nations d'Europe, dont l'aire d'intérêt stratégique est limitée, et qui n'ont pas à leur disposition les possibilités techniques des Grands.

Observons une fois de plus que le GPS, sous sa version protégée de grande précision, satisfait l'ensemble des besoins militaires, ce qui est peu surprenant puisqu'il est conçu précisément dans ce but. Il les satisfait même bien au-delà des normes nécessaires à l'usage maritime, du fait qu'il doit aussi satisfaire les usagers aériens, beaucoup plus exigeants que les marins suivant certains critères (positionnement 3 D, performances dynamiques, etc.). Les systèmes autres que le GPS sont à réexaminer du point de vue des différents critères du tableau, et notamment de la précision et de la couverture, et à examiner du point de vue des critères spécifiques militaires.

En ce qui concerne le positionnement à grande distance, il apparait que le besoin militaire est plus exigeant en précision que le besoin civil, au moins dans certaines zones et pour certains usages. Le système OMEGA n'est pas suffisant pour y faire face, bien que ses performances en matière de précision différentielle puissent lui assurer une bonne efficacité pour certaines applications (précision de rendez-vous). D'autres solutions doivent être envisagées. L'une d'elle pourrait être une extension des réseaux LORAN C actuellement existants, si des sites convenables peuvent être trouvés sous contrôle ami. Les procédés de localisation par horloge ultrastable paraissent également susceptibles à terme d'offrir une réponse à ce besoin. Le positionnement à moyenne portée ne fait pas apparaître de différence notable par rapport au besoin civil, sinon sur les critères spécifiquement militaires.

Pour le positionnement à courte portée, il apparait qu'en sus des zones d'intérêt économique, on doit envisager la couverture précise de zones d'intérêt militaire : approches des ports militaires et des bases navales, et de leurs chenaux d'accès, d'une part, approches des sites de dégagement ou de repli éventuel, d'autre part. Il serait d'un haut intérêt que l'ensemble soit couvert à l'aide de systèmes identiques, afin de rendre possible l'emploi d'un matériel unique sur tous les types de navires militaires, auquel le personnel de mise en oeuvre et d'entretien serait bien entrainé. Ceci conduit à concevoir des systèmes à implantation légère, dont les chaines fixes couvriraient les zones d'intérêt permanent, et dont des chaînes mobiles pourraient être installées très rapidement en une région quelconque du littoral, en fonction des besoins opérationnels (dragage de mine, établissement et maintien de chenaux de sécurité, etc.). Le critère de discrétion n'est pas critique pour une telle fonction, et on peut parfaitement envisager des systèmes VHF ou UHF à mesure de distance si leurs performances de portée et de saturabilité sont conformes au besoin. Idéalement, ces systèmes devraient être identiques ou compatibles avec ceux qui assurent la couverture du besoin civil dans les "zones difficiles" d'accès aux ports importants, étudiés plus haut.

4.5. Navigateurs de plaisance

Nous avons dans les paragraphes précédents passé en revue l'essentiel des besoins des différentes catégories d'utilisateurs professionnels de la radiolocalisation maritime. Cette revue était nécessairement sommaire, et conduite sur un plan général, alors qu'il conviendrait, avant d'étudier la solution appropriée à l'un ou l'autre type de besoin, de connaître avant tout sa répartition géographique, que nous avons seulement évoquée, et l'effectif réel de la population d'usagers à servir, dont nous n'avons pas parlé. Ceci ne veut pas dire que les solutions les meilleures résideraient dans des choix régionaux différents, mais bien au contraire que cette étude régionale permettrait de trouver les solutions de compromis applicables sur la base géographique la plus large possible, avec les meilleures chances de succès auprès des usagers, et pour le meilleur profit des autorités responsables de la sécurité de la navigation maritime.

Je voudrais en guise de conclusion, et presque pour mémoire, évoquer une catégorie de navigateurs dont je n'ai pas parlé dans les pages qui précèdent : Il s'agit des plaisanciers, dont le nombre sans cesse croissant est attesté en France, et certainement dans les autres pays. Les bateaux de grande croisière, utilisés par des propriétaires fortunés, ne se différencient pas des bâtiments professionnels de taille comparable, et pour eux le critère de coût est souvent peu critique. Mais pour le plus grand nombre, constitué par les bateaux de croisière modestes, c'est évidemment ce critère qui est fondamental. Le coût des équipements, à l'heure actuelle, leur interdit pratiquement l'accès à la radionavigation, si l'on excepte bien sûr l'emploi des radiogoniomètres et l'utilisation de systèmes tels que le CONSOL. Cette catégorie d'utilisateurs mérite assurement que ses besoins soient pris en considération. Le problème est donc posé, il n'est pas simple, et sa solution n'est pas en vue.

5. REFERENCES

- "Marché Mondial de la localisation par le système GEOLE". Centre National d'Etudes Spatiales - Société SEMA - Paris 1974 - (non publié).

- "Etude de l'Utilité des systèmes de position" - Service des Phares et Balises - Société OPEFORM - Paris 1976 - (non publié).

CRITERES D'EVALUATION / EVALUATION CRITERIA → / CATEGORIE DU BESOIN / CATEGORY OF REQUIREMENT ↓	Couverture Coverage	Précision Accuracy	Type Kind of fixes	Disponibilité Availability	Qualité Global quality	Coût Cost
Navigation océanique (tous navires civils) Ocean navigation (civil ships)	Mondiale World	2 à 3M	P/R	Bne	Exc.	Mod
Navigation océanique courante (navires de guerre) Current ocean navigation (military ships)	Mondiale à zones Stratégiques	500m à 2M.	P/R	Exc.	Exc	Psens
Navires de recherche (hydrographie, océanographie, géophysique en) mer profonde Deep sea research vessels (hydrography, oceanography, geophysics)	Mondiale	20m à 1M	P/R	Bne	Cour	Psens
Navires de recherche sur le plateau continental Research vessels on continental shelf	100 à 300M	50 à 100m	P	Exc	Bne	Psens
Navires câbliers Cable ships	Mondiale	500 à 1000m	P/R	Bne	Cour	Psens
Atterrissage et navigation courante en zone côtière (tous navires) Landfall and current coastal navigation (all ships)	25 à 50M.	500m	P	Bne	Bne	Mod
Navigation près de terre en zones sensibles (tous navires) Coastal navigation in difficult areas (all ships)	Variable gén.t < 25M.	100 à 200m	P	Exc	Exc	Moy
Bateaux de pêche en opération Fishing boats operating	100 à 300M	20 à 100m	P	Bne	Exc	Mod
Navigation dans les chenaux d'approche (tous navires) Channel and fairway navigation (all ships)	Chenaux et eaux intérieures	20m	P	Exc	Bne	Mod
Navigation dans les dispositifs d'organisation et routes DW (très grands navires) Navigation in ships routering schemes and DW routes (very large ships)	Variable jusqu'à 150M	200m	P	Exc	Bne	Moy
Guerre des mines, surveillance rapprochée (navires de guerre) Mine sweeping, coastal patrol (military ships)	Variable gén.t < 25M.	20 à 50m	P	Exc	Bne	Psens
Hydrographie-géophysique en zone moyenne Hydrography-geophysics at median distance from coast	Variable gén.t < 200M	10 à 50m	P	Bne	Cour	Psens
Hydrographie à la côte, dragages, travaux maritimes Coastal hydrography, dredging, coastal engineering	10M	3 à 10m	P	Bne	Cour	Psens

M Mille marin Nautical mile — m Mètre Meter

P Positionnement permanent Continuous positioning required — R Positionnement récurrent à cadence lente ordre de l'heure admissible Low rate of fixes admissible hourly basis

Exc Excellente Excellent — Bne Bonne Good

Cour Courante Common — Moy Moyenne Average

Mod Modéré Moderate — PSens Peu sensible Small weight

DISCUSSION

S.Horowitz, Ionospheric Radio Physics Branch, Electromagnetic Sciences Division, L G Hanscom AFB, MA 01731, USA

In answer to the questions raised in the speaker's presentation, LORAN C Transmitters are in continuous operation with special effort to maintain pulse shape and reliability. The coding delays inserted into the slave transmitters assures that there is no ambiguity in the identification of the hyperbolic time difference line for a given network. As one moves into a different network or chain, the pulse group rate is changed to eliminate the ambiguity problem. LORAN C receivers are being manufactured which cost under $1000.00, are simple to use and highly reliable. There are anti-jam and ECCM provisions in the receivers and transmitters for military use. To achieve high accuracy the LORAN C utilizes only the first three cycles of the transmitted pulse (and the sixth cycle in LORAN D). The remainder of the pulse has no useful function and unfortunately only serves as interference.

E.R.Swanson, Naval Electronics Laboratory Center, San Diego, CA 92152, USA

Thank you for an excellent paper. You properly noted that there are many different aspects of technical considerations. I wish to elaborate in the area of accuracy and reliability especially as related to safety. Nominal accuracy as specified at the 10 or at the 95% level is useful in evaluating typical error. Part of the reliability consideration should be the probability of *blunders*, i.e., gross deviations. Most systems have built-in blunder modes. For example, loss of the "to-from" indication on a VOR may yield an error of precisely 180°. There is a probability of this specific error occurring. Blunders may have a dominant effect in evaluating safety. They can be handled as the probability of errors of a certain size or greater. That is, they are represented by a probability not a distance such as a c.e.p. Often safety is of more importance than accuracy. We must not equate the two. Improved accuracy may be harmful to safety.

Author's Reply

Je partage absolument le point du vue du Dr Swanson. Le fait que les écarts de position des suprèmes de radionavigation ne suivent pas en général une loi normale rend nécessair d'adjoindre a la donnée ECP un parametre supplémentaire caractérisant le risque d'erreur grossière, si l'on veut étudier le problème d'une façon fine et complète. L'évaluation de la qualité du système de lever de doute (comme identification procedure) des systèmes hyperboliques peut être considéré comme un aspect de cette nécessité de caractériser les sources d'erreur grossières.

SYLEDIS, A RADIOPOSITIONING SYSTEM

P. LAURENT and G. NARD
SERCEL
Av. Bel Air - 44 Nantes-Carquefou
FRANCE

SUMMARY

Radiopositioning systems usually implemented for accurate maritime surveys are based on phase comparison or on radar pulse techniques. They present several well known drawbacks such as ambiguity, sensitivity to soil conductivity, line of sight limited range.

A pulse system has been developed, named SYLEDIS, which does not rely on radar techniques. The use of compression of long pulses by correlation methods, and of a carrier wave in the 420 - 450 MHz band, has resulted in a highly accurate system, insensitive to soil conductivity, exhibiting no ambiguity problem, capable of ranges of 2 or 3 times the line of sight and not limited by obstacles of reasonable size.

SYLEDIS is highly flexible and operates in range or hyperbolic modes.

SYLEDIS equipment is fully operational, in factory production and commercialized.

I - INTRODUCTION

A tremendous expansion of maritime works has been observed during the last thirty years, covering many fields of activity : oil industry, harbour building and maintenance, hydrography, etc...

The necessity of continuously and accurately knowing the instantaneous position of crafts involved in these operations has become more and more acute.

At first, existing radionavigation aids, such as DECCA, have been used. But these large fixed installations intended to cover definite areas are seldom available where needed ; and when available, their accuracy is generally too poor.

It has thus been necessary to create positioning systems specially adapted for these applications. These radiolocation systems involve accuracies of one to a few tens meters and ranges seldom exceeding 200 kilometers. They must be easy to transport, to install, to operate and to service.

2 - POSITIONING SYSTEMS

A dozen of different systems presently meet these requirements. From the technical point of view, to which the operational characteristics are tightly connected, the concerned radiolocation systems can be considered as roughly relevant of two categories : phase comparison systems and pulse systems.

2.1. Phase comparison systems

Phase comparison systems measure the travel time of waves from shore transmitters to the ships by comparing the phases of the incoming waves at the ship receiver.

Phase comparison systems easily meet accuracy and range requirements, but they display two major drawbacks : ambiguity and sensitivity to soil conductivity.

Ambiguity results from the physical impossibility of determining a phase difference which would not be in the interval $0 - 2\pi$; a value of the phase difference corresponds to a series of possible positions separated by intervals or "lanes". Ambiguity can be solved by lane counting. But this method is ineffective in case of poor reception, or interference conditions ; besides it does not allow to enter an insert into the coordinate pattern, coming from the high sea. Lane identification can be carried out using the radiopositioning system itself, but this implies the use of extra frequencies, increasing the size, weight and cost of the equipment.

Phase comparison radiolocation systems generally operate in the 2 Megahertz band. Propagation of these waves is highly dependant on soil conductivity.

Heterogeneity of soil conductivity along the path of the waves causes distorsion of coordinate patterns, requiring calibration and introduction of reading corrections. Moreover, climatic and meteorologic variations deteriorate repeatability. One knows how difficult it is sometimes to select shore station location to achieve a favourable geometric configuration of the coordinate patterns, while at the same time avoiding paths of the waves over heterogeneous soils.

Yet, waves in the 2 MHz band are fairly insensitive to obstacles of reasonable size, and so range is not limited to line of sight.

Poor conductivity of soil also causes high attenuation of the waves, reducing the range and practically preventing the use of these systems over land, for instance for airborne surveys.

Skywave interference is often associated with this frequency band. It must be pointed out that in fact it is not an acute problem for the ranges considered here, as generally dangerous skywave interference only appears at night for distances above 200 kilometers.

It might also be mentioned that designing sound direct range measuring systems in that frequency band encounters severe difficulties, particularly if the system is intended to be multiparty. But this point of view has lost much of its importance, as distances now involved necessitate introduction of cartographic projection corrections : lines of coordinates are neither circles nor hyperbolae, and circular coordinates are not relevant of the simple use of compasses for accurately plotting positions.

Moreover, progress in electronics allows to dispose on board even of small crafts of low cost specialized computers or computing circuits which transform circular or hyperbolic coordinates in real time so that positions are now directly displayed in X and Y UTM, Lambert or geographical coordinates.

Nevertheless, it remains that coverage of rectilinear or convex coasts is much easier with a range system than with a hyperbolic system.

2.2. Pulse systems

Distances are determined with pulse systems, by measuring the travel time of a short wave pulse along these distances.

Pulse systems are generally range systems involving an interrogator on the mobile vehicle and responder beacons on shore. Measured time corresponds to the travel of the pulse from the interrogator to the beacon and back.

Pulse systems can be easily designed so that practically no bothersome ambiguity problem occurs.

Accuracy and range of these systems are not dependant on soil conductivity as they operate in the microwave band.

Pulse systems used for application considered here are based on radar techniques.

In order to achieve the required accuracy, pulses must be very narrow (.1 microsecond or so). This involves a large frequency spectrum, implying the use of microwaves for its full transmission.

The energy required to achieve the wanted range is entirely contained in the narrow pulse which must be consequently of high amplitude. Microwaves and high peak power techniques imply the use of such devices as klystrons and magnetrons involving frequent and costly maintenance operations, and do not allow to beneficiate of the high reliability and negligible maintenance associated with fully solid state technology.

A major drawback of pulse systems using microwaves is their poor range, limited to line of sight. Transmission beyond the horizon can only be achieved by increasing the peak power, leading thus to weight, cost, and complication of equipment incompatible with the applications considered here. Besides, hills, houses, woods, dykes, ships are absolute obstacles for microwaves.

3 - THE SYLEDIS SYSTEM

It has thus been felt more and more necessary to dispose of a system that would not display the drawbacks of phase comparison and usual pulse systems, while meeting the range and accuracy requirements, and which would also be of reasonable price, easy to operate, and reliable.

SYLEDIS seems to be a sound answer to the problem.

3.1. Operation of SYLEDIS

SYLEDIS makes use of compression of long coded pulses, by correlation techniques. A long pulse means here 2.66 milliseconds. The energy is spread in the pulse, the peak power being thus kept relatively low : 20 watts. The compression ratio being 5.000 the energy contained in the SYLEDIS long pulse would correspond for a classical system using .5 micro-second pulses to a peak power of 100 kw.

Due to the low peak power involved, SYLEDIS equipment is entirely solid state and requires no maintenance operation.

Each 2.66 ms. pulse is composed of 40 repetitions of a 127 bits pseudo-random code. Each bit has a duration of .5 microsecond. Such a code is generated by a shift register with logic feedback.

SYLEDIS makes also use of special time discrimination methods which highly increase the accuracy of distance determination, and the distance separation capability, that is the rejection of unwanted paths.

Operation of SYLEDIS will be understood by reporting to figure 1.

1. The interrogator at point A transmits a coded pulse generated by its code generator (code 1).

2. The pulse is received at beacon B where code 1 is correlated with a code (2) generated by a code generator identical to code generator A. Code (2) is synchronized with code (1).

3. Code (2) is then transmitted by B, and received at A.

4. At A, code (2) received from B is correlated with a code (3) generated by a second generator of A. Code (3) is synchronized with code (2).

5. Time difference between the origins of codes (1) and (3) is measured : it is the time taken by the waves to travel from A to B and back to A. If wave velocity is known the interrogator unit computes and displays distance A B .

Ambiguity in the SYLEDIS system is determined by the duration of the elementary code, that is 66 microseconds, or 10 kilometers. No problem results from this residual ambiguity in the practical daily use of SYLEDIS.

3.2. Organization of the system

Transmission and reception are time-shared using the same frequency, each slot corresponding to one pulse.

SYLEDIS can operate as a range or as a hyperbolic system.

When operating as a range system, an interrogator is placed on the mobile vehicle, and responder beacons are set on shore.

The cycle of transmission and reception is so arranged that one mobile interrogator can operate with 27 shore beacons.

Four mobile interrogators can operate simultaneously, and in that case, the number of shore beacons is limited to 6.

When operating in hyperbolic mode, the number of simultaneous user is not limited, as the mobile equipment is then purely passive.

In a same installation, some users may operate in the range mode, while others operate in the hyperbolic mode.

The cycle of transmission and reception slots in synchronized at each interrogator and beacon by a synchronizing pulse which is sent at the beginning of each cycle either by one of the interrogator, or by one of the shore beacons in the case of hyperbolic operation.

3.3. Frequency

The frequency spectrum of the pulse has a bandwidth of $\pm$ 2 MHz. A relatively low frequency can so be used for the carrier wave.

The 420 - 450 Megahertz band has been adopted. This band is officially devoted to radiolocation.

The carrier wave is $0 - \pi$ phase modulated by the bits of the coded pulse.

A complete SYLEDIS installation only requires one frequency.

In the same area, up to eight chains can operate simultaneously, using frequencies differing only by 5 kilohertz.

Adjacent SYLEDIS chains can be integrated in a single unit covering long distances along coasts, all operating on the same frequency.

3.4. Propagation - Range - Stability

Propagation of waves in the concerned frequency band is now fairly known. Results of investigation and experience have been gathered by K. A. NORTON in a publication of the Bureau of Standard, so that they are easy to use, and allow prediction of transmission conditions, and choice of antenna type and elevation.

Ground wave plays the leading part at short distance, while propagation by tropospheric scatter is preponderant at high ranges.

Near the horizon, attenuation is proportional first to $\frac{1}{D^6}$ then to $\frac{1}{D^8}$. Then diffusion in the low atmosphere, and farther, tropospheric scatter take place. Beyond the horizon, these different waves combine, resulting in an important short term fading effect. A slow fading effect due to meteorologic and climatic variations is also observed.

But if one disposes of a system capable of a high power reserve, the probability of establishing a transmission is high enough to allow the use of this type of propagation in a reliable way for distances affording sufficient area coverage.

This is the case for SYLEDIS, as a consequence of the high power reserve afforded by techniques involved in the system.

Experience of daily use of SYLEDIS has fully confirmed the range possibilities as deduced from NORTON's survey.

Ranges currently achieved with the 20 watts peak power SYLEDIS lie between two and three times the line of sight.

Height of antennae is very important not only because it determines the line of sight distance, but also it favours propagation slightly beyond the horizon. Yet, for very long ranges, height of antennae is much less important, and distance and height become even independant one from each other (Fig. 2).

Operation is possible down to 100 meters of a 20 watts shore beacon.

Whenever only short distances are involved, transmitted peak power of SYLEDIS stations can be reduced to 100 milliwatts, resulting in lower drain from the power source, and reduction of the risks of interferences with other radio transmissions in the area. In that case, operation of SYLEDIS is still possible up to the horizon. For instance, if the elevation of the antenna is 10 meters for the interrogator and 40 meters for the shore beacon, a situation easily and often met in practice, the range of the SYLEDIS

operating on 100 mW will be 34 km. This is a fairly sufficient coverage for most of harbour areas.

As an example, following table shows results of stability tests conducted in spring 1975 by the french Hydrographic Service of the Navy.

Three beacons, A, B and C, were installed on lighthouses or on signal-stations with respective antenna heights 84, 40 and 74 meters. The interrogator antenna was erected successively at elevations 11, 22 and 61 meters.

Beacon antenna height (meters	Interrogator antenna height (meters)	Line of sight (km)	Distance (km)	Ratio range/line of sight	Accuracy (σ, meters) Day	Night
A	11	43.70		2.33	5.62	-
(84)	22	48.50	101.9	2.10	4.02	4.39
	61	59.40		1.72	3.28	1.94
B	11	33.74		2.90	6.69	-
(40)	22	38.55	97.8	2.54	6.74	6.46
	61	49.47		1.98	4.06	2.09
C	11	41.64		1.64	2.59	-
(74)	22	46.46	68.3	1.47	0.91	0.90
	61	57.39		1.19	1.48	1.43

Adopted velocity of waves 299, 695 km/sec. corresponds to a mean refractive index at the surface of the sea of 1.000 325. The variation of the index, as deduced from the observed pressure, temperature and hygrometry during the experiments, was found to contribute for a maximum error of $\pm$ .5 meters.

Propagation of V H F waves in the considered frequency band is not influenced by soil conductivity. Range and accuracy are in no way deteriorated whether waves travel over sea or fresh water, dry or humid land. Repeatability is not affected by climatic or meteorological variations, such as rain and snowfalls, frost, seasonal modification of salinity of estuaries, etc...

Contrarily to what happens with systems using microwaves, the range of SYLEDIS is not or is only slightly reduced by the presence on the path of the waves of obstacles of reasonable size, such as dunes, dykes, ships, villages, woods. Line of sight condition is not at all a must.

The result of these features is a high grade of flexibility for the system.

Location of land stations can be selected by only bearing in mind the best coverage of the operation area from the point of view of the geometrical configuration of coordinate patterns. Beacons can be installed, far inland, if convenient, and elevated points such as hills, towers, tall buildings can be used to increase the range and hence the covered surface. Coverage of estuaries, rivers, sinuous accesses to harbour is eased.

An example of such an implementation of the SYLEDIS system is displayed by the ship conning system of the harbour of ANTIFER, near LE HAVRE in FRANCE. This system has been in operation since one year. ANTIFER harbour is a newly built oil harbour intended to accomodate large tankers up to 300.000 tons. The total length of the access channel is about 30 km. Three SYLEDIS shore stations are set up on shore. They are received on board the ship by a transceiver which is carried by the pilot. The instantaneous position of the ship is thus determined on shore. The position data are processed by a computer which computes parameters : along track and off track distances of the ship relative to the channel axis and corresponding speeds, heading and rate of turn when two transceivers are placed on the ship. These data are transmitted in digital form to the ship, using the SYLEDIS carrier wave according to a time sharing arrangement of position determination and parameter transmission.

The parameter values are displayed on a console carried by the pilot.

Four inbound or outbound ships can be simultaneously accomodated. An unlimited number of other ships, e.g. service ships, can receive the signals from the SYLEDIS shore stations and use the system as a hyperbolic positioning system.

3.4. SYLEDIS equipment

3.4.1. Shore beacon

The shore station is composed of a beacon unit and of an antenna.

The beacon unit is contained in a waterproof cabinet weighing 15 kg. Sizes are 38 x 16 x 45 cm.

Power supply requirement is 11 to 14 volts, or 22 to 30 volts DC, 40 watts.

Different types of antenna are used according to required range. Most commonly used is a mast carrying four vertical folded dipoles. They are orientated along the mast either for omnidirectional radiation, or for 110° directivity. Length and overall dimensions of the antenna are respectively 160 cm and 17 cm. For short ranges, a single dipole is used contained in a rod of fibre glass 80 cm in length and 2,5 cm in diameter.

MTBF of beacon unit is 4 000 hours.

No maintenance is required.

3.4.2. Interrogator

The interrogator is a single drip proof unit weighing 15 kg - Sizes are 50 x 22 x 40 cm.

Power supply is 22 to 30 volts DC, 5 Amps.

Receiving and transmitting circuits of the interrogator are composed of the same modules as those of the beacon.

Three lines of 7 LED's display three ranges in hundredths of kilometers to decimeters, or three hyperbolic LOP's.

Ship antenna is always omnidirectional, of the types already described for the shore beacon.

MTBF of the interrogator is 1.200 hours.

No maintenance is required.

4 - CONCLUSION

While keeping the advantages of range and accuracy of phase comparison and pulse systems, SYLEDIS seems to eliminate several acute problems encountered with these systems, as ambiguity, influence of soil conductivity and sensitivity to obstacles.

The resulting increase in flexibility opens the way to a new field of application, as the coverage of difficult waterways such as sinuous rivers and estuaries, a problem to which ship guiding and mine sweeping are often faced. Traditional application of accurate radiolocation systems are also solved in an easier way : geophysical exploration, plateform positioning, pipe laying, sounding and dredging operations in harbour access channels, bottom mapping, etc...

The fact that range and accuracy are not degraded on land opens the field of aeronautical applications as airborne surveys, guiding of helicopters, calibration of landing systems.

5 - REFERENCES

LAURENT, P. 1975 : New radioelectric pilot aid for Harbour Entrance. Proceedings of the Offshore Technology Conference, n° 2175. HOUSTON (TEXAS) - U.S.A. -

NARD, G., LAURENT, P. 1975 : Le SYLEDIS, Système de radiolocalisation précise à courte et moyenne distance, Navigation, revue de l'Institut Français de Navigation, XXIII, n° 89, January 1975, p. 18 to 34.

NORTON, K.A., 1959 : Transmission Loss in Radio Propagation - 2 - National Bureau of Standard - Technical note n° 12.

VOISARD, J., 1975 : Dispositif de guidage des navires dans les chenaux d'accès portuaires. IALA conférence, OTTAWA.

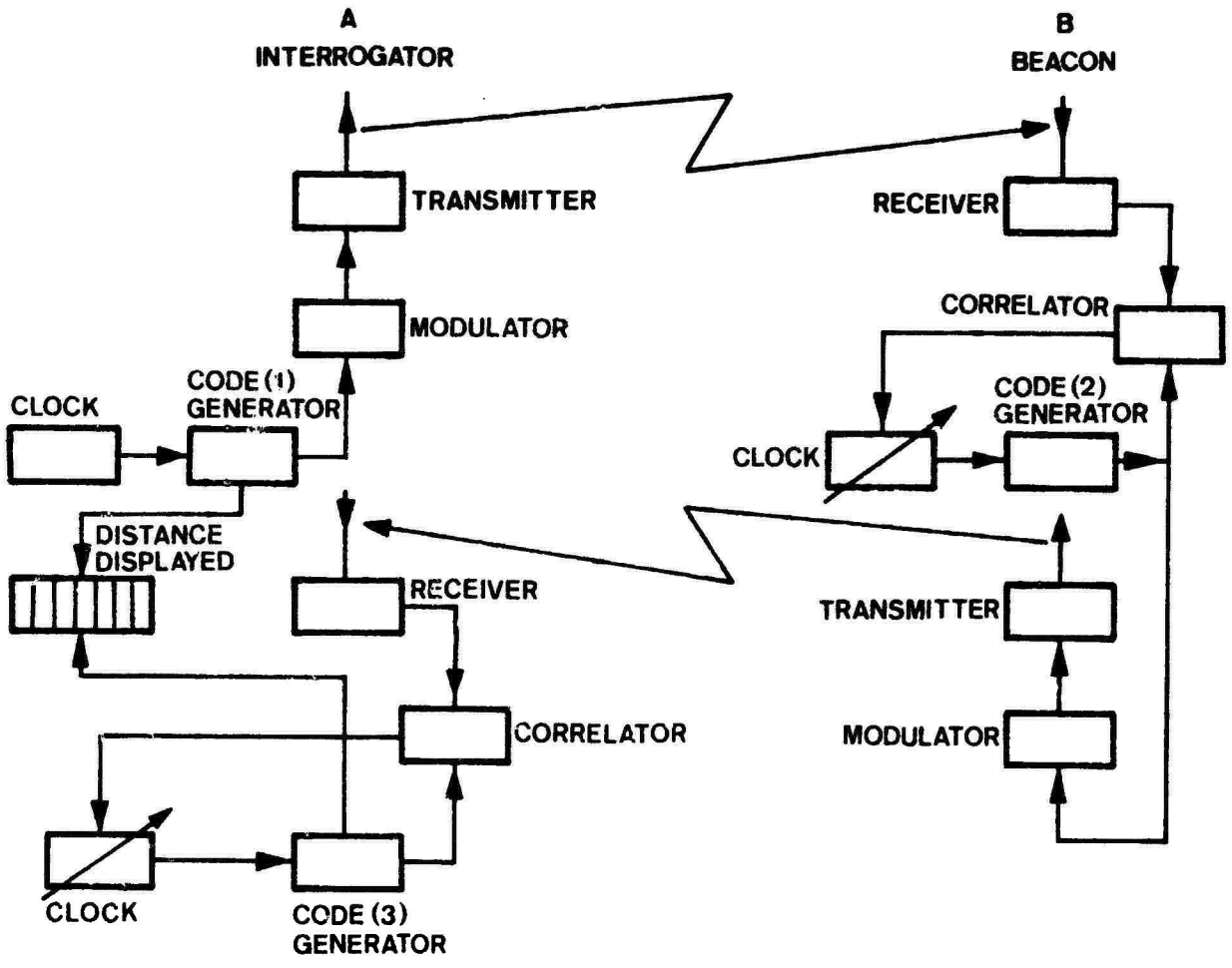


Figure 1 - Operation principle of SYLEDIS.

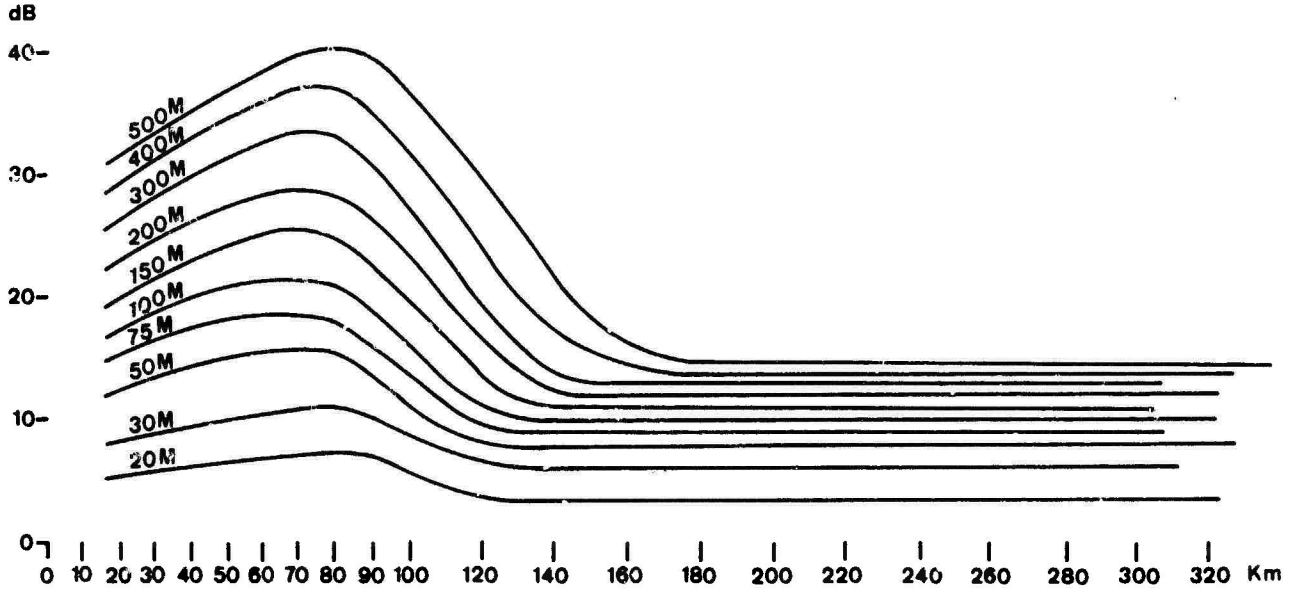


Figure 2 - Improvement of power reserve achieved by increasing the height of an antenna for U H F waves.

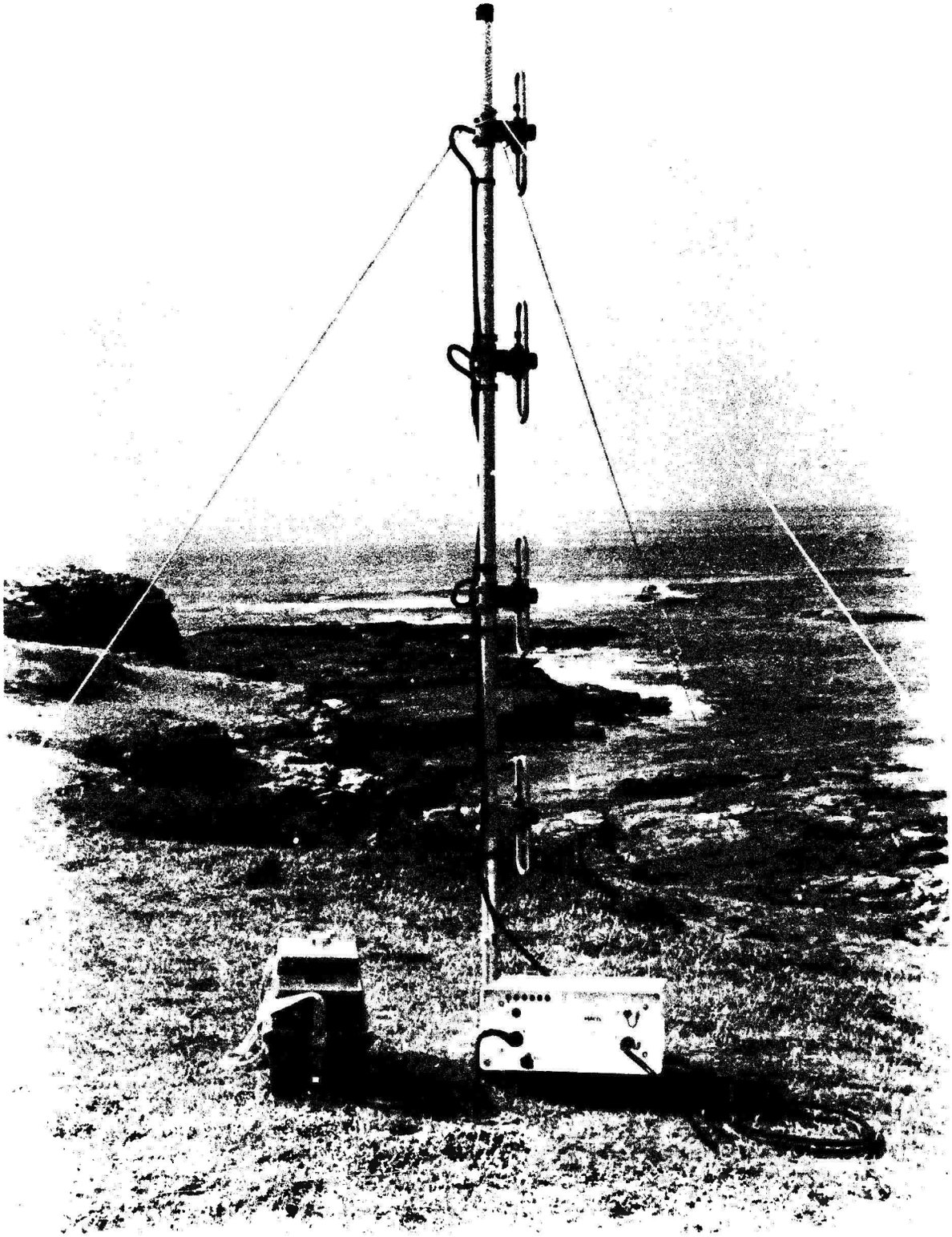


Figure 3 - Complete SYLEDIS Shore Station :
- right : the beacon unit.
- left : the battery power supply.

Figure 4 - The Interrogator.

C.Goutelard

Je sachaiterais aborder deux points: Tout d'abord, une question: Pourquoi répéter 40 fois le même code plutôt que d'utiliser un seul code qui soit sans effet de bond, tel qu'un code de Gday? Ensuite, il apparait que l'utilisation de signaux à spectres étalés est une solution interessante qui peut, dans l'avenir apporter des solutions pratiques compte tenu des moyens que l'on possède desormais avec des techniques numériques de traitement.

Author's Reply

Le code utilise l'a été parce qu'il permet un traitement et une generation simple et compatible avec l'economie recherchèe dans la realisation du système.

The indicated code has been adopted because it allows simple and easy generation and processing, which is compatible with the simplicity and low cost required for the considered positioning system.

A.Shuval, National Commitee for Space Research, Radio Observatory, P.O.B. 4655, Haifa, Israel

What is the bandwidth of the transmitted pulse?

Author's Reply

The bandwidth of the transmitted pulse is ± 2 MHz.

APPLICATIONS OF THE DOPPLER TECHNIQUE AS AN AID TO BEARING MEASUREMENT

T.B. Jones and C.T. Spracklen
Physics Department
University of Leicester, U.K.

SUMMARY

It is now possible to design H.F. direction finders with an instrumental accuracy of better than 0.1 degrees standard deviation. Unfortunately, this high accuracy cannot be realised in practice because of the perturbations and tilts which are present in the reflecting medium. These allow signals to be received in off great-circle path directions and consequently an error is produced in the measured bearing. Particularly troublesome sources of errors are the so-called travelling ionospheric disturbances (TID's).

As the disturbances propagate, they produce changes in the reflection point of the signals and consequently a small "Doppler-like" frequency change occurs in the reflected wave. A Doppler system is described which provides an indication of the presence, or absence, of TID's even when a short-duration nonfrequency stable transmission is monitored. Thus, a good/bad bearing indication can be given and some assessment of the likely variances in the measured bearing predicted. When more than one propagation mode is present multiple frequency shifts are observed in the received signal. These are easily identified and the system is an efficient monitor of multimode conditions.

The application of this Doppler technique has led to an improvement in the performance of a wide aperture direction finder; moreover, it is possible to estimate the expected variances from the Doppler observations.

1. INTRODUCTION

Wide aperture direction finding systems (WADF) can now be constructed with instrumental errors such that the standard deviation of the error curve, taken over all azimuths and a 2 octave frequency range, is as low as 0.1 degrees. The instrumental errors are usually measured by radiating signals, either from a helicopter within line of sight, or from a vehicle-borne mobile transmitter within ground wave range. When measuring the azimuths of ionospherically propagated signals, however, the accuracy of the system is found to be considerably less than that obtained for ground wave signals and a typical figure for the variance would be 2 to 4 degrees squared depending on the range. Following the early pioneering work of Bramley and Ross using interferometer DF systems (ROSS, 1947; BRAMLEY, 1951, 1953, 1955), the most significant sources of error have now been identified (WHALE, 1959). These are the solar-related lateral tilts of the ionosphere (LAT's) and the perturbations of the ionosphere which have become known as travelling ionospheric disturbances (TID's) and are almost certainly due to acoustic-gravity waves (HINES, 1959a,b, 1960, 1965; CAMPBELL and YOUNG, 1963; TOLSTOY, 1963; WICKERSHAM, 1965). The errors in WADF due to both these causes have over the last few years been measured in detail and, in general, one-third of the variance can be attributed to LAT's, or SIT's (systematic ionospheric tilts) as they are sometimes called, and about two-thirds to TID's. This ratio is, however, very variable depending on the state of the ionosphere, the azimuth of the arriving waves, the time of day, the season and the geophysical conditions. On some occasions most of the variance is due to TID's whilst on others the errors due to SIT's are predominant. Because TID's are normally the greatest source of DF error, recent HFDF work has been concentrated mainly in studying TID's and on devising schemes to overcome TID-induced bearing errors. It is difficult to predict either the magnitude or the time of occurrence of a TID and efforts have therefore been concentrated on detecting these irregularities by experimental procedures.

A system is described for detecting ionospheric disturbances by means of the small Doppler frequency shifts they produce in the received HF signal. The distortions of the iso-ionic contours associated with these events allow off great-circle path propagation and hence incorrect bearings are recorded by the direction finder. The system described can establish the presence or absence of TID's and hence the probability that bearing errors are present. This can be achieved even when the signal is available for only a short duration (20 sec), is modulated and originates from a transmitter whose frequency is unstable.

2. EXPERIMENTAL ARRANGEMENT

Three Doppler receivers, R1, R2 and R3, are spaced around the direction finder as indicated in Fig. 1. A large spacing between receivers was used initially but this was later reduced and the smaller triangle produced the better indication of TID activity. The difference in the frequencies recorded at each pair of receivers is measured, thus providing "difference Doppler" frequency. The frequency difference between the local standard and the received signal is measured at each receiving site and the local standards are adjusted so that frequency differences of $f_2 = 43$ Hz, $f_3 = 46$ Hz and $f_1 = 52$ Hz are obtained at R1, R2 and R3, respectively, when no TID's are present. These Doppler frequencies are relayed to the central DF site for comparison. The Doppler frequencies from each pair of receiving sites were compared to produce the "difference Doppler" frequencies $(f_1 - f_2)$, $(f_2 - f_3)$ and $(f_3 - f_1)$. The three frequencies are combined to form one signal which was recorded by an FM tape recorder and also fed directly to the spectrum analyser for real time display. The complete system is shown diagrammatically in Fig. 2.

Before starting the experiment the synthesizers used to provide the local standards at R1, R2 and R3 were carefully compared so that their frequency standards were identical within the frequency resolution of the measuring system (0.01 Hz). In this way the true offset at each receiver for any nominal frequency was known accurately. Conditions of zero Doppler shift, i.e. no TID activity, can be recognised because the spectrum of the combined signal will consist of the preset frequency differences (3, 6 and 9 Hz) and these are easily identified. When travelling ionospheric disturbances (TID's) are present, the frequency of the signals received at the three locations will differ and f_1, f_2 and f_3 will change from their nominal values. The magnitude of the change depends on the characteristics of the TID and also on the spacing between the receiving sites.

3 DETECTION OF TID ACTIVE AND MULTIMODE CONDITIONS

The frequency spectrum of the combined signal is expected to consist of three single valued peaks located at the three difference frequencies (3, 6 and 9 Hz) if no Doppler shifts due to TID's are present. These conditions are illustrated in Fig. 3 in which spectra obtained at consecutive 1 min intervals have been overlaid to produce a three-dimensional plot. Single valued peaks are produced at the expected frequencies and these show little frequency deviation as time progresses. Fig. 4 is a similar plot for TID active conditions and although the frequency peaks are still single value their position on the frequency axis varies as time progresses. This change in the difference frequencies arises because the TID produces different rates of change in the reflection heights of the signals received at R1, R2 and R3.

The presence of multimode propagation can also be detected since the reflection point of each mode will in general be displaced differently by the passage of the TID. Thus, in place of a single peak at a particular difference frequency, the spectrum is likely to contain more than one frequency component depending on the number of modes involved. This situation is illustrated in Fig. 5, where the single spectral peak due to the dominant mode becomes progressively multi-valued as other modes are received. It should be noted that the various spectral peaks do not necessarily correspond to actual modes since frequency differences are being examined. Thus, if two modes are received at R1 and at R2, up to four "difference frequencies" can occur and four peaks might be present in the vicinity of $(f_1 - f_2)$. The number of peaks obtained will depend on the relative amplitude of the component modes.

A simpler presentation of the data is obtained by suppressing the amplitude information and displaying only the frequency variations as a function of time. This has been done in Figs. 6, 7 and 8. Stable propagation conditions with no evidence of TID activity are illustrated in Fig. 6. This probably corresponds to E_s propagation but this could not be confirmed since oblique ionograms were not available over the path. Spectra for a TID active situation are reproduced in Fig. 7 and typical oscillations with periods of about 20 mins are evident. An example of multimode conditions is reproduced in Fig. 8. Multiple frequency components occur in each of the difference frequencies and in these circumstances low figures of merit are associated with the bearing measurements.

Both presentations illustrate that TID active periods and multimode conditions can be detected by the "difference Doppler" method, even when the signals are available for only one or two scans of the frequency spectrum. The time for one scan is 20 secs.

4. DIFFERENCE FREQUENCIES AND BEARING ERROR

The difference frequency experiments were designed to detect the presence or absence of TID's. No attempt has been made at bearing correction since this required integration of the signal (JONES and SPRACKLEN, 1975). Nevertheless, it is of considerable interest to compare the 'difference Doppler' results with the measured bearings. To aid this comparison, the deviation of each 'difference Doppler' frequency from its nominal value has been measured and the modulus of each of the three values added together to give a Doppler index $\Sigma \Delta f = 0$. When TID's produce rapid changes in the ionosphere large values of $\Sigma \Delta f$ will be recorded. A further refinement was introduced by multiplying each 'difference Doppler' by the sine of the angle which the line joining the two receiving sites makes with the transmitter direction. This angle is obtained with sufficient accuracy from the direction finder. In this way emphasis is given to TID's moving at right angles to the propagation path since these produce the maximum bearing error.

Data for quiet conditions, when no TID's are present, are reproduced in Fig. 9. The bearing measurements indicate little error and no fluctuations are evident. The three difference Doppler frequencies are all very close to their nominal offset values and no multiple frequency peaks are present, thus the index $\Sigma \Delta f$ is approximately zero for this period. In these circumstances it would be easy to conclude from the Doppler index that no TID activity was present and that good quality bearings would be obtained.

Data for two TID active periods are reproduced in Figs. 10 and 11. Well-defined 20 min oscillations are noted in the bearing and errors of up to 2 degrees are recorded. $\Sigma\Delta f$ is non-zero and takes on very large values when the bearing changes rapidly. Since the magnitude of $\Sigma\Delta f$ is proportional to the rate of change of the ionospheric reflection levels, it might be expected that a maximum frequency deviation would correspond to a maximum rate of change in the ionosphere tilt and hence to a period of rapid change in bearing error. This effect is evident in the data reproduced in both Figs. 10 and 11.

5. CONCLUSIONS

Conditions when TID's are absent ('good' bearings) and when TID's are active ('bad' bearings) can be recognised by the difference Doppler technique with some confidence. This can be achieved with a non-frequency stable transmitter radiating modulated signals for a time of less than 1 min. It is not possible to make a bearing correction but a good/bad indication can be obtained. Some difficulty arises in

establishing the magnitude of the frequency change necessary to declare a bearing 'bad' and further studies are being undertaken in an attempt to resolve this problem. The behaviour of the bearing error during multimode conditions is also being investigated.

REFERENCES

BRAMLEY, E.N., 1953, "Direction finding studies of large scale ionospheric irregularities", Proc. R. Soc. 220, p.39.

BRAMLEY, E.N., 1955, "Some comparative directional measurements on short radio waves over different transmission paths", Proc. IEE 102, p.544.

BRAMLEY, E.N. and ROSS, W., 151, "Measurements of the direction of arrival of short radio waves reflected from the ionosphere", Proc. R. Soc. 207, p.252.

CAMPBELL, W.H. and YOUNG, J.M., 1963, "Auroral zone observations of infrasonic pressure waves related to ionospheric disturbances and geomagnetic activity", J. Geophys. Res. 68, p.5909.

HINES, C.O., 1959a, "Motions in the ionosphere", Proc. Inst. Radio Engrs. 47, p.176.

HINES, C.O., 1959b, "An interpretation of certain ionospheric motions in terms of atmospheric waves", J. Geophys. Res. 64, p.2210.

HINES, C.O., 1960, "Internal atmospheric gravity waves at ionospheric heights", Can. J. Phys. 38, p.1441.

HINES, C.O., 1965, "Motions in the neutral atmosphere", In Physics of the Earth's Upper Atmosphere, C.O. Hines, I. Paghis, T.R. Hartz and J.A. Fejer (Eds.), Chapter 6, Prentice-Hall, New Jersey.

JONES, T.B. and SPRACKLEN, C.T., 1975, "The correction of errors produced in H.F. direction finders by travelling ionospheric disturbances", Proc. Agard Conf. on Radio Systems and the Ionosphere, No. 173, Ref. 34.

ROSS, W., 1947, "Fundamental problems in radio direction finding at H.F.", Proc. IEE 94, Pt. IIIA, p.154.

TOLSTOY, I., 1963, "The theory of waves in stratified fluids including the effects of gravity and rotation", Rev. Mod. Phys. 35, p.207.

WHALE, A.H., 1959, "Effects of ionospheric irregularities and the auroral zone on the bearings of short-wave radio signals", J. Atmos. Terr. Phys. 13, p.258.

WICKERSHAM, A.F., 1965, "Comparison of velocity distributions for acoustic gravity waves and travelling ionospheric disturbances", J. Geophys. Res. 70, p.4875.

TABLE I

```
      SUBROUTINE DIPOLY     73/72   OPT=1 TRACE MANTRAP              FTN 4.5+410        78/11/01. 13.19.34        PAGE      1

      1                 SUBROUTINE DIPOLY                                                                    DIPOLY      2
                        COMMON /CONST/ PI,PIT2,PID2,DUM(5)                                                   DIPOLY      3
                        COMMON /YY/ MODY,Y,PYPR,PYPTH,PYPPH,YR,PYRPR,PYRPT,PYRPP,YTH,PYTPR                   DIPOLY      4
                       1,PYTPT,PYTPP,YPH,PYPPP,PYPPT,PYPPP                                                   DIPOLY      5
      5                 COMMON R(6) /WW/ ID(10),W0,W(400)                                                    DIPOLY      6
                        EQUIVALENCE (EARTHR,W(2)),(F,W(6)),(FH,W(201))                                       DIPOLY      7
                        DATA (MODY=6HDIPOLY)                                                                 DIPOLY      8
                        ENTRY MAGY                                                                           DIPOLY      9
                        SINTH=SIN(R(2))                                                                      DIPOLY     10
     10                 COSTH=SIN(PID2-R(2))                                                                 DIPOLY     11
                        TERM9=SQRT(1.+3.*COSTH**2)                                                           DIPOLY     12
                        T1=F*(EARTHR/F(1))**3/F                                                              DIPOLY     13
                        Y=T1*TERM9                                                                           DIPOLY     14
                        YR= 2.*T1*COSTH                                                                      DIPOLY     15
     15                 YTH= T1*SINTH                                                                        DIPOLY     16
                        PYPPR=-3.*YR/R(1)                                                                    DIPOLY     17
                        PYRPT=-2.*YTH                                                                        DIPOLY     18
                        PYTPR=-3.*YTH/R(1)                                                                   DIPOLY     19
                        PYTPT=.5*Y                                                                           DIPOLY     20
     20                 PYPP=-3.*Y/R(1)                                                                      DIPOLY     21
                        PYPTH=-3.*Y*SINTH*COSTH/TERM9**2                                                     DIPOLY     22
                        RETURN                                                                               DIPOLY     23
                           END                                                                               DIPOLY     24


          SYMBOLIC REFERENCE MAP (R=1)

ENTRY POINTS
    1  DIPOLY           1  MAGY

VARIABLES     SN  TYPE              RELOCATION
    53  COSTH     REAL
    14  EARTHR    REAL                        WW
   323  FH        REAL                        WW
     0  MODY      INTEGER                     YY
     2  PID2      REAL                        CONST
     4  PYPPH     REAL                        YY
    16  PYPPR     REAL                        YY
     2  PYPR      REAL                        YY
    10  PYRPP     REAL                        YY
     7  PYRPT     REAL                        YY
    12  PYTPR     REAL                        YY
     0  R         REAL          ARRAY   / /
    54  TERM9     REAL
    13  W         REAL          ARRAY         WW
     1  Y         REAL                        YY
     5  YR        REAL                        YY
     3  DUM       REAL          ARRAY         CONST
    20  F         REAL                        WW
     0  ID        INTEGER       ARRAY         WW
     0  PI        REAL                        CONST
     1  PIT2      REAL                        CONST
    20  PYPPP     REAL                        YY
    17  PYPPT     REAL                        YY
     3  PYPTH     REAL                        YY
     6  PYRPR     REAL                        YY
    14  PYTPP     REAL                        YY
    13  PYTPT     REAL                        YY
    52  SINTH     REAL
    55  T1        REAL
    12  W0        REAL                        WW
    15  YPH       REAL                        YY
    11  YTH       REAL                        YY

EXTERNALS          TYPE   ARGS
    SIN            REAL     1 LIBRARY          SQRT      REAL      1 LIBRARY

COMMON BLOCKS    LENGTH
    CONST             8
    YY               17
    / /               6
    WW              411

STATISTICS
  PROGRAM LENGTH                  550       46
  CM LABELED COMMON LENGTH       6F48      436
  CM BLANK COMMON LENGTH           68        6
```

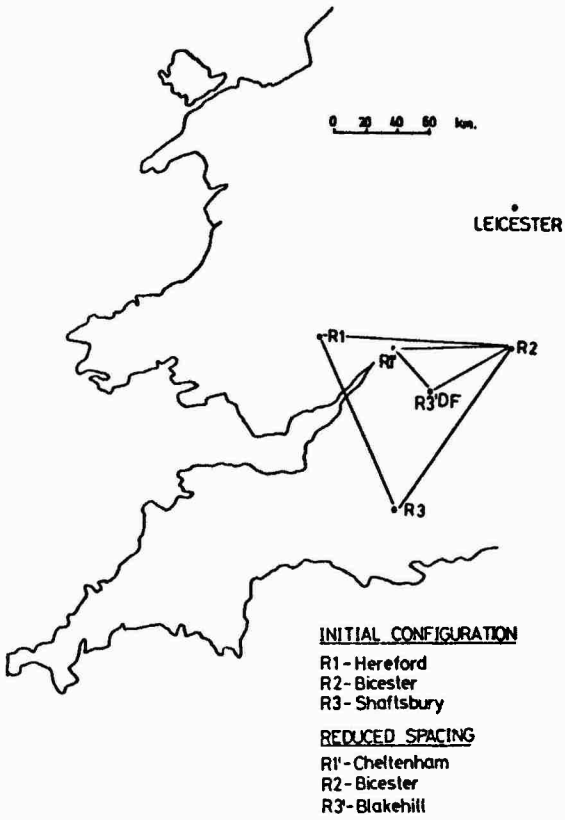


Fig.1 Location of receiving sites in 'difference Doppler' experiment. The large triangle R1, R2, R3 was used initially but the separation was later reduced by using R_1' R_2 and R_3'. DF is the direction finder

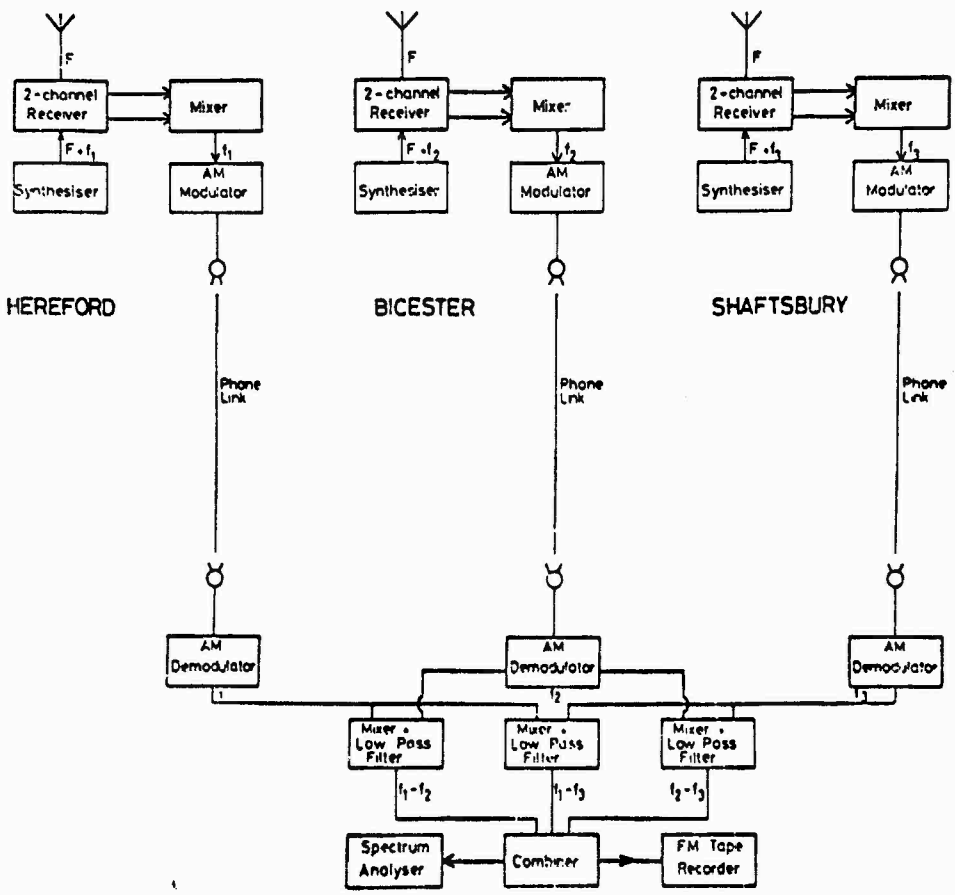
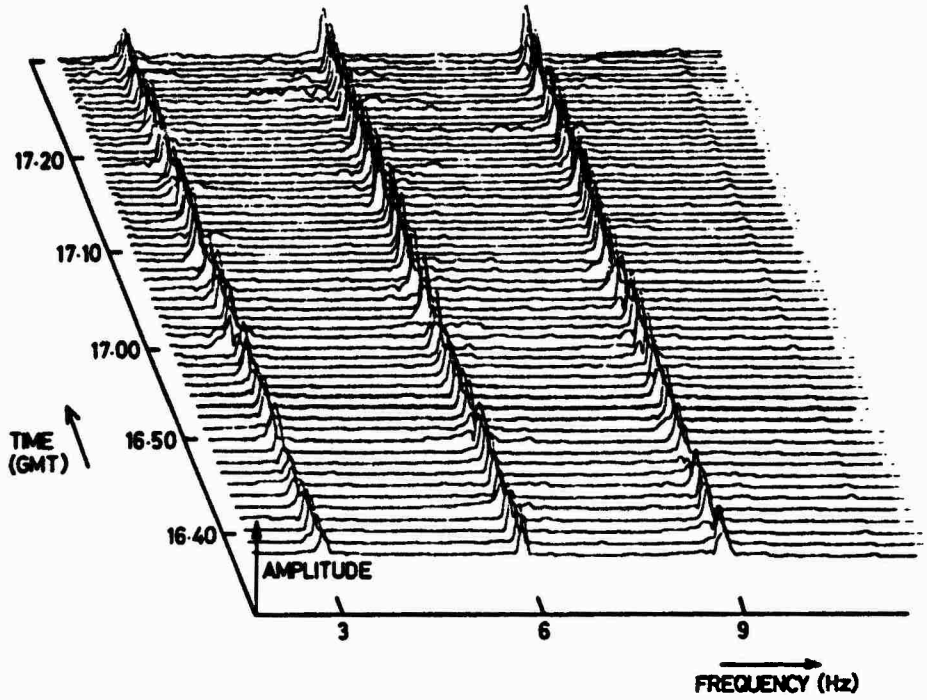
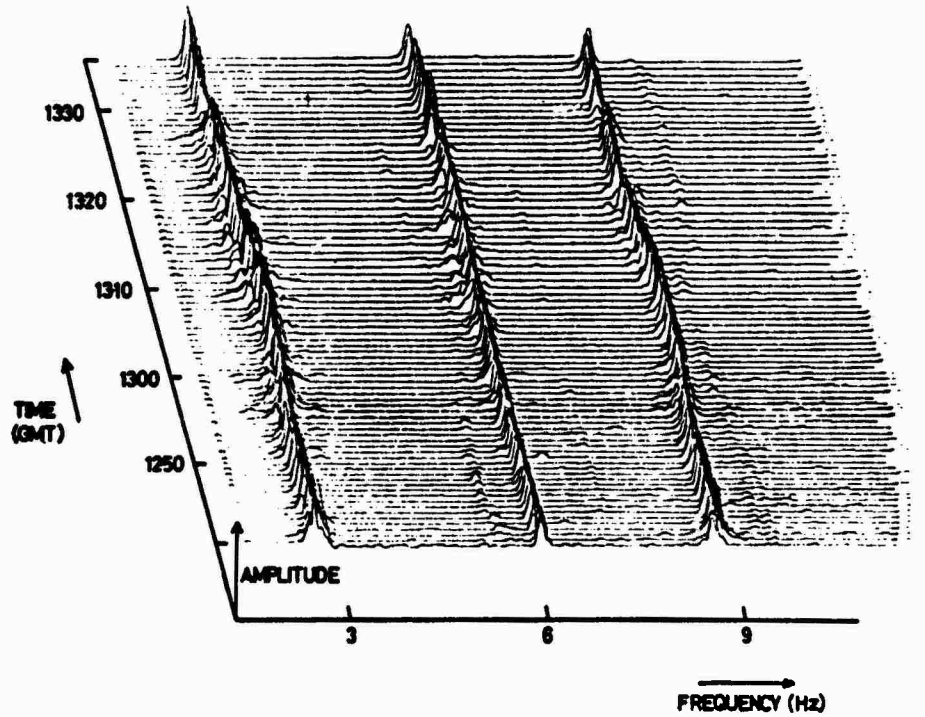


Fig.2 Block diagram of difference Doppler experiment



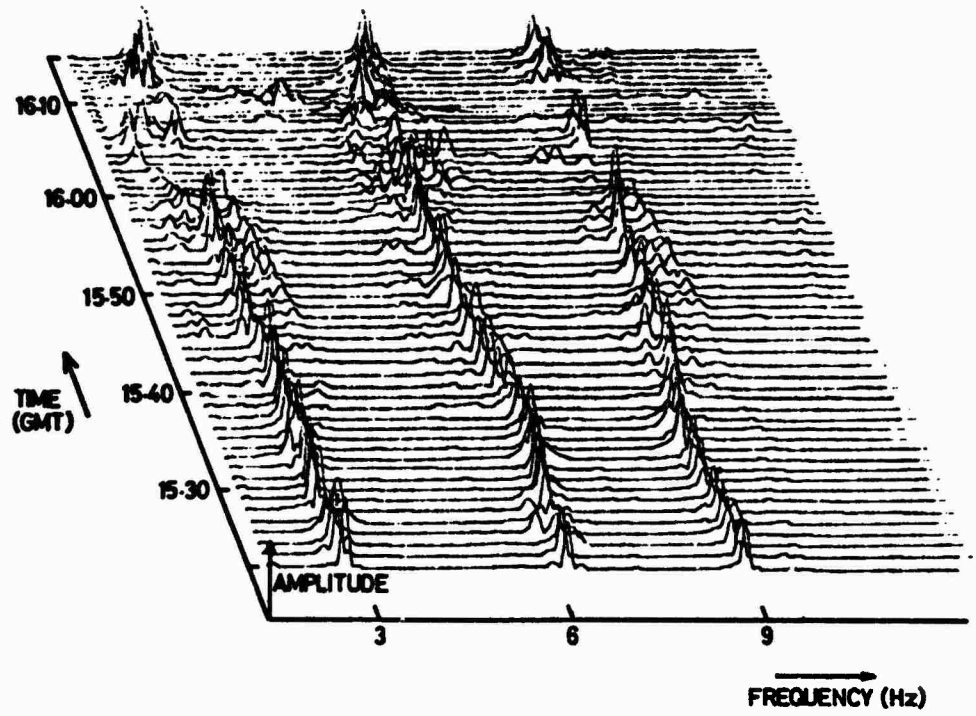
DIFFERENCE DOPPLER 13 NOV. 75 1638 - 1730 GMT
E_s CONDITIONS 9·535 MHz. BEROMUNSTER

Fig.3 Frequency spectra showing variation of 'difference' frequencies with time. Quiet conditions – no frequency deviations



DIFFERENCE DOPPLER 12 NOV. '75 1241 - 1336 GMT
TID ACTIVE CONDITIONS 9·535 MHz. BEROMUNSTER

Fig.4 Frequency spectra showing variation of 'difference' frequencies with time. TID active conditions – frequency deviations evident



DIFFERENCE DOPPLER 12 NOV '75 1522–1613 GMT
MULTI MODE CONDITIONS 9·535 MHz. BEROMUNSTER

Fig.5 Frequency spectra showing variation of 'difference' frequencies with time. Multimode conditions – frequency deviations and multiple frequency peaks present

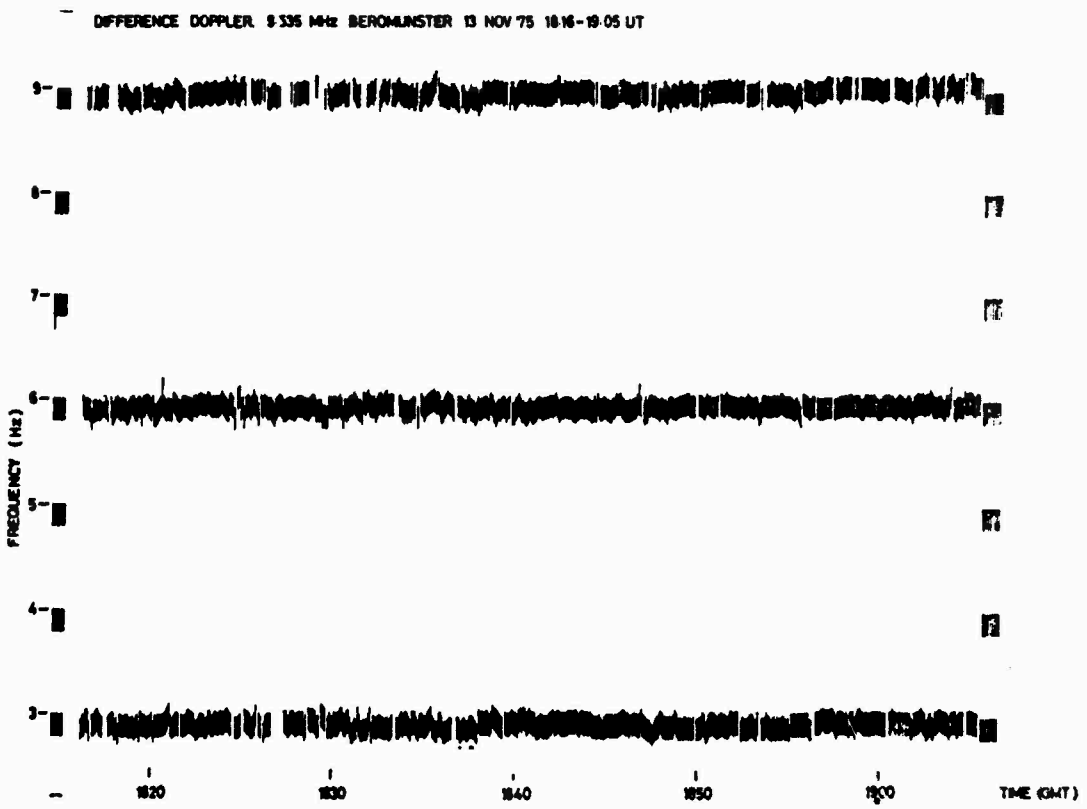


Fig.6 Time variation of 'difference' frequencies. Quiet conditions – no TIDs

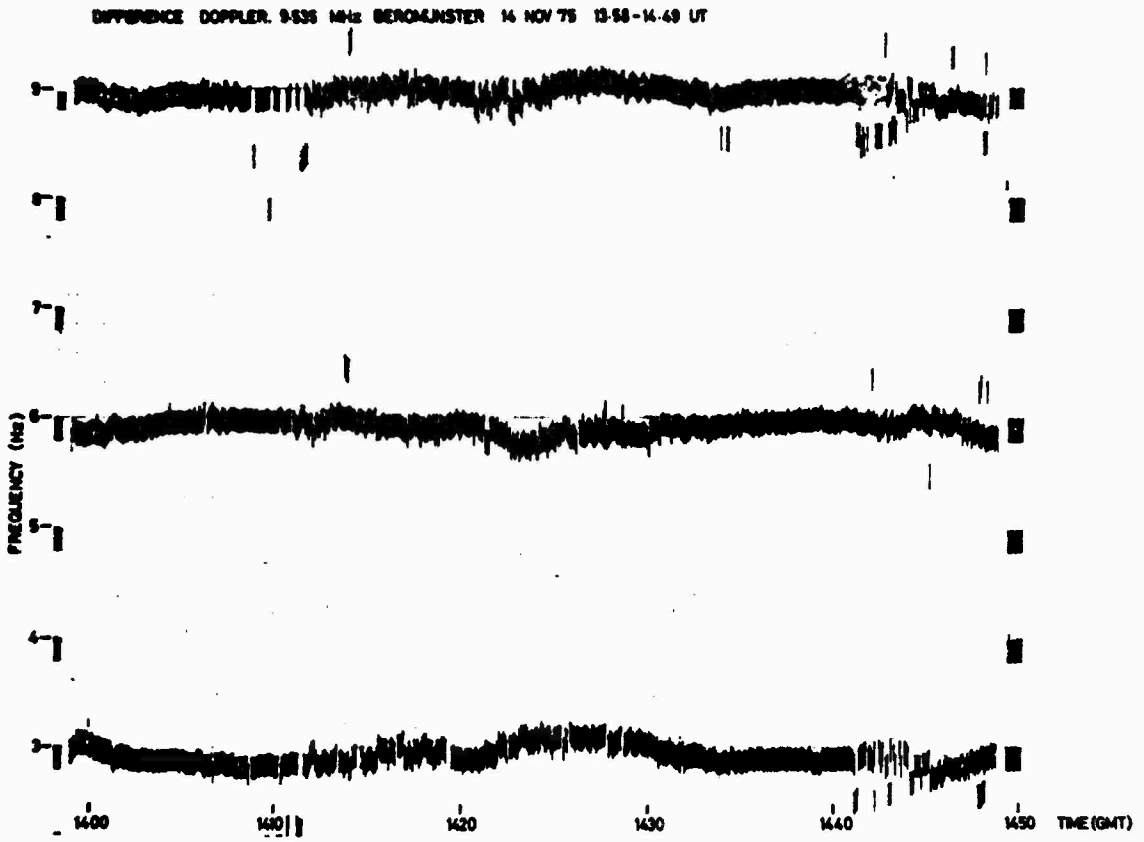


Fig.7 Time variation of 'difference' frequencies, 20 min period oscillations

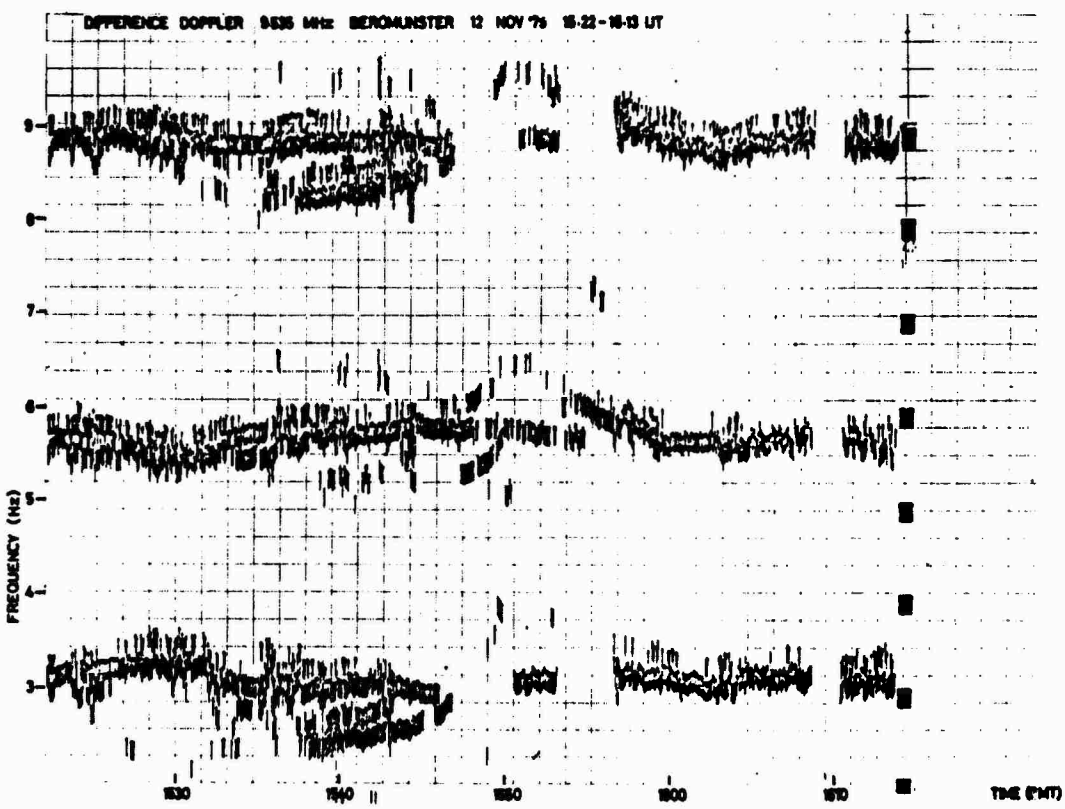


Fig.8 Time variations of 'difference' frequencies, multimode conditions

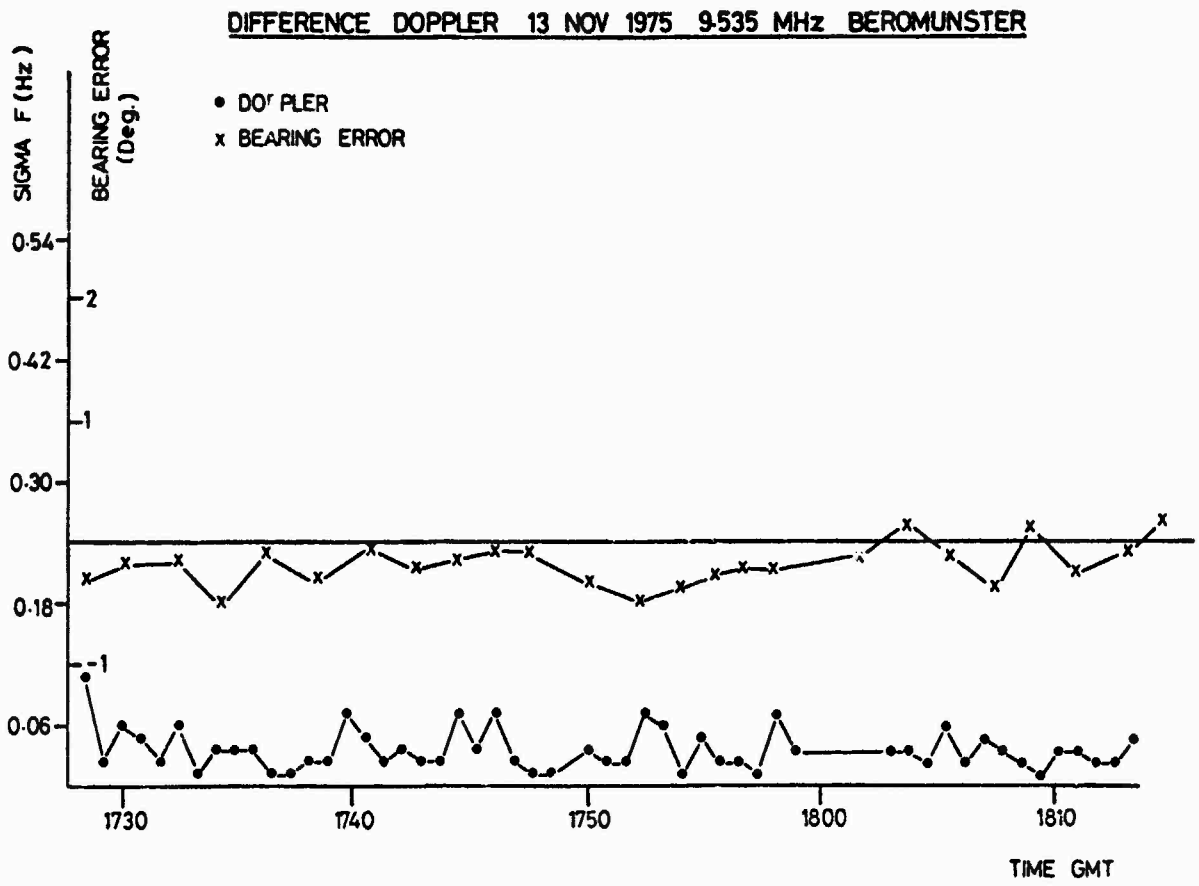


Fig.9 Comparison of Doppler index $\Sigma\Delta f$ and bearing error. No TIDs

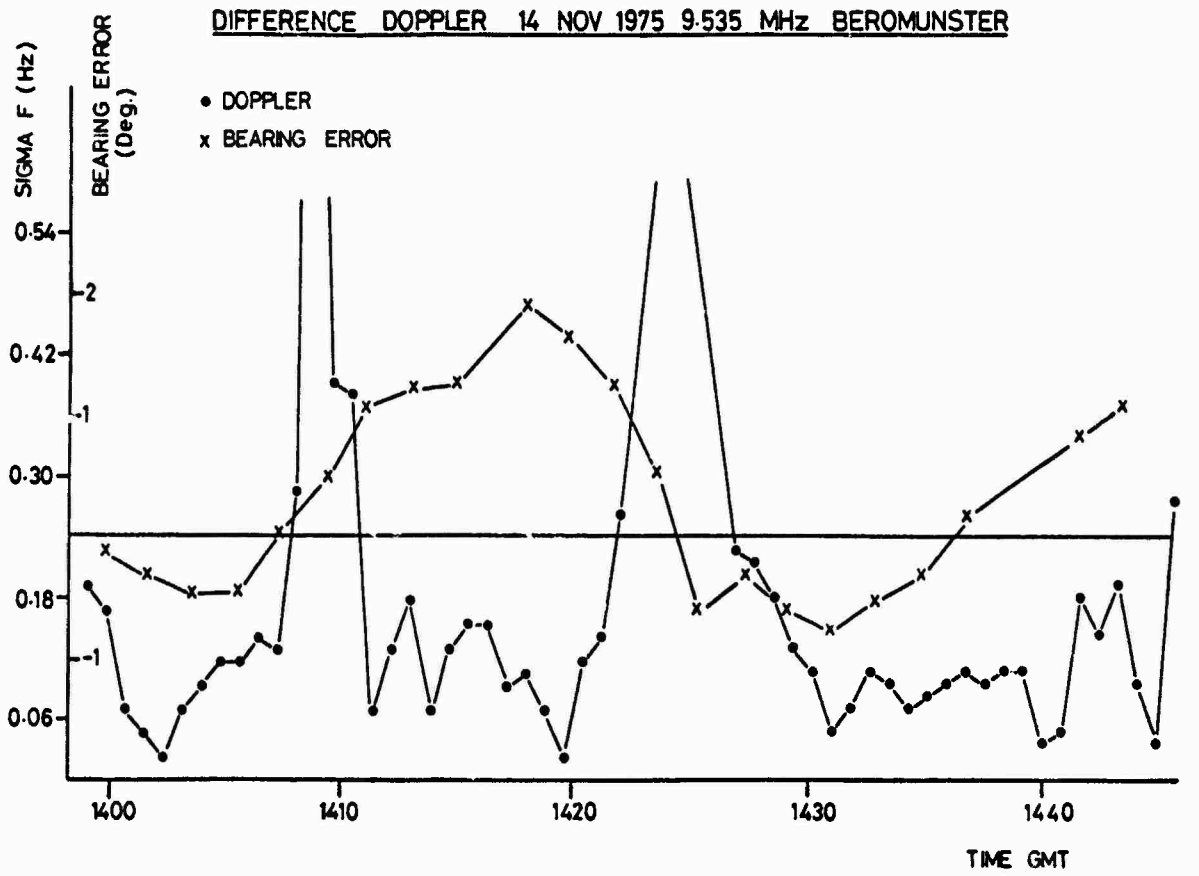


Fig.10 Comparison of Doppler index $\Sigma\Delta f$ and bearing error. TID active

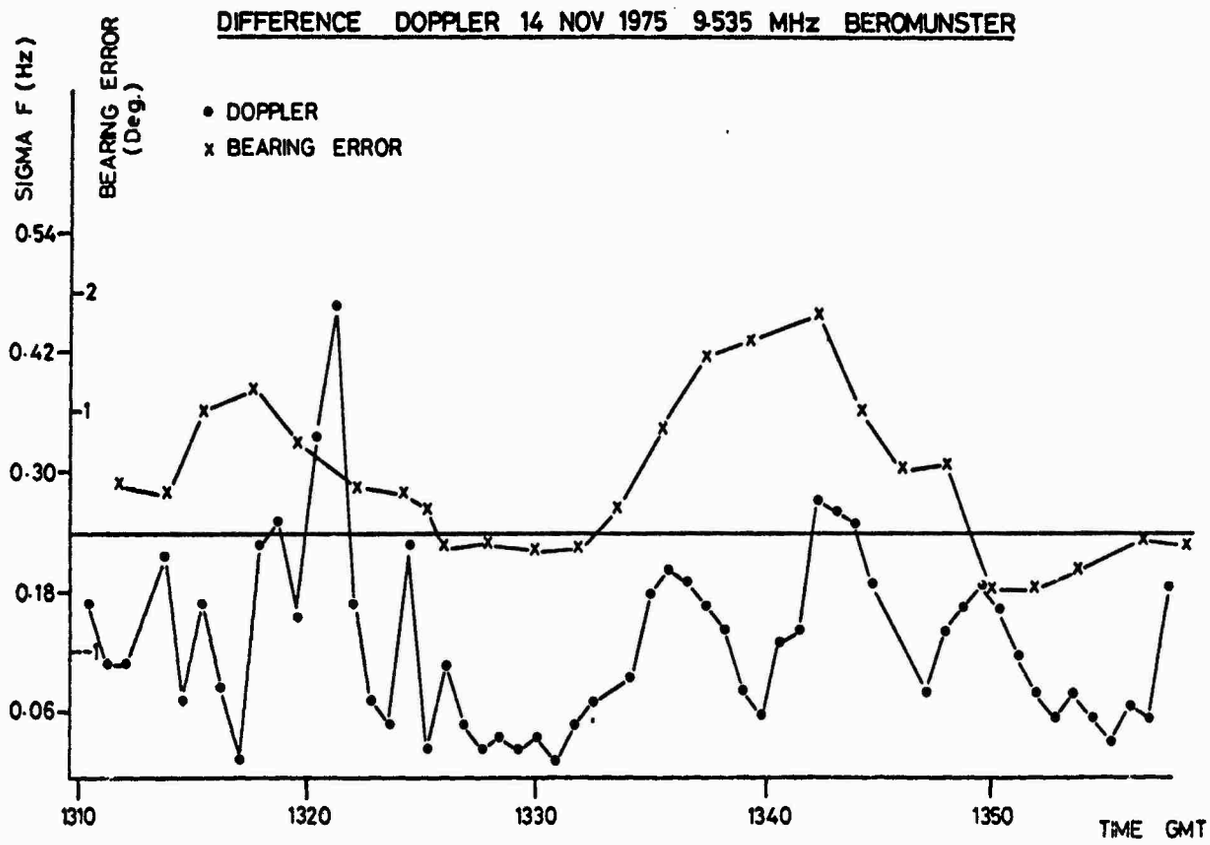


Fig.11 Comparison of Doppler index $\Sigma\Delta f$ and bearing error. TID active

DISCUSSION

R.H.Doherty, CRPL$_i$, P.O.Box 1056, Boulder, Colorado 80306, USA

In view of your experiences in direction finding systems and also in LF and VLF phase monitoring systems and in view of the fact that Russian LORAN C transmitters exist, but with no accurate locations given, how well do you think we could locate these stations, particularly if sky wave signals only were to be used?

Author's Reply

It is not possible at this time to give a definite answer to your question except that we probably have the understanding and the technical expertise to solve this problem. I think that this would be a fruitful area for further research.

C.Goutelard

L'exposé du Docteur T.B.Jones est trés intéressant et m'a tout particuliérement captivé car utilisant une methode de soundage par retrodiffusion avec une antenne orientable nous trouvons les mêmes phénomênes que lui, notamment sur les fréquences doppler. Cependent, les perturbations ionosphériques agissent comme des lampes placées dans l'ionosphere et agissent sur l'amplitude. L'utilisation de l'amplitude s'est avérée, pour nous trés difficile. Le Docteur T.B.Jones peut-il nous dire s'il l'a fait et le cas échéant comment il a procédé.

Author's Reply

The Doppler frequency changes are accompanied by amplitude variations. We have not made a detailed study of these since they are very complicated phenomena and more difficult to interpret than the frequency changes. It is perhaps worth noting that amplitude changes can be produced by mechanisms other than TIDs which are the principal objective of our study.

H.Soicher, US Army Electronics Command, Port Monmouth, NJ 07703, USA

Is the propagation mode nearly vertical in your observation?

Are you concerned with actually locating an unknown transmitter? In such case you need bearing and elevation which necessitates a sounding system!

Author's Reply

No, the Doppler observations are made on the signals received from the distant transmitters whose position is being determined by the direction finders. This means that no sounding of the ionosphere is necessary and the technique is entirely 'passive'. We are not concerned with locating the unknown transmitter by means of the Doppler method but in assessing the quality of the measurements obtained on a wide aperture DF system.

POSITION-FINDING OF FIXED HF-TRANSMITTERS BY MEANS OF TRAVELING IONOSPHERIC STRUCTURES

J. Röttger
Max-Planck-Institut für Aeronomie
D-3411 Katlenburg-Lindau 3, F.R.G.

SUMMARY

An inversion method is described which represents an alternative approach to the common direction-finding systems to determine the unknown position of a remote transmitter under sky-wave conditions. The technique bases on the measurement of signal parameters which are influenced by the reflection at traveling ionospheric structures, and evidently makes use of a fact considered to introduce errors in common direction-finding techniques.

It is assumed that the parameters of a traveling ionospheric structure or disturbance, like velocity and azimuth of propagation, are deduced by a measurement of the Doppler shift using a known array of three transmitters and one common receiver. The cross-correlation of the signals of the known array with the unknown signal yields the position of the unknown transmitter by applying some geometrical calculations. The basic method to evaluate the signal variations by means of the cross-correlation and cross-spectrum technique is described. Calculations basing on appropriate models of traveling ionospheric structures are carried out to prove the applicability of the introduced method. Some implications on the accuracy and reliability of the position determination are considered.

1. INTRODUCTION

HF-CW-Doppler arrays are commonly in use to investigate the amplitude, velocity and propagation direction as well as the wavelength of traveling ionospheric disturbances (TIDs) (e.g. DAVIES and BAKER, 1966; DAVIES and JONES, 1972; HERRON, 1973). In these experiments, mostly three transmitters at the corners of a triangle at a distance of up to 100 km are used to transmit highly frequency-stable HF-CW signals. After reflection at the ionosphere, the signals are recorded at a common receiving station in the centre of the triangle. From the Doppler shift of the three signals, the vertical velocity of the ionospheric reflection areas is calculated. By means of evaluating the time delay between the Doppler frequency variations of the three signals, the horizontal velocity vector and the wavelength of the TIDs propagating through the network can be deduced.

Another kind of experiments are the ionospheric drift measurements making use of the spaced receiver technique (KENT and WRIGHT, 1968). The method bases on the principle that one measures the interference pattern of a source signal reflected at drifting ionospheric irregularities. From this pattern the parameters of the irregularities are obtained.

While in these commonly applied experiments the locations of the transmitters and the receiver are known whereas the parameters of the TIDs or the velocities of drifting irregularities are to be deduced, an inverted system can be used to find the location of a transmitter if the TID or drift parameters are known, e.g. by means of a separate measurement. This method makes use of the influence of the time-varying propagation medium which commonly is regarded as a disadvantage of direction-finding systems where ionospheric variations cause a considerable error in the position determination (e.g. GEORGE, 1972; MORGAN, 1972; JONES and SPRACKLEN, 1976).

Suppose we have measured the horizontal velocity vector of traveling ionospheric structures (e.g. drifting irregularities or TIDs) in the ionosphere which influence the phase path of a reflected signal, and provide that the coherence length of these structures is larger than the dimensions of the position-finding network consisting of three transmitters and one central receiver. A signal of a fourth transmitter with unknown position which is received at the common receiving station will also be influenced due to the traveling ionospheric structures (TISs) like the signals of the known transmitters. By measuring the time shift (by means of cross-correlation or cross-spectrum techniques) between the signal variations of the three known and the one unknown transmitter, one can deduce the horizontal propagation direction of the TISs and the position of the unknown transmitter.

The introduced system for position finding under sky-wave conditions, which is described in this paper, bases on a rather different principle than the commonly used direction-finding systems. The method simply evaluates the image of a transmitting source reflected at a distorted two-dimensional mirror (i.e. the reflecting ionosphere disturbed by traveling ionospheric structures). Since the time-varying distortion of the reflector is known by a measurement, the source can be located by a cross-correlation or cross-spectrum analysis. In this paper, we present the basic features of this method by applying a model introducing traveling ionospheric structures or disturbances to calculate the signal distortions from which then the original data are reconstructed. The dependence on the signal-to-noise ratio as well as some conditions to minimize the error sources are outlined.

2. BASIC EQUATIONS TO CALCULATE THE POSITION OF AN UNKNOWN TRANSMITTER

The HF-CW-Doppler technique makes use of frequency variations introduced by a time- and space-varying reflector of radio waves which, in the case of HF sky-wave propagation, is the ionosphere. From the measured Doppler shift of the CW signal the time and space variations of the reflector can be reconstructed and an inversion method can be applied to reconstruct the observational network.

In an array of four transmitters TX0, TX1, TX2 and TX3 (see Fig. 1) and one common receiving station RX, we find the projection of the sky-wave propagation paths in the horizontal x-y plane indicated by the

straight lines TX0-RX, TX1-RX, TX2-RX and TX3-RX. The projection of the ionospheric reflection points is given by the circles number 0,1,2 and 3, assuming simply one-hop ionospheric propagation without ray-path unsymmetries due to horizontal electron density gradients of the background ionosphere. It is furthermore implied that all four transmitters operate at approximately the same frequency so that the ionospheric influence is comparable on all propagation paths.

Suppose two characteristic structures, e.g. stretched sporadic-E clouds, propagate through this array. The moving structures which cause phase-path variations to be measured by the Doppler technique are characterized by their horizontal velocity component v_j and the propagation of azimuth α_j. The index $j = 1,2$ shall define two different traveling ionospheric structures TIS 1 and TIS 2. When these structures move through the network, we will observe them passing the different reflection points at different times t. We, for instance, will first observe TIS 1 passing through reflection point 0 at the time t_{0j}. Somewhat later it passes point 3. At the time t_{1j} it will be at reflection point 1 between TX1 and RX, and finally at point 2. A similar picture will be observed when TIS 2 moves through the system. If we assume that the shape, velocity and azimuth of the TISs do not change during an observational time interval, we measure the time differences τ_j ($j = 1,2$) when the two TISs are observed at points 2 and 3 with respect to point 0, and find from the basic formula

$$v\,t = x \sin\alpha + y \cos\alpha \tag{1}$$

the azimuths α_j and the velocities v_j of the TISs to be:

$$\alpha_j = \arctan \frac{y_2 \cdot \frac{\tau_{3j}}{\tau_{2j}} - y_3}{x_3 - \frac{\tau_{3j}}{\tau_{2j}} \cdot x_2} \tag{2}$$

and

$$v_j = \frac{-x_3 \sin\alpha_j - y_3 \cos\alpha_j}{\tau_{3j}} \tag{3}$$

with $\tau_{2j} = t_{2j} - t_{0j}$ and $\tau_{3j} = t_{3j} - t_{0j}$ $(j = 1,2)$. (4)

The four reflection points are defined by their coordinates x_i, y_i with $i = 0,1,2,3$. Thus, the first index of τ_{ij} always indicates the reflection point and the second index the TIS. In the case of determining the azimuth and velocity of propagation, we need only measurements at the points 0,2 and 3, where the transmitter locations X_i, Y_i ($i = 0,2,3$) as well as the receiver location and therefore also the locations of the ionospheric reflection points are known. The location of the unknown transmitter TX1 can be deduced from the additional measurement of the time differences $\tau_{1j} = t_{1j} - t_{0j}$ after reduction of the azimuths α_j and velocities v_j.

Knowing v_j and α_j from the basic measurement (equations (2),(3) and (4)) and measuring the time difference $\tau_{11} = t_{11} - t_{0j}$ which TIS 1 needs to travel from the known reflection point 0 to the unknown reflection point 1, we find this unknown point 1 to be somewhere on the straight line I. This can be understood since we measure τ_{11} to be constant for all locations of TX1 (e.g. <TX1>) with the reflection point lying on line I. A second measurement basing on TIS 2 with a different azimuth of propagation leads to the straight line II. The crossing point of these lines I and II consequently is the ionospheric reflection point 1 of the path from the unknown transmitter TX1 to the receiver RX. Assuming symmetrical ray paths, thus, also the unknown position X_1, Y_1 can be calculated.

The position finding also can be achieved using a hybrid system determining the straight line I by using one TIS as an indicator and using a direction finder to measure the angle β (see Fig 1). The crossing of these two lines is the unknown reflection point 1 from which the unknown position X_1, Y_1 also can be found.

Returning to the original method, the coordinates of the unknown ionospheric reflection point x_1, y_1 are deduced from (1),(2) and (3) to be:

$$x_1 = \frac{\frac{v_1 \cdot \tau_{11}}{\sin\alpha_1} - \frac{v_2 \cdot \tau_{12} \cdot \cos\alpha_1}{\cos\alpha_2 \cdot \sin\alpha_1}}{1 - \frac{\cos\alpha_1 \cdot \sin\alpha_2}{\sin\alpha_1 \cdot \cos\alpha_2}}, \tag{5a}$$

$$y_1 = \frac{v_2 \cdot \tau_{12} - \frac{x_1 \cdot \sin\alpha_2}{\cos\alpha_2}}{\cos\alpha_2} \tag{5b}$$

with $\tau_{1j} = t_{1j} - t_{0j}$ $(j = 1,2)$.

From these we consequently reduce the coordinates of the unknown transmitter location:

$$X_1 = 2 \cdot x_1 \,, \tag{6a}$$

$$Y_1 = 2\, y_1 + Y_o \,, \tag{6b}$$

where Y_o is the ordinate of TX0.

It should be kept in mind that these values are exact only for the applied model of TISs which significantly can be discriminated from noise fluctuations and do not change their characteristics throughout the intire interval of measurement. The calculations, furthermore, base on geometrical optics calculating the ray paths which are assumed to be symmetrical and do not suffer any azimuthal deviations. Another necessary condition to apply this method is that the TIS pattern has to be two-dimensional, which for instance can be achieved by two different one-dimensional TISs (used in the explained example and extended in Chapter 3.1) or by periodic structures of spatially interfering traveling ionospheric disturbances (Chapter 3.2).

It is anticipated that the accuracy of the introduced basic method can be improved by applying more sophisticated ray-tracing methods. This, however, does not exclude the necessity to check the temporal and spatial coherence of the TISs as well as applying proper data-processing techniques to improve the signal-to-noise ratio by filtering out significant TISs to measure the time shifts τ with high accuracy.

3. DATA PROCESSING TECHNIQUES FOR TIS AND POSITION DETERMINATION

The equations to calculate the TIS characteristics, azimuth and velocity, as well as the position of the unknown transmitter base on the measured values of the time delays τ caused at different locations in the ionosphere. We, thus, find it necessary to develop a procedure to measure the time delays at a most high accuracy and reliability.

In order to clarify the measuring technique, we present in Fig. 2 an example of a Doppler display recorded in Huancayo/Peru, where a three-point HF-CW-Doppler array was set up to investigate traveling ionospheric disturbances in equatorial regions (RÖTTGER and BECKER, 1977). The characteristic sizes of this network are between 20 and 50 km. The notations MAT, PAM and GRB in Fig. 2 indicate the three records from the three paths TX0 (Matahuasi), TX2 (Pampas) and TX3 (Gran Bretañe) to the common receiving station at the Geophysical Observatory of Huancayo. Besides some pulse-like distortions, clearly wave-like Doppler-frequency variations with periods between about 10 and 20 minutes are observed. These oscillations, which are caused by medium-scale TIDs, indicate a time-delay shift between the three paths which incidently points to the fact that the TID structure moved through the network. Since these non-periodical and periodical TIS patterns often are observed on Doppler records taken at all latitudes (e.g. review of FRANCIS, 1975), we are confident to present in the following part two different methods to deduce the time shift for pulse-like TISs as well as periodic TIDs, and calculate by these means the position of an unknown transmitter.

As pointed out in Chapter 2, we need to measure the time series of parameters characterizing the signals transmitted via the four different paths. In the following parts we confine to the signal parameter, Doppler frequency variation, since this parameter appears to be most sensitive to motions in the ionosphere. The time series of Doppler frequency variations show typical events like demonstrated in the sample records of Fig. 2. In the case, a TIS moves through the network, these records indicate a phase or time shift between the variations measured on the different paths. To deduce the time shifts at a most high accuracy, both, the cross-correlation and the cross-spectrum technique are accepted as appropriate tools. Whereas the cross-spectrum technique is most adaptive to periodic variations allowing to filter significant oscillations, thus, improving the signal-to-noise ratio, we find the cross-correlation technique most suitable for pulse-like variations.

After pointing out the basic items to calculate the time shifts by both techniques, a model, deduced from deterministic signals covered by noise, is applied to show the applicability of these techniques. The model calculations allow a trustable check, since we calculate the signal distortion by introducing a known TIS moving through a known network, add artificial noise and apply either the cross-correlation or the cross-spectrum technique. Finally, we close the circuit calculating the TIS parameters and the transmitter location from the results of the cross-correlation and cross-spectrum technique computations and compare the results with those values originally introduced.

A time-varying signal parameter A_i of each transmitter TX_i ($i = 0,1,2,3$) measured at the receiver RX is given by

$$A_i(t) = S_i(t,\underline{r}_i) + N_i(t) \,, \tag{7}$$

where t is the time and $\underline{r}_i = (x_i, y_i, z_i)$ is the radius vector to a location (number $i = 0,1,2,3$) in the ionosphere where an influence of the signal parameter due to a TIS has to be considered. Basing on the introducing assumptions of geometrical ray-path calculations, we consider $z_i = \text{const}$. A_i may stand for any signal parameter, like incidence angle, amplitude or frequency. In the applied case of the HF-CW-Doppler technique, the parameter A_i is the measured Doppler frequency which consists of the contribution of the signal S_i and the noise N_i received in the frequency band under evaluation. The noise also can be due to some arbitrary modulation (e.g. FM modulation if we apply the HF-CW-Doppler technique), distorting the signal to be evaluated for the purpose of position determination.

The signal-Doppler frequency S_i consists of the transmitted signal frequency S_i^o distorted by ionospheric variations, which shall be defined by a distortion parameter $D(t,\underline{r}_i)$:

$$S_i(t,\underline{r}_i) = S_i^o(t) \cdot D(t,\underline{r}_i) \,. \tag{8}$$

The signal distortion $D(t,\underline{r}_I)$ is assumed to be due to a TIS propagating through the reflection points 0,1,2 and 3 so that variations of all measured signals S_I occur, which indicate a time shift as a function of the TIS and the array parameters.

After this basic definition of the signals influenced by traveling ionospheric structures and noise, we introduce the cross-correlation and the cross-spectrum technique for the evaluation of the data time series.

3.1. Cross-correlation technique

For any signals defined by (7), estimates of the cross-correlation functions

$$CCF_i(\tau) = \frac{1}{T}\int_T A_i(t)\, A_o(t+\tau)\, dt \qquad (i = 1,2,3) \tag{9}$$

can be calculated in the observational time interval T. From the maxima of these cross-correlation functions CCF_I, the significant time-delay shifts τ with reference to reflection point 0 are to be deduced. If two TISs are traveling through the network (see Fig. 1), we find from the time shifts τ_{IJ} of the maxima of the cross-correlation functions:

$$\tau_{ij} = \tau(\max_j(CCF_i))$$

with i = 1,2,3 (indicating the reflection points) and j = 1,2 (indicating the TISs). These values τ_{IJ} then can be used to calculate by means of (2),(3) the azimuths α_j, velocities v_j, and from (5a),(5b) the coordinates of the unknown reflection point x_1, y_1.

The outlined procedure shall now be elucidated by an example. For this purpose, we assume the HF-CW-Doppler network to be defined by the following locations of the ionospheric reflection points using the coordinate system introduced in Fig. 1:

$$x_o = 0.0 \text{ km}, \quad x_1 = 17.5 \text{ km}, \quad x_2 = -25.0 \text{ km}, \quad x_3 = +25.0 \text{ km},$$

$$y_o = 0.0 \text{ km}, \quad y_1 = -20.0 \text{ km}, \quad y_2 = -45.0 \text{ km}, \quad y_3 = -45.0 \text{ km}.$$

Two pulse-like TISs with equal horizontal velocities $v_{1,2} = 150 \text{ m s}^{-1}$ are assumed to travel approximately simultaneously through the network at the different azimuth angles $\alpha_1 = 150^o$ and $\alpha_2 = 210^o$. The TISs are straight lines not changing their characteristics throughout the entire network (see Fig. 1). Basing on these assumptions, evidently the exact times can be calculated when the TISs influence the signals propagating along the four paths TX0-RX, TX1-RX, TX2-RX and TX4-RX. To discriminate in the calculations between the two TISs, TIS 1 is introduced to cause twice the amplitude in frequency variation than TIS 2. For simplicity, the frequency variations are set to be step functions with a duration of 1 min, which corresponds to typical TIS dimensions of some kilometers. TIS 1 first causes a negative frequency shift followed by a positive shift, whereas TIS 2 acts just vice versa. The additive noise component is generated by equally distributed random numbers. The signal-to-noise ratio is determined by the ratio of the peak amplitude of TIS 1 and the peaks of the noise. Since typical times, the TISs need to travel through the network, are in the order of several minutes, samples from the time series $A_I(t)$ are taken every 30 seconds for digital computation of the estimates of the cross-correlation functions CCF_I.

A sample from the time series generated by means of the described procedure is shown in the upper right half of Fig. 3. The pulse response 1 is due to TIS 1 and the response 2 to TIS 2. The noise contributes to the signal as an additive component. From the time series A(t), the estimate of the cross-correlation function CCF is calculated using equation (9) and setting the integration time interval T = 100 min, which is approximately ten times the maximum expected time shift τ_{IJ}. T can be reduced to a smaller value which, however, decreases the signal-to-noise ratio of the CCF, so that the significance of the peaks at τ_{IJ} will be diminished. The greater the signal-to-noise ratio of the original time series the shorter the integration time T can be chosen. It is proposed to apply a procedure to minimize T, which has to take into account the signal-to-noise ratio, the shape of the pulse response due to the TISs as well as the maximum time shift. Since the noise contribution not only is from external sources but may be due to an arbitrary modulation of the transmitters (in our first approach we have assumed a noise-like time variation of the modulation signal), one also should investigate in succeeding analyses how the method reacts on different modulation techniques.

If the signals $S_I^o(t)$ contain any correlated component, if the common receiver RX introduces signal variations (e.g. due to a non-stable oscillator), or if the variable ionospheric influence is equal at all reflection points (e.g. due to very large-scale disturbances with dimensions an order of magnitude greater than the characteristic dimensions of the TISs), the signals have to be expressed by

$$S_I^o(t) = S_I^s(t) \cdot D^o(t), \tag{10}$$

where $S_I^s(t)$ are the original stable signals and $D^o(t)$ is a correlated disturbance. Due to the influence of $D^o(t)$, we will find a peak of the cross-correlation functions near $\tau = 0$. We will call this in the following part an instrumental correlation. Since this peak occurs in all three CCFs, we can calculate a mean $\overline{CCF} = \frac{1}{3}\sum_{i=1}^{3} CCF_I$, and use it for calculating modified cross-correlation functions $MCCF_I = CCF_I - \overline{CCF}$. We, thus, can eliminate the instrumentally correlated, non time-shifted constituents of the signals and improve the significance of the determination of the time shifts τ_{IJ}. A different approach than the cross-correlation technique to improve the significance of the method, i.e. the reliability of the position determination, is given by the filtering technique which is presented in Chapter 3.2.

An example of the cross-correlation function calculated from the time series $A_0(t)$ and $A_2(t)$ is shown in Fig. 3. At $\tau = 0$ we find the peak due to an instrumental correlation (here the peak-to-peak

variation of $D^o(t)$ is assumed to be equal to the peak-to-peak variation of the uncorrelated noise $N_i(t)$). The significance level to find the relevant peaks 1 and 2 can be defined by the level which is exceeded by the uncorrelated noise peaks with a 5% probability (or any other percentage value). Another way to find the relevant peaks and the corresponding time shifts τ_{ij} is to search for the two highest maxima of the MCCFs and assume these to be due to two TISs. Since this procedure, which shall be called "maximum amplitude criterion", strongly depends on the signal-to-noise ratio, we have to find out the limiting conditions for the significance of the parameters deduced.

The depicted method was applied to compute the picture shown in Fig. 4. We find that the amplitudes of the cross-correlation functions more clearly exceed the noise level with increasing signal-to-noise ratio S/N. Also the calculated ionospheric position, indicated in the upper part of the diagram by the coordinates x and y, more and more approaches to the real values x_1 and y_1. Due to the fact that the digital samples from the time series were not taken continuously but at discrete times every 30 seconds, we find a systematic quantization error of the final position. However, this error can be minimized by increasing the sampling rate. It is found from this computation - provided that the above mentioned conditions, e.g. spatial and temporal coherence etc. are fulfilled - that the position of an unknown transmitter can be deduced with a sufficient accuracy if the signal-to-noise ratio is greater than about two and the sampling rate is high enough to reduce the quantization error to an acceptable value.

3.2. Cross-spectrum technique

It was shown that the cross-correlation technique appears to be applicable to signal distortions caused by non-periodic, pulse-like TISs. Because more often periodic TISs, i.e. traveling ionospheric disturbances (TIDs), are observed than pulse-like TISs, we find that the cross-correlation technique should be exchanged by the more appropriate tool of the cross-spectrum technique. This fact we anticipate since in periodic systems a spectrum analysis in any case is more advantageous due to its property to filter out noise components. (It should be mentioned here that this spectrum analysis must not be mixed up with the spectrum analysis carried out during the preparatory processing procedure when deducing the Doppler spectrum of the signal.)

The wave-like distortion due to a TID is assumed to be

$$D(t,\underline{r}) = D_o \sin(\underline{k} \cdot \underline{r} - \omega t), \tag{11}$$

where D_o is the disturbance amplitude,
$\underline{k}$ is the wave vector of the TID,
ω is its angular frequency,
t is the time.

The radius vector $\underline{r}$ is defined, as in Chapter 3.1, pointing to the location where the TID influences the ray path. Introducing the parameters of the coordinate system given in Fig. 1 and knowing that the horizontal phase velocity of a TID is $v = \omega/k$ and the horizontal wavelength is $\lambda = 2\pi/k$, the distortions due to two TIDs ($j = 1,2$) are:

$$D_{ij} = D_j(t,\underline{r}_i) = D_{oj} \sin\left(\frac{2\pi}{\lambda_j}(x_i \sin\alpha_j + y_i \cos\alpha_j) - v_j \cdot t\right). \tag{12}$$

Introducing (12) into (8), we find the signal time series $A_i(t)$ varying periodically if TIDs with horizontal wavelengths λ_j are moving through the network. When assuming the coordinates x_i, y_i, the velocities v_j and the propagation azimuths α_j as well as the integration time T to be the same like those used in the cross-correlation computations, and assuming two TIDs of the medium-scale class with the periods $T_{TID\,1} = 20$ min and $T_{TID\,2} = 33$ min (which at the horizontal phase velocity $v = 150$ m s^{-1} ($v = \lambda/T$) correspond to $\lambda_1 = 180$ km and $\lambda_2 = 300$ km), we find the cross-correlation function CCF 2 shown in Fig. 5. It is evident from this picture that different peaks like in the case of the double-pulse TISs can no more be distinguished.

Since we find that the cross-correlation technique appears to be not appropriate in the case of periodic disturbances, the spectrum analysis has to be introduced. The complex spectrum, calculated at the integration time interval T, is given by its estimate

$$A^*(\omega) = \frac{1}{T} \int_T A(t) \exp(-i\,\omega\,t)\, dt \tag{13}$$

from which we can deduce the amplitudes A_i and phases φ_i ($i = 0,1,2,3$), measured as a function of the TID frequency ω, respectively period T_{TID}, on the propagation paths 0,1,2 and 3. When introducing a significance level to find non-statistical amplitudes in the spectrum, again the amplitudes of relevant TIDs can be determined (e.g. by means of the maximum amplitude criterion). Consequently also the phases of the TID distortions which are observed in the time series are found in the spectrum. From these phases, the time shifts

$$\tau_{ij} = \frac{\varphi_{ij} - \varphi_{oj}}{2\pi} \cdot T_{TID\,j} \tag{14}$$

are calculated. These time shifts τ_{ij} then are used to compute the azimuths α_j, the velocities v_j as well as the coordinates x_i and y_i, by means of equations (2),(3) and (5a),(5b).

An example of this technique is given in Fig. 6. In the diagram the time series of the signal oscillations due to the influence of the TIDs at the reflection points 0,1,2 and 3 are shown. Here TID 1 travels at an azimuth angle $\alpha_1 = 240^o$ and TID 2 at an angle $\alpha_2 = 120^o$. The amplitudes of the disturbances are $D_{o1} = 2.00$ and $D_{o2} = 1.75$. To clarify the picture of the time-shifted oscillations, the noise amplitude is assumed to be zero. We evidently find in the time series of Fig. 6 the interfering oscillations at the periods $T_{TID\,1} = 20$ min and $T_{TID\,2} = 33$ min.

To get an impression on the ionospheric corrugation which is caused by two TIDs spatially interfering over the network, we calculated the print-out shown in Fig. 7. The print-out density contours can be understood to be corrugations of height levels of constant plasma frequency in the ionosphere (which

act as a corrugated mirror for HF waves). Like in the coordinate system of Fig. 1, the upper part of the picture in Fig. 7 is north. Depending on the horizontal phase velocities $v_{1,2} = 150$ m s^{-1}, the interference pattern takes about one hour to move from the northern part to the southern part of the map, where the two wave disturbances independently propagate approximately into WSW and ESE direction.

In Fig. 8, the influence of the signal-to-noise ratio on the accuracy of the spectrum analysis as well as the consecutive position determination are demonstrated. Basing on the integration time $T = 100$ min and the sampling frequency $1/30$ s^{-1}, the complex spectrum can be calculated for the periods $T = 100$ min/n ($n = 1,2,3,\ldots,100$). Figure 8 shows the harmonic analysis of the signal time series as observed on the path TX1-RX. This diagram does not show the periods shorter than 16.6 min since in the case of $T_{TID\,1} = 20$ min and $T_{TID\,2} = 33$ min the higher order spectral lines are due to noise contributions only. As expected from the introduced model, the amplitudes A_{1j} of the spectral lines at $T_{TID} = 20$ min and 33 min clearly exceed the noise level (e.g. at $T_{TID} = 100$ or 50 min). For signal-to-noise ratios greater than 0.75 (S/N with reference to D_{01}) the amplitudes almost exactly approach the correct values. The corresponding phases φ_{1j} even reach this condition at a much lower signal-to-noise ratio, whereas the noise phases at $T_{TID} = 100$, 50, 25 and 17 min indicate random variations only.

In the lower part of Fig. 8 the velocity $\bar{v} = (v_1 + v_2)/2$ and azimuth of the TID phase-fronts $\overline{|\alpha'|} = (|\alpha'_1| + |\alpha'_2|)/2$ as well as the coordinates of the unknown reflection point 1 are shown as a function of the signal-to-noise ratio. These parameters are deduced from the complex cross-spectra applying the maximum amplitude criterion. We find that at a signal-to-noise ratio greater than two the values of the parameters v, α, x_1 and y_1 rather exactly are the original values, which in Fig. 8 are indicated by the horizontal lines.

The results presented in Fig. 8 depend peculiarly on the condition that the periods of the TIDs are integer portions, i.e. harmonics of the integration time interval. Since this condition mostly cannot be achieved in praxis, we ask for the error which is introduced by a truncation of the time series. It should be treated in succeeding investigations, however, if a cross-power spectrum analysis, which combines the cross-correlation and the cross-spectrum technique, would be more convenient than the described harmonic analysis.

In Fig. 9 the results of computations are shown, starting with the truncation of the time series at $t = 20$ min (see Fig. 6), which is one-fifth of the integration time $T = 100$ min used to calculate the results given in Fig. 8. The velocity v and the azimuth angle α as well as the coordinates x,y approach almost exactly the original values for increasing integration time ($T \to t = 100$ min). It is to note that after one hour observation, i.e. integration time, and at the rather low signal-to-noise ratio of two, the accuracy of the azimuth measurement as well as the accuracy of the position determination appear to be very acceptable. To determine the velocity with a similar accuracy, we, however, find it necessary to integrate almost 100 minutes.

To use this method for applications, it is proposed to start the integration, i.e. the spectrum analysis, as soon as a few minutes of observations are available. With increasing integration time, the parameter values should converge more and more if a TID structure moves through the network. By taking statistical averages from different samples of observational intervals, one should in praxis find the statistics of traveling ionospheric disturbances and the position of an unknown transmitter with an appropriate reliability.

4. CONCLUSION

The technique for position-finding under sky-wave conditions, which is described in this paper, bases on a rather different principle than the commonly used direction-finding systems. The presented method is to evaluate the image of a transmitting source reflected at a corrugated mirror (i.e. the reflecting ionosphere disturbed by a traveling ionospheric structure). After the time-varying distortion of the reflector is determined by a measurement using a three-point HF-CW-Doppler system, the source is located by the cross-correlation or the cross-spectrum technique.

The applied model calculations prove that the basic technique operates adequately to determine the position of an unknown transmitter. It appears that this method indicates some advantages in comparison to common direction-finding systems using goniometers, provided that those basic conditions of spatial and temporal coherence of the traveling ionospheric structure (TIS) as well as sufficient integration time are fulfilled. It is understood that the technique operates most appropriate under strong TIS conditions (even the background ionospheric drift might be used), which is just opposite to the direction-finding technique, which operates properly under undisturbed conditions. Thus, one should think about a hybrid system consisting of a goniometer direction-finder and an HF-CW-Doppler array to provide operation under undisturbed (no TISs) and disturbed (TISs) conditions. On the other hand, a supplementary HF-CW-Doppler system can - besides determining the position of a transmitter - give valuable information on the traveling ionospheric structures or disturbances to correct the bearing error of the direction-finding system. Another possibility would be to use one direction-finder to measure the azimuth of the unknown transmitter and deduce one crossing line by means of the cross-spectrum technique using an HF-CW-Doppler system to determine the transmitter location. This hybrid system would be appropriate if the TIS pattern is one-dimensional only and one does not find in a further measurement a succeeding TIS with different propagation azimuth.

We find another advantage of the HF-CW-Doppler technique in using very simple antennas instead of the sophisticated direction-finder antennas. The cross-correlation and the cross-spectrum technique allow under specified conditions even to localize transmitters with an unstable signal frequency or arbitrary modulation. Multi-path errors (e.g. 1F- and 2F-propagation) can be eliminated in case of TID signal distortions. The fact that TIDs cause different Doppler shifts on different paths can be used to discriminate the paths by their different Doppler shift and amplitude. It is understood that the method operates more reliable at very steep than at low elevation angles. Due to the fact that common direction-finding systems show increasing error at steep ionospheric incidence angles, the introduced method appears to perform a reasonable supplement to the direction-finding system in use.

It should be stressed, however, that supplements and improvements of the HF-CW-Doppler technique to use TIS patterns for position-finding should be applied in order to minimize the error rate and improve the reliability when operating the system in praxis. As pointed out earlier, checks with ray-tracing computations, taking into account different ionospheric models, are desirable as well as more expanded investigations on the influence of the frequency stability, i.e. different modulation techniques, the influence of the signal-to-noise ratio, the integration time, the optimum quantization of the sampling rate, optimum significance levels and the temporal and spatial coherence of the moving TIS pattern. A most desirable continuation of the procedure presented in this paper is expected to be a comparison of the applied model calculations with measurements using a four-point HF-CW-Doppler network as well as to compare the obtained results with the goniometric direction-finding.

In this paper the features of the cross-correlation and the cross-spectrum technique are investigated by means of model calculations taking into account the influence of TIDs on the propagation of HF-waves and v.v. deducing the TID parameters from the observations. It appears that the cross-spectrum technique is most appropriate to be applied for the evaluation of periodic time series recorded by means of HF-CW-Doppler systems. We expect as a consequence an improvement of the evaluations of HF-CW-Doppler signals for position determination and for the investigation of the characteristics of travelling ionospheric disturbances.

REFERENCES

DAVIES, K. and D.M. BAKER, 1966, "On frequency variations of ionospherically propagated HF radio signals", Radio Sci. 1, No. 5, 545-556.

DAVIES, K. and J.E. JONES, 1972, "Evidence for waves and winds in the ionospheric F region", Space Research XII, 2, 1149-1156.

FRANCIS, S.H., 1975, "Global propagation of atmospheric gravity waves: A review", J. Atmos. Terr. Phys. 37, 1011-1054.

GEORGE, P.L., 1972, "H.F. ray tracing of gravity wave perturbed ionospheric profiles", AGARD-CP-115, paper 32.

HERRON, T.J., 1973, "Phase velocity dispersion of F-region waves", J. Atmos. Terr. Phys. 35, 101-124.

JONES, T.B. and C.T. SPRACKLEN, 1976, "The correction of errors produced in H.F. direction finders by travelling ionospheric disturbances", AGARD-CP-173, paper 34.

KENT, G.S. and R.W.H. WRIGHT, 1968, "Movements of ionospheric irregularities and atmospheric winds", J. Atmos. Terr. Phys. 30, 657-691.

MORGAN, A.D., 1972, "The effect of ionospheric disturbances on the bearings of incoming sky waves", AGARD-CP-115, paper 28.

RÖTTGER, J. and H. BECKER, 1977, "Die HF-CW-Dopplermethode und ihre Anwendung in der Ionosphärenforschung", Kleinheubacher Ber. 20 (to be published).

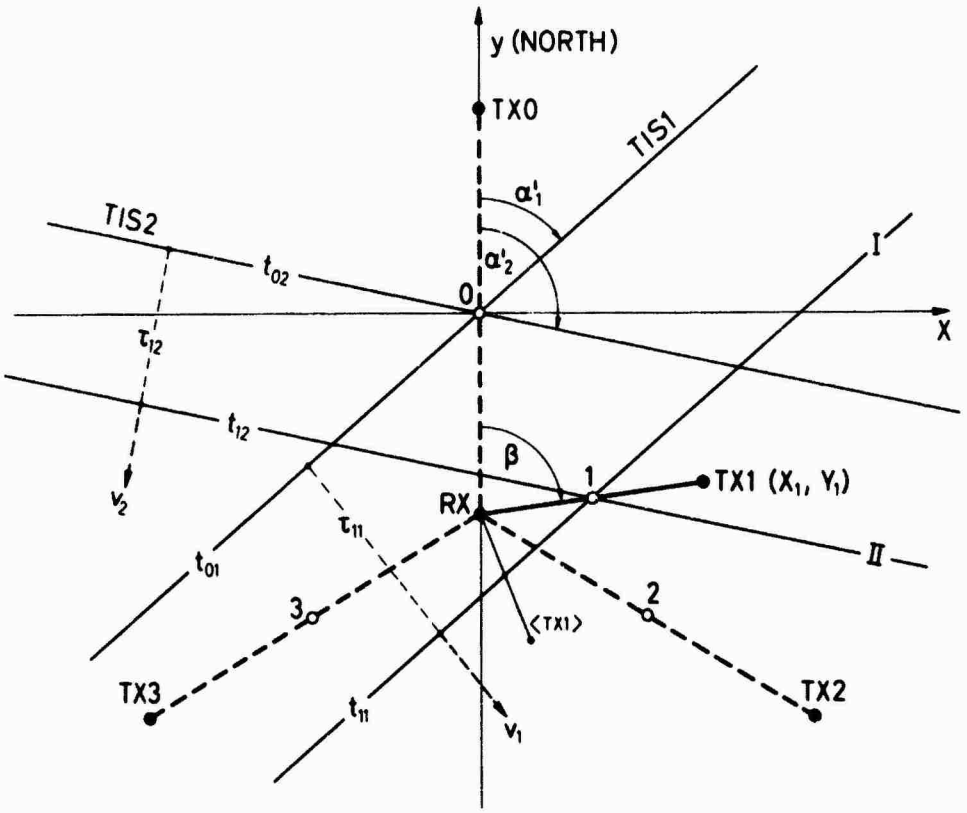


Fig. 1 Array of four transmitters TX0, TX1, TX2 and TX3 and one common receiver RX, located in the horizontal x,y plane. The origin of the coordinate system is the projection of the ionospheric reflection point of the path TX0–RX. The azimuthal propagation angle α of a traveling ionospheric structure TIS is $\alpha = \alpha' + 90^{\circ}$ measured with reference to the north-direction, where α' is the angle of the phase-front of a TIS. Further parameters are described in the text.

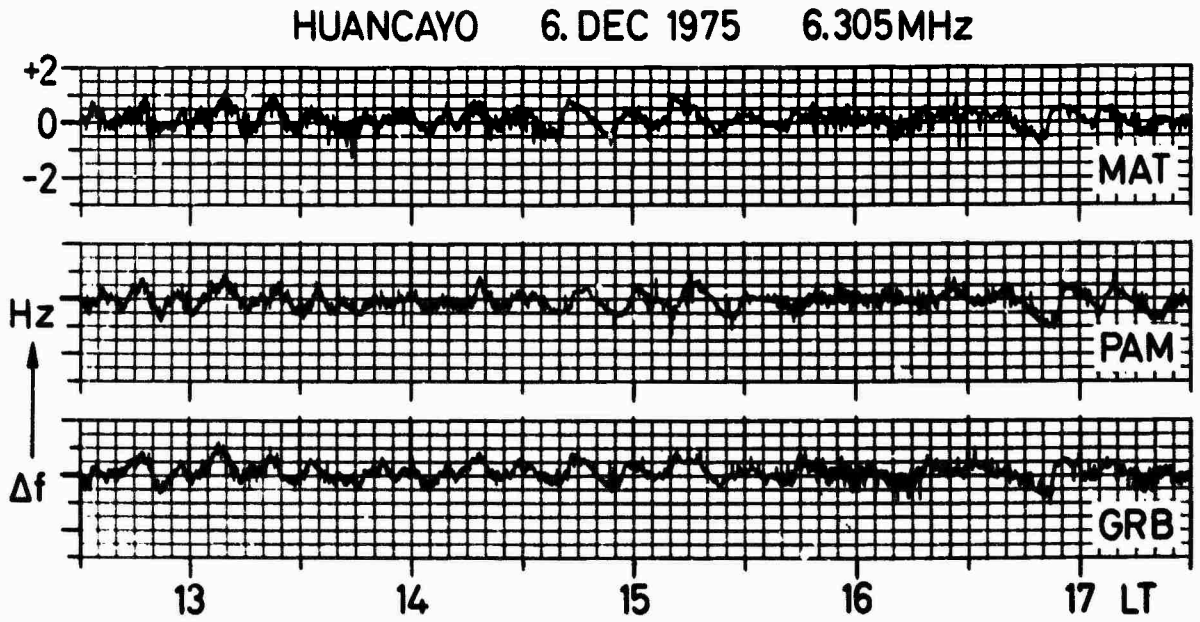


Fig. 2 HF-CW-Doppler records from the Huancayo network. The oscillations of the Doppler frequency Δf measured as a function of local time LT on the three paths MAT, PAM and GRB are caused by medium-scale traveling ionospheric disturbances propagating through the network.

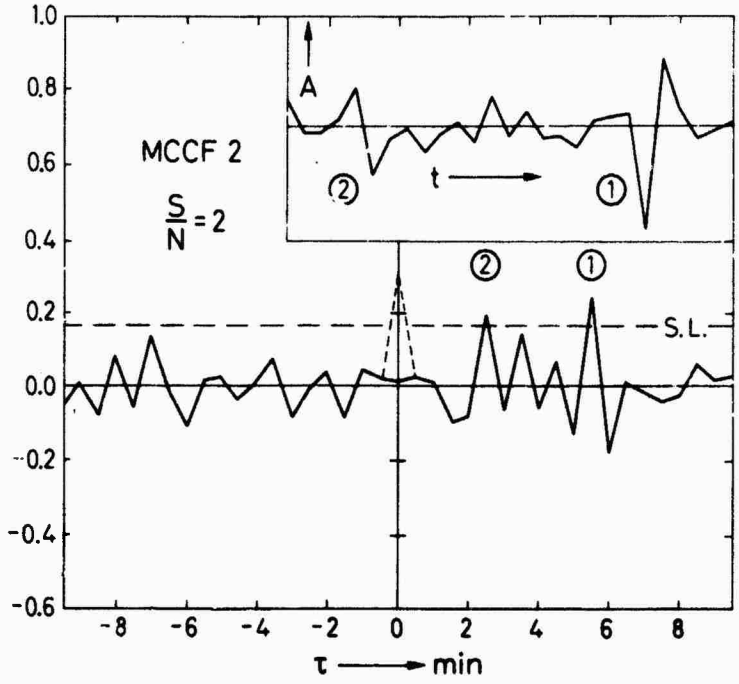


Fig. 3 Time series (upper right half) and modified cross-correlation function (elimination of the peak at $\tau = 0$) of double-pulse signals at a signal-to-noise ratio of two. The time scale t is equal to the scale of the time shift τ.

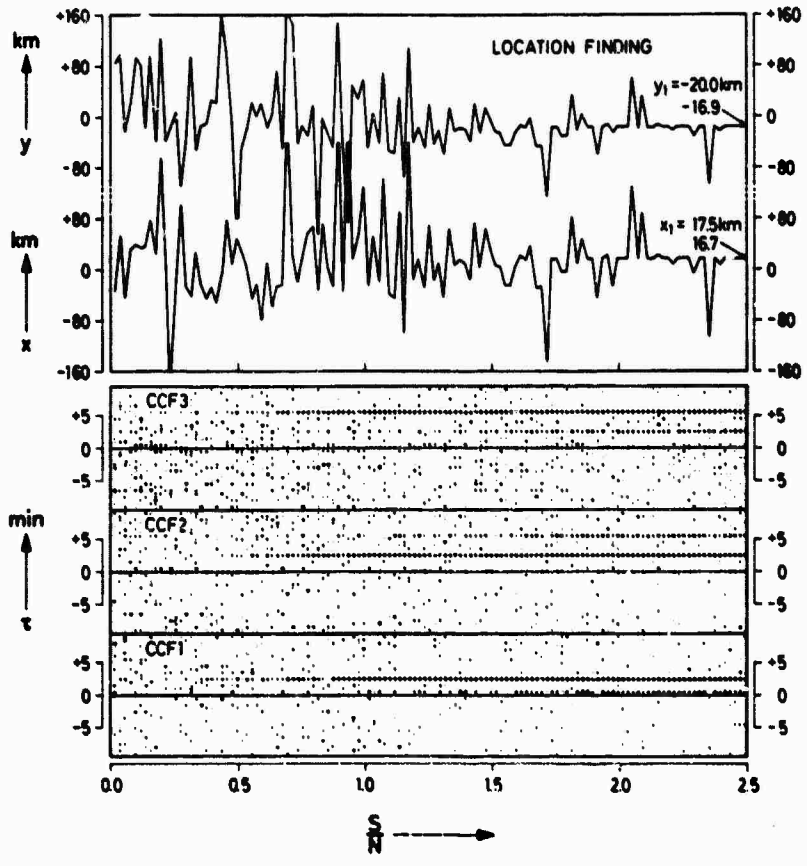


Fig. 4 Location-finding by means of pulse-like TISs (parameters are given in the text).
In the lower part of the graph the cross-correlation functions are printed in a three-dimensional density display as a function of the signal-to-noise ratio S/N.
From the set of the three cross-correlation functions the coordinates x_1, y_1 of the unknown ionospheric reflection point are calculated and plotted in the upper diagram.
Each sample of S/N is from one different set of random numbers which introduces the statistical scatter of the calculated results.

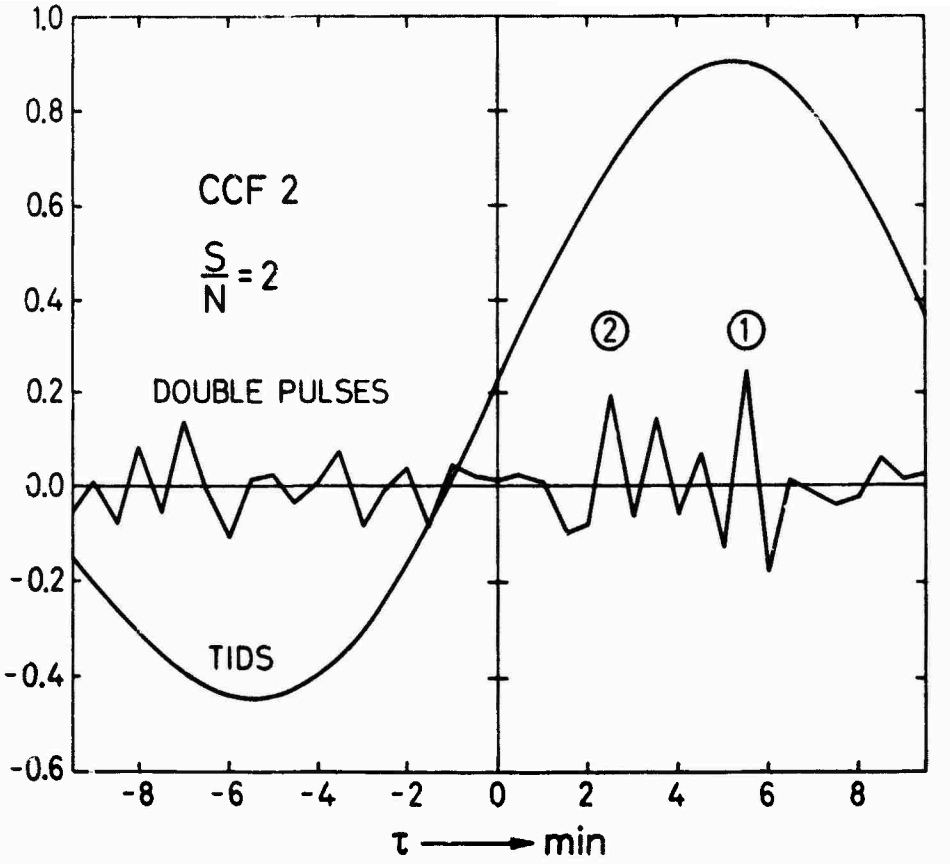


Fig. 5 Cross-correlation functions of double-pulse TISs and TIDs.

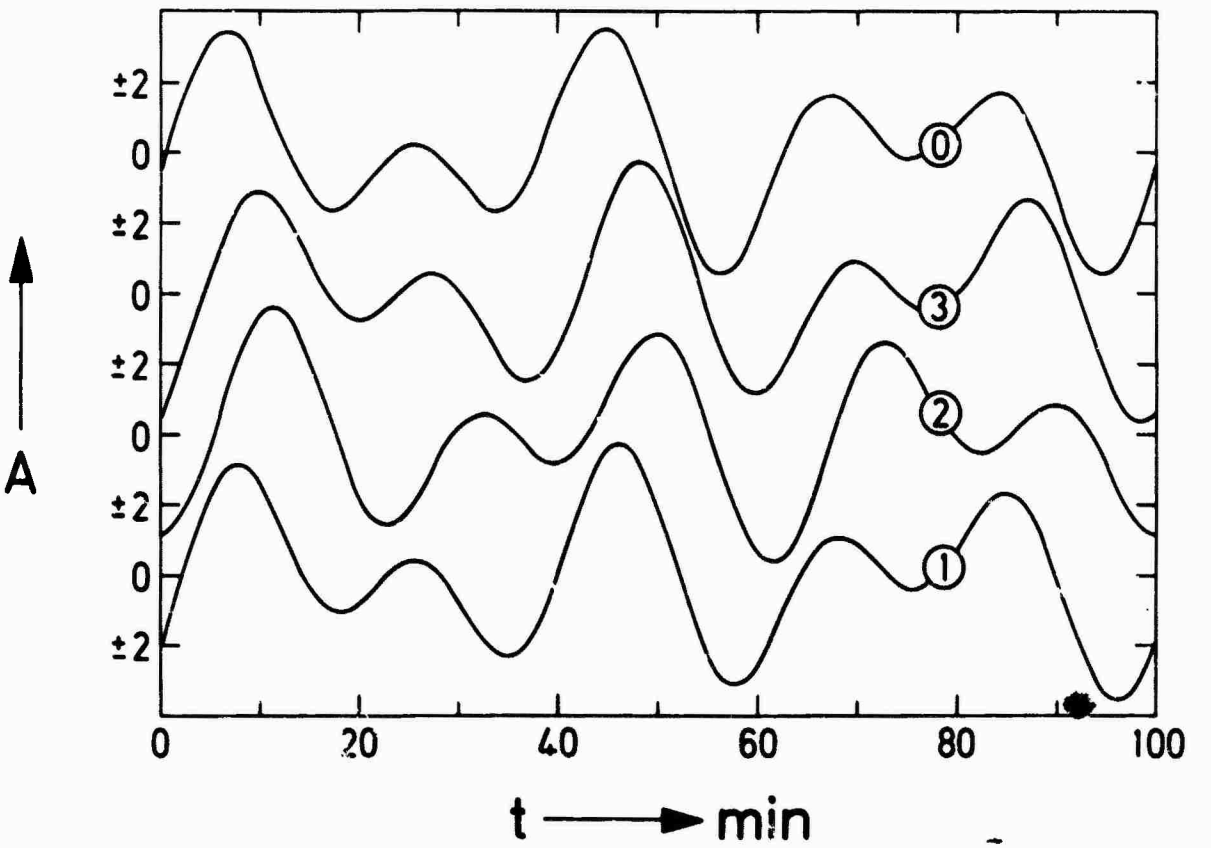


Fig. 6 Time series A(t) measured on the four paths 0,1,2 and 3. The phase-shifted oscillations are due to two interfering TIDs with the periods $T_{TID\,1}$ = 20 min and $T_{TID\,2}$ = 33 min.

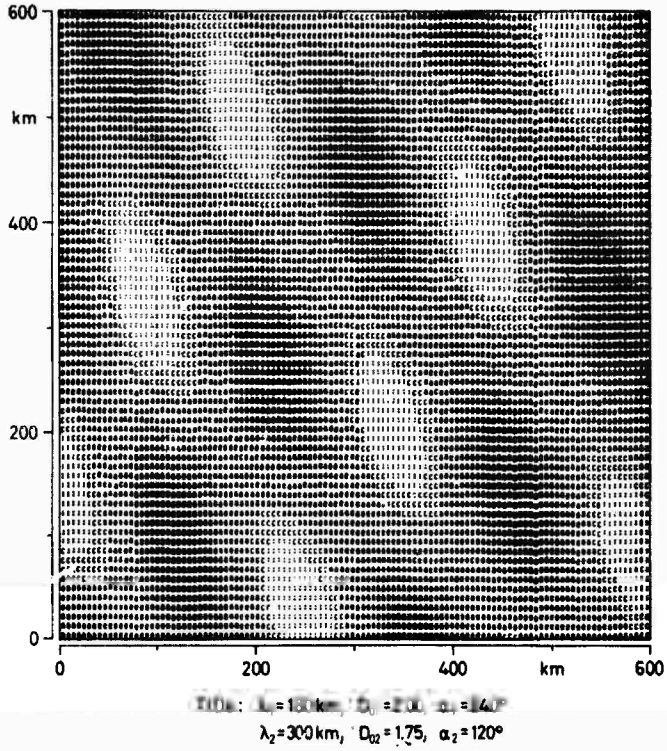


Fig. 7 Density display showing the ionospheric corrugation due to two spatially interfering TIDs with different wavelength, amplitude and propagation direction.

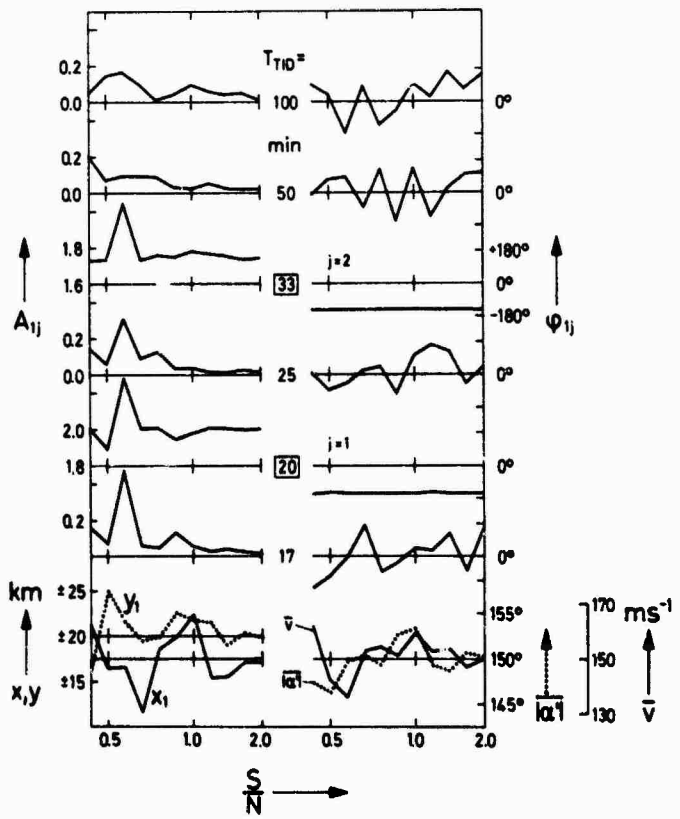


Fig. 8 Result of a spectrum analysis showing the amplitudes A_{1j} and phases φ_{1j} at the periods T_{TID} = 100 to 17 minutes as a function of the signal-to-noise ratio S/N.
From the significant amplitudes and phases observed at T_{TID} = 20 and 33 min, the velocity $\bar{v}$ and the phase-front angle $|\alpha'|$ as well as the location x_1, y_1 of the unknown ionospheric reflection point are calculated.
Each sample of S/N is from one different set of random numbers which introduces the statistical scatter of the calculated results.

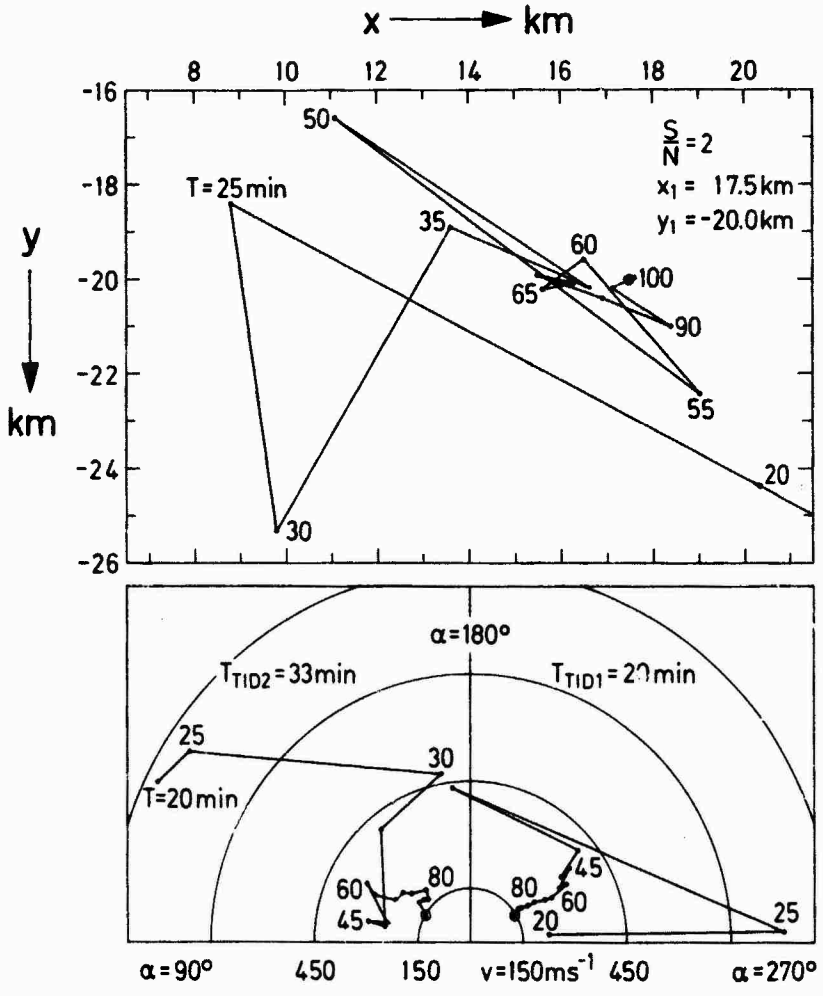


Fig. 9 Determination of TID velocity v and azimuth α (lower diagram: $\alpha_1 = 240^\circ$, $\alpha_2 = 120^\circ$, $v_{1,2} = 150$ m s^{-1}) and the location of the unknown ionospheric reflection point (upper diagram: $x_1 = 17.5$ km, $y_1 = -20.0$ km) for various integration times T.
Each sample of the time series truncated at T is from one different set of random numbers which introduces the statistical scatter of the calculated results.

DISCUSSION

T.B.Jones, Physics Department, University of Leicester, UK

(1) At night TID's are large and you have good correlation over the array; however, the time differences measured are very small and difficult to measure accurately. Would this lead to errors in your position finding system?

(2) During daytime the TID are small and greater time delays are obtained. However, UK results indicate considerable dispersion of the TID even over small distances of about 40 km. This limits the usefulness of the 3 station techniques for DF work and TID studies.

(3) The use of the 'cross correlation' technique is not very satisfactory and does not lead to accurate measurement of the time displacement (t). A Fourier analysis appears to be much better for determining t but it does tend to emphasize the higher frequency components of the TID. Can one be certain that one has included enough of the 'record' to obtain a time spectrum of the TID?

Author's Reply

(1) The introduced method can be applied either to non-periodic or to periodic disturbances. To reduce the error in position determination, we have to assure that the time of measurement is comparable to the characteristic time of the disturbance, i.e. the time a non-periodic TIS needs to travel through the network or the period of a TID.

(2) A further condition for a high accuracy of the system is that the temporal as well as the spatial coherence of the structures is sufficiently great, i.e. the TISs/TIDs should not decorrelate over the distance of the network. TIDs are observed to correlate rather properly over distances of some ten kilometers, which appear to be the optimum dimensions of the position finding network. This incidentally is an advantage of the system operating more reliably at steep than at low incidence angles.

(3) As it is pointed out in the paper, in case of periodic disturbances the Fourier analysis is much more accurate than the cross-correlation technique and should be applied much more often than up to now (in many cases it even seems to be better than a power spectrum analysis). Proper fitting of the integration interval to the periodic TIDs as well as reliable significance tests surely have to be applied in order to deduce all information from the data without preference of some spectral components.

A Simple Multipath Error Reduction Method for Single Site DF Systems

M. Böhm
Standard Elektrik Lorenz AG (ITT)
Stuttgart, Germany

SUMMARY

Multipath reduction for precise df systems is performed using a variety of methods. Besides frequency diversity and Doppler schemes, space diversity schemes as well as directional diagrams have found wide applications.

It is well known that using wide apertures is one of the best methods to reduce multipath errors. However, low frequency systems that also are required to be mobile are not very suited for wide apertures.

The paper describes a simple method of reducing multipath errors by aperture sampling using at least three antenna elements. This method is derived from a simple error model that is independent from frequency. At each of the three elements phase and amplitude of the electromagnetic field are sampled simultaneously, which is made possible by providing each antenna element with its own receiver. All receivers, however, are using the same local oscillator to be coherent.

The phase/amplitude pairs are digitized and fed into a computer which basically - in the case of one reflector only - has to solve the equation

$$A_1 \cos(\Delta\alpha_1 - \Delta\beta) - A_3 \cos(\Delta\alpha_2 - \Delta\beta) = \frac{A_3^2 - A_1^2}{2 A_2}$$

A_1, A_2 and A_3 are the amplitudes of the field at the three antenna elements, $\Delta\alpha_1$ is the measured, distorted phase difference between the first and second element, $\Delta\alpha_2$ is the measured, distorted phase difference between the second and third antenna element. $\Delta\beta$ is the undistorted phase difference between two neighboured antenna elements and leads to the undistorted angle of incidence φ by using the formula

$$\varphi = \arcsin \frac{\lambda \Delta\beta}{2\pi d}$$

where λ is the wavelength and d the distance between two neighboured elements.

The basic method is applicable to linear as well as circular arrays. In order to evaluate the method an L-band test system was built, featuring a five λ linear array.

Test results are presented and discussed, as well as potential and limitations of the method. The paper concludes with some indications of further refinements of the basic method.

Introduction

DF systems have been an important tool for military surveillance since decades in all branches of the armed forces in many nations. Since the first patents were granted to radio df proposals at the beginning of this century, numerous designs have been used, one of the most popular being the Adcock type equipment.

As accuracy requirements increased, however, the limitations of Adcock df equipment in a multipath environment led to widebase systems, one of the best-known being circular Doppler df developed in World War II. While this approach provides good results with apertures of 5 - 10 λ its application is difficult for mobile systems in the VHF and lower frequency bands because of too large mechanical dimensions. There is also a problem for multipath situations with a small spatial angle between direct and reflected signals. In this case the simple phase measurement technique requires intolerably large apertures.

The purpose of this paper is to discuss a method of signal processing that provides sufficient accuracy while requiring considerably less aperture than conventional processing schemes. The basic approach relies on aperture sampling techniques as sketched in Fig. 1. This technique, which is increasingly also being advised for low flyer detection radar, provides more information than a conventional sum-difference pattern of the same aperture. This additional information is the distribution of rf signal phase and amplitude over aperture.

To fully exploit this information, mathematical processing is required. Due to the technological progress of digital processors, be they mini or micro, simple, low cost equipment has become feasible. However, despite the performance of modern processors, one has to search for simple algorithms in order to keep the processor effort relatively low especially when it comes to high speed applications (measurement time a fraction of one second). The algorithms discussed in this paper are very simple and flexible, as will be shown in the following sections.

Error Generation Model

The basic configuration which generates multipath errors is sketched in Fig. 2. Assuming a 2-element interferometer with a baseline d one can determine the direction of arrival φ (DOA) of radio energy transmitted from a source by measuring the phase difference $\Delta\beta$ of the 2 signals arriving at the coherent receivers E1 und E2.

If a specular reflector is introduced there are two signals at each of the receivers. These signals are combined vectorially in the receivers. As their travelling paths are different ($b_1+c_1 \neq b_2+c_2$), the distorting signals S1 und S2 arrive at E1 and E2 with different phases, if not with different amplitudes, and therefore distort the phase difference $\Delta\beta$. As a result, the DOA measured contains an error.

For a very simple linear array comprising three antenna elements equidistantly spaced the received signals with their relative phase relationship are sketched in Fig. 3a, which is rather self-explanatory.

In order to extract the distorting phase component from this sketch, a normalization is performed by substracting the undistorted phase β from the measured phase α of the signals at each two neighboured antenna elements. The result is given in Fig. 3b. It is apparent that the distorting phase components Φ_i and Φ_{i+1} are different from each other. It is also apparent that the locus of the signal A resulting from the geometrical adding of direct signal N and multipath signal S is a circle in the simple case of specular reflection and one plain reflector only.

If there are several multipath signals from various reflectors in different directions the resulting locus of the normalized signal diagram is, of course, no longer a circle, but completely unpredictable, as is indicated in Fig. 4.

For the case of 2 multipath components arriving with a $+10^o$ and -10^o spatial offset from the direct signal, and having 0.7 and 0.2 resp. the amplitude of this signal, a phase plot for a 5λ linear array with 32 elements is given in Fig. 5. From this plot it is apparent that enormous phase errors result from employing small apertures. For example, if one measures the difference of phase between the elements 10 and 11 and between 17 and 18, these two results are completely unusable. However, if one measures phase difference between elements 1 and 32, one gets a much better result due to the averaging effect of the wider aperture.

This, of course, is standard knowledge, and wide aperture systems are being used extensively. These systems, however, provide mechanical problems at lower frequencies for mobile applications, as was mentioned earlier. It is therefore desirable to search for more efficient multipath error reduction methods than just wide aperture phase processing.

Basic Error Reduction Approach

Based on the normalized signal diagram sketched in Fig. 6 an error reduction approach has been evaluated that is based on a rather simple algorithm. This approach is easily understood for the case of a circular locus of the normalized signal diagrams, i.e. one multipath signal only.

Fig. 6 shows two possible antenna configurations, a 3-element and a 4-element array, with the related normalized signal diagrams. It can easily be seen that the 3-element case is a special one of the more general 4-element case.

If the antenna elements are spaced equidistantly as sketched, from the very geometry of the normalized signal diagram it is apparent that each diagram includes two distances or lines z of equal length. Therefore, in the case of the 3-element approach, the following is valid:

$$A_1^2+A_2^2 - 2A_1A_2 \cos(\Delta\alpha_1-\Delta\beta) = A_2^2+A_3^2 - 2A_2A_3 \cos(\Delta\alpha_2-\Delta\beta) \qquad (1)$$

A_1, A_2, A_3 are the measured signal amplitudes, $\Delta\alpha_1$ and $\Delta\alpha_2$ the measured phase differences, and $\Delta\beta$ is the undistorted phase necessary to determine the DOA. A similar algorithm can be derived in the case of the 4-element array.

Formula (1) leads to a quadratic equation, one solution of which is the desired $\Delta\beta$. The other solution refers to the DOA of the multipath component involved.

In the described case one always gets the exact DOA independent from aperture. In the real world, however, with its more complex multipath and receiver noise, there are limitations. In order to evaluate what the limitations are and how they might be neutralized, a test system was designed and evaluated. This test system, comprising a 5 λ array with 32 fixed positions of a movable antenna, was used to collect phase/amplitude values in a digital, storable form. These 32 pairs of data for each "scan" were used to process them using the 3-element or 4-element approach, as will be described later.

Test Set-Up

The test set-up that was used for collecting real, clearly defined signal samples of a given aperture is shown in Fig. 7. In a truck the 1 GHz, pulsed transmitter feeding 3 directional antennas was operated. Signals feeding the antennas could be varied in phase and amplitude. The 3 antenna elements were set equidistantly.

Another truck about 200 m away carried two receivers that were connected to two antennas, one of which was fixed and used as a reference, while the other was movable on a rail over a 5 λ range. The movable antenna was used to sample over aperture the signal resulting from the interference of the three signals transmitted from the 3 directional antennas of the first truck. The measurement results for a chosen interference situation were stored on a punched tape.

There were required several provisions to stabilize this test set-up and to avoid drifts during the measurement of one set of 32 signal samples that could be taken. The two most important were crystal-controlled transmitter and receivers, and a warm-up time of about 30 minutes before taking measurements.

The movable antenna could be locked in 32 positions, defined by precision holes in the rail. Thus each position was exactly reproducible at any time.

Fig. 8 gives a more detailed block diagram of transmitter and receivers. It should be emphasized that the receiver for the movable antenna had to be operated not with an automatic, but a manual gain control, in order to detect the amplitude variations of the interfered radio signal over aperture.

Of course, the receivers had to be carefully calibrated with regard to amplitude and phase response at various signal levels.

Test Results

The measured signal samples (32 phase/amplitude pairs for each "scan") were taken 3 each from equidistant positions and submitted to the algorithm (1) given before.

The phase/amplitude pairs were measured with 1 % error in amplitude and 1° error in phase. This led to partly significant errors of the computed result for a single 3-element array as soon as there were 2 multipath components. Therefore several results of different 3-element arrays were averaged using various methods.

Fig. 9 shows how the method of averaging influences final accuracy. Although this is not very satisfactory, it can lead to an improved performance of df systems with limited aperture compared to conventional systems using phase processing only.

No averaging was performed of results derived from different "scans". Thus time "diversity" offers an additional method of averaging, which is very familiar and effective, particularly when there are moving sources or multipath signals varying because of scatter effects, for example.

Application of Principle to Adcock DF

The 3-element and 4-element approach can easily be applied to Adcock df systems provided each antenna element of such a system is equipped with its own receiver, and the receivers are coherent. Assuming there is a 4-element Adcock group with middle antenna then the 3-element and the 4-element algorithm (see Fig. 5) can be applied twice each. Doing so one gets two sets of solutions the ratio of which is the DOA, as is shown below.

If 1, 2, 3, 4, m are the designations of the 5 antennas, then A_1, A_2, A_3, A_4, A_m are the measured signal amplitudes at each antenna. The measured phase differences are $\alpha_1-\alpha_2 = \alpha_{12}$, $\alpha_2-\alpha_3 = \alpha_{23}$, $\alpha_3-\alpha_4 = \alpha_{34}$, $\alpha_4-\alpha_1 = \alpha_{41}$. In addition the phase differences α_{1m}, α_{m3}, α_{2m}, α_{m4} can be measured (antennas 1-m-3 and 2-m-4 form a 3-element array each).

These measured data are used to compute respective phase differences $\Delta\beta_i$ via following formulae

$$A_1^2+A_m^2 - 2A_1A_m \cos(\alpha_{1m}-\Delta\beta_1) = A_3^2+A_m^2 - 2A_3A_m \cos(\alpha_{m3}-\Delta\beta_1) \quad (2)$$

$$A_2^2+A_m^2 - 2A_2A_m \cos(\alpha_{2m}-\Delta\beta_2) = A_4^2+A_m^2 - 2A_4A_m \cos(\alpha_{4m}-\Delta\beta_2) \quad (3)$$

$$A_1^2+A_2^2 - 2A_1A_2 \cos(\alpha_{12}-\Delta\beta_3) = A_3^2+A_4^2 - 2A_3A_4 \cos(\alpha_{43}-\Delta\beta_3) \quad (4)$$

$$A_1^2+A_4^2 - 2A_1A_4 \cos(\alpha_{14}-\Delta\beta_4) = A_2^2+A_3^2 - 2A_2A_3 \cos(\alpha_{23}-\Delta\beta_4) \quad (5)$$

These formulae provide $\Delta\beta_1$, $\Delta\beta_2$, $\Delta\beta_3$, $\Delta\beta_4$. The DOA is

$$\varphi = \text{arc tg}\,\frac{\Delta\beta_1}{\Delta\beta_2} = \text{arc tg}\left(\frac{\Delta\beta_3}{\Delta\beta_4}\right) + \pi/4 \quad (6)$$

The two results can be averaged to get the most probable value if they differ from each other.

Fig. 10 shows the block diagram of such an upgraded Adcock DF. The two most important features - besides the multiple coherent receivers - are the phase/amplitude measuring device and the minicomputer or - even better - microprocessor. Existing Adcock stations thus can be upgraded at reasonable cost to provide both conventional operator-based CRT display and unattended automatic, computer-based multipath resistent operation.

Modifications of Basic Approach

In order to find out whether other algorithms than the ones already discussed might provide improved results a variety of algorithms based on ratios and similarities of distances and triangles in the normalized signal diagram were evaluated. Though in the case of one specular reflector only there is no better than the 3- or 4-element algorithm as discussed, it appeared to be advantageous in a multiple interference environment to apply algorithms like the one given in Fig. 11. Based again on the measurable values A_1, A_2, A_3, α_{12}, α_{23} of a 3-element linear antenna array applying the equation

$$\frac{z_1^2}{z_2^2} = \frac{x^2 + \Delta A_1^2}{y^2 + \Delta A_2^2} \quad (7)$$

leads again to an "undistorted" phase difference β_{12}. Of course, this solution contains an error, as the equation above is only correct if, in Fig. 11, the shaded triangles are really similar ones. This, however, normally is not the case, though with every locus this can be assumed as an approximation if the points 1, 2 and 3 are close enough to each other. Anyway, the discussed approach never leads to an exact solution, but comprises an estimation only.

This estimation, however, can improve results by a factor of 2 to 3 compared to pure phase processing in a multipath environment, as can be seen in Fig. 12a and 12b. In these figures simulation results for a 17.5λ, 40 element linear array and a 10.6λ, 32 element circular array are shown, which give a comparison of phase processing vs. phase/amplitude processing using the estimator approach of Fig. 11.

The multipath environment is very similar in both cases (linear and circular array). Parameters are the initial phase and the DOA of the stronger (0.5) interfering signal, while the weaker (0.2) interfering signal always has the same initial phase and DOA. In Fig. 12a for the linear array there is shown no error between zero and 1° elevation. This is due to a program property and should not be considered as real.

It is apparent from both figures that better results are given by the estimation approach particularly at small differences of DOAs of direct and strong multipath signal. This property may be advantageous for low flyer detection as well as long distance, over-the-horizon DF.

Anyway the improvement potential of such estimator approach can be used either to increase locating accuracy or to reduce aperture.

It might be interesting to consider to more depth the relation between the approaches discussed here and the usual imagination of directional diagrams, when phase/amplitude processing is involved. At the first glimpse, at least, there is no apparent relation visible.

Conclusions

Most severe limitation of single site df systems is a multipath environment. While overcoming this limitation by use of wide apertures is a well-known technique, cost and mobility considerations often demand a different approach.

Sampling i.e. phase/amplitude signal measurement at discrete points of apertures, even of small ones, appears to provide a tool for more effective signal processing using mini- or microcomputers.

As one of a variety of possible processing approaches a normalized signal diagram was discussed that allows for new, simple processing algorithms, which can be implemented at very low hardware cost.

Using the discussed methods Adcock df upgrading and low flyer detection equipment may be two of a variety of advantageous applications.

References

1 Zur Eliminierung und Reduzierung reflexionsbedingter Winkelmeßfehler in Funkortungssystemen mit kleinen effektiven Meßbasen
M. Böhm; Thesis, TU Berlin, 1974

2 Low-Angle Radar Tracking in the Presence of Multipath
W.D. White; IEEE Transactions, Vol. AES-10, No. 6, Nov. 1974

3 Radar Multipath
D.K. Barton; Microwave Journal, April 1976, pages 35 ff

4 Array Aperture Sampling Technique for Multipath Compensation
F.G. Willwerth, J. Kupiec; Microwave Journal, Vol. 19, No. 6, June 1976, pages 37 ff

5 Keep Track of that Low-Flying Attack
P.R. Dax; Microwaves, April 1976, pages 36 ff

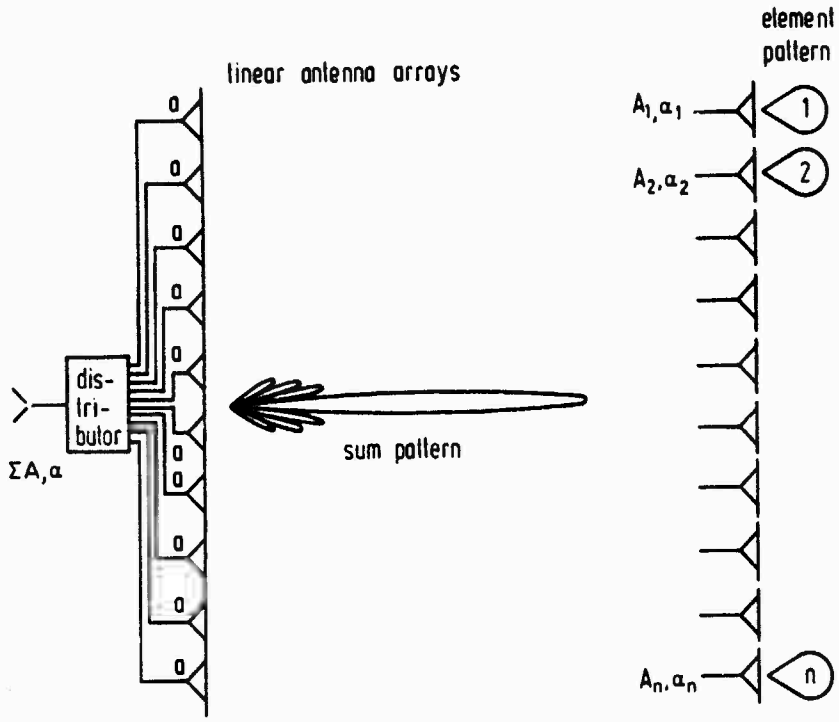


Fig.1 Aperture Sampling

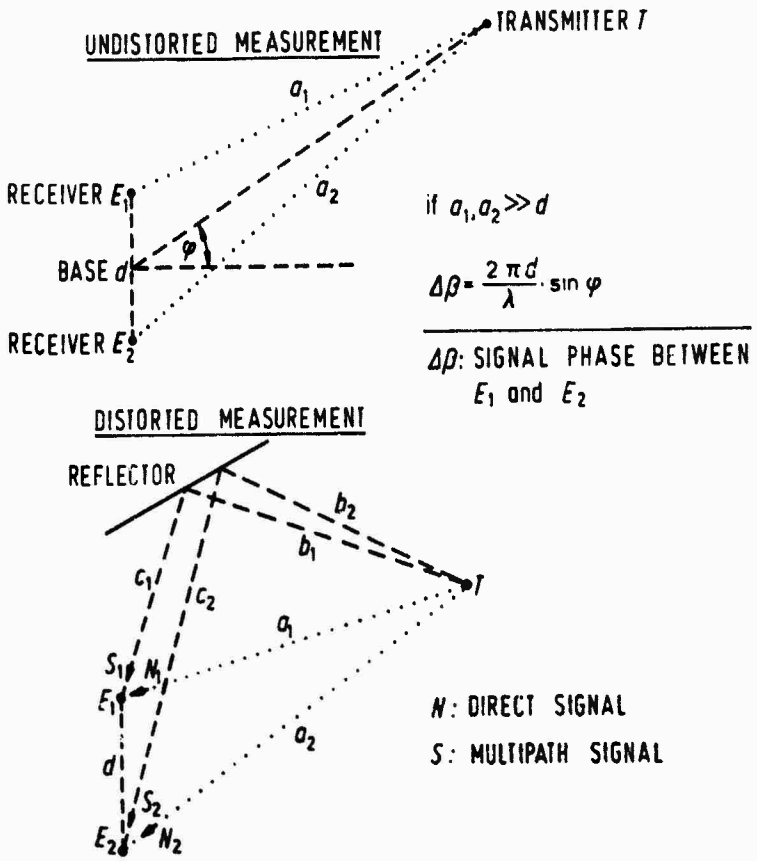


Fig. 2 Basic Generation of Multipath Errors in Interferometers

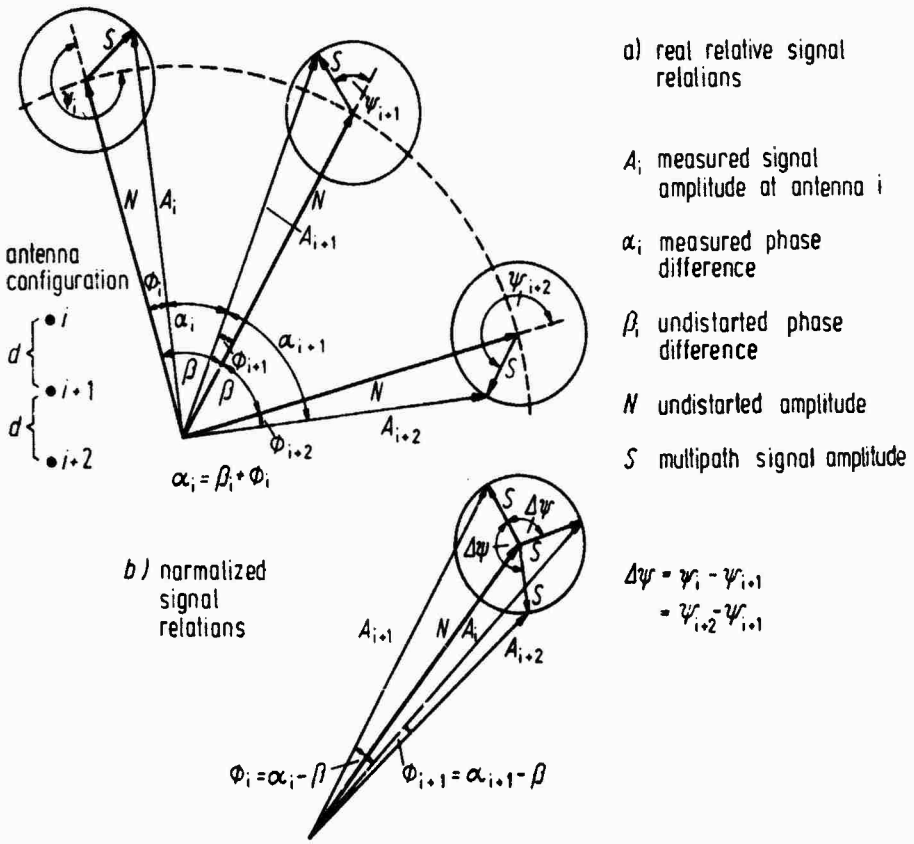


Fig. 3 Signal Vector Diagram

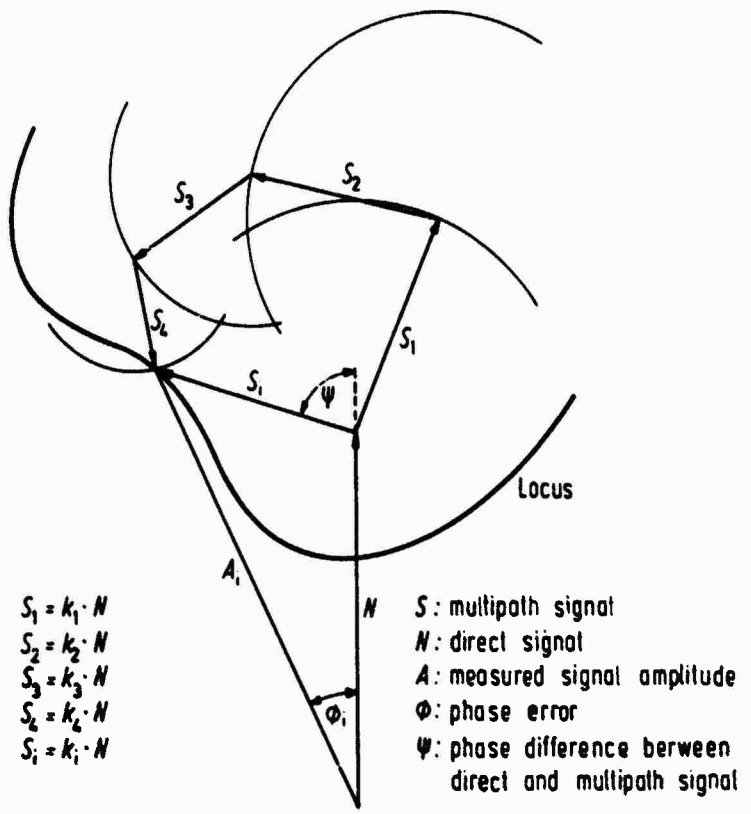


Fig. 4 Normalized Locus With 4 Multipath Components

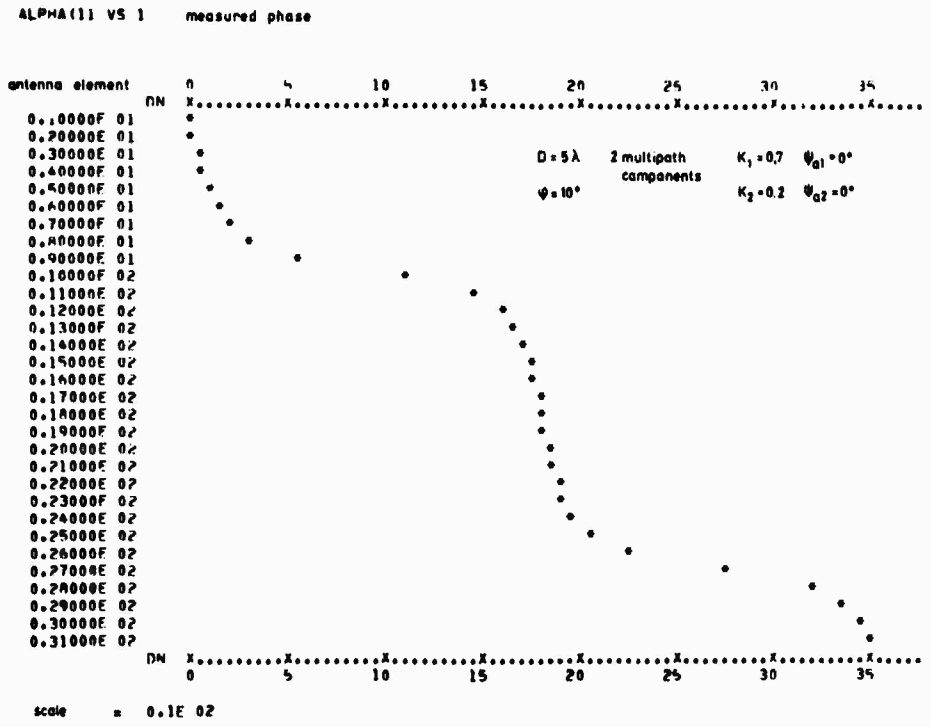
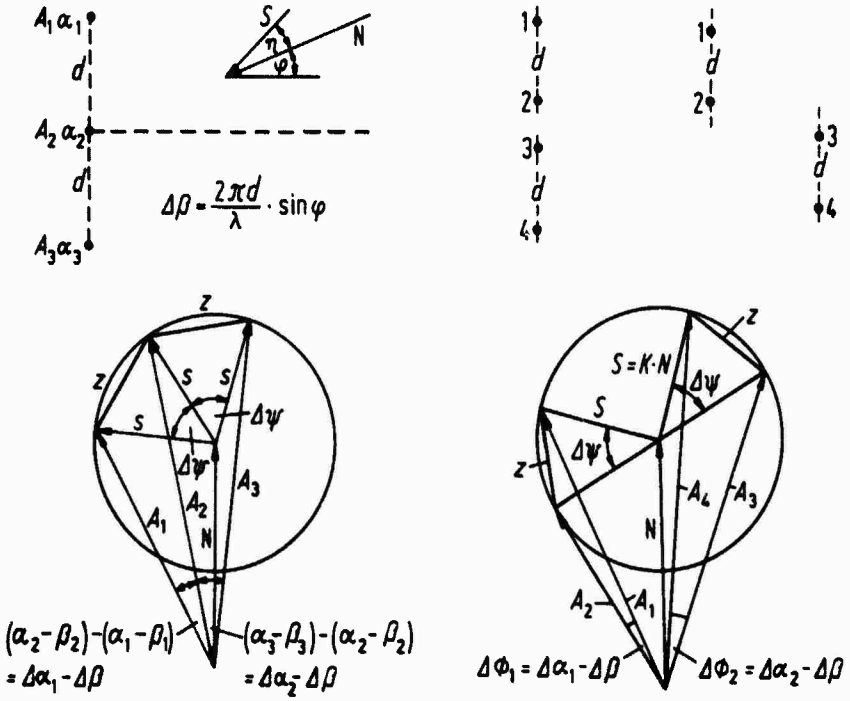


Fig. 5 Typical Phase Function Distorted By Multipath



$$\text{3-El.-APPR}: A_1^2 + A_2^2 - 2A_1A_2\cos(\Delta\alpha_1 - \Delta\beta) = A_2^2 + A_3^2 - 2A_2A_3\cos(\Delta\alpha_2 - \Delta\beta)$$

$$\text{4-El.-APPR}: A_1^2 + A_2^2 - 2A_1A_2\cos(\Delta\alpha_1 - \Delta\beta) = A_3^2 + A_4^2 - 2A_3A_4\cos(\Delta\alpha_2 - \Delta\beta)$$

Fig. 6 Basic Multipath Error Reduction Approach Using 3 or 4 Antenna Elements

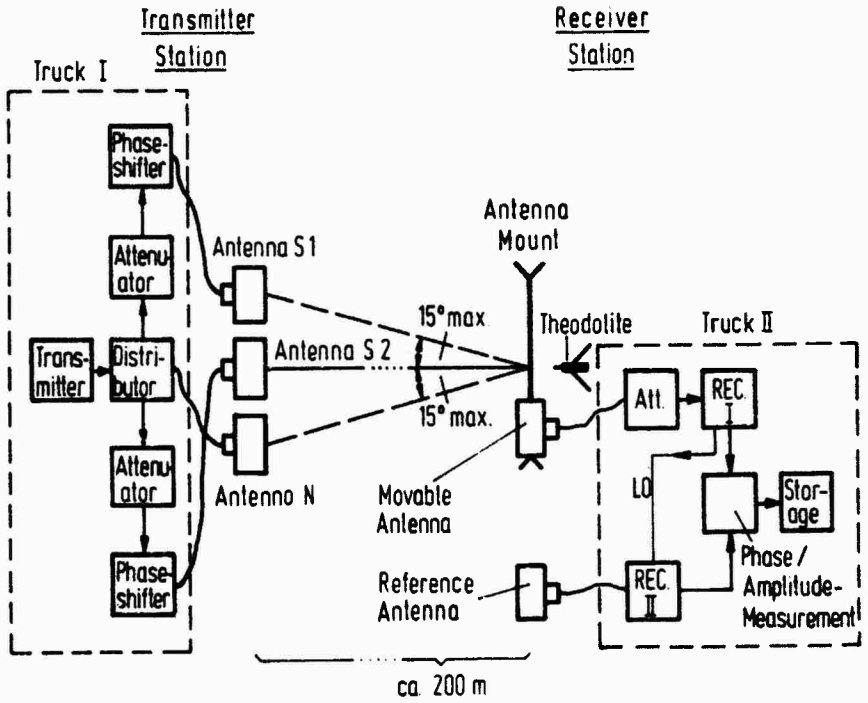


Fig. 7 Test Set-up For Linear Array

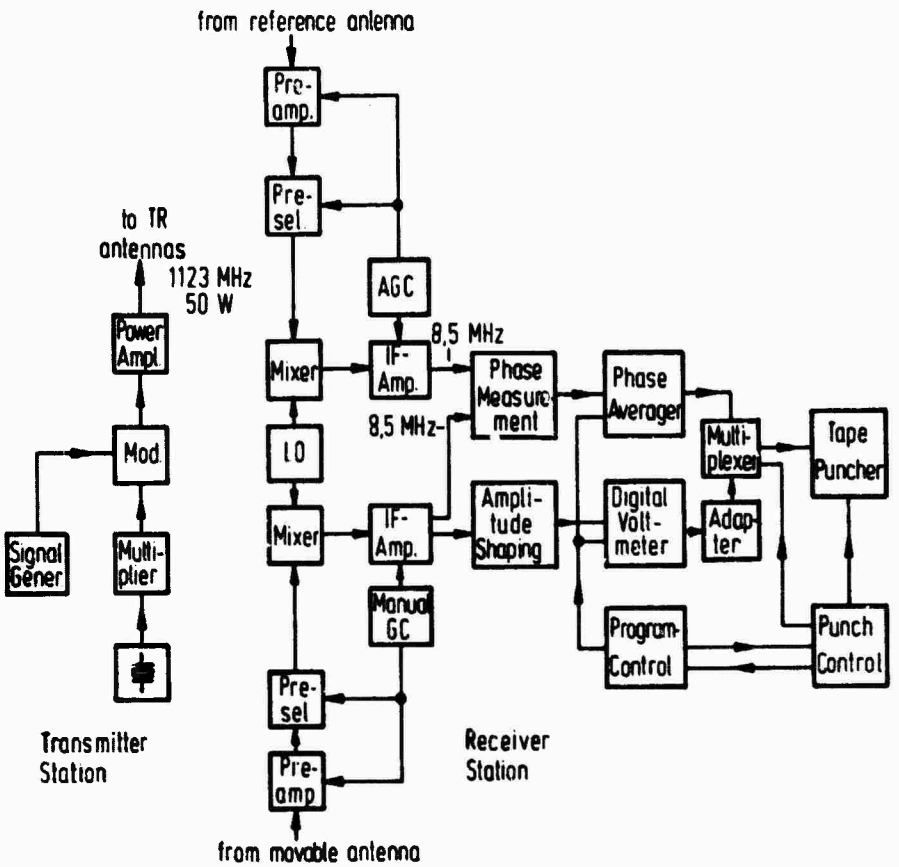
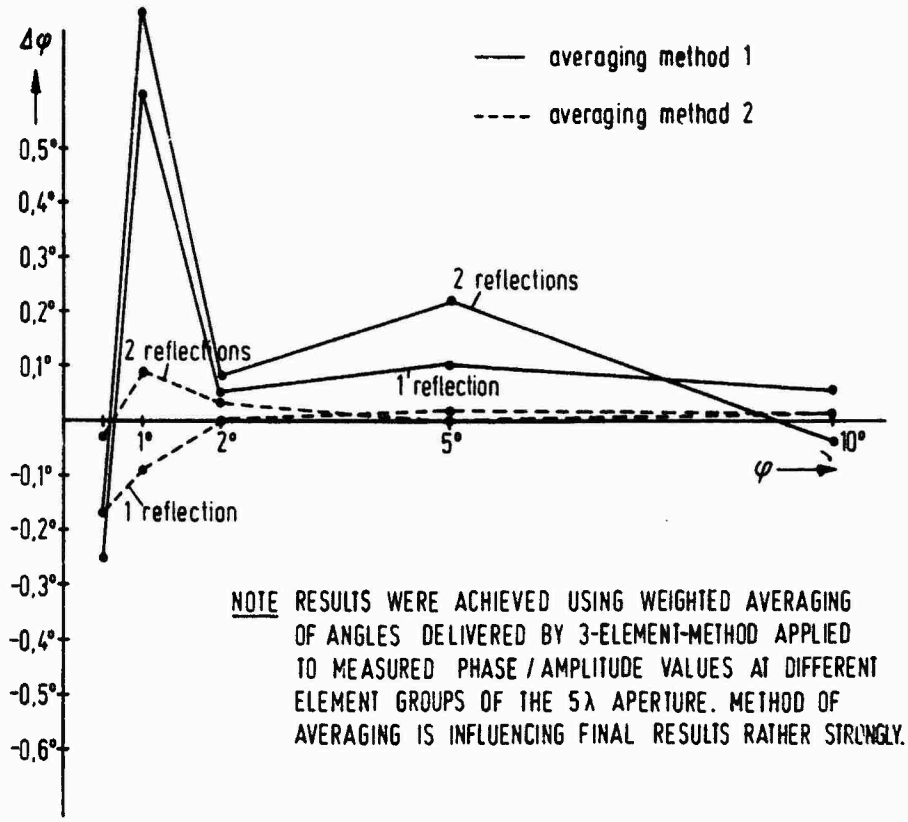


Fig. 8 Detail of Test Set-up



NOTE RESULTS WERE ACHIEVED USING WEIGHTED AVERAGING OF ANGLES DELIVERED BY 3-ELEMENT-METHOD APPLIED TO MEASURED PHASE / AMPLITUDE VALUES AT DIFFERENT ELEMENT GROUPS OF THE 5λ APERTURE. METHOD OF AVERAGING IS INFLUENCING FINAL RESULTS RATHER STRONGLY.

Fig. 9 Samples of Test Results

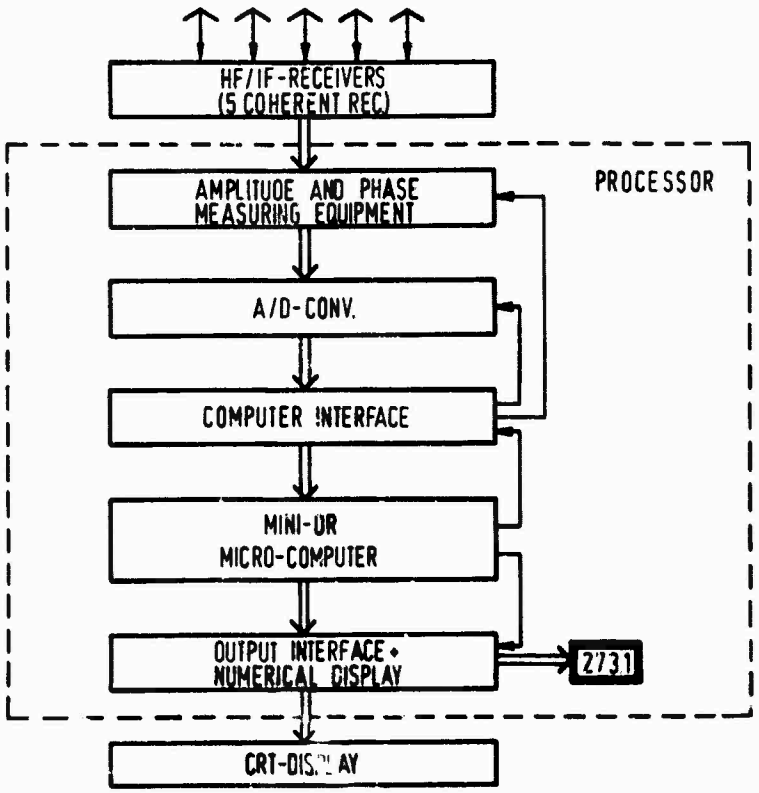


Fig. 10 Block Diagram of 3/4-Element Adcock DF

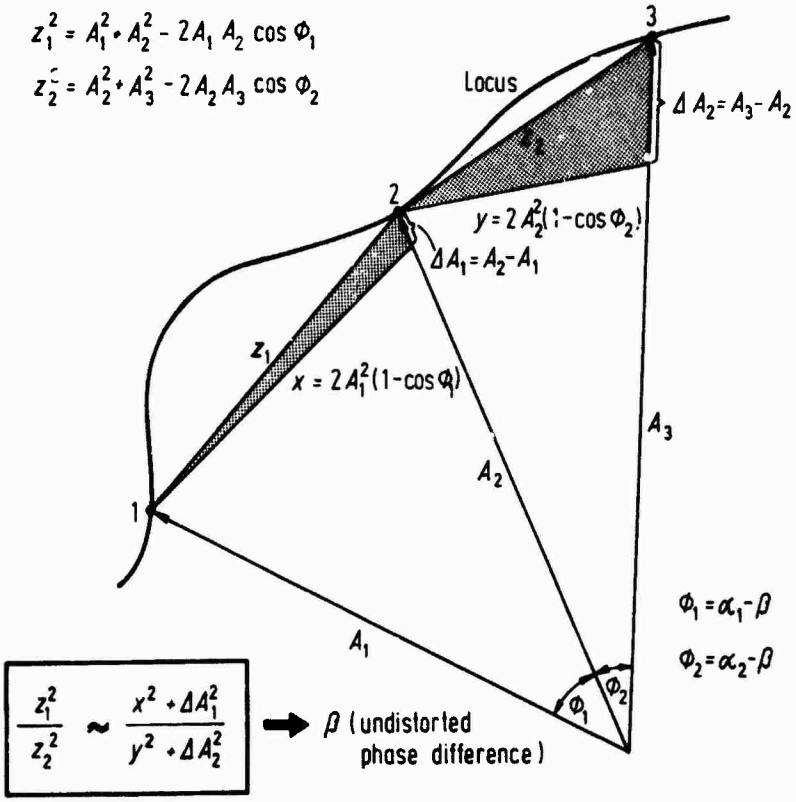


Fig. 11 Modified 3-Element Approach

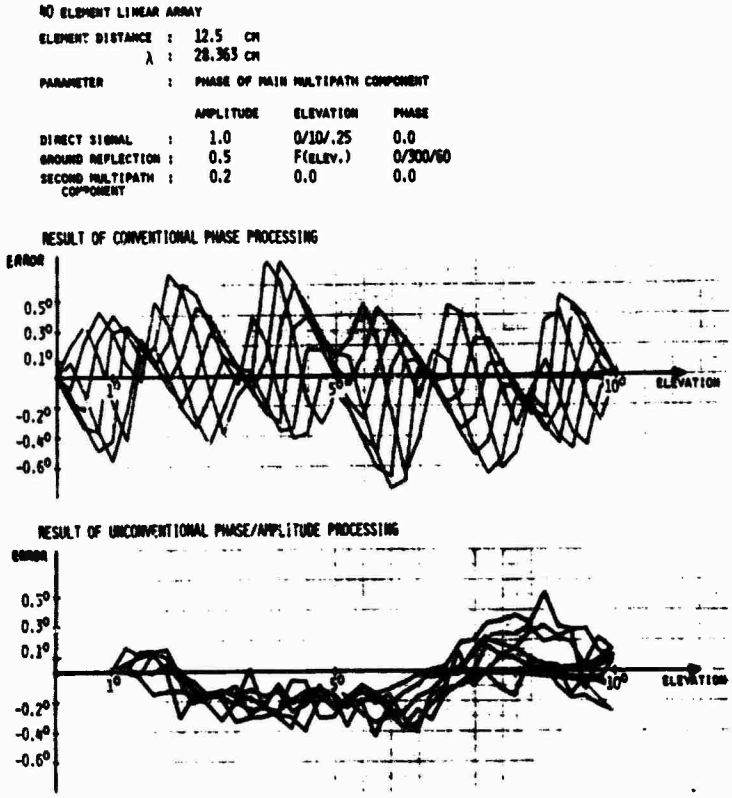


Fig. 12a Some Simulation Results

CIRCULAR 52 ELEMENT ARRAY

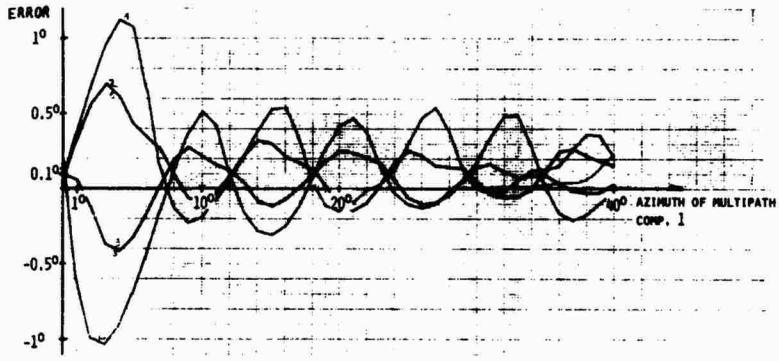
DIAMETER : 300 CM

λ : 28.363 CM

PARAMETER : PHASE OF MULTIPATH COMPONENT 1

	AMPLITUDE	AZIMUTH	PHASE
DIRECT SIGNAL .	1.0	0.0	0.0
MULTIPATH COMP. 1 :	0.5	0/90/1	0.60,120
MULTIPATH COMP. 2 :	0.2	25.0	0.0

RESULT OF CONVENTIONAL PHASE PROCESSING



RESULT OF UNCONVENTIONAL PHASE/AMPLITUDE PROCESSING

Fig. 12 b Some Simulation Results

SINGLE FREQUENCY USE OF THE NAVY NAVIGATIONAL SATELLITE SYSTEM

by

Abraham Shuval and Jonathan Mass
National Committee for Space Research
Radio Observatory, P.O.B. 4655
Haifa, Israel

ABSTRACT

In the US Navy Navigational Satellite System (NNSS) simultaneous transmission and reception on two frequencies is normally used in order to overcome the navigation error caused by the ionosphere.

The navigation error caused by the ionosphere while using one frequency only is investigated.

Several ways to reduce the navigation errors, while receiving on one frequency only, are discussed.

It is shown that one frequency only (400 MHz) could be used if a small increase in average time between navigation fixes and an accuracy of about 200 m are tolerable.

1. INTRODUCTION

The use of satellites for world-wide accurate navigation especially for ships is by now well established.

The US Navy is now operating a system of 6 satellites in low altitude polar orbits – the US Navy Navigational Satellite System – NNSS[1].

These satellites transmit continuous radio transmissions on about 150 and 400 MHz, and also time signals and accurate information on their orbital parameters.

A user on earth measures the Doppler shifts of these two frequencies and from these, together with the transmitted information, can determine his position on earth to an accuracy of 20–40 meters (for a stationary user)[2].

The ionosphere is one of the sources of errors in the NNSS – due to the influence of it's refractive index on the phase velocity of the radio waves.

The method used to correct for the first order ionospheric error is the two-frequency method based on the fact that the Doppler shift distortion caused by the ionosphere is a function of frequency, thus allowing to eliminate that distortion by comparing the Doppler shift curves at two frequencies[1,3].

In many cases however, the extreme accuracy of the system is not needed and cheaper receiving equipment (one frequency only) might be used. The navigation error caused by the ionosphere while receiving one frequency only will be calculated, and several ways to decrease this error will be discussed. These include mainly using a model ionosphere and relative navigation.

It is shown that one could use one frequency only together with a model ionosphere, or use a mode of relative navigation, and achieve navigation accuracy of about 200 m if one could tolerate a small increase in the average time between passes.

2. THE NNSS

This navigation system consists today (1975) of 6 satellites in an almost circular polar orbit, about 1000 km height. Each satellite transmits a CW wave on about 150 and 400 MHz. Also, each satellite transmits it's orbital parameters and time signals by phase modulation. The orbital parameters transmitted are taken from a memory inside the satellite.

The information into the memory comes from an Injection Station which receives the orbital data from a computing center, and transmits it to the satellite once every 12–14 hours. The computing center receives the results of measurements of several tracking stations around the earth, and from these data computes the current orbital data and predicts the orbit 12–16 hours ahead. This information is transmitted to the Injection Station and from it to the satellite. The Injection Station also corrects the satellite's clock which controls the time signals (every 2 minutes). For obtaining a complete navigation fix, reception of one satellite only is needed.

Figure 1 illustrates the passage of a satellite transmitting a frequency f_T above a receiving station on earth.

The satellite is in an almost circular polar orbit, about 1000 km height, Points 1; 2; 3; 4; 5 are the points where the satellite had been at times t_1; t_2; t_3; t_4; t_5 correspondingly.

S_i – are the distances between satellite and station at the times t_i, and the angles i_i are the angles between the ray's direction and the vertical at the "center" of the ionosphere, i.e., the zenith angle of the satellite at the "center" of the ionosphere (the "center" of the ionosphere is usually between 350 to 400 km height, as will be mentioned later).

The navigating station receives the satellite's transmissions, decodes the orbital data and performs what is called the Doppler Integral, i.e., counting cycles of the received frequency between each two successive time marks. This is better explained in Figure 2.

In the figure, a typical curve of the received frequency f_r as a function of time is shown. The received frequency f_r is shifted from the transmitted frequency f_T due to the Doppler effect (dotted line). The influence of the ionosphere is shown by the solid line and is manifested by the fact that the slope of the curve is smaller, which means that the satellite seems to be at a larger distance from the station than it actually is.

The actual instrumentation on the ground beats the received signal with a locally produced signal of frequency f_G, and counts the number of cycles of the frequency of difference $(f_G - f_r)$ between each two successive time points. That is – the number of cycles counted between the times t_1 and t_2 – N_{12} is given by

$$N_{12} = \int_{t_1}^{t_2} (f_G - f_r)dt \qquad (1)$$

(ignoring at this time the influence of the ionosphere).

N_{12} corresponds to A_{12} in Figure 2, which in turn is the complementary of B_{12} to the area $(f_G - f_r)\cdot(t_2 - t_1)$.

B_{12} is the Doppler integral and is equal to the distance difference $(S_2 - S_1)$ expressed in wavelengths (as will be shown later).

The navigation procedure requires the measurement of a number of areas – A_{ij}. The orbital data and time information (from which the satellite's position at time t_i can be determined) are given to a computer optimization program which varies the station's coordinates until a best fit (in the least square sense) to the cycle counts – $N_{i(i+1)}$ is achieved. The result of the optimization is of course the coordinates of the station.

The measured accuracy of the system for a *stationary* station is quoted as 20–40 meters[2]. This means that all sources of navigation errors – uncertainty of satellite position, influence of ionosphere, instrumentation noise errors etc., have been eliminated almost completely.

In many cases however, the extreme accuracy of the system is not needed and one is ready to trade off between accuracy and price, i.e., one is ready to accept a somewhat lower accuracy system at lower prices.

The costly parts of the equipment are, of course, the dual frequency receiver and the computer.

Hence, using one frequency receiver (if possible) and other means of calculations could reduce the price of the equipment needed for navigation.

The practical solution for the computer problem is to perform the calculations in a computing center which can serve a whole area. The computing center receives the results of the measurements from the navigating stations by radio telephone or any other means of communication, performs the calculations and sends this information back.

Reception on one frequency only means not correcting the influence of ionosphere and therefore introducing a navigation error.

This work deals with computation of the navigation error due to the uncorrected influence of the ionosphere and some ways to overcome this source of navigation error.

3. INFLUENCE OF THE IONOSPHERE ON NNSS TRANSMISSIONS

The frequency received on the ground f_r is given by

$$f_r = f_T + \Delta f \tag{2}$$

where Δf is the Doppler frequency shift introduced by the relative motion of the satellite to the navigating station

$$\Delta f = -\frac{f_T}{c} \cdot \frac{d}{dt} \int_s n ds \tag{3}$$

n is the refractive index of the ionosphere, c the velocity of light in vacuum. The integral is the optical path length of the ray from the satellite to the station, and would be equal to the geometrical distance S (Fig.1) if no ionosphere were present, e.g., if $n = 1$ all along the path. It is assumed when differentiating the integral that the satellite's movement is negligible during the time it takes the transmission to cover the distance from satellite to ground. Were it not for the ionosphere, Δf would measure the relative velocity between satellite and ground.

Using the very frequency approximation of the Appleton-Hartree formula for the refractive index as

$$n = 1 - \frac{e^2 N_e}{8\pi\epsilon_0 \cdot m f_r^2} = 1 - aNe \tag{4}$$

where N_e is the density of free electrons in the ionosphere, e – electron charge, m – electron mass and ϵ_0 – the permittivity of space. From Equation (4) it can be seen that a is a constant proportional to f_r^{-2}. Substituting into (3), and then substituting the result into (2) and then to (1) one gets[4]

$$N_{12} = (f_G - f_r)(t_2 - t_1) + \frac{ft}{c}(S_2 - S_1) - \frac{f_T}{c} a \left[\int_{s_2} N_e ds - \int_{s_1} N_e ds \right]. \tag{5}$$

The first member of the right hand side is a constant because f_T, f_G, t_2 and t_1 are known.

The second member is the number of wavelengths in the geometrical difference between distances $(S_2 - S_1)$ and represents the "pure" Doppler effect.

The third member represents the influence of the ionosphere and is manifested by the difference between the two integrals. In order to calculate the integrals one substitutes into (5) the expression

$$ds = dh \sec i \tag{6}$$

where dh is an element of height, and i is the zenith angle of the ray at each point along the ray's path.

The result is

$$N_{12} = (f_G - f_T)(t_2 - t_1) + \frac{f_T}{c}(S_2 - S_1) - \frac{f_T}{c} \cdot a \cdot [\langle \sec i_2 \rangle I_2 - \langle \sec i_1 \rangle I_1] \tag{7}$$

where I is the Total Electron Content (TEC) and is given by the expression

$$I = \int_0^{hs} Ne ds \tag{8}$$

where hs is the satellite's height.

The TEC is therefore the number of free electrons in a column of 1 m^2 cross-section area and height of hs.

$\langle \sec i \rangle$ – is the value of sec i at the "center" of the ionosphere, usually taken at the height of 350 to 400 km. The assumption that sec i is constant and the "center" of the ionosphere is of course an approximation, and it introduces an error (depending on the profile of Ne as a function of height) that can be neglected[5].

The use of $n = 1 - aNe$ represents a drastic abbreviation of the full Appleton-Hartree formula[6] and is justified for frequencies above 100 MHz (References 4 and 5).

Another approximation tacitly used is the neglect of the bending in the ray's path due to refraction which would require corrections to Equation (5) which was based only on the change in phase velocity and assumed a straight line ray path from satellite to ground.

Tucker and Fannin[3] have shown that at the frequencies of interest – the bending of the ray may be neglected.

4. NAVIGATION ERROR IN THE NNSS WHILE RECEIVING ONE FREQUENCY

In order to understand the calculation of the navigation error, a little more information on the optimization program used in the calculation of position of the ground station is needed.

The user in the navigating station performs Doppler counts (integrals) and can write about $8 \div 9$ equations.

$$N_{i(i+1)} = (f_G - f_T)(t_{i+1} - t_i) + (f_T/c)(S_{i+1} - S_i) \qquad (9)$$
$$1 < i < 8, 9 .$$

Equation (9) is derived from Equation (7) by neglecting the influence of the ionosphere (as is actually done in the NNSS after the two frequency correction).

The left hand side of (9) (for all i) are the results of the Doppler counts during the times $(t_{i+1} - t_i)$. The time difference $(t_{i+1} - t_i)$ is a constant of the system controlled by the satellite's clock (120 sec.).

In every equation there are three unknowns – longitude and latitude of the ground station (as manifested in $S_{i+1} - S_i$ when the satellite's position is known from the information that it transmits) and the frequency difference $(f_G - f_T)$ (taking into account the drifts of the satellite's clock and of the local oscillator in the navigation receiver).

For the solution of these equations an initial solution, not too far from the real solution (up to about 100 km) is assumed. From this initial solution the right hand side of (9) can be calculated – $N^{*}_{i(i+1)}$, and the merit function – M is calculated.

$$M = \sum_i (N_{i(i+1)} - N^{*}_{i(i+1)})^2 . \qquad (10)$$

The optimization program changes the three variables until M is minimized.

The optimization procedure depends on the computer available and the required time to obtain results.

For the calculations of the navigation errors a computer program was written to determine the Doppler frequency counts of a satellite's pass when the coordinates of the station, frequency of transmission and the 6 orbital elements describing the satellite's orbit are known.

The orbital elements used were those given together with their derivatives by NASA, and published in the "Five Line Orbital Elements Sets" as a regular service.

Another input to the program is the TEC. The total ionospheric electron content has been studied quite extensively since satellite techniques became available[7,8,9,10] and it's behaviour in low and medium latitude is well known. It shows a strong diurnal variation with about 10:1 day to night ratios, strong solar cycle variations ($\sim$ 5:1) and also strong seasonal and latitudinal dependence. Not only the TEC contributes to the navigation errors but also the horizontal changes in the TEC, usually called horizontal gradients. There are longitudinal changes which correspond fairly accurately to the changes with local time at a fixed location. The latitude changes depend mostly on solar zenith angle, but there is also a considerable influence to the earth's magnetic field which is asymmetric with respect to the equator.

The TEC itself and it's horizontal gradients – the north-south (N–S) gradient and the east-west (E–W) gradient are also given as inputs of the computer program.

The output of the program includes the satellite's coordinates as function of time, and the results of the Doppler integral that whould have been measured at the navigating station. The Doppler counts are calculated for two cases, with and without the influence of the ionosphere according to Equation (7).

If the results of the Doppler measurements that were calculated without the influence of the ionosphere, are given to the optimization program, then the result would be the real location of the navigating station, but if Doppler Integrals that were calculated while taking into account the influence of the ionosphere are given, then the results would be another location. The difference between this location and the real one is the navigation error.

It should be noted that for these calculations the optimization parameters were longitude and latitude of navigating station, while the frequency difference $(f_G - f_T)$ was assumed to be constant. This was done because only the influence of the ionosphere was sought.

These calculations were performed for one NNSS satellite (1968 – 012 A). It does not really matter which one because the orbits of the satellites are very similar. The results are shown in Figure 3 (Ref. 11).

In the figure, the total navigation errors caused by the ionosphere while receiving on one frequency only (400 MHz) are shown as a function of ground distance (ground distance is defined as the distance, on ground, from the navigating station to the satellite ground path at the closest approach point) and TEC as a parameter.

It can be seen that, the closer the station is to the satellite's ground track, and the higher is the TEC – the greater are the navigation errors.

For this particular case the horizontal gradients of TEC were assumed to be zero.

5. CORRECTION OF IONOSPHERIC ERRORS

If, for the practical reasons mentioned above, one prefers to use one frequency receivers only, then there are several ways (less powerful than the 2-frequency method) to reduce the ionospheric navigation error:

(a) Navigation using estimated ionospheric corrections, in which the ionospheric TEC is estimated using model ionospheres.

(b) Simple differential navigation, in which the navigation results are compared to those obtained at almost the same time at a fixed and known ground station, not too far away (up to approximately 500 km) from the navigating station. The *same* corrections in latitude and longitude are applied at the navigating station, as those required at the fixed and known station.

(c) Differential navigation, in which the received signals are retransmitted to a fixed and known master station, there to be compared to the locally received signals, and treated to obtain directly the position of the navigating station relative to the master station. A similar method was suggested, for instance, by Farkas[12], for 2-frequency reception. It does not seem to be practical for a simplified one frequency system.

(d) Relative navigation in which a fixed and known master station measures the ionospheric parameters influencing the satellite signal, estimates those applicable at the site of the navigating station for correcting it's ionospheric error.

In the first method (a), the error is due to the variability of the ionospheric TEC relative to the model. At large TEC values this is approximately ±25% (Ref.13). It can therefore be assumed that the total error due to the ionosphere is cut to about 25%. It can be seen on Figure 3 that the error is practically linear with TEC (see also Figures 11, 12 in Reference 14). In Figure 4 (Ref.15) the total ionospheric error at 400 MHz is plotted for a TEC of 4.10^{17} el/m^2 as function of ground distance – curve 1 (actually this is the curve for that particular TEC value in Figure 3, but drawn in a linear scale).

Curve 2 in Figure 4 represents the error reduced to 25% by the estimated ionospheric correction, as in paragraph (a).

The error in the simple differential navigation (para.(b)) is shown in Figure 4, curve 3. This curve is obtained simply by deducting the ionospheric error of the fixed station from the error of the navigating station on Figure 3 (this can be done since the navigation errors are mainly longitudinal[11,14], so that the simple deduction in one direction is a good approximation). It was assumed that the fixed station is 400 km closer to the satellite ground path than the navigating station. For smaller distances between the two stations the errors will be proportionally smaller.

Curve 3 begins to exceed even the total ionospheric errors at 700 km because then the *fixed* station gets too close to the satellite's path.

Curve 5 is the same as curve 3, shifted 400 km to the left and represents the error when the fixed station is 400 km farther away from the satellite ground path than the navigating station.

The error in relative navigation depends of course on the accuracy of the TEC measurement and also on the extrapolation of the TEC from the fixed station to the point relevant to the navigating station. There are several ways to measure the TEC, but for practical reasons it is better to measure the TEC with a dual frequency navigation receiver at the fixed station[11].

The measuring procedure is as follows: the fixed station receives the NNSS transmission on the two frequencies $f_{T1} = 399968$ KHz and $f_{T2} = 149988$ KHz, beats them with two ground frequencies $f_{G1} = 400\,000$ KHz and $f_{G2} = 150\,000$ KHz respectively, and performs *two* Doppler counts $-N^1_{i(i+1)}$ and $N^2_{i(i+1)}$ between successive time marks at the two frequencies.

The expressions for the Doppler integrals are the same as (7) – after inserting there the appropriate frequencies (not neglecting the – a term according to (4)). It can easily be seen that

$$N^1_{i(i+1)} - \frac{8}{3} N^2_{i(i+1)} = -\frac{55}{72} \cdot \frac{e^2}{\pi \epsilon_0 mcf_{T1}} \cdot [<\sec i_{i+1}> I_{i+1} - <\sec i_i> I_i] \,. \qquad (11)$$

For each measurement there are 3 unknowns – the TEC at the beginning and the end of the measurement interval (according to the geometry between satellite and station) and the height of the "center" of the ionosphere.

But, because the satellites are in polar orbits, they pass the observer's horizon only in a North-South (N–S) direction and the signals they transmit are influenced by the TEC and it's changes in that direction only.

Hence the number of unknowns can be reduced to 3 – namely – the TEC at the point relevant to the fixed station, it's north-south changes that can be represented by a linear N – S gradient (this holds for most cases, especially in mid-latitudes and relatively small distances – up to 500 km) and the height of the center of the ionosphere.

By the above mentioned procedure the several equations like (11) obtained at the fixed station can be solved by a suitable optimization procedure.

The east-west (E–W) changes can be represented by an E–W linear gradient that can be estimated from average graphs showing the diurnal variation of TEC as function of local time (one degree of longitude is equivalent to 4 minutes of time on the diurnal variation curve).

Estimating the accuracy of the TEC measurement by the 2-frequency method as 5% (Refs 8, 16), and a maximal error of 5% due to the N–S gradient extrapolation (gradient variability due to Travelling Ionospheric Disturbances and other sources[13,16], one gets for a TEC of $4 \cdot 10^{17}$ el/m^2 a remaining worst-case TEC error of $4 \cdot 10^{16}$ el/m^2.

A small error in estimating the E–W gradient is about $2 \cdot 10^{16}$ el/m^2. (11)

So, in the relative navigation mode, the uncorrected ionospheric TEC is about $6 \cdot 10^{16}$ el/m^2 (for TEC = $4 \cdot 10^{17}$ el/m^2). This residue of the TEC causes navigation error shown in Figure 3 (for TEC = $6 \cdot 10^{16}$ el/m^2). This curve is shown in a different scale in Figure 4, curve 4.

6. DISCUSSION

It can be seen from Figure 4 (curve 2) that even when receiving the NNSS satellite at the single frequency of 400 MHz (actually 399.968 MHz) the navigation errors due to the ionosphere can be reduced to below 200 m., when the data is corrected, using a model ionosphere (or ionospheric prediction) for all passes, the ground path of which (ground distance) is 300–2000 km away from the station. Curve 4 shows that a more accurate evaluation of the ionosphere using actual measurements at a fixed station can reduce the errors by about 50% and extend the zone of 200 m navigation error to passes as near as 170 km.

The simple differential method which is plotted for a fixed station 400 km nearer to the satellite's ground path (curve 3) is seen to pass the 200 meter mark already at about 880 km. Even when the fixed station is 400 km farther away from the satellite's ground path (curve 5), the 200 m error is exceeded already at 480 km.

There seems to be little advantage to this correction method, except maybe for large distances (above 1000 km) to the satellite's ground path. It should be noted, though, that if the distance between the fixed and navigating stations is reduced, the 200 m error will be exceeded at smaller distances from the ground path.

The practical problem is what penalty in waiting time one has to pay if one wishes to use a simpler navigation.

In order to answer that question the following statistical computations were performed: a table of the times of passes of the six NNSS satellites was computed, from the orbital elements that were published by NASA in the "Two Lines Orbital Elements Sets". (These are essentially the "Five Line Orbital Elements Sets" mentioned above in connection to the computations for Figure 3, only that some of the parameters that were published in the former case have to be computed by the user as shown in Reference 17.)

The arbitrary period chosen for the statistical computations was 100 days (June 29, 1975 – October 6, 1975) and the table was computed for Haifa (32.857 N, 324,906 W) as an example of mid-latitudes. The times between two successive suitable satellites passes was considered as the random variable, and a suitable pass was one that was within the ground distance (from Haifa) that is allowed for each particular correction method.

The cumulative distributions of this random variable was computed from the table of passes and the results are shown in Figure 5. These graphs represent the probabilities that within the time that elapsed from the last suitable pass, there will be at least another one – i.e. the probabilities of maximum waiting time.

The actual average waiting time is about half the maximum waiting time assuming a nearly uniform distribution of the times of the passes along the time axis.

Four cases are shown – assuming a navigation error of less than 200 m.

(1) All passes up to 2000 km ground distance from Haifa to the satellite's ground path (farther away passes have too low evaluation angle and are rejected in the NNSS).

This corresponds to the case of the regular NNSS navigation procedure using the 2-frequency ionospheric correction. There were 1923 passes, and the average time between the passes was about 75 min.

(2) Passes at gound distances from 170 and up to 2000 km.

This corresponds to the relative navigation method (paragraph (d), Section 5).

There were 1760 passes and the average time between them was about 82 min.

(3) Passes at ground distances from 300 to 2000 km.

This corresponds to navigation using an estimated ionosphere (paragraph (a), Section 5).

There were 1619 passes and the average time between them was about 89 min.

(4) Passes at ground distance from 480 to 2000 km at one side of Haifa east or west) and passes from 880 to 2000 km on the other side of Haifa.

This corresponds to the simple differential navigation (paragraph (b), Section 5).

There were 1264 passes and the average time between passes was about 114 min.

It can be seen from Figure 5 that there is a sharp rise of the curves between 100–110 min. which is caused by the fact that the periods of revolutions of the satellites are in that range, and when a satellite passes at a suitable distance east of the receiving station, it is likely that after one more revolution in space, the satellite will pass west of the station – also at a suitable distance.

7. CONCLUSIONS

Trying to simplify the navigating stations equipment by using a one-frequency receiver (and a communication link to a fixed master station), one has to pay a penalty mainly in waiting time for a navigation fix, because one cannot use the near zenithal passes in such a case.

Figure 5 summarizes this situation and shows that for a 99% probability of obtaining an error smaller than 200 m, one has to practically double the approximate waiting times. The relative navigation methods of ionospheric error correction can improve the situation slightly. For a 95% probability one pays a penalty of 20% increase only in waiting time (for the relative navigation method).

The above is correct for the influence of the ionosphere; other sources of error may introduce different considerations.

All the errors were calculated for a TEC of $4 \cdot 10^{17}$ el/m^2. At other TEC's the corresponding errors will be almost linearly proportional.

At Haifa for example[9], midday TEC values of $8 \cdot 10^{17}$ el/m^2 correspond to equinox days of high solar activity, while $4 \cdot 10^{17}$ el/m^2 is approximately an average midday and afternoon value. At all other times of the day, especially at night and early morning, the TEC and also the associated errors are of course much smaller.

BIBLIOGRAPHY

1. Stanzel, T.A., Jr — *The Navy Navigational Satellite System: Description and Status*. Navigation, J. of Inst., of Navigation, Vol.15, No.3, pp.229-243, Fall 1968.

2. Piscane, V.L. Holland, B.B. Blake, H.D. — *Recent (1973) Improvements in the Navy Navigation Satellite System*. Navigation, J. of Inst. of Navigation, Vol.20, No.3, pp.224-229, Fall 1973.

3. Tucker, A.J. Fannin, B.M. — *Analysis of Ionospheric Contributions to the Doppler Shift of CW Signals from Artificial Earth Satellites*. J. Geophys. Res., Vol.73, pp.4325-4334, 1968.

4. Mass, J. — *Total Ionospheric Electron Content by Doppler Integral of Satellite Signal*. J. Atmos. and Terres. Phys., Vol.24, pp.549-553, 1962.

5. Da Rosa, A.V. — *Propagation Errors in VHF Satellite-to-Aircraft Ranging*. IEEE Trans. Ant. and Prop., Vol. AP-17, No.5, pp.628-634, September 1969.

6. Budden, K.G. — *Radio Waves in the Ionosphere*. The University Press, Cambridge, 1961.

7. Yeh, K.C. Swensen, G.W. — *Ionospheric Electron Content and it's Variations Deducted from Satellite Observations*. J. Geophys. Res., Vol.66, pp.1061-1067, 1961.

8. Garriot, O.K.
Da Rosa, A.V.
Ross, W.J.
Electron Content Obtained from Faraday Rotation and Phase Path Length Variation. J. Atm. Terr. Phys., Vol.32, pp.705-725, 1970.

9. Mass, J.
Houminer, Z
Ionospheric Research Using Satellites. Final Scientific Report, February 1969, Contract AF 61(052)-837, National Committee for Space Research Radio Observatory, Haifa, Israel.

10. –
Joint Satellite Studies Group, *Ionospheric Electron Content and Scintillation Studies at Widely Spaced Low Latitude Station.* Plant. Space Sci., Vol.16, pp.1277-1287, 1968.

11. Shuval, A.
Relative Navigation Using Satellite Transmissions, M.Sc. Thesis at the Technion, Israel Institute of Technology, Haifa, Israel, 1975.

12. Farkas, L.
A Position Difference System Using Satellite Doppler Signals. Proc. of 5th Mil-E-Con, June 1961.

13. –
AGARDograph No.66, on *Total Electron Content and Scintillation Studies of the Ionosphere.* Edited by J.Aarons, p.55, March 1973.

14. Mass, J.
Shuval, A.
Calculation of the Ionospheric Errors of the Navy Navigational Satellite System (NNSS). In Final Scientific Report, Contract F 61052-0017, USAF, February 1972. NCSR Radio Observatory, Haifa, Israel.

15. Mass, J.
Shuval, A.
Ionospheric Errors in the NNSS Satellite Navigation System. In Final Scientific Report, Contract F 44620-72-C-0059, USAF, June 1975, NCSR Radio Observatory, Haifa, Israel.

16. Rao, N.N.
Yeh, K.C.
Comparison of Faraday and Doppler Methods of Obtaining Ionospheric Electron Content. J. Geoph. Res., Vol.73, No.7, pp.2447-2458, April 1968.

17. –
Explanation of the Two-Line Orbital Elements, from February 20, 1974. Published by NASA, Goddard Space Flight Center, Code 512, Greenbelt Maryland 20771, USA.

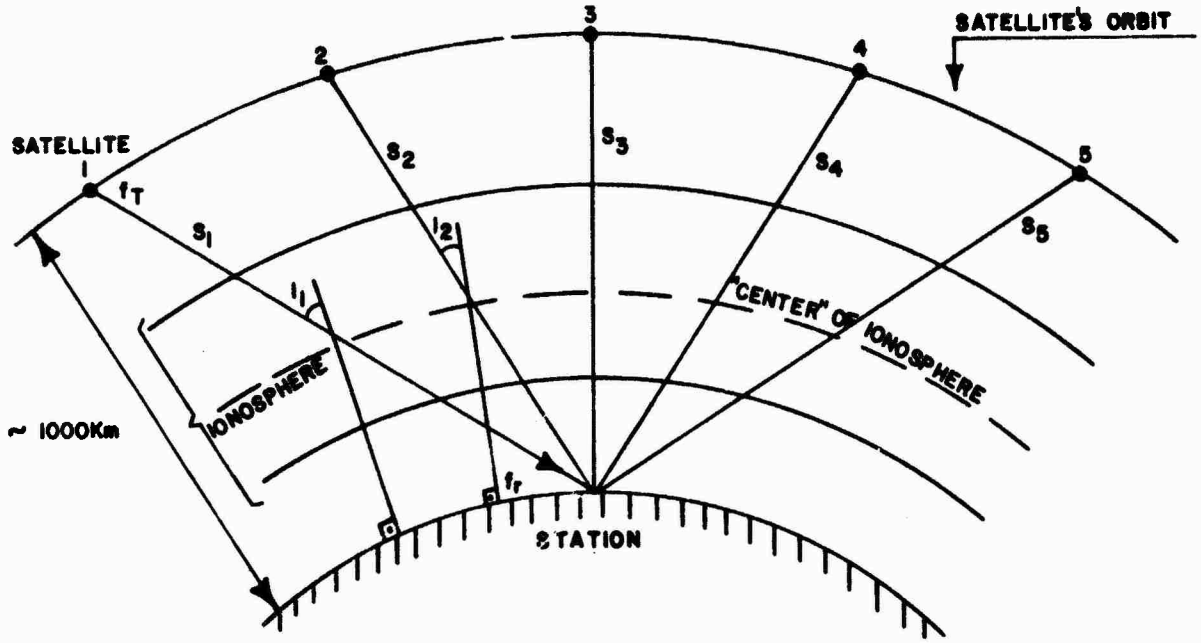


Fig.1 Simplified geometry for an NNSS satellite. The numbered points on the satellite's orbit, correspond to time points in Figure 2

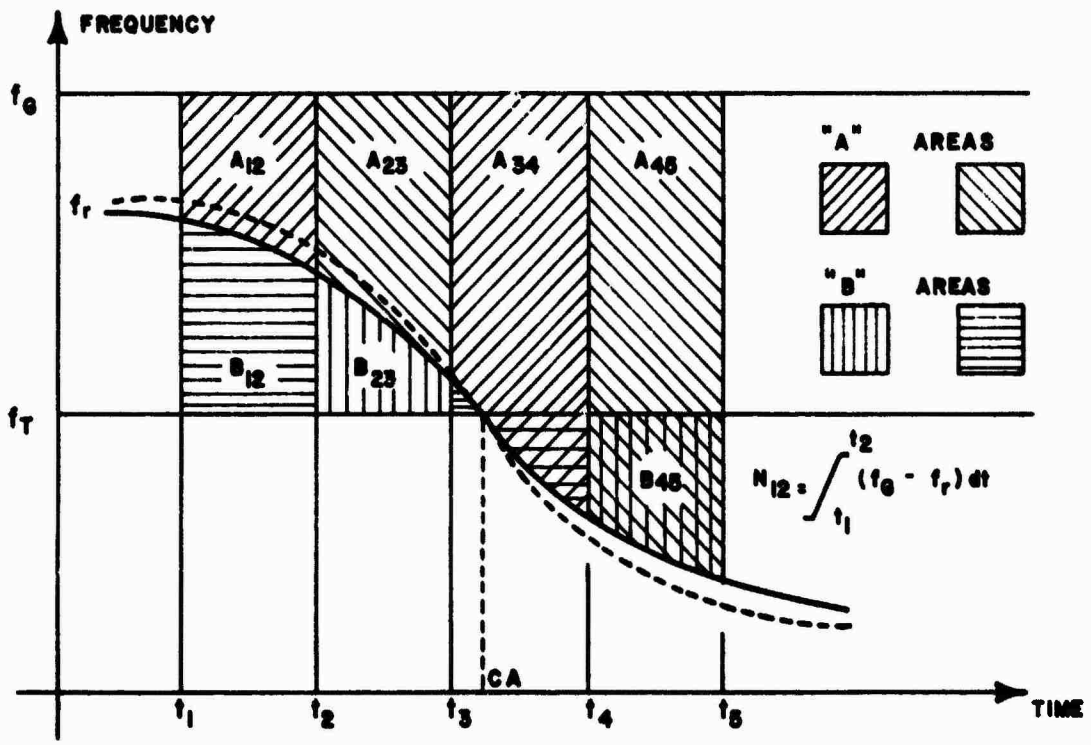


Fig.2 A typical curve of the received frequency f_r as function of time. The received frequency is shifted due to the Doppler effect. The dotted curve illustrates the received frequency if the ionosphere was not present. The shaded areas correspond to the Doppler integral (or cycle counts) measured on ground. CA-closest approach point

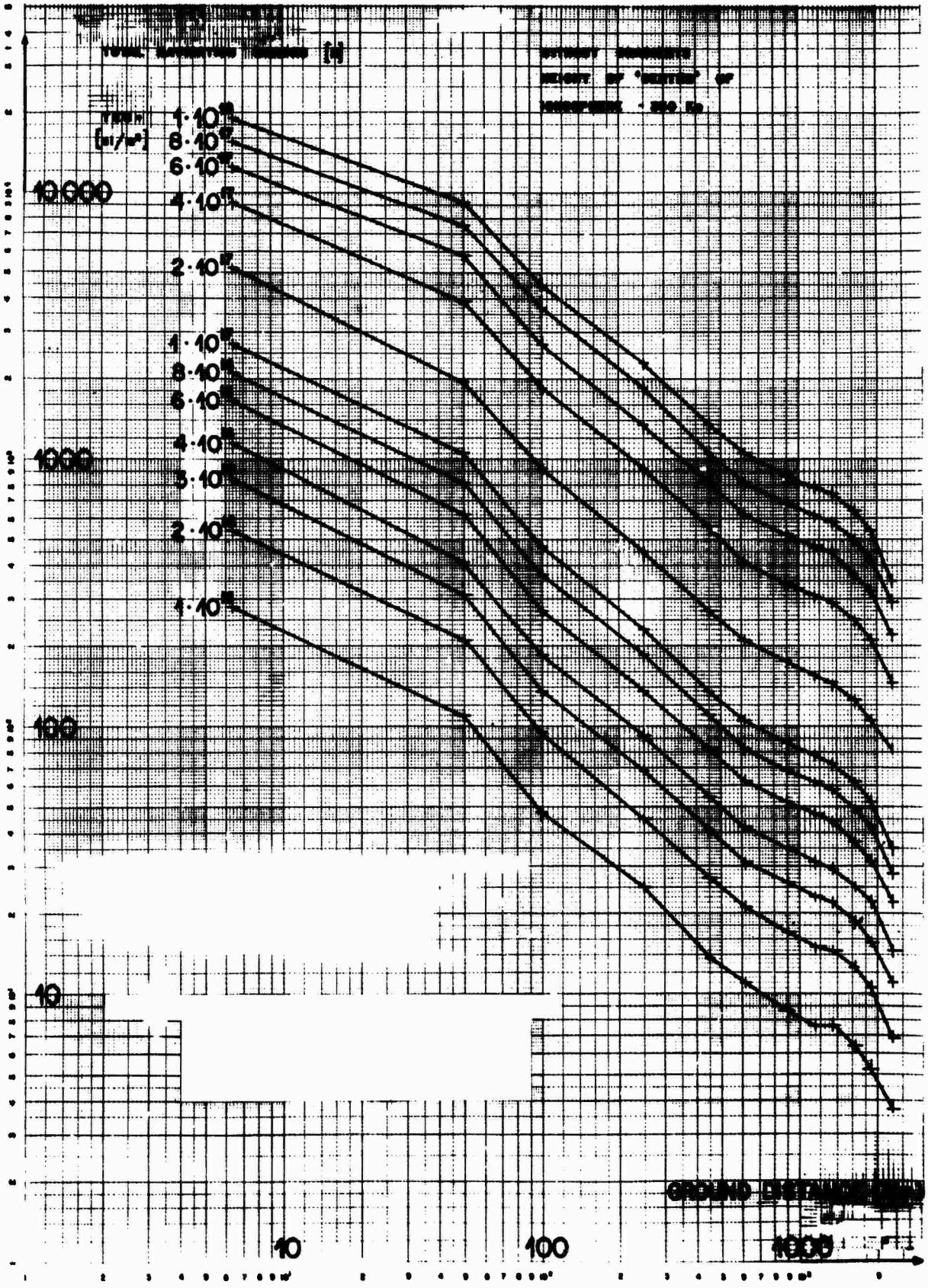


Fig.3 Total navigation errors caused by the ionosphere to NNSS satellite passes at various ground distances. The errors shown are those that were received by using information on the Doppler integral from the whale pass (12 min). The parameter is the TEC

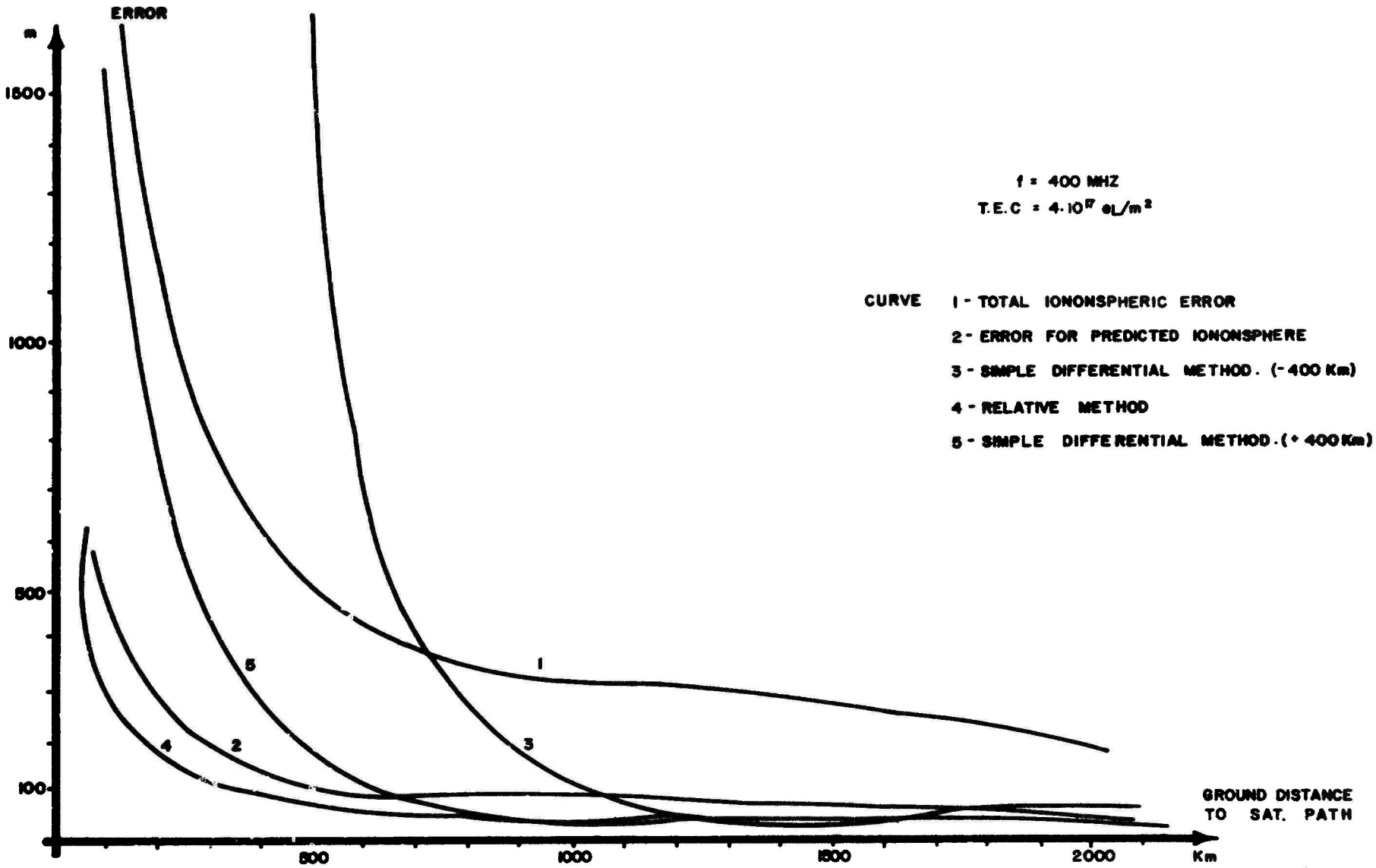


Fig.4 Total navigation errors for different ionospheric correction techniques

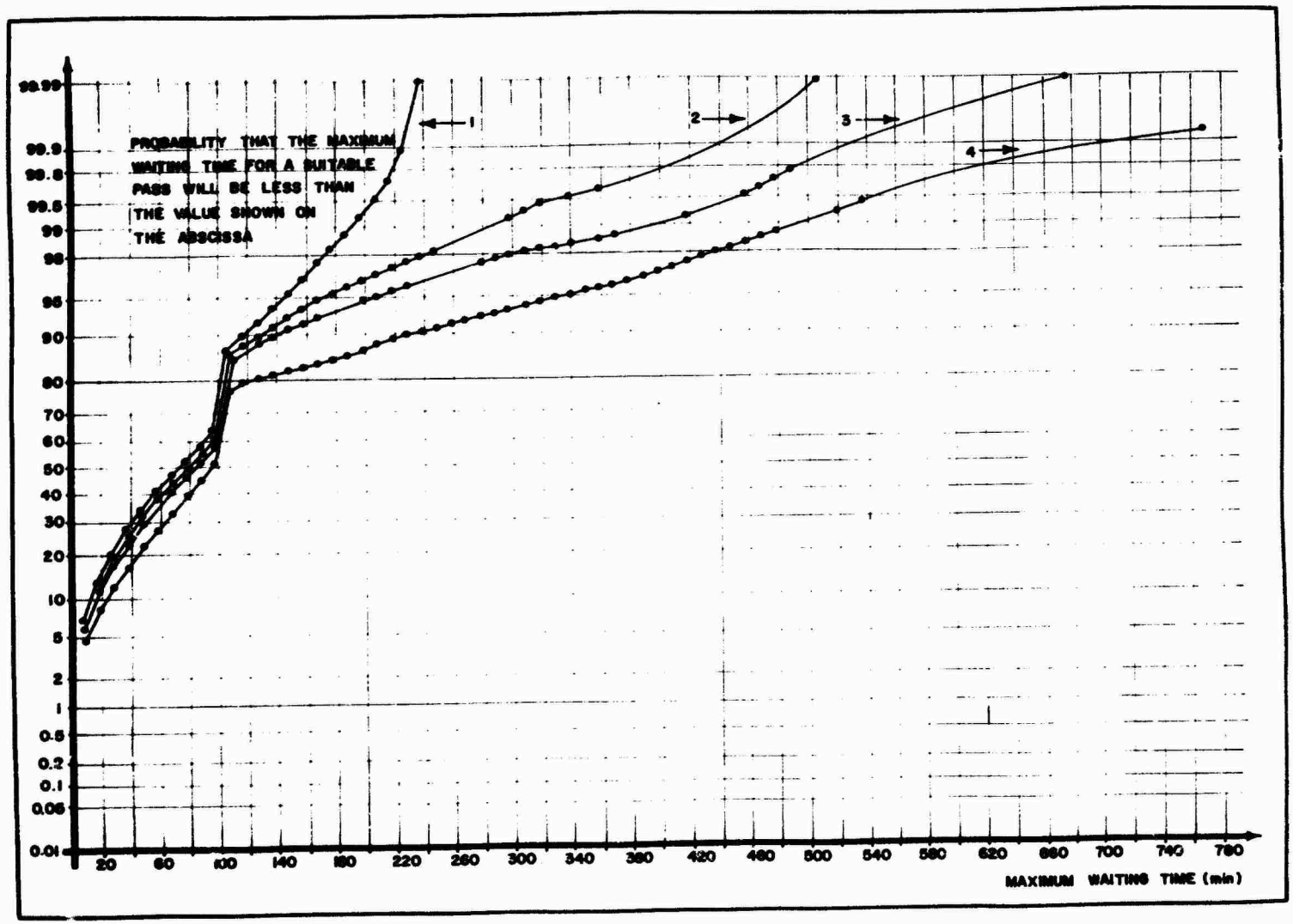


Fig.5 Probabilities of maximum waiting time for a suitable pass, (for navigation errors less than 200 m)

1. Regular NNSS (2 frequency correction)
2. Relative navigation
3. Navigation using estimated ionosphere
4. Simple differential navigation

A STUDY OF SUDDEN IONOSPHERIC DISTURBANCES AND THEIR EFFECT ON VLF POSITION FIXING ACCURACY.

M.E.Perry
Operational Research and Computer Systems Department
The Decca Navigator Company Limited
New Malden,Surrey,England.

SUMMARY

With the concentration of air traffic over the North Atlantic Organised Track Structure area it has become necessary to ensure that aircraft do not stray from the specified track by more than given amounts. This study is to determine the positional errors that may be encountered by an aircraft using the OMEGA system in this area during periods of high solar activity. This has been achieved by the collection of OMEGA phase records from many sources. These records have been analysed to give an approximate frequency and propagation path dependence. These dependencies have been used to convert the collected data to a normalised form from which it can be used to determine the approximate errors that would be observed at any point in the North Atlantic area.

1. DESCRIPTION AND COLLECTION OF SUDDEN IONOSPHERIC DISTURBANCE DATA.

The data used in the preparation of this study came from two main sources:

(a) Phase recordings of Omega Trinidad received at RAE Farnborough from 1966 through until 1971. This data was mainly at 10.2 KHz but also included some at 11.3 KHz and 13.6 KHz.

(b) Phase recordings of Omega Trinidad received at RGO Herstmonceux from November 1969 until June 1973. This data was at 12.0 KHz.

In addition, small amounts of data were used from the following propagation paths:

(c) Forestport to RAE Farnborough
(d) Aldra to RAE Farnborough
(e) Trinidad to Ottawa

} all at 10.2 KHz

This was used for the determination of the propagation path dependence.

Availability of the above records, and their specific use for determination of propagation path or frequency dependence, is shown in Figure 1.

Although the actual parameters noted depended on the type of recording used, sufficient data was collected on each disturbance to determine maximum deviation, time of event and a measure of the length of the event.

Only those disturbances which occurred when the entire propagation path was illuminated were used.

2. THE FREQUENCY DEPENDENCE OF SIDs.

The size of a phase shift during a SID is dependent on the frequency in use. For the reduction of all the data collected it was necessary to determine the relationship which exists between the different frequencies.

Fortunately there was a reasonable amount of overlapping data on 10.2 KHz, 11.3 KHz, 12 KHz and 13.6 KHz on the paths Trinidad to RAE Farnborough and Trinidad to RGO Harstmonceux. These paths are sufficiently similar for the path dependence to be ignored.

Three diagrams were drawn showing the ratios of the sizes of the SIDs on these frequencies:-

Fig.2a shows the ratio size (10.2 KHz) / Size (11.3 KHz)
Fig.2b shows the ratio size (10.2 KHz) / Size (12.0 KHz)
Fig.2c shows the ratio size (10.2 KHz) / Size (13.6 KHz)

The mean value and standard deviation of each ratio was calculated. The mean ratio and standard deviation for SIDs with phase deviations greater than 10μs at 10.2 KHz were also calculated separately.

Fig.3 shows the resultant frequency relationship for

(a) All SIDs and (b) Large SIDs (> 10μs at 10.2 KHz)

It can easily be seen that the standard deviations are much larger in the case of the small SIDs. This is probably due to the SID size being only slightly larger than the noise on the recordings.

With the results shown, a linear relationship of SID size to frequency was chosen although it appears that this would probably not hold over a larger range of frequencies.

3. THE PATH DEPENDENCE OF SIDs.

SIDs are caused by increased ionization in the D layer, reducing the effective reflection height for VLF waves. This results in an apparent reduction in phase for each VLF signal depending on its propagation path.

The causes of ionization of the upper atmosphere are as follows (MITRA,A.P., 1974):-

(i) The far ultraviolet 1350-1750 Å responsible for the dissociation of O_2.

(ii) The whole spectrum from 100-1350 Å and in particular $L\alpha$ at 1216 Å, mainly responsible for the quiet day D region.

(iii) X-rays above 8 Å affecting the lower E region.

(iv) X-rays below 8 Å which are negligible under quiet conditions but at flare times are the main source of D region ionization.

In addition, particle flux from the Sun, and cosmic rays, can also affect the ionosphere. Charged particles emitted from the Sun can cause very large disturbances in the ionosphere lasting for many hours, and even days. This type of disturbance, known as a PCA (Polar Cap Absorption) can be very serious for propagation paths crossing the Auroral Zone but are very small outside that zone. Very little data exists on disturbances of this type so this report will deal more with the disturbances caused by the electromagnetic radiation from the Sun.

The actual effect caused by a flare will be a very complicated function of the present state of ionization of the ionosphere and of the detailed spectrum and time dependence of the flare radiation. Such comprehensive data exists for only a very small number of flares and so this report will consider only two sets of wavelengths, 0.5 - 3.0 Å and 1.0 - 8.0 Å. A list of the outstanding events at these wavelengths is published monthly (Solar Geophysical Data). There is very close temporal agreement between the lists of outstanding events and the list of observed SIDs.

The records used to determine propagation path dependence were 10.2 KHz phase records for March '69, April '70 and May '70 as shown in Fig.1. Only those SIDs which were definitely associated with a known X-ray flare were used in this part of the study. Each observed SID was given a two letter name. The X-ray flux of the associated flare was noted as was the mean value of the cosine of the solar zenith angle along the path. In addition, the Hardness Ratio was calculated for each X-ray flare.

The Hardness Ratio is given by $\dfrac{\text{X-ray flux}(0.5 \rightarrow 3\ \text{Å})}{\text{X-ray flux}(1.0 \rightarrow 8\ \text{Å})}$

In Fig.4 the phase deviations due to SIDs on the Trinidad to RAE Farnborough path can be seen plotted against the mean incident X-radiation in the range 0.5 - 3 Å. For greater clarity the data has been divided into three subsets as follows:-

(i) low hardness ratio 0 - .019 shown as .... dj
(ii) middle hardness ratio .020 - .035 shown as .... DD
(iii) high hardness ratio .036 - .054 shown as .... EO

As can be seen, the low hardness flares tend to produce larger SIDs for the same hard X-ray illumination. This is caused by the ionizing effect of the larger amount of soft X-rays present. Similarly, high hardness flares, being relatively deficient in soft X-rays, tend to produce smaller SIDs for a given hard X-ray illumination.

Even with this differentiation between various types of X-ray flares some SIDs appear above or below the mean for their type. This can often be shown to be due to the flare being particularly long lasting giving a larger SID, or particularly short lasting giving a smaller SID.

Now consider the subset of SIDs of a given hardness ratio. In each case the plot shows that a certain threshold value of incident hard X-radiation is required before any disturbance is noticed. Then there exists a fairly linear dependence between SID size and incident hard X-radiation, followed by a progressively slower rise in SID size with increasing incident hard X-radiation. Presumably, given infinite incident hard X-radiation the size of the observed SID would tend to a limit which would be dependent on the hardness of the X-radiation.

Hence, we might expect to obtain an expression like:

$$y = y_0 \ (1 - e^{k(x_0 - x)}) \quad \ldots\ldots(1)$$

where y is SID size

y_0 is Max SID size = function of path length and hardness

x is mean incident radiation

x_0 is threshold mean incident radiation

k is some constant, possibly a function of hardness.

Such an expression is shown in Fig.5

The equation (1) is somewhat cumbersome to use for normalising the data for different paths under different illuminations. It does, for instance, require to know the hardness of a particular X-ray flare before it can do any conversion. As it is impossible to enter all of the X-ray information relevant to all of the observations a simpler approach must be found. A simple power law was tried as in equation(2) below.

$$y = C(h,pl)\ x^n \quad \ldots\ldots (2)$$

where y is SID size

x mean incident radiation

C is some constant = function of path length and hardness

n is some power

h is hardness

pl is pathlength

An investigation was carried out into the validity of equation (2)

We can deduce

$$\log y = \log C + n \log x$$

So four graphs (one for each propagation path for which data existed) were plotted of the above relationship for the middle subset of hardness (ratio between .020 and .035).

In each case a value of n = .65 was found to provide a reasonable fit to the bulk of the data with the exception of the very small SIDs where the threshold effect would cause a departure (a bend down effect) from a strict power law. These graphs are shown in Figs.6(i)-(iv)

From these graphs the value of C can be determined for each path. In Fig.7 these values for C were plotted against pathlength and we can see that the value of C is almost linear with pathlength. Hence, we may assume the following:

$$c(h,pl) = c(h) * pl. \quad \ldots\ldots (3)$$

We now have all the tools we need to normalise the data from several paths to a fixed path of length 1 radian, subjected to normal illumination.

We now have $y = c(h).pl.x^n$ from (2) and (3)

$$= c(h).pl.\ flux^n.\ (\overline{\cos Z})^n \quad \ldots\ldots (4)$$

since mean incident radiation $x = flux.\ \overline{\cos Z}$

where Z is the zenith angle of the Sun

Now y observed $= c(h).\ pathlength.\ flux^n.\ (\overline{\cos Z})^n$ $\ldots\ldots (5a)$

y normalised $= c(h).\ 1.\ flux^n.\ 1^n$ $\ldots\ldots (5b)$

$$y\ normalised = \frac{y\ observed}{pathlength * (\overline{\cos Z})^n} \quad \ldots\ldots (6)$$

The pathlength is known, the mean solar zenith angle can be calculated using the data and time of observation so the normalised value can be calculated (assuming a factor has already been applied to correct for the frequency in use.)

4. THE DISTRIBUTION OF PHASE ERRORS DUE TO SIDs NORMALISED TO A PATH SUBTENDING 1 RADIAN AT THE CENTRE OF THE EARTH, WITH NORMAL ILLUMINATION AND FREQUENCY OF 10.2KHz.

Computer programmes were written to analyse the collected data. Each SID was normalised separately using equation (6). The resulting normalised value was used to compile a graph showing the probability of a certain normalised phase error being exceeded. Examples of such graphs for the year 1970 are shown in Fig.8. A separate graph has been compiled for each month, the number of SIDs analysed being shown on that graph. Note should be made of the graph for November 1970. This was the worst monthly period encountered during the study and was the set of data used to determine flare time position fixing errors (see later).

5. PROBABLE POSITION FIXING ERRORS AND THEIR VARIATION WITH TIME.

Having determined the probability of observing certain phase errors on a normalised path, we can now use this information to determine the probable phase error observed at a given point using a given transmitter at a given time. Moreover, if we do these calculations for three transmitters (Liberia, North Dakota and Norway) we can determine what the position fixing error is likely to be. Fig.9 shows the probable "error footprint" observed at 60N 55W under conditions similar to the month of November 1970 (already mentioned). The curves shown represent the errors one might expect to get with probabilities 10^{-2}, 10^{-3} and 10^{-4} had these conditions arisen during the months of June or December. The curves also show the variation of the size and direction of the error with time of day (times in GMT).

As can be seen, in both the results for June and December, the error is very great in roughly the direction of the Liberian transmitter when that propagation path is illuminated. The error is in roughly the direction of the North Dakota transmitter when only that path is illuminated, and roughly in the direction of the Norway transmitter when only that path is illuminated (i.e. in June with the effect of the Midnight Sun).

At no time does the position error exceed 15 nautical miles.

6. CONCLUSION.

This study was undertaken with a specific, practical aim in mind, namely to determine approximately probable position errors at points where no actual phase records exist. Although the frequency and path dependencies used would not fully satisfy theoretical considerations, it is believed that they are sufficiently realistic to be applied in the way described to obtain practical results.

ACKNOWLEDGEMENTS.

The Author wishes to thank the National Air Traffic Services in conjunction with the Ministry of Defence, Royal Aircraft Establishment, Farnborough, for commissioning this study, and also Dr.D.Allison for his measurements of the thousands of phase records used during the study.

REFERENCES:

Mitra,A.P. 1974, "Ionospheric Effects of Solar Flares", D. Reidel Publishing Company.

"Solar-Geophysical Data", Environmental Data Service, National Oceanic and Atmospheric Administration, U.S.Department of Commerce.

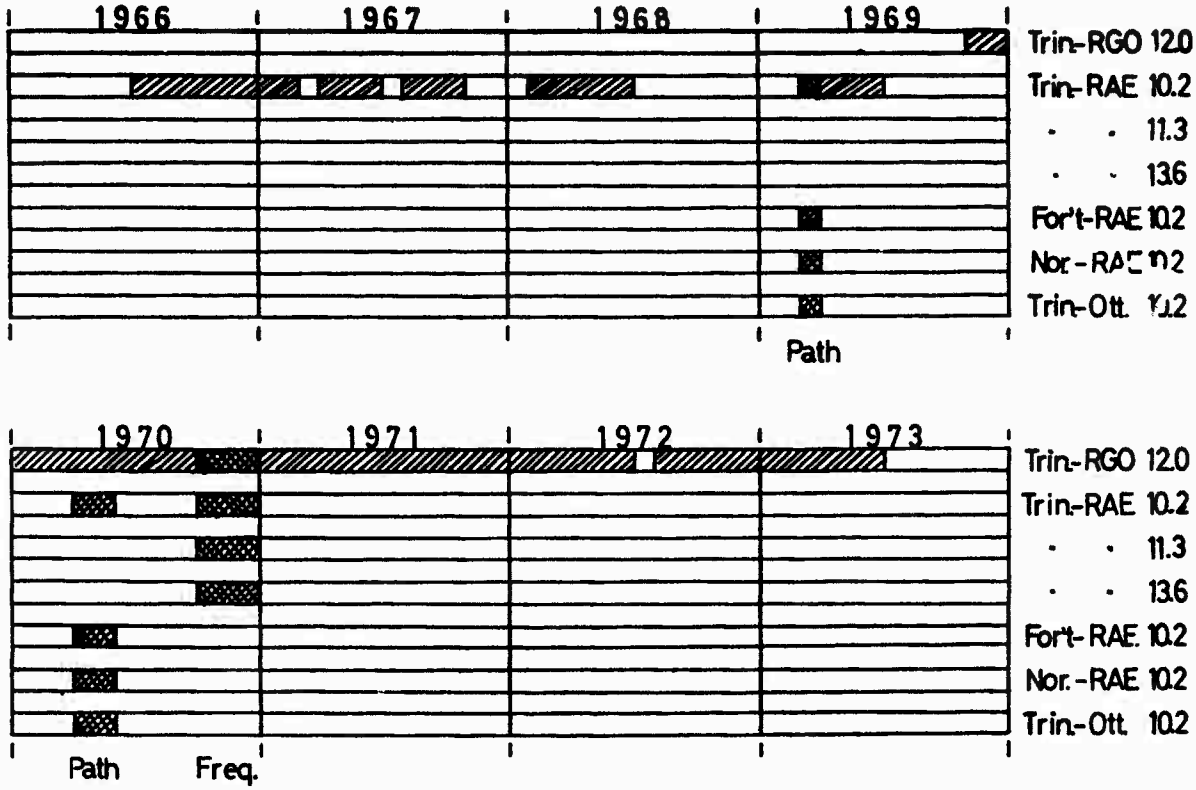


FIG. 1 shows which records were used

(a) for the disturbance statistics (single hatched)
(b) for the determination of frequency and propagation path dependence (cross hatched)

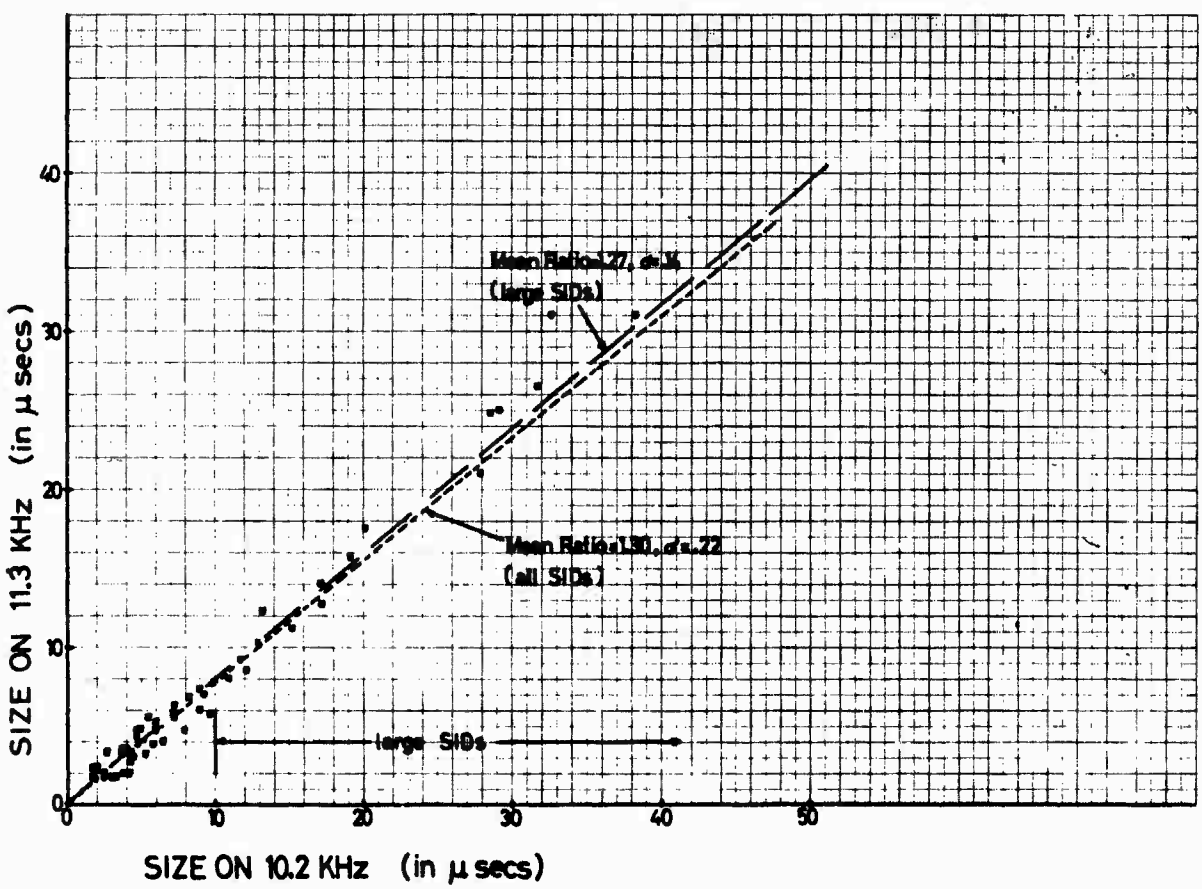


FIG. 2a shows the relationship between size on 10.2KHz and size on 11.3KHz

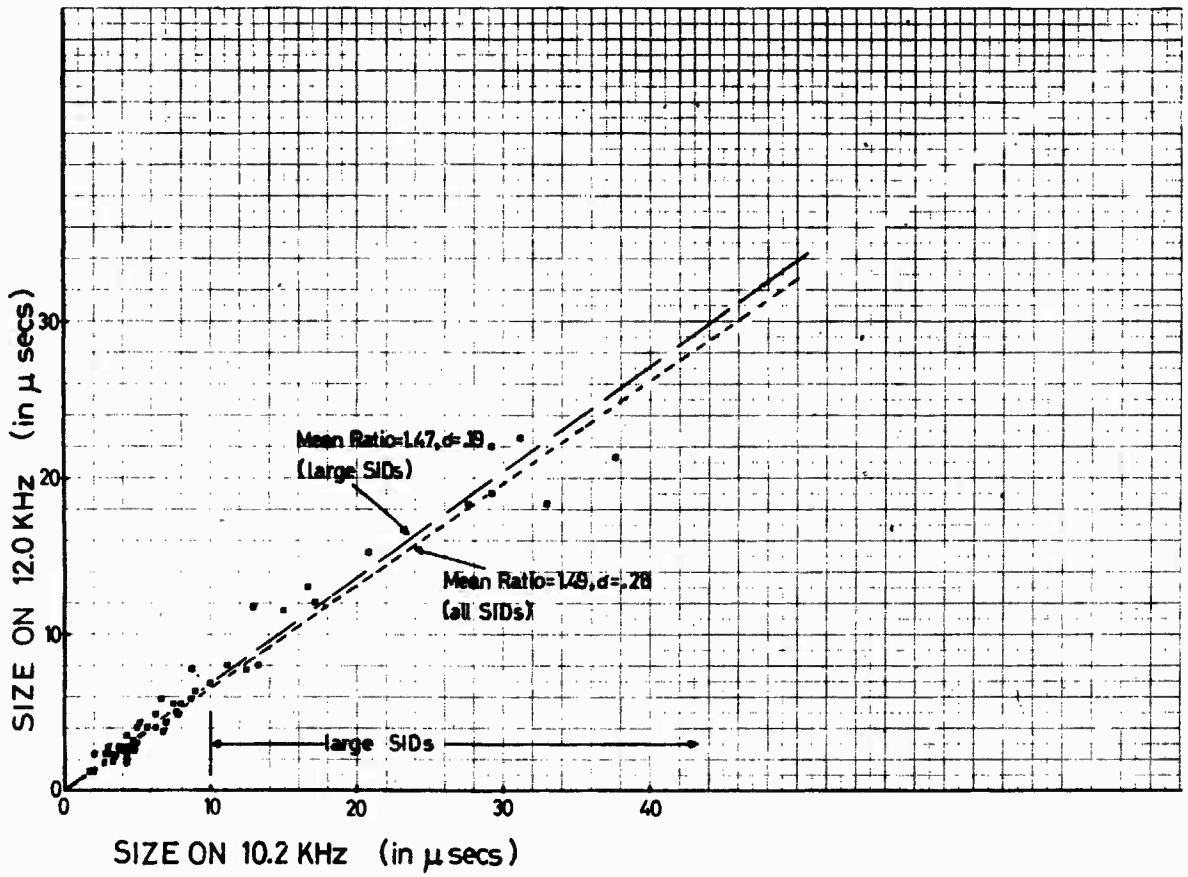


FIG.2b shows the relationship between size on 10.2KHz and size on 12.0KHz

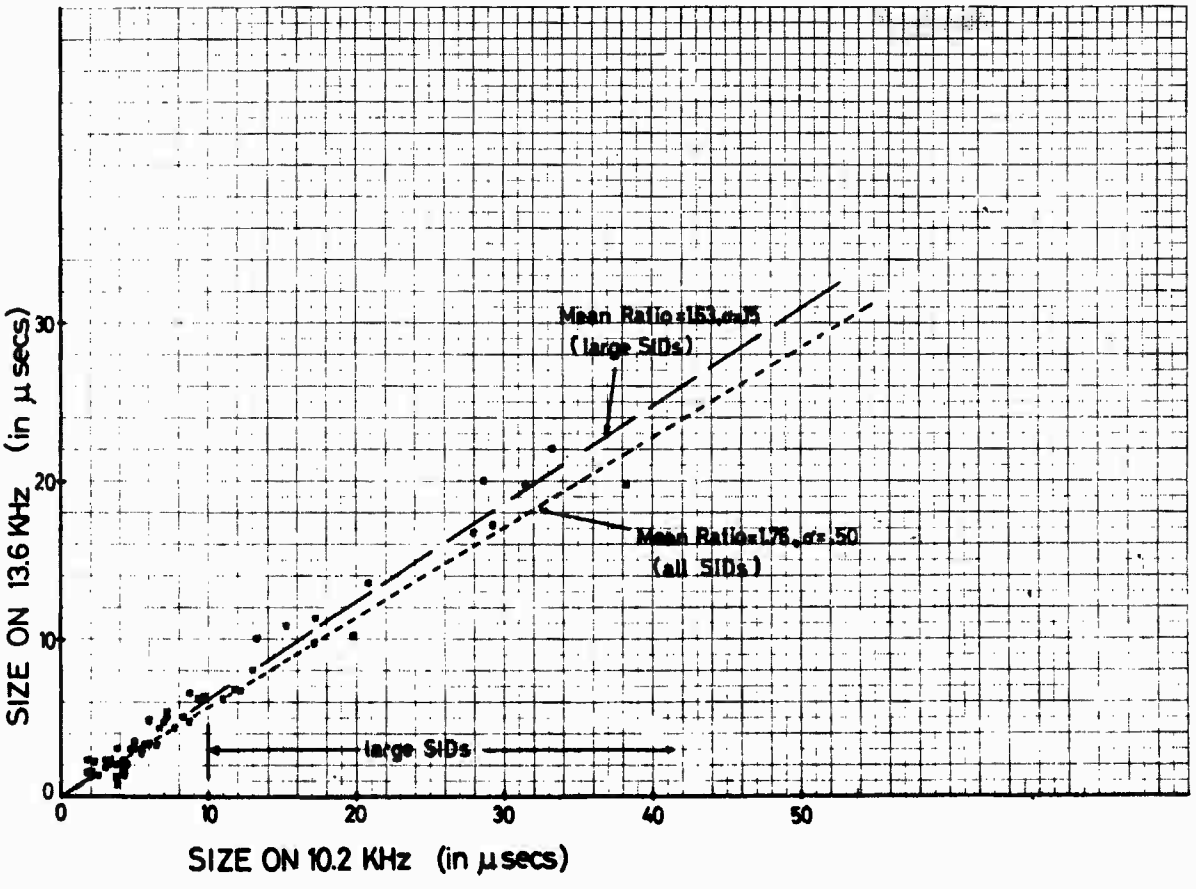


FIG.2c shows the relationship between size on 10.2KHz and size on 13.6KHz

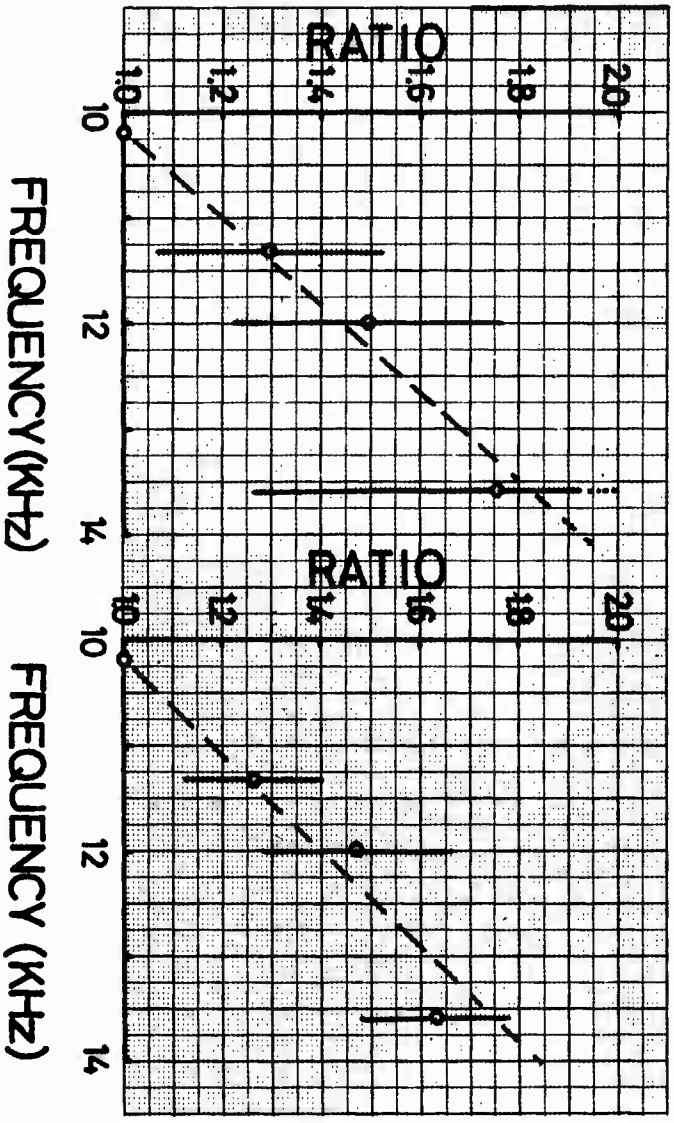


FIG.3 shows the frequency relationship in graphical form, the circles representing the mean values of the ratios with error bars of 1 standard deviation.

The relationship on the left is drawn for all SIDs and that on the right is for large SIDs (> 10μs at 10.2KHz) only.

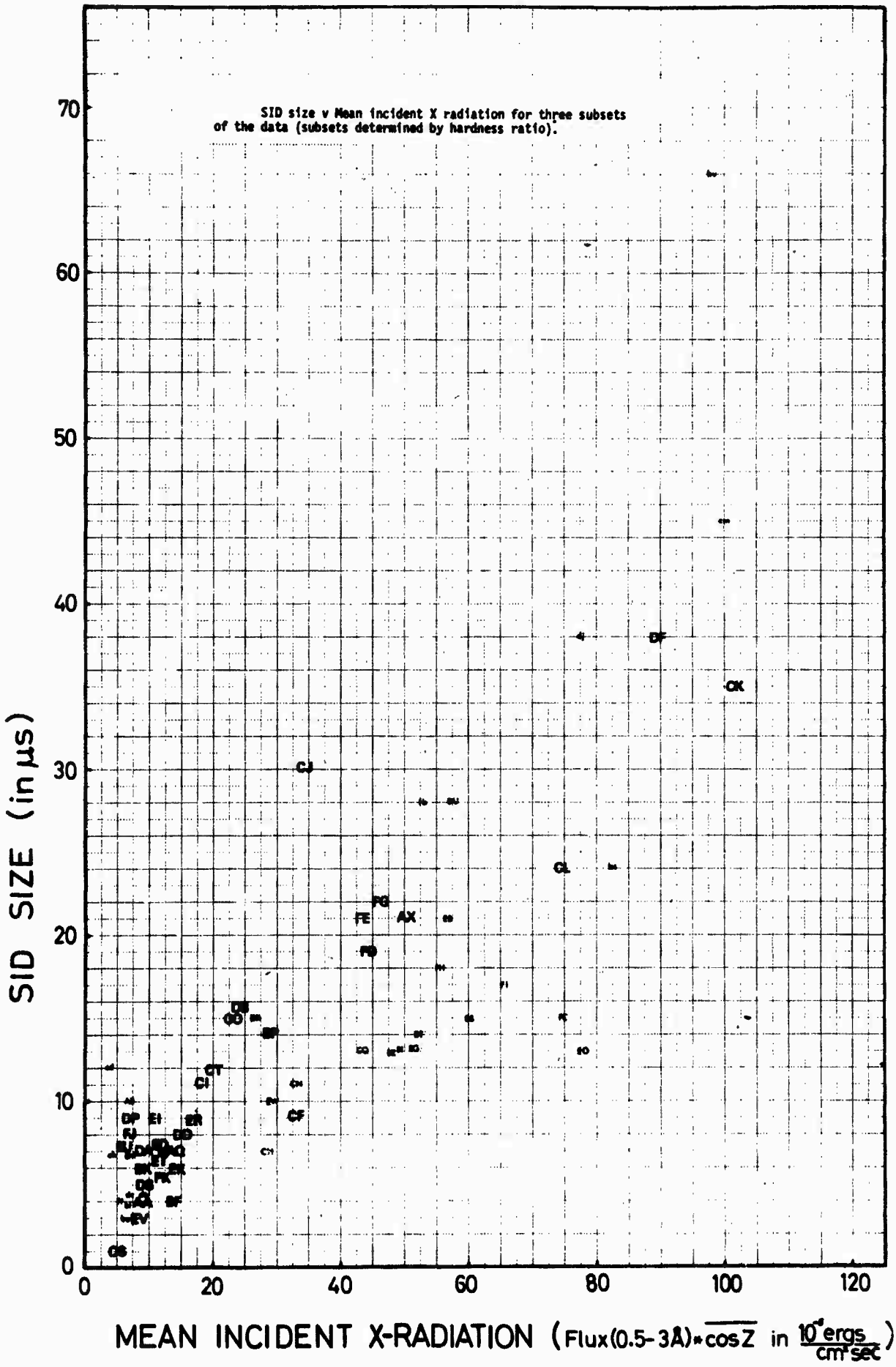


FIG.4 shows the relationship between observed SID size and mean incident X-radiation.

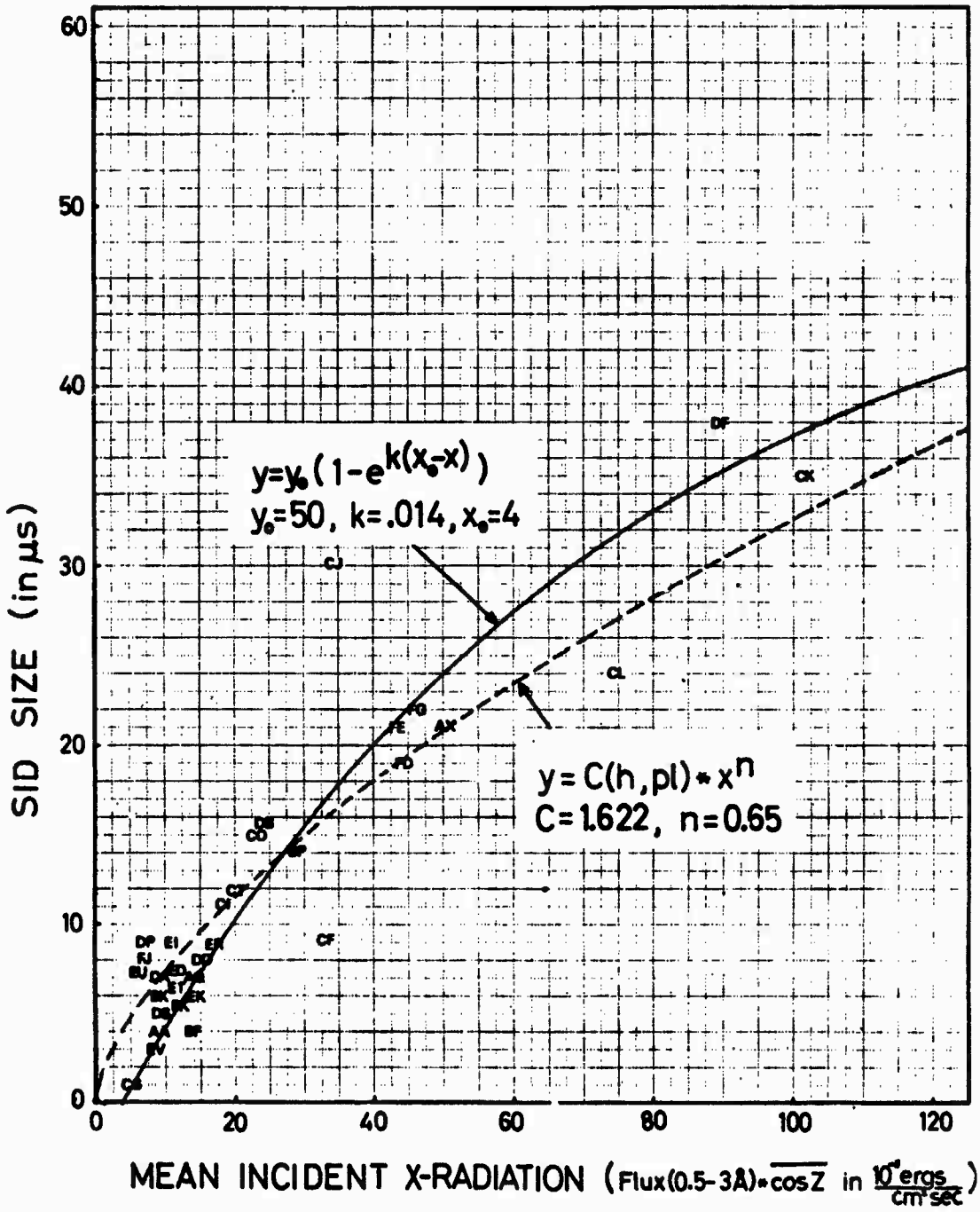
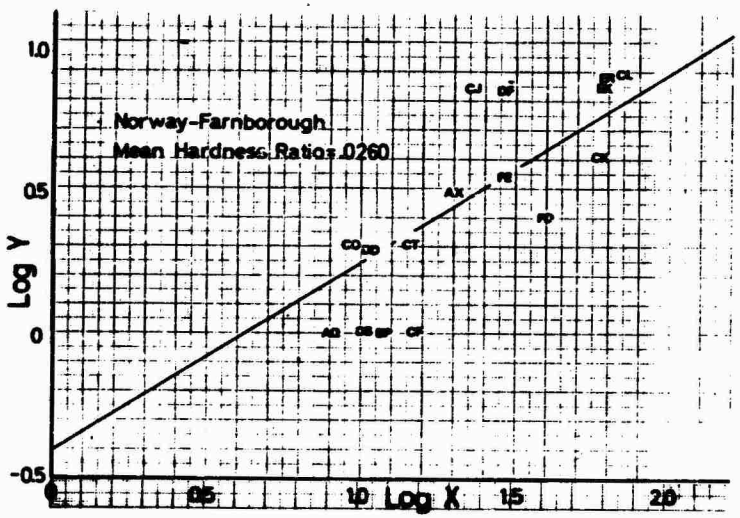
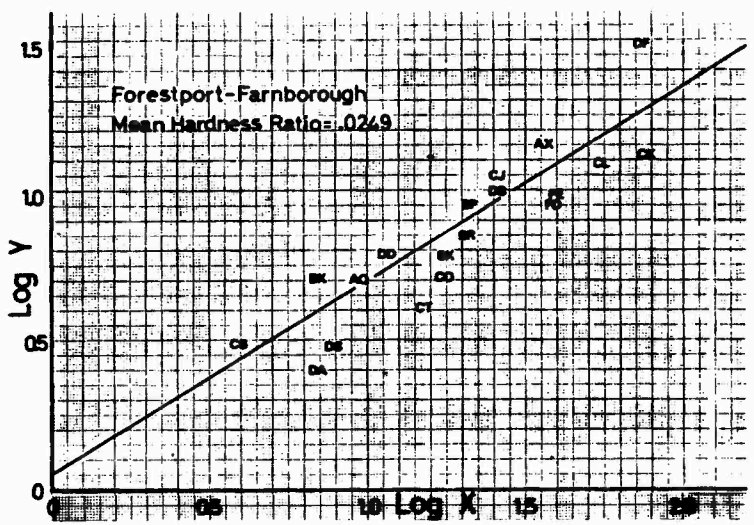
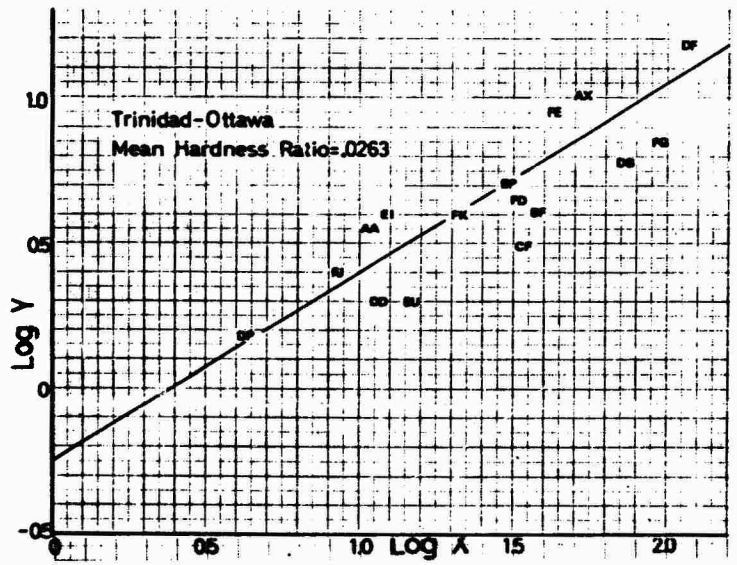


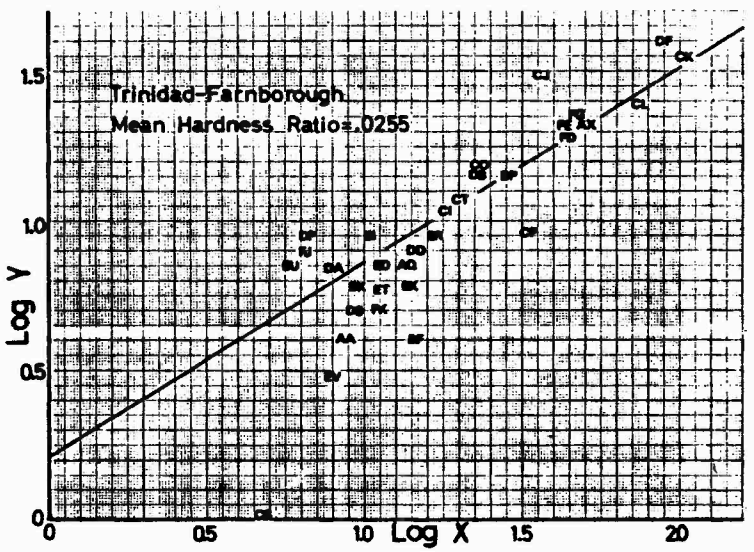
FIG.5 Two possible expressions to model the behaviour of the SIDs with hardness ratic .020 - .035.





Log(SIDsize) v Log(Incident X ray flux)





Log(SIDsize) v Log(Incident X ray flux)

FIGS.6 (i) - (iv). The relationship between log (SID size) and log (Incident X-ray flux).

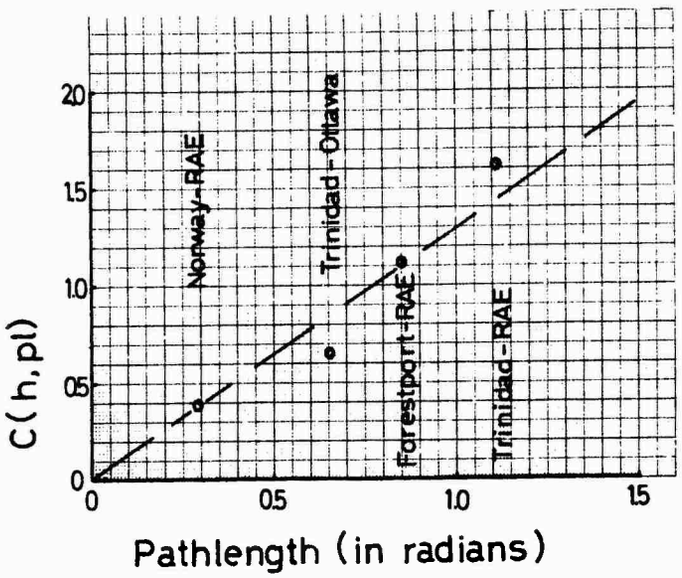
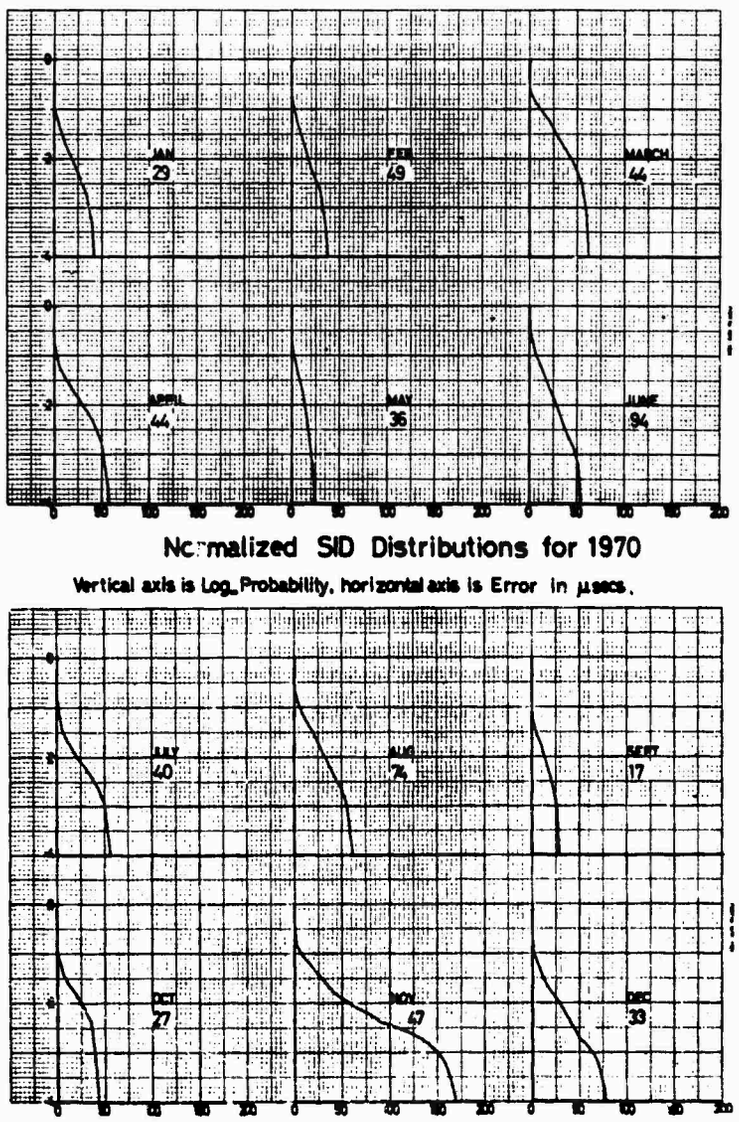


FIG. 7 The variation of C(h,pl) with pathlength.
Note: The line shown is only one approximation. There may be a better one through the right hand three points only which would give rise to a normalization based on a reduced path length.



Normalized SID Distributions for 1970

Vertical axis is Log_Probability, horizontal axis is Error in μsecs.

FIG. 8 Normalized SID Distribution for 1970 (Derived from records made at RGO Herstmonceux of the 12.0KHz transmissions from Trinidad).

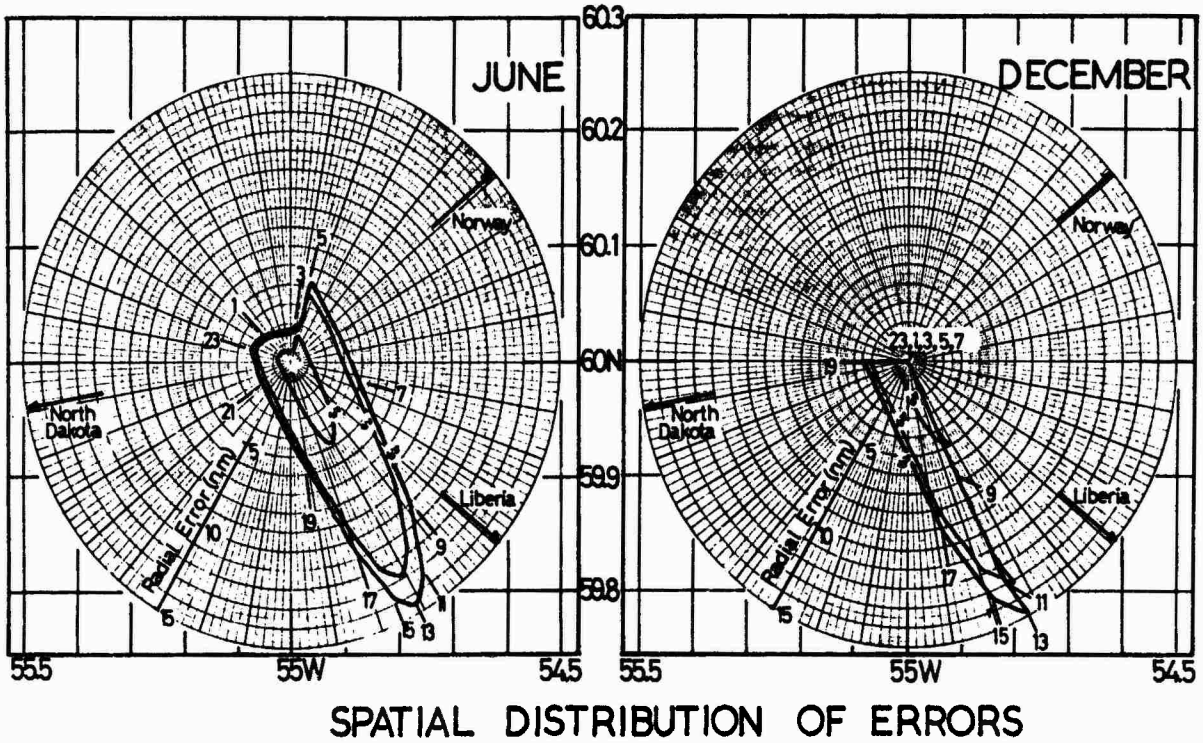


FIG.9 Typical "Error Footprints" obtained at 60N 55W during the months of June and December. Contour lines are with probabilities 10^{-2}, 10^{-3} and 10^{-5}; times shown are GMT.

(9) Conference proceedings.

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's Reference (14)	3. Further Reference	4. Security Classification of Document
	AGARD-CP-209	ISBN 92-835-0189-6	UNCLASSIFIED

5. Originator Advisory Group for Aerospace Research and Development
North Atlantic Treaty Organization
7 rue Ancelle, 92200 Neuilly sur Seine, France

6. Title (6) PROPAGATION LIMITATIONS OF NAVIGATION AND POSITIONING SYSTEMS.

7. Presented at the Electromagnetic Wave Propagation Panel Specialists' Meeting held in Istanbul, 20–22 October 1976.

8. Author(s)	9. Date (11)
Various	Feb[illegible]77
10. Author's Address	**11. Pages**
Various	350 (12) 343p.

12. Distribution Statement This document is distributed in accordance with AGARD policies and regulations, which are outlined on the Outside Back Covers of all AGARD publications.

13. Keywords/Descriptors

Navigation	Navigation satellites	Limits
Electromagnetic wave transmission	Hyperbolic navigation	
Position (location)	Radio navigation	

14. Abstract

In order to assess the requirements for navigation and positioning systems, the current status of these systems, and the limitation the propagation medium places on systems in being or planned, the AGARD EPP held a Specialists' Meeting in Istanbul, October 20–22 1976, on this subject. The concept of the meeting was to outline requirements and progress in systems and to plan programs for future studies to reduce any propagation limitations on navigation systems.

Accordingly, the meeting reviewed various subjects including civilian requirements for sea, air and ground navigation systems both short and long range, and propagation study needs of Loran C, Omega, NAVSTAR, and Aerostat. In addition, requirements for other systems were outlined and possible propagation problems discussed. Proposed future research and development programs was the subject of a round table discussion. The aim was to mitigate the propagation problems that systems will and have encountered.

CDS 400 043

AGARD Conference Proceedings No.209
Advisory Group for Aerospace Research and Development, NATO
PROPAGATION LIMITATIONS OF NAVIGATION AND POSITIONING SYSTEMS
Published February 1977
350 pages

In order to assess the requirements for navigation and positioning systems, the current status of these systems, and the limitation the propagation medium places on systems in being or planned, the AGARD EPP held a Specialists' Meeting in Istanbul, 20–22 October 1976, on this subject. The concept of the meeting was to outline requirements and progress in systems and to plan programs for future studies to reduce any propagation limitations on navigation systems.

P.T.O.

AGARD-CP-209

Navigation
Electromagnetic wave transmission
Position (location)
Navigation satellites
Hyperbolic navigation
Radio navigation
Limits

AGARD Conference Proceedings No.209
Advisory Group for Aerospace Research and Development, NATO
PROPAGATION LIMITATIONS OF NAVIGATION AND POSITIONING SYSTEMS
Published February 1977
350 pages

In order to assess the requirements for navigation and positioning systems, the current status of these systems, and the limitation the propagation medium places on systems in being or planned, the AGARD EPP held a Specialists' Meeting in Istanbul, 20–22 October 1976, on this subject. The concept of the meeting was to outline requirements and progress in systems and to plan programs for future studies to reduce any propagation limitations on navigation systems.

P.T.O.

AGARD-CP-209

Navigation
Electromagnetic wave transmission
Position (location)
Navigation satellites
Hyperbolic navigation
Radio navigation
Limits

AGARD Conference Proceedings No.209
Advisory Group for Aerospace Research and Development, NATO
PROPAGATION LIMITATIONS OF NAVIGATION AND POSITIONING SYSTEMS
Published February 1977
350 pages

In order to assess the requirements for navigation and positioning systems, the current status of these systems, and the limitation the propagation medium places on systems in being or planned, the AGARD EPP held a Specialists' Meeting in Istanbul, 20–22 October 1976, on this subject. The concept of the meeting was to outline requirements and progress in systems and to plan programs for future studies to reduce any propagation limitations on navigation systems.

P.T.O.

AGARD-CP-209

Navigation
Electromagnetic wave transmission
Position (location)
Navigation satellites
Hyperbolic navigation
Radio navigation
Limits

AGARD Conference Proceedings No.209
Advisory Group for Aerospace Research and Development, NATO
PROPAGATION LIMITATIONS OF NAVIGATION AND POSITIONING SYSTEMS
Published February 1977
350 pages

In order to assess the requirements for navigation and positioning systems, the current status of these systems, and the limitation the propagation medium places on systems in being or planned, the AGARD EPP held a Specialists' Meeting in Istanbul, 20–22 October 1976, on this subject. The concept of the meeting was to outline requirements and progress in systems and to plan programs for future studies to reduce any propagation limitations on navigation systems.

P.T.O.

AGARD-CP-209

Navigation
Electromagnetic wave transmission
Position (location)
Navigation satellites
Hyperbolic navigation
Radio navigation
Limits

Accordingly, the meeting reviewed various subjects including civilian requirements for sea, air and ground navigation systems both short and long range, and propagation study needs of Loran C, Omega, NAVSTAR, and Aerostat. In addition, requirements for other systems were outlined and possible propagation problems discussed. Proposed future research and development programs was the subject of a round table discussion. The aim was to mitigate the propagation problems that systems will and have encountered.

ISBN 92-835-0189-6

Accordingly, the meeting reviewed various subjects including civilian requirements for sea, air and ground navigation systems both short and long range, and propagation study needs of Loran C, Omega, NAVSTAR, and Aerostat. In addition, requirements for other systems were outlined and possible propagation problems discussed. Proposed future research and development programs was the subject of a round table discussion. The aim was to mitigate the propagation problems that systems will and have encountered.

ISBN 92-835-0189-6

Accordingly, the meeting reviewed various subjects including civilian requirements for sea, air and ground navigation systems both short and long range, and propagation study needs of Loran C, Omega, NAVSTAR, and Aerostat. In addition, requirements for other systems were outlined and possible propagation problems discussed. Proposed future research and development programs was the subject of a round table discussion. The aim was to mitigate the propagation problems that systems will and have encountered.

ISBN 92-835-0189-6

Accordingly, the meeting reviewed various subjects including civilian requirements for sea, air and ground navigation systems both short and long range, and propagation study needs of Loran C, Omega, NAVSTAR, and Aerostat. In addition, requirements for other systems were outlined and possible propagation problems discussed. Proposed future research and development programs was the subject of a round table discussion. The aim was to mitigate the propagation problems that systems will and have encountered.

ISBN 92-835-0189-6

NATO OTAN
7 RUE ANCELLE · 92200 NEUILLY-SUR-SEINE
FRANCE
Telephone 745.08.10 · Telex 610176

DISTRIBUTION OF UNCLASSIFIED AGARD PUBLICATIONS

AGARD does NOT hold stocks of AGARD publications at the above address for general distribution. Initial distribution of AGARD publications is made to AGARD Member Nations through the following National Distribution Centres. Further copies are sometimes available from these Centres, but if not may be purchased in Microfiche or Photocopy form from the Purchase Agencies listed below.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Coordonnateur AGARD – VSL
Etat-Major de la Force Aérienne
Caserne Prince Baudouin
Place Dailly, 1030 Bruxelles

CANADA
Defence Scientific Information Service
Department of National Defence
Ottawa, Ontario K1A OZ2

DENMARK
Danish Defence Research Board
Østerbrogades Kaserne
Copenhagen Ø

FRANCE
O.N.E.R.A. (Direction)
29 Avenue de la Division Leclerc
92 Châtillon sous Bagneux

GERMANY
Zentralstelle für Luft- und Raumfahrt-
dokumentation und -information
Postfach 860880
D-8 München 86

GREECE
Hellenic Armed Forces Command
D Branch, Athens

ICELAND
Director of Aviation
c/o Flugrad
Reykjavik

ITALY
Aeronautica Militare
Ufficio del Delegato Nazionale all'AGARD
3, Piazzale Adenauer
Roma/EUR

LUXEMBOURG
See Belgium

NETHERLANDS
Netherlands Delegation to AGARD
National Aerospace Laboratory, NLR
P.O. Box 126
Delft

NORWAY
Norwegian Defence Research Establishment
Main Library
P.O. Box 25
N-2007 Kjeller

PORTUGAL
Direcção do Servico de Material
da Forca Aerea
Rua de Escola Politecnica 42
Lisboa
Attn: AGARD National Delegate

TURKEY
Department of Research and Development (ARGE)
Ministry of National Defence, Ankara

UNITED KINGDOM
Defence Research Information Centre
Station Square House
St. Mary Cray
Orpington, Kent BR5 3RE

UNITED STATES
National Aeronautics and Space Administration (NASA),
Langley Field, Virginia 23365
Attn: Report Distribution and Storage Unit

THE UNITED STATES NATIONAL DISTRIBUTION CENTRE (NASA) DOES NOT HOLD STOCKS OF AGARD PUBLICATIONS, AND APPLICATIONS FOR COPIES SHOULD BE MADE DIRECT TO THE NATIONAL TECHNICAL INFORMATION SERVICE (NTIS) AT THE ADDRESS BELOW.

PURCHASE AGENCIES

Microfiche or Photocopy
National Technical
Information Service (NTIS)
5285 Port Royal Road
Springfield
Virginia 22151, USA

Microfiche
Space Documentation Service
European Space Agency
10, rue Mario Nikis
75015 Paris, France

Microfiche
Technology Reports
Centre (DTI)
Station Square House
St. Mary Cray
Orpington, Kent BR5 3RF
England

Requests for microfiche or photocopies of AGARD documents should include the AGARD serial number, title, author or editor, and publication date. Requests to NTIS should include the NASA accession report number. Full bibliographical references and abstracts of AGARD publications are given in the following journals:

Scientific and Technical Aerospace Reports (STAR),
published by NASA Scientific and Technical
Information Facility
Post Office Box 8757
Baltimore/Washington International Airport
Maryland 21240, USA

Government Reports Announcements (GRA),
published by the National Technical
Information Services, Springfield
Virginia 22151, USA

Printed by Technical Editing and Reproduction Ltd
Harford House, 7–9 Charlotte St, London W1P 1HD

ISBN 92-835-0189-6