BEARING ERRORS PRODUCED BY PROPAGATION EFFECTS IN HF DIRECTION FINDING

BERNARD LAWRENCE TEDD

Being a dissertation submitted to the Faculty of Science, Leicester University, in candidature for the degree of Philosophiae Doctor

Department of Physics
University of Leicester
November, 1982
ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to:

a. Professor T.B. Jones for suggesting the area of research presented in this thesis, and for his excellent advice throughout the studentship.

b. Dr. H.J. Strangeways for his efforts in solving the major hardware and software problems involved in tasking the DF, and his keen interest in the analysis and interpretation of the acquired data.

c. The Science and Engineering Research Council for financing my grant during this period of study.

d. GCHQ, Cheltenham for providing much of the equipment used in this investigation, and for their help in the acquisition of the bearing data.

e. My colleagues in the ionospheric physics group, Leicester for their efforts on my behalf, but particularly for their encouragement.

f. The Head of the Physics Department, Leicester for the opportunity of working at the University.

g. The many others who have assisted me by providing ionospheric data, computing facilities, information and advice.
To my family and friends

I love you ....
CONTENTS

GLOSSARY

CHAPTER ONE: Introduction
1.1 THE NEUTRAL ATMOSPHERE 1
1.2 THE IONOSPHERE AND RADIO PROPAGATION 3
   1.2.1 Brief history 3
   1.2.2 Structure of the ionosphere 4
   1.2.3 Remote sensing of the ionosphere 5
1.3 HF RADIO COMMUNICATIONS 6
   1.3.1 The influence of the ionosphere 6
   1.3.2 HF direction finding 7
1.4 THE PRESENT INVESTIGATION 12
1.5 SUMMARY 16

CHAPTER TWO: Data acquisition
2.1 INTRODUCTION 17
2.2 THE DIRECTION FINDER 17
   2.2.1 The aerial array 17
   2.2.2 Beam-forming network 18
   2.2.3 Receiver and Processing Device (DF6) 20
2.3 TASKING THE DIRECTION FINDER 20
   2.3.1 Overview 20
   2.3.2 Communication considerations 22
   2.3.3 The PDP 11/03 Tasking program 24
   2.3.4 System reliability 25
2.4 GRAPHICS OPTION 26
   2.4.1 The graphics terminal 26
2.4.2 Real-time display of data
2.4.3 Displaying bearing data on a map of Europe
2.5 THE BEARING DATA
2.5.1 The recording periods
2.5.2 The monitored signals
2.5.3 Example of the bearing data
2.5.4 Quality of the bearing data
2.6 CONCLUSIONS

CHAPTER THREE : Large-scale Travelling Ionospheric Disturbances
3.1 INTRODUCTION
3.2 PROPERTIES OF ATMOSPHERIC GRAVITY WAVES (AGWs)
  3.2.1 Ideal, isothermal, dissipationless atmosphere
  3.2.2 Realistic atmosphere
3.3 AGW GENERATION
3.4 MID-LATITUDE LARGE-SCALE AGW RESPONSE
  3.4.1 Influence of source characteristics
  3.4.2 The large-scale AGW characteristics
3.5 IONOSPHERIC RESPONSE TO AGWs
3.6 OBSERVATIONS OF LARGE-SCALE TIDs
  3.6.1 Experimental techniques
  3.6.2 Characteristics of large-scale waves
3.7 CONCLUSIONS

CHAPTER FOUR : Experimental observations and HF raytracing of large-scale TIDs
4.1 INTRODUCTION
4.2 TID OBSERVATIONS
  4.2.1 Data analysis
CHAPTER SIX : An Investigation of the Effect of Systematic Ionospheric Tilts (SITs) on HF Bearing Measurements

6.1 INTRODUCTION 144

6.2 SIT-INDUCED BEARING ERROR OBSERVATIONS 144
   6.2.1 Measurements for E-W paths 144
   6.2.2 Measurements for SE-NW paths 150
   6.2.3 Measurements for other paths 150

6.3 RAYTRACING ANALYSES 155
   6.3.1 The 3-D ionospheric model based on predictions (MODEL 1) 160
   6.3.2 A simple model based on a plane effective reflection surface (MODEL 2) 181
   6.3.3 Realistic raytracing investigation 194

6.4 INTERPRETATION OF EXPERIMENTAL OBSERVATIONS IN TERMS OF IONOSPHERIC PROCESSES
   6.4.1 The photochemical regime 201
   6.4.2 Transport of ionization 204
   6.4.3 Anomalous SIT effects 206

6.5 CONCLUSIONS 208

CHAPTER SEVEN : The Reduction of Propagation Bearing Errors in HF Direction Finding

7.1 INTRODUCTION 211

7.2 METHODS OF CORRECTING FOR TID-INDUCED BEARING ERRORS 211
   7.2.1 Doppler Techniques 211
   7.2.2 Reference to bearings of known transmitters 214

7.3 METHODS OF CORRECTING FOR SIT-INDUCED BEARING ERRORS 217
   7.3.1 Regression analysis 217
7.3.2 HF raytracing 218
7.4 WAVE-INTERFERENCE EFFECTS 219
7.5 ASSIGNING VARIANCES TO BEARING MEASUREMENTS 223
  7.5.1 Reference transmitter technique 223
  7.5.2 Quality factor 234
7.6 CONCLUSIONS 237

CHAPTER EIGHT: Conclusions 239
  Suggestions for further study 242

REFERENCES
Glossary

Except where otherwise defined in the text, the meanings of the symbols most often employed are listed below.

A amplitude
AGW atmospheric gravity wave
c speed of sound
DF direction finder
f frequency (signal)
f\text{c} F layer critical frequency
fo\text{E} E layer critical frequency
h height (reflection surface)
hb (real) height of base of F layer
h'F virtual height of base of F layer
h'F\text{2} virtual height of base of F\text{2} layer
hm (real) height of F layer peak
H scale height
HF high frequency
i angle of incidence
io angle of incidence at base of layer
I geomagnetic dip angle
k wave vector
K Kelvin (temperature)
Kp planetary magnetic index
LT local time
M(3000)F2 maximum usable frequency factor for a distance of 3000 km for transmission by the F2 layer
MUF maximum usable frequency
n number, refractive index
N electron density
N-S  north to south (transmission propagation direction)
0 and X  magneto-ionic components
P  phase path
R  path length
Ro  earth's radius
Rx  receiver
S-N  south to north
SIT  systematic ionospheric tilt
TID  travelling ionospheric disturbance
Tx  transmitter
UT  universal time
V  Df variance
ym  semi-thickness of F layer
z  reduced height, height, f/MUF
zo  height of maximum perturbation

Greek symbols
\( \lambda \)  wavelength
\( \Delta \)  elevation angle
\( \alpha \)  tilt
\( \chi \)  solar-zenith angle
\( \beta \)  bearing, perturbation
\( \Theta \)  latitude
\( \phi \)  longitude, phase
\( \Theta' \)  angle between signal path and TID direction
\( \delta \beta \)  bearing error
\( \omega \)  angular frequency
\( \omega_b \)  Brunt-väisälä frequency
\( \omega_a \)  acoustic cut-off frequency
\( \beta_r \)  transmission azimuth (initial)  \( \beta \approx \beta_r -180 \)
\( \sigma \)  standard deviation
\( \delta \)  wave perturbation amplitude
BEARING ERRORS PRODUCED BY PROPAGATION EFFECTS IN HF DIRECTION FINDING

- BERNARD LAWRENCE TEDD

ABSTRACT

The intrinsic high accuracy of modern HF direction finders cannot be realized in practice due to the perturbations and tilts which exist in the ionosphere. Particular attention is given in this dissertation to the effect of Travelling Ionospheric Disturbances (TIDs) since their occurrence is difficult to predict and the magnitude of the induced bearing error extremely variable.

TID activity is quantified by monitoring simultaneously the bearings of a number of transmitters whose locations and frequencies are well known. This information is then employed to determine the expected bearing error on a target transmission reflected in the same area of the ionosphere. The major limitations of this correction scheme are investigated.

The large-scale temporal and spatial variations of the quiet ionosphere (Systematic Ionospheric Tilts, SITs) produce bearing errors of similar magnitude to those due to TIDs. Rapid calculation of these errors for various path geometries and ionospheric conditions is possible by raytracing through 3-D ionospheric models. The limitations of a correction scheme which employs a 3-D model based on ionospheric predictions are examined.

Variance is a particularly useful statistic in assessing the 'reliability' of bearings made on a particular HF circuit. A scheme is proposed in which the bearings of a network of reference transmitters are monitored continuously allowing a variance to be assigned to any target transmission, even if it is only present for a very short time. These variance values are of considerable value in practical DF applications.

Whenever possible the experimental observations are related to the underlying physical processes in the ionosphere which generate these DF errors.
CHAPTER ONE

INTRODUCTION

1.1 THE NEUTRAL ATMOSPHERE

The neutral atmosphere is subdivided according to various criteria based on temperature, composition and state of mixing, as illustrated in Fig. 1.1. The temperature classification is generally the most useful one. The coldest part of the atmosphere has a temperature of $\sim 180^\circ K$ and is located at the mesopause (80-85 km), while the hottest part (over $1000^\circ K$) occurs at altitudes above about 400 km in the thermosphere. The well-mixed part of the atmosphere, in which composition does not change with height, is called the turbosphere (or the homosphere). The composition is similar to that found at ground level. Above this region is the heterosphere where mixing is absent and the composition changes with height. Lighter gases become relatively more abundant at greater heights due to diffusive separation.

The atmosphere is in a continual state of motion due to the existence of a variety of wind and wave motions. At thermospheric heights, the winds are driven by horizontal gradients of air pressure, which are in turn caused by horizontal temperature gradients. These gradients are produced by the thermal expansion of the atmosphere during the day (the diurnal bulge). A large number of different wave motions can also exist in the atmosphere. Planetary waves have wavelengths
Fig. 1.1  Nomenclature of the upper atmosphere  
(after Hargreaves, 1979)

Fig. 1.2  Typical vertical electron density profiles of the mid-latitude ionosphere  
(after Hargreaves, 1979)
of thousands of kilometers and are known to play an important role in global weather systems. An atmospheric tide is a perturbation due to an external agent such as the Sun, to which it maintains a fixed relationship while the earth rotates beneath. A solar tide has a period of 24 hours or a submultiple (12, 8, 6 etc.). Another important class of wave is the acoustic-gravity wave which is influenced by both compressional and gravitational forces. These waves have wavelengths of tens to several thousands of kilometers, periods of minutes to hours, and travel at speeds related to the speed of sound. At the long-period end of the acoustic-gravity wave spectrum, the waves are called internal gravity waves and are of particular interest to the present investigation.

The upper atmosphere (at heights greater than 50-70 km above the earth's surface) is ionised by energetic EUV and X radiations from the Sun. Once produced, the free electrons and ions tend to recombine and a balance is established between electron production and loss. The net concentration of free electrons is greatest at heights of several hundred kilometers. Although the electron concentration at these heights amounts to only one percent of the concentration of neutral gases, the ionised component introduces important features into the upper atmosphere. Such phenomena include electric currents and fields, plasma processes and the reflection of radiowaves.

1.2 THE IONOSPHERE AND RADIO PROPAGATION

1.2.1 Brief History

The existence of electromagnetic waves was predicted
around 1860 by James Clerk Maxwell (1873). The validity of this theory remained in doubt until 1887, when it was established through the experiments of Hertz (1893). By this time, Balfour Stewart, in a celebrated article in Encyclopedia Britannica in 1878, provided evidence for the existence of the earth's ionosphere. Stewart suggested that the most probable cause of the temporal variations in the earth's magnetic field was the presence of electric currents flowing in the upper atmosphere.

The practical implications of the experiments of Hertz were realized at the turn of the century by Guglielmo Marconi. On December 31, 1901 Marconi succeeded in transmitting and receiving signals from Cornwall, England to Newfoundland in the most famous of all radio-propagation experiments. Kennelly (1902) and Heaviside (1902) independently suggested that radio waves were returned to the earth's surface after reflection by an electrically conducting layer in the upper atmosphere.

The realization of the importance of high-frequency radio propagation in the early 1920's led to renewed interest in the properties of the Kennelly-Heaviside layer, as the ionosphere was then called. Appleton and Barnett (1926) in England and Breit and Tuve (1926) in the U.S.A. performed experiments to determine the height of the conducting layer. Their work inferred that the ionosphere was stratified and Appleton introduced the letters E and F, and later D to denote the different ionospheric layers.

1.2.2 Structure of the ionosphere

The existence of these layers has since been confirmed and
typical vertical profiles of electron density are reproduced in Fig. 1.2. For a detailed explanation of the ionospheric structure, account must be taken of the atmospheric composition, of the solar spectrum of radiation and of the detailed photochemistry. A simple theory of layer formation was developed by Chapman (1931), and although the basic assumptions are not strictly valid, the theory has been an invaluable guide in the analysis of experimental data.

The temporal and spatial behaviour of the ionospheric E and F1 layers is subject to close solar control. However, the F2 layer has a number of anomalous features, and to explain the existence of the layer at all, the vertical diffusion of plasma needs to be considered (not included in the Chapman theory). In order to interpret the F2 layer behaviour during both quiet and geomagnetically disturbed conditions, a number of other processes need to be examined. These include, i) chemical changes, e.g. variations in the relative concentrations of the atomic and molecular neutral species, ii) diurnal heating and cooling of the ionospheric plasma, iii) neutral wind effects, e.g. winds tend to lower the F layer during the day and raise the layer at night, iv) electric field effects.

1.2.3 Remote sensing of the ionosphere

The earliest techniques for observing the ionized regions of the upper atmosphere made use of the propagation of radio waves. In the 1930s the ionosonde technique was developed and has since played an important role in the systematic study of the long-term variations of the ionosphere. The principle of
ionospheric sounding is to transmit a pulse of radio waves vertically and to measure the time which elapses before the echo is received from the reflecting layer. The time of flight provides a measure of the height of reflection, and the frequency of the reflected radiowave is equal to the plasma frequency \[ \propto N^\frac{1}{2} \] at the height of reflection. An important variation on ionospheric sounding is the HF Doppler technique. Perturbations of the reflecting layer produce a change in the phase path of the wave which can be monitored by measuring the time-varying Doppler shifted frequency of the reflected radio wave. Many other techniques have been employed in the study of the ionosphere, the latest of these is based on the incoherent scatter of VHF or UHF radio waves. The EISCAT facility is such a system which has recently been constructed to study the complexities of the ionosphere at high latitudes.

1.3 HF RADIO COMMUNICATIONS

1.3.1 The influence of the ionosphere

Despite the development of satellite communications and improvements to submarine cables, the total volume of HF radio traffic is larger than ever before, and a substantial research effort is still being devoted to the development of improved ionospheric predictions. A basic task of a prediction scheme is to estimate the frequency spectrum available to the operator at any given time on a particular HF circuit. Computer programs have been developed to provide predictions of circuit parameters such as path MUFs (maximum usable frequencies),
signal strengths and circuit reliabilities. These programs employ numerical maps of monthly median ionospheric parameters which are based upon data collected over many years from vertical-incidence ionosondes.

Greater problems exist in forecasting the effect of the short-term variability of the ionosphere on an HF communications circuit. Disturbances of solar origin, e.g. simultaneous ionospheric disturbances (SID), polar cap absorption events (PCA) or ionospheric storms, can greatly disrupt communications. Nevertheless, short-term propagation forecasts are routinely made by a number of organizations and are based on experience of recent radio conditions and the trend of solar-geophysical activity. Besides the regular forecasts, rapid warnings can also be provided on the occurrence of particular kinds of events, e.g. solar flares.

Irregularities such as sporadic E, scintillations, spread F and Travelling Ionospheric Disturbances (TIDs) also degrade the performance of HF communication circuits by focusing or de-focusing the signal and by other forms of fading. It is difficult to predict these effects in detail, but some form of compensation for them may be achieved by monitoring the ionosphere in real time.

1.3.2 HF direction finding

HF direction finding (Ross, 1949; Hopkins and Pressey, 1958; Gething, 1966, 1978; Jones and Reynolds, 1975) has applications in both the civil and military sectors, e.g. search and rescue operations following ship or aircraft disasters at sea. Many
problems are more severe in direction finding than in navigational position finding systems since transmissions on which bearings (defined as the azimuth angle of arrival of a ray, measured clockwise from north) are taken are usually non-co-operative, e.g. the signal may be received for a very short time and other signal parameters such as frequency may not be optimum at all receiving locations.

The ideal direction finder (DF) is capable of working over a wide frequency band, $360^\circ$ of azimuth and $90^\circ$ of elevation, of dealing with all forms of modulation, and of giving accurate and reliable bearings, even on very short transmissions (Gething, 1966). Unfortunately, no one type of DF satisfies all the above requirements and the choice of the system employed depends on the research or operational requirements.

A simple vertical loop can be rotated about a vertical axis to find a position giving a signal minimum in a receiver connected to it. If the waves from the target transmitter (of unknown location) were to travel in a horizontal plane, the normal to the loop would lie along the great-circle path to the transmitter (or in the reciprocal direction, i.e. $\pm 180^\circ$). However, early workers realised that such an arrangement is unsuitable for use on skywave signals, which arrive from unknown elevation angles and of mixed polarisation after their passage through the ionosphere. The position of a signal minimum can give an indicated bearing grossly in error (typically $30^\circ$, where bearing error is defined as the difference between the measured and the true great-circle bearings) due to polarisation effects (Smith-Rose and
These authors also recognised that the Adcock DF, patented in 1919, would be very much less susceptible to polarisation error. The early type of Adcock system consisted of four vertical elements at the corner of a square (aperture of about six metres), with buried feeders leading to a central hut. Adcock systems were employed extensively during the Second World War and as instrumental errors were identified and reduced, the importance of errors arising from site imperfections and wave-interference effects became greater. Interest grew in systems working on a larger portion of the arriving wavefront, i.e. wide aperture systems which have apertures comparable to or greater than one wavelength.

Most wide-aperture systems employ circular rings of identical vertical elements and three main types can be distinguished. In the Wullenweber type, the outputs from a sector of elements are combined to form a lobe pattern of desired shape. A Doppler DF is a system in which a frequency modulation is imposed on the signal by moving a single receiving element rapidly round the perimeter of a circle (Whale, 1954). In practice, a simulated Doppler effect can be obtained by smooth commutation round a ring of fixed elements. The third type is the Earp and Godfrey (1947) system in which the bearing information is obtained from phase comparisons between successive aerials in the circular ring of aerials. The disadvantage of the last two types is that they suffer from a 'capture' effect, which tends to suppress evidence of the existence of weaker signals in the presence of strong
interfering signals.

Interferometer systems have also been developed, where angle of arrival information is obtained by the measurement of the phase difference between two spaced aerials. A single phase measurement defines a cone on which the normal to the phase front must lie. Two intersecting cones, derived from two orthogonal arrays of aerials, will define the angle of arrival in both azimuth and elevation. Very wide aperture (greater than ten wavelengths) linear arrays have been constructed mainly for research purposes, and can provide instant resolution of $0.1^\circ$ at the lower end of the HF band (Sweeney, 1970; Rice, 1971). In some direction finding systems, mode selection and rejection techniques have been suggested to improve the quality of received signals by reducing multipath distortion and interference. These systems have various disadvantages, however, such as restrictions of the azimuth and/or frequency coverage, moreover, ambiguities in determining cone angles arise if the spacing of the aerial pairs is large compared with the wavelength of the signal.

In view of these difficulties, the Wullenweber has attracted increasing attention and seems to offer the closest approach to the ideal DF (when only azimuth angles are to be determined). Systems with apertures of about 300 m have been built mainly for research (e.g. Hayden, 1961). Such instruments installed on a good site can measure the azimuth of arrival of a single ray to an accuracy of about $0.1^\circ$. Operational Wullenwebers, such as the one employed in the present investigation, have smaller apertures (about 100 m) and
therefore have reduced accuracies (typically $0.5^\circ$). These systems provide a useful suppression of site errors, but apertures of 4-5 wavelengths are required before a worthwhile reduction in wave interference is achieved.

The intrinsic high accuracy of modern HF DFs cannot generally be realized due to the nature of radio wave propagation through the ionosphere. The major types of propagational bearing error are as follows:

a. **Magneto-ionic deviations**

A radio wave entering the ionosphere is split into ordinary and extraordinary components which then follow independent paths. Lateral path deviations are experienced by these components and bearing errors are produced in the received signal. These errors are generally very small ($< 0.1^\circ$), except when the signal path length is shorter than about 500 km or when the signal frequency is within a few percent of the path MUF (Rao, 1969; Gething, 1978).

b. **Wave-interference errors**

Wave interference occurs when more than one wavefront arrives simultaneously at the DF via different 'paths' or modes of propagation (Ross, 1949; Bain, 1953, 1956). Large fluctuations in the bearing (of many degrees in time scales of a fraction of a second to a minute) can occur as the relative phase between modes change. This type of error is associated with the normal fading of a multimoded signal.

c. **TID-induced bearing errors**

The passage of atmospheric gravity waves through the ionosphere produces wave-like distortions in the isoionic contours,
referred to as Travelling Ionospheric Disturbances (TIDs). As a TID moves across a HF signal reflection region the observed bearing can exhibit quasi-periodic variations of several degrees (Bramley and Ross, 1951; Bramley, 1953; Morgan, 1972; Jones and Reynolds, 1975). TIDs are often classed either as medium- or large-scale waves and a comparison of their properties is reproduced in Table 1.1.

d. Systematic Ionospheric Tilt (SIT) errors
SITs are the large-scale temporal and geographical variations of electron density in the quiet ionosphere. They have magnitudes of several degrees and can induce similar variations in the observed bearing. Errors of this type commonly occur at sunrise and sunset and are reproduced from day to day in a consistent manner (Ross and Bramley, 1947; Titheridge, 1958; Morgan, 1974).

1.4 THE PRESENT INVESTIGATION

The University of Leicester and GCHQ, Cheltenham have co-operated over a number of years in the study of propagation bearing errors and the present investigation is a continuation of this joint effort.

In operational DF some form of correction is required for the bearing errors produced by TIDs and SITs, but failing this some estimate of the bearing reliability and the expected bearing error limits would be of value. The major topics addressed in this thesis are outlined below.

a. Correction for TID-induced errors
In the past, attention has largely been confined to the
<table>
<thead>
<tr>
<th>Property</th>
<th>Medium scale</th>
<th>Large scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed</td>
<td>$&lt; 300 \text{ m s}^{-1}$</td>
<td>$&gt; 300 \text{ m s}^{-1}$</td>
</tr>
<tr>
<td>Period</td>
<td>10 to 40 min</td>
<td>$&gt; 30 \text{ min}$ often exceeds 1 hour</td>
</tr>
<tr>
<td>Occurrence</td>
<td>mainly day time</td>
<td>day and night, mostly at night</td>
</tr>
<tr>
<td>Relationships with geophysical events</td>
<td>none</td>
<td>associated with geomagnetic activity</td>
</tr>
</tbody>
</table>

**Characteristic features:**

(a) Number of cycles  
- single events or trains 2 or 3 cycles only
- of waves

(b) Wave front tilt  
- $\sim 45^\circ$  
- nearly horizontal

(c) Shape  
- changes over distances of about 100 km  
- retained for thousands of kilometres

(d) Velocity/period relationship  
- velocity increases with period  
- velocity increases to a maximum then decreases with increasing period

---

Table 1.1 Comparison of the properties of medium- and large-scale TIDs (after Jones and Reynolds, 1975)
correction of TID-induced bearing errors and considerable success has been achieved in this area by means of Doppler techniques. These error correction schemes require equipment in addition to the DF, whereas a minimum of auxiliary equipment is desired in a practical DF system.

The present study considers a correction scheme based on the assumption that TID activity can be quantified by monitoring the bearings of a number of transmitters whose locations, schedules and operational frequencies are well known (reference transmitters). By comparing the time delays in the observed bearing error signatures on the various paths, the speed and direction of a TID is determined and this information is then employed to determine the expected error on a target transmission reflected in the same area of the ionosphere. The feasibility of this scheme is examined by:

1. The acquisition of a large bearing data set. An HF DF is interfaced to a PDP 11/03 minicomputer to enable the bearings of a number of HF signals to be measured on a sequential basis, under program control.

2. The examination of the degree of correlation of TID-induced bearing error time signatures measured for geographically spaced paths. The observed correlation (in the case of large-scale TIDs) is interpreted in terms of present theories of the generation and propagation of AGW/TIDs and the response characteristics of the HF direction finding technique.

3. The identification of the limitations involved in monitoring transmitters not under the direct control of the DF authority.
b. Correction for SIT-induced errors
An examination of the data collected has indicated that SITs produce errors at least as great as those due to TIDs. In the past, a number of theoretical studies have been performed to investigate the lateral path deviation effects caused by SITs. These include raytracing analyses which employ 3-D ionospheric models based either on vertical-sounder ionospheric data or on numerical prediction maps of ionospheric coefficients. The present study considers in detail the feasibility of employing these techniques to correct for SIT-errors in practical DF.

Previous explanations of the diurnal behaviour of SIT-induced bearing errors have been based solely on the solar-zenith angle control of the ionosphere. Many other processes are important in determining the large-scale temporal and geographical variations of the quiet ionosphere and these are discussed with reference to the experimental observations.

c. Assigning a variance
The concept of variance is particularly useful in assessing the 'reliability' of bearings made on a particular HF circuit. Bearing accuracy can be correlated with the modes of propagation and the ionospheric conditions. A scheme for computing variances is proposed in the present study. By continuously monitoring the bearings of a network of reference transmitters, variances corresponding to different propagation conditions can be determined. A variance can then be
immediately assigned to any target transmission even if it is only present for a very short time. These variance values are of considerable value in practical DF applications.

1.5 SUMMARY

The work presented in this dissertation is directed at a better understanding of the propagation errors present in HF direction-finding systems. Particular attention is given to the effect of TIDs since their occurrence is difficult to predict and the magnitude of the induced bearing error extremely variable. The influence of the large-scale systematic ionospheric tilts is also considered and the feasibility of correcting for these errors is determined. The reliability of the bearings likely to be encountered on any particular propagation path is quantified by a variance value. Whenever possible the experimental observations are related to the underlying physical processes in the ionosphere which generate the DF errors.
CHAPTER TWO

DATA ACQUISITION

2.1 INTRODUCTION

The experimental arrangement employed in a series of experiments conducted between October 1979 and April 1981 is illustrated schematically in Fig. 2.1. An HF direction finder is interfaced to a PDP 11/03 minicomputer to enable the bearings of a number of HF signals to be measured on a sequential basis, under program control. This chapter describes how the DF is controlled by the PDP 11/03, the details of the transmissions monitored and the data acquired. The bearing data are stored on floppy disk and are later transferred to the Leicester University mainframe computer to facilitate analysis. The option of displaying the bearing data on a graphics terminal during data acquisition is also discussed.

2.2 THE DIRECTION FINDER

2.2.1 The Aerial Array

A circular wide aperture array (Wullenweber) situated at Benhall, UK was employed in the experimental investigations. The array consists of two sets of twenty-four aerials (elevated feed monopoles) equispaced around concentric circles of 50 and 150m diameter respectively. Satisfactory radiation patterns from a single ring of aerials are difficult to maintain over
the frequency range 1-30 MHz, therefore, two rings are employed; the inner and outer rings cover the 10-30 and 1-10 MHz bands respectively. Each ring of aerials is connected to its own beam-forming network by feeder cables of equal electrical length, so that the relative phases of any signal received at the aerials are preserved at the inputs of the network.

2.2.2 Beam-forming network

The network synthesises twenty-four independent beams; each beam is derived from the signals of a group of eight adjacent aerials around the circle. The formation of a typical beam (signals from aerials 3 to 10) is illustrated in Fig. 2.2. The beam is arranged to give maximum response to signals arriving from the direction shown, that is, perpendicular to the chord drawn through the outermost aerials of the array. The maximum response occurs if the signals at all eight aerials are in phase, however, as the aerials lie on a arc of a circle, the signals will not be in phase. The response is optimized by introducing time delays into the signal paths by means of suitable lengths of coaxial cable. These delays are equal to the various times that a wave, travelling in the direction indicated in Fig. 2.2, takes to traverse the distances D1, D2 and D3.

The beam-forming network consists of a spinning goniometer, a capacitative switch, that selects the aerials employed at any instant. A 'difference' network can be employed to generate a goniometer 'rabbits-ears' pattern, as illustrated.
Fig. 2.1 Schematic representation of the experimental arrangement.

Fig. 2.2 Schematic plan of the aerial system (after Starbuck, 1969)
in Fig. 2.3. A 'sum' radiation pattern is similarly generated by means of a 'sum' network. In practice, the number of rotor blades is larger than shown in Fig. 2.3, in order to achieve smoother commutation.

2.2.3 Receiver and Processing Device (DF6)

To obtain a bearing measurement the receiver is tuned to the desired frequency and the goniometer output is automatically processed to compute the maximum of the 'sum' pattern for each bearing. A second algorithm computes a 'quality-factor' which effectively tests the closeness with which the output from the goniometer fits the expected polar diagram. The maximum and minimum values of the quality factor are normalised to 0 and 99 respectively (a small value corresponds to a good fit, a large value corresponds to a poor fit).

The DF system described above (also see Starbuck, 1968; Hockley, 1973 and Gething, 1978) has an instrumental accuracy of approximately 0.5 degrees. Larger aperture arrays have a better performance (instrumental accuracy of approximately 0.1 degrees) but suffer from their high cost and size.

2.3 TASKING THE DIRECTION FINDER

2.3.1 Overview

A Fortran program has been developed for the PDP 11/03 minicomputer (Strangeways, 1979) that allows the DF to be
Fig. 2.3. Formation of rotating 'rabbit-ears' radiation pattern with goniometer (from Gething, 1978)
tasked under program control. An environment was created (through a system generation to create a PDP 11/03 foreground/background monitor with multi-terminal access) to enable the PDP 11/03 to communicate with five remote terminals (via five lines or channels). The lines are assigned as in Table 2.1. Control of the tasking program is achieved through a teletype (decwriter). Separate input and output lines with the DF6 were found necessary to prevent unwanted interaction between these lines.

2.3.2 Communication Considerations

A number of communication problems between the DF6 and PDP 11/03 computer systems were encountered due to the differences in their normal working modes (Strangeways, 1979).

a. The PDP 11/03 is interfaced to a 'Master' DF6 which also communicates to a 'Slave' DF6 at 75 baud. The PDP 11/03 4-channel asynchronous serial interface board (DLV11-J) is unable to be set at 75 baud and is conveniently set at 300 baud. Unfortunately, the replacement of the respective DLV11 single-channel boards (set at 75 baud) in the Master and Slave DF6s by DLV11-F single-channel boards (set at 300 baud) was only possible in the case of the Master DF6. Nevertheless, communication was achieved by incorporating a 'sleep' routine in the PDP 11/03 tasking program which suspends program execution for 1/12 second between each character sent to the DF6.

b. A difference exists in both the coding and the character length employed by the DF6 (5 bits/character BAUDOT)
<table>
<thead>
<tr>
<th>CHANNEL OR LINE</th>
<th>COMMUNICATION PURPOSE</th>
<th>BAUD RATE</th>
<th>BITS PER CHARACTER</th>
<th>INTERFACE BOARD</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Teletype / Decwriter</td>
<td>300</td>
<td>8</td>
<td>DLV11-J</td>
</tr>
<tr>
<td>3</td>
<td>Redundant</td>
<td>300</td>
<td>8</td>
<td>DLV11-J</td>
</tr>
<tr>
<td>2</td>
<td>Line from DF6</td>
<td>300</td>
<td>5</td>
<td>DLV11-J</td>
</tr>
<tr>
<td>1</td>
<td>Line to DF6</td>
<td>300</td>
<td>5</td>
<td>DLV11-J</td>
</tr>
<tr>
<td>4</td>
<td>Graphics Terminal</td>
<td>9.6K/300</td>
<td>8</td>
<td>DLV11-F</td>
</tr>
</tbody>
</table>

Table 2.1 Assignment of channels (lines)
and the PDP 11/03 (generally 8 bits/character ASCII). The tasking information is sent to the DF6 in binary coded decimal and the bearing data are received back from the DF6 in Baudot code (converted to ASCII by the tasking program). For compatibility channels 1 and 2 of the PDP 11/03 DLV11-J interface board were customised from the normal 8 bit channels to produce 5 bit channels. In addition, characters are sent to the DF6 as one character words (achieved by defining the characters as 'logical*1' quantities in the tasking program) since Baudot code comprises one character words.

c. The 'normal mode' of operation employed between the PDP 11/03 and a remote terminal (DF6) is such that a current line of characters (bearing data from the DF6) is unavailable to the PDP 11/03 tasking program until a carriage return (or a line feed) is received. However, the PDP 11/03 does not recognise a carriage return in Baudot code. This difficulty is overcome by employing the 'immediate' mode of communication (by suitably altering the PDP 11/03 Job Status Word) enabling the characters, sent by the DF6, to be immediately available to the tasking program.

2.3.3 The PDP 11/03 Tasking Program

The tasking program instructs the direction finder to determine the bearings of a number of HF signals on a sequential basis. The input parameters of the program are:

a. The required signal frequencies.

b. The receiver bandwidths for the respective signals (default=750 Hz).

c. The start and stop times (UT) of the data acquisition
Several options are incorporated in the program enabling:

a. The set of signal frequencies to be changed during program execution at specified times. This option is particularly useful when no operator is present and the frequencies need to be altered to accommodate changes in ionospheric conditions (for example, lower frequencies are required at night because of the lower MUFs).

b. The insertion of a specified time delay between each cycle of bearings for each of the transmitters monitored. This enables the data collection rate to be reduced and thus allows collection to be continued beyond the maximum of 20 hours (imposed for continuous data collection by the storage capacity of the floppy disk).

c. The display of bearing data on a graphics terminal during data acquisition (see section 2.4).

2.3.4 System reliability

Once the initial communication problems were overcome the system proved to be reliable during data aquisition. A few minor hardware faults occurred during the recording periods but were easily rectified. Small amounts of bearing data were lost due to the system crashing which was attributed to the DF6 calibration routine causing the tasking program to fail. This problem was remedied by increasing the time allowed for the DF6 to calibrate itself and this delay was incorporated into the tasking program. The system also crashed on several other occasions due to local power failures, however, these events were out of our control.
2.4 THE GRAPHICS OPTION

2.4.1 The graphics terminal

The raster-scan Sigma 5670 graphics unit (resolution 768 by 512) possesses a refresh facility and is able to store information on up to three pixel planes. Each plane can be viewed separately or superimposed on one or both of the other pixel planes. A split-baud rate (DLV11-F interface board) between the PDP 11/03 and the graphics terminal is employed. The faster rate to the graphics terminal (9.6 kbaud) allows the rapid display of the information sent to the graphics terminal.

2.4.2 Real-time display of data

The graphics terminal enables the bearings of a number of signals to be displayed in real-time. These are represented as displacements from the respective true great-circle bearings as a function of time. The bearing data for three signals are stored on each of the three pixel planes so that the bearings of a total of nine signals can be displayed. A maximum of four hours data are stored on each pixel plane, after which time the graphics unit is commanded to output these four hours of data to a hardcopy device. A new time axis is then drawn four hours advanced from the previous start time and plotting is continued from the left-hand side of the graphics screen. Displaying bearing data in real time provides the operator with information concerning the propagation conditions (for example, transmitters ceasing to transmit, the presence of TID activity, interference conditions). If circumstances dictate, the
operator can then change the signals being monitored.

2.4.3 Displaying bearing data on a map of Europe

A routine for drawing World maps (provided by the Leicester University Computer Laboratory) was incorporated into a program developed for the PDP 11/03. The program draws bearing data (stored on a file) on a map of Europe. A gnomic projection is employed to ensure that great-circle paths appear as straight lines. The centre of projection of the map is chosen to coincide with the location of the direction finder, to ensure that bearing angles are conserved. Bearing data, for any number of signals, can be displayed for a stipulated time interval. The bearings are continuously updated by drawing the most recent bearing and deleting the least recent bearing. The variability of bearing data during a given time interval can be clearly discerned by means of this program. A thirty minute period of data plotted on the graphics unit is displayed in Fig. 2.4. The program could easily be adapted in the future to enable bearing data to be displayed in this form directly from the DF6.

2.5 THE BEARING DATA

2.5.1 The recording periods

Approximately 700 hours of data were recorded during the following periods:

a. 16-24 October 1979
b. 7-21 May 1980
c. 12-20 June 1980
A summary of the data recorded is presented in Fig. 2.5. This figure also indicates the large-scale TID events observed in the bearing data (see chapter 4) and the geophysical conditions during the recording periods ($R_z$ is the Zurich sunspot number, $K_p$ is the planetary magnetic index). Data were obtained during a wide range of geophysical conditions (as indicated by the range of values for $K_p$ and $R_z$). The March recording session was selected to coincide with geomagnetically disturbed conditions.

The system was generally run at 300 baud, except for the June recording session, when the system was run at 75 baud due to operating restrictions. A few modifications were necessary to run the system at 75 baud but the basic configuration remained the same. When monitoring eight signals, 48 and 32 bearings per hour can be determined for each signal at the faster and slower baud rates respectively. A total of 10-11 seconds are required for each bearing determination comprising 64 goniometer revolutions (8 seconds) and the time taken to task the direction finder (2-3 seconds).

2.5.2 The monitored signals

Over 100 different frequencies were monitored at different times during the six recording periods. Table 2.2 lists the details of the signals that were most frequently monitored and Fig. 2.6 indicates the locations of some of these signals. The
Fig. 2.5a  Summary of data recorded:
Fig. 2.5b: Summary of data recorded.
<table>
<thead>
<tr>
<th>TRANSMITTER</th>
<th>FREQUENCY (MHz)</th>
<th>PATH LENGTH (km)</th>
<th>TRUE GREAT-CIRCLE BEARING</th>
<th>TYPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luxembourg</td>
<td>6.09</td>
<td>633</td>
<td>109.3</td>
<td>B</td>
</tr>
<tr>
<td>Issoudun</td>
<td>6.175</td>
<td>603</td>
<td>148.6</td>
<td>B</td>
</tr>
<tr>
<td>Swiss Radio International (S.R.I)</td>
<td>9.535</td>
<td>931</td>
<td>122.4</td>
<td>B</td>
</tr>
<tr>
<td>Madrid</td>
<td>9.57</td>
<td>1303</td>
<td>185.3</td>
<td>B</td>
</tr>
<tr>
<td>Geneva</td>
<td>9.943</td>
<td>860</td>
<td>131.8</td>
<td>FSK</td>
</tr>
<tr>
<td>Vienna</td>
<td>10.1185</td>
<td>1393</td>
<td>100.7</td>
<td>FSK</td>
</tr>
<tr>
<td>Prague</td>
<td>10.125</td>
<td>1220</td>
<td>92.3</td>
<td>FSK</td>
</tr>
<tr>
<td>Moscow</td>
<td>10.92</td>
<td>2586</td>
<td>64.8</td>
<td>FSK</td>
</tr>
<tr>
<td>Berlin</td>
<td>11.459</td>
<td>1074</td>
<td>81.3</td>
<td>FSK</td>
</tr>
<tr>
<td>Stockholm</td>
<td>11.705</td>
<td>1279</td>
<td>48.5</td>
<td>B</td>
</tr>
<tr>
<td>Wertachtal</td>
<td>11.72</td>
<td>1050</td>
<td>113.6</td>
<td>B</td>
</tr>
<tr>
<td>Madrid</td>
<td>11.92</td>
<td>1333</td>
<td>184.9</td>
<td>B</td>
</tr>
<tr>
<td>Copenhagen</td>
<td>15.165</td>
<td>1038</td>
<td>60.3</td>
<td>B</td>
</tr>
<tr>
<td>Lisbon</td>
<td>16.281</td>
<td>1563</td>
<td>203.5</td>
<td>FSK</td>
</tr>
</tbody>
</table>

B — BROADCAST  FSK — FREQUENCY SHIFT KEYING

Table 2.2 Some of the more frequently monitored transmitters
transmitters that were monitored are in the frequency range 6-18 MHz and have ranges of 600 to 3000 km. The midpoints of the signal paths lie in the geographic area 46 to 56 degrees North and -3 to 17 degrees East corresponding approximately to 1500 km in the North-South and East-West directions respectively. The separation of the path midpoints of the signals range between 100 and 1500 km.

The criteria for choosing the signals for a particular recording session were:

a. The transmitters are situated in Western Europe and their locations are known.

b. The signal frequencies are high enough to ensure dominant single-moded propagation conditions.

c. The signals are reliably received for a reasonable length of time (of the order of hours) since changing the frequencies destroys the continuity of the bearing data.

In practice, it was difficult to find a large number of signals that satisfied all of these criteria, particularly during the night-time recording sessions. This difficulty is discussed in chapter 7 in relation to the feasibility of correcting for TID-induced bearing errors.

2.5.3 Example of the bearing data

Eight signals were monitored in the recording session of March 26, 1981 and the data recorded between 1700-2200 UT is reproduced in Fig. 2.7. This figure is referred to in later chapters, but is presented here to illustrate the nature of the bearing data obtained. The bearing data are plotted about the
FIG. 2.7a Untiltered bearing data 26 March 1981 (event 12)

Monitored different signal
mean of the respective data sets for each of the eight signals. The temporal interval between each datum point, denoted by a circle, is 90 seconds. The horizontal dashed line represents the true great-circle bearing of the transmitter at the DF. Bearings that lie outside the ordinate range are represented as a 'W' indicating a 'wild' bearing. An 'X' denotes a time for which the DF failed to determine a bearing.

The locations of two of the transmitters, TX1 and TX2, are unknown. Initially, signals of known location are monitored but when reception of these ceases (the transmitters may be turned off) the DF picks up the next strongest signal at each respective frequency. At 2045 UT the frequency set was changed by program control. However, on this occasion only the 11.92 and 9.535 MHz signals were changed. The choice of the new frequencies were chosen many hours earlier when the system was set up for the recording session. Unfortunately, due to operating regulations, it was not possible to monitor the data acquired after 1700 UT. Hence, during the night-time recording sessions, changing the frequencies to better suit the ionospheric conditions was only possible by program control and without direct reference to the quality of the data currently being recorded.

2.5.4 Quality of the bearing data

The great variability of the quality of the bearing data is evident from Fig. 2.7. The cause of the poor data for the 11.995 and 9.77 MHz transmissions is uncertain since their transmitter locations are unknown. The three major types of bearing error produced by propagation effects are particularly
evident in the record for the Luxembourg 6.09 MHz transmission as indicated below:

a. A noise-like fluctuation of the bearing which is produced by wave-interference effects.

b. An increasing negative trend in the bearing error which results from the presence of systematic ionospheric tilts.

c. Oscillations with a quasi-period of about 50 minutes which result from TID activity.

2.6 CONCLUSION

The control of the DF by means of the PDP 11/03 minicomputer allowed the bearings of a number of signals to be monitored simultaneously over extended periods of time. The bearing data were transferred to a more powerful computer to facilitate analysis. In the following chapters the results of the study of this data set are presented.
CHAPTER THREE

LARGE-SCALE TRAVELLING IONOSPHERIC DISTURBANCES (TIDs)

3.1 INTRODUCTION

Travelling ionospheric disturbances (TIDs) are the ionospheric response to atmospheric gravity waves (AGWs) in the neutral atmosphere (Hines, 1960). Many authors have presented review papers on AGW theory and observational results, e.g. Georges, 1967; Testud, 1973; Yeh and Liu, 1974; Hines, 1974; Francis, 1975; Richmond and Roble, 1979; Hunsucker, 1982. This chapter describes the characteristics of large-scale disturbances which are generated in the high-latitude ionosphere during geomagnetically disturbed conditions. They are long-period, high-speed waves that travel from high to low latitudes along a broad front and are generally dissipated before they reach the opposite hemisphere.

3.2 PROPERTIES OF AGWS

3.2.1 Ideal isothermal dissipationless atmosphere

Hines (1960) determined the properties of plane acoustic-gravity waves by considering the linearized equations of motion of an ideal isothermal dissipationless neutral atmosphere. Perturbations to the ambient atmosphere are assumed to have a space and time dependence proportional to \( \exp(iwt-ik_x x-ik_z z+z/2H) \) where \((x,y,z)\) is a rectangular
cartesian coordinate system, x and z represent the horizontal and vertical axes respectively. The dispersion relation is given by

$$k_z^2 = \left( \frac{\omega_b^2 - \omega^2}{\omega^2} \right) k_x^2 - \left( \frac{\omega_a^2 - \omega^2}{c^2} \right)$$

3.1

where \( k_x \) and \( k_z \) are the horizontal and vertical wavenumbers and \( \omega \) is the (angular) wave frequency. The parameters \( H, c, \omega_a \) and \( \omega_b \) are the atmospheric scale height, the speed of sound and the acoustic cut-off and Brunt-Väisälä frequencies respectively. The dispersion relation is illustrated by Fig. 3.1. The periods \( 2\pi/\omega_a \) and \( 2\pi/\omega_b \) are about 4 and 5 minutes in the lower atmosphere and about 14 and 15 minutes respectively in the thermosphere.

The spatial dependence of a wave for which the dispersion relation holds involves oscillations as a function of x and z coupled with an exponential dependence as \( \exp(z/2H) \). An upward propagating wave travels through an increasingly rarefied atmosphere in which there are fewer and fewer particles to carry the energy of the wave. Consequently, each particle must carry more of the energy as the particle number decreases; therefore as the wave progresses its amplitude must increase. For \( \omega < \omega_b \), buoyancy forces are important and balance the vertical pressure gradient (compressional forces) giving rise to internal gravity waves (freely propagating waves). The nature of a gravity wave in the low frequency limit (\( \omega \ll \omega_b \)) for which the buoyancy forces dominate is depicted schematically in Fig. 3.2. The wave is composed of transverse velocity oscillations which, by displacing fluid particles
Fig. 3.1 Curves Relating Normalized Horizontal Phase Velocity \( \left( \omega / c_k x \right) \) to Normalized Wave Period \( \left( \omega_h / \omega \right) \) for Various Values of \( \phi \), the Angle Between the Group Velocity and the Horizontal \( \text{after Francis, 1973b} \)

Fig. 3.2 Schematic Representation of an Elementary Internal Atmospheric Gravity Wave \( \text{[Hines, 1960]} \)
above and below their equilibrium altitudes subject them to a buoyancy force which acts as the restoring force of the wave. The phase velocity is perpendicular to the group velocity, and waves whose energy is propagating obliquely upward at a shallow angle have phase fronts propagating downward at a steep angle.

Following the size classification nomenclature suggested by Georges (1968), AGWs and their associated TIDs can be divided into medium- and large-scale waves (see Table 1.1 of chapter one). The principal distinction between the two is that the horizontal speeds of large-scale waves are substantially greater and those of medium-scale waves are substantially less than the lower atmospheric sound speed $c (=300 \text{ m/s})$.

3.2.2 Realistic atmosphere

An isothermal atmosphere, in which $\omega_a$, $\omega_b$, $c$ and $H$ are all constant, is homogeneous with respect to the direction of wave propagation. Waves propagate along straight-line paths undergoing no reflection or refraction. However, the earth's atmosphere is far from homogeneous due to temperature, dissipative and neutral wind effects. These are now discussed:

a. Temperature

The large vertical temperature variation of the earth's atmosphere induces corresponding variations in $\omega_a$, $\omega_b$ and $c$ and can cause the refraction and reflection of AGWs (Francis, 1973b). Ducting has been suggested to play an important role in the propagation of large-scale waves by Francis (1973a). The ducting mechanism causes a surface wave to propagate along the steep temperature gradient at the base of
the thermosphere (Thome, 1968). In the real atmosphere, the sound speed continues to increase with height throughout the thermosphere, transforming these surface waves into internal waves with tilted wavefronts. The AGW characteristics (e.g. phase speed) are strongly influenced by temperature changes in the thermosphere caused by day-night and solar cycle variations (Francis, 1973a).

b. Dissipative effects
Waves in the atmosphere are acted on by dissipative forces of which ion drag, viscosity and thermal conduction are the most important. At ionospheric heights, AGWs set the ions into motion via collisions with the neutrals. The earth's magnetic field prevents these ions from conforming freely with the neutral motions induced by the wave, since they experience a retarding force in any direction other than parallel with the geomagnetic field direction. A relative motion therefore exists between the ions and the neutrals which tends to dissipate the gravity wave energy. This 'ion drag' effect (Hines, 1960; Gersham and Grigor'yev, 1965; Hines and Hooke, 1970) is particularly important for the dissipation of waves with periods greater than one hour.

The effects of viscosity and thermal conduction are produced by the random thermal motion of the individual gas molecules. As a wave propagates into the upper atmosphere, dissipation becomes increasingly strong (arising from the exponential decrease in the ambient neutral density) which ultimately attenuates the wave, its energy being converted to heat over a range of altitudes (Pitteway and Hines, 1963;
Francis, 1973a). In the case of large-scale waves, ion drag is of greater importance than molecular dissipation during the day, but of smaller importance during the night (Richmond, 1979a).

The somewhat surprising result that viscosity appears to enhance gravity wave ducting has been reported by Francis (1975). This author proposed that the rapid increase of viscosity with height can strongly reflect upward propagating waves of long wavelength. However, Richmond (1978a) suggested that a ducting mechanism does not operate in the earth's atmosphere. The thermospheric modes computed by Francis (1973a) can instead be explained in terms of internal waves undergoing total reflections in the lower thermosphere and partial reflections in the upper atmosphere due to dissipative effects.

c. Neutral wind effects

In the thermosphere, neutral winds can be as high as 200 m/s (Rishbeth, 1972) and their effect on AGWs can be very important (Hines and Reddy, 1967). Cowling et al. (1971) established that the horizontal background winds act as a selective filter allowing only certain waves to reach the thermosphere. Waves can be grouped into three types, i) those that penetrate the F region of the thermosphere, ii) those that are reflected by winds blowing against the waves, and iii) those that are critically coupled to the winds blowing in the same direction as the waves (for which the horizontal phase speed of a wave equals the horizontal wind speed).

Generally, winds will have a much greater effect on medium-scale waves whose horizontal speeds are of the same
order of magnitude as the winds. Tilbrook and Jones (1979) reported that vertical winds also have a significant influence on the propagation of AGWs, but again primarily for medium-scale waves.

3.3 AGW GENERATION

The excitation of AGWs requires an energy input to the atmosphere. Medium-scale waves are generated by both tropospheric and auroral processes, but large-scale waves are produced only by upper-atmospheric sources (E region and above) as found in the high-latitude ionosphere.

Surges in the auroral electrojet have been considered by many authors as probable AGW sources. The following two mechanisms introduce into the neutral atmosphere a local time-dependent disturbance whose low-frequency components travel outwards from the source region as a AGW train.

a. A surge in the current can provide a force on the charged particles which is transferred to the neutral gas via collisions. This Lorentz force $\mathbf{F}$ (Wilson, 1969) is given by

$$\mathbf{F} = \frac{1}{\rho} (\mathbf{j} \times \mathbf{B}) \quad 3.2$$

where $\mathbf{j}$ is the ionospheric current density, $\mathbf{B}$ is the magnetic flux density and $\rho$ is the charged particle density.

b. A localized heat source can be produced by a current surge in the form of Joule heating $S$, (Davis and da Rosa, 1969)

$$S = \frac{1}{\rho} (\mathbf{j} \cdot \mathbf{E}) \quad 3.3$$

where $\mathbf{E}$ is the electric field.
The relative importance of these two mechanisms in AGW excitation is a subject of great debate, e.g. Chimonas and Hines, 1970; Brekke, 1979; Richmond, 1978b, 1979a; Maeda, 1982. Present evidence indicates that both the Lorentz force and Joule heating can be effective in generating AGWs, but Joule heating is more important in generating waves capable of propagating to middle latitudes, particularly for large-scale events (Richmond, 1978b).

AGWs are probably generated by a number of other auroral processes, e.g. the localized heating produced by intense auroral particle precipitation, but these are likely to be of lesser importance (Hunsucker, 1982).

3.4 MID-LATITUDE LARGE-SCALE AGW RESPONSE TO AURORAL SOURCES

Details of some of the major investigations of AGW generation by auroral sources and their propagation to midlatitudes are listed in Table 3.1.

3.4.1 Influence of source characteristics

The nature of the auroral source will have an important influence on the AGW response at midlatitudes. The important theoretical results appropriate to this thesis are now summarized:
a. The amount of AGW energy generated depends on the electric field strength and on the electron density enhancement in the auroral region. The energy is proportional to the square of either the heat input or the momentum input of the source, so large amplitude disturbances are particularly effective in generating AGW energy (Richmond, 1978b).

b. The relatively rapid variation of the local sound speed $c$ with altitude in the lower thermosphere and the small values of $c$ below 110 km tend to reflect waves upward and these reflected waves interfere above the source constructively or destructively with waves originally launched upward. The effective source altitude for the combination of reflected and directed waves can be considered to be approximately 110 km. This has an important influence on the distribution of AGW velocities generated (Richmond, 1978b).

c. Horizontal structure of the electrojet currents of scales of 100 km or less, and temporal variations shorter than a few tens of minutes are unimportant for the generation of large-scale high-speed waves capable of travelling to middle latitudes (Richmond and Matsushita, 1975).

d. The spectrum of AGWs generated depends in a complicated manner on the horizontal dimensions and temporal scales of the source. The characteristic vertical width of the source is limited to 2-4 scale-heights, thus placing constraints on the vertical wavelengths and horizontal velocities of the generated waves (Richmond, 1978b).

e. The midlatitude AGW response is strongly dependent on the longitude/Local Time (LT) and the Universal Time (UT) of the auroral substorm (Rees et al., 1980).

Source terms representing a force and/or heat input are introduced into the hydrodynamic equations.

Isothermal, dissipationless atmosphere assumed.

Equations are linearized, fourier-transformed in time and combined to yield a single equation for which a Green's function solution is obtained.

Response to a finite source is computed by integrating the Green's function (for a point impulse source) over the entire source region.

II. Testud(1973), Francis(1974), Richmond(1978b)

As in (I), except long distance propagation treated using the WKB approximation.

Dissipative processes included.

Francis(1974) also takes into account a rigid earth's surface.
NUMERICAL TREATMENTS

III. Richmond and Matsushita (1975), Richmond (1979a), Hernandez and Roble (1978), Roble et al. (1978)

Simulates the global wind and temperature response to a large geomagnetic substorm.
Solves non-linearized hydrodynamic equations over a 2-D grid of points in altitude and latitude between the pole and the equator.
Longitudinal symmetry and dipolar magnetic field assumed.
Roble et al. (1978) also computed global ionospheric response.

IV. Fuller-Rowell and Rees (1981), Rees et al. (1980)

As in (III), but employs 3-D time dependent model.
Solves momentum and energy equations simultaneously.
Rees et al. (1980) also incorporates non-coincident geographic and geomagnetic poles.

Table 3.1b Details of the major theoretical investigations of AGW generation/propagation
3.4.2 The large-scale AGW characteristics

The large-scale AGW response at midlatitudes is now discussed:

a. Dispersion
The AGW response cannot be described in terms of a simple wave of a fixed period and wavelength. The combination of AGWs of different periods and wavelengths generated in the auroral oval superimpose to give the appearance of a disturbance whose period increases with decreasing altitude and with increasing horizontal distance from the source (Row, 1967; Chimonas and Hines, 1970; Testud, 1973; Richmond and Matsushita, 1975).

The thermospheric response to a large isolated magnetic substorm (Richmond and Matsushita, 1975) indicates that disturbances propagate both equatorward and poleward from the source region. The front of the equatorward disturbance moves at a nearly constant speed of 750 m/s reaching 50° latitude approximately fifty minutes after the storm onset. The time history of the equatorward wind velocity, reproduced in Fig. 3.3, illustrates the dispersive nature of the AGW response. At 350 km altitude, the apparent period of the disturbance head is about one hour at 50° latitude and two hours at 30° latitude. At lower altitudes the apparent period can be significantly longer.

b. Attenuation
Richmond and Matsushita (1975) reported a relatively small attenuation with both time and horizontal distance from the source. In Fig. 3.3, the amplitude of the initial impulse above
300 km is as great at 30° as at 50° latitude. This behaviour seems to conflict with the analytic results of Francis (1973a) who found an exponential decrease in amplitude with horizontal distance for a single gravity mode. The apparent discrepancy arises because the broad spectrum of gravity waves which are generated in the Richmond and Matsushita (1975) simulation interfere in different fashions at different locations and times. Close to the source the longest period waves tend to cancel each other at high latitudes, and at short times after the storm onset. Remote from the source, the short period waves are largely dissipated and only the longer period waves make their appearance. Francis (1973a) found that AGWs of periods less than two hours are generally dissipated before travelling 5000 km.

c. Amplitude

The vertical variation of apparent phase of an AGW disturbance decreases with increasing altitude, and the vertical growth of amplitude is much slower than the exponential growth determined for an individual wave in a dissipationless atmosphere (Richmond and Matsushita, 1975). Two causes of this behaviour are acting simultaneously; they are, i) the influence of dissipative effects (see section 3.2.2), and ii) the superpositioning of the various waves, as discussed above.

d. Oscillatory nature

The dissipationless theoretical treatments of Chimonas and Hines (1970) and Testud (1970) determined that the substorm response consisted of a long-lasting train of AGW oscillations, however, Testud (1973) demonstrated that the dissipative forces
rapidly suppress the oscillatory tail of the disturbance so that only a single impulse is observed at the head of the disturbance. This behaviour is evident in the neutral horizontal velocity reproduced in Fig. 3.4. The opposite polarity of the dayside and nightside currents induces a corresponding inversion of the dayside and nightside wave forms produced by the Lorentz force.

Fuller-Rowell and Rees(1981) predicted that the midlatitude response to an individual substorm consists of a series of wave-like disturbances. A wave from the nearby auroral oval is observed first followed about 90 minutes later by a second wave originating in the dayside auroral oval that has been focused and guided by the enhanced mean anti-sunward thermospheric winds (the mean flow). A third wave may arrive even later due to the dayward propagating wave from the nearby auroral oval being reflected, refocused and guided by the mean flow. These authors also predicted that the largest wave amplitudes and the most complex response occurs between 0000 and 0600 LT.

e. Non-linear effects

Richmond and Matsushita(1975) report that the AGW disturbance onset is everywhere rapid and increasing the source strength increases the rapidity of the onset. This behaviour is suggested to be caused by non-linear wave-steepening effects which generate the higher frequency waves locally. Similar results were presented by Richmond(1979a).
Fig. 3.3 Time histories of the equatorward wind velocity at 30° and 50° latitude at various altitudes. Also indicated at 350 km are the temperature perturbations. At the bottom is given the total power input into the auroral oval of each hemisphere (north and south) due to currents and particles injected from the magnetosphere.

(after Richmond and Matsushita, 1975)

Fig. 3.4 Response of neutral horizontal velocity at 2500 km from source and at an altitude of 300 km, showing the effects of three successive source terms. Curve 1: Lorentz force due to an eastward auroral electrojet current; curve 2: Lorentz force for a westward auroral electrojet current; and curve 3: Joule heating effect of either of the two currents (from Testud, 1973).
3.5 IONOSPHERIC RESPONSE TO AGWs

The influence of AGWs on the ambient ionisation of the ionosphere has been examined in detail by Hooke (1968, 1970) and Nelson (1968). Below the F1 ledge dynamic, photoionisation and chemical loss processes all play a part in determining the ionospheric response, however, in the F2 region where neutral particle number densities are low and ionospheric chemistry is slow, dynamic processes are by far the most important. Clarke et al. (1971) determined that ambipolar diffusion will also be significant but mainly above the F layer peak. These authors reported that non-linear effects are also important since AGWs of perturbation amplitude may produce TIDs not of perturbation amplitude. Dynamic processes are now discussed since TIDs are primarily observed in the F2 region.

The neutral atmosphere is coupled to the ions by collisions with a collision period much shorter (less than one second) than AGW periods. The coupling is thus effectively instantaneous, and in the absence of the geomagnetic field, the ions and electrons would have the same fluid velocity as the neutrals \( u_i = u_e = u \). However, the motion of the charged particles is influenced by the presence of the earth's magnetic field \( B \). The ion gyrofrequency is much larger than the ion-neutral collision frequency in the thermosphere, and therefore ions and electrons are constrained to move along the geomagnetic field with velocity \( u_e \)

\[
    u_e = \frac{B (B \cdot u)}{B^2}
\]
A schematic representation of the coupling between AGWs and electrons is reproduced in Fig. 3.5. The neutral velocity fluctuations are aligned nearly parallel to the phase fronts (that is, the wave is composed of transverse velocity oscillations). Because the wave is predominantly transverse, its neutral density fluctuations are quite small (the motion is nearly noncompressive). However, the electron motions are constrained to be parallel to the magnetic field and therefore compressions and rarefactions of the electron gas are induced by the AGW even though the wave itself is virtually non-compressive (Francis, 1973b). Hooke (1970) solved the linearized electron continuity equation to determine the variation of electron density Ne' about the unperturbed value Neo

\[ A = \frac{\text{Ne}'}{\text{Neo}} = \left( \frac{\text{u} \cdot \mathbf{k} + \mathbf{Q}_b \cdot \mathbf{z}}{\omega} \right) \left\{ \left( k \cdot \mathbf{Q}_b \right) + i \left( \mathbf{Q}_b \cdot \mathbf{z} \right) \left[ \frac{1}{2H} + \frac{d}{dz} \left( \ln \text{Neo} \right) \right] \right\} \] 3.5

where \( \mathbf{Q}_b \) is a constant unit vector directed parallel to the magnetic field, \( \mathbf{Q}_z \) is a unit vector in the direction of the wavevector \( \mathbf{k} \). An investigation of equation 3.5 (Hooke, 1970) reveals that the ionospheric response is highly anisotropic depending on the AGW parameters, the geomagnetic dip and the prevailing vertical ionisation gradient.

a. Geomagnetic dip

Near the F2 layer peak the term \( d/dz(\ln \text{Neo}) \) vanishes and the term \( 1/2H \) can be neglected under the condition \( 4k_z > 1/H \) (which holds for AGW propagation, Hines, 1960). Hence, the amplitude \( A \) can now be represented by
and depends on the two factors: $(u \cdot \mathbf{b})$, the component of gravity wave-induced neutral gas velocity, and $(k \cdot \mathbf{b})$ the component of the wave vector, in the direction of the magnetic field. The effects of these two factors on the ionospheric response to AGW propagation is illustrated in Fig. 3.6.

Hooke(1970) concluded that the tilt of the wave front (which is related to the wave period) relative to the magnetic dip angle $I$ is primarily responsible for determining the amplitude of the ionospheric response to AGWs. For very nearly transverse AGWs propagating meridionally Sterling et al.(1971) obtained the following expression for $A$

$$A = \frac{U \tau_g}{\lambda} \left( \frac{\tau}{\tau_g} \right) \sin \left\{ 2\sin^{-1} \left( \frac{\tau_g}{\tau} \right) - 2I \right\}$$

where $\tau_g$ is the Brunt-Väisälä period, $\tau$ is the wave period, $\lambda$ is the wavelength of the wave.

Francis(1973a) computed the TID amplitude as a function of dip $I$ for an undamped equatorward propagating 90 minute gravity mode (see Fig. 3.7). At the magnetic equator, the geomagnetic field is horizontal so that the electrons are only allowed to move with the small neutral fluid fluctuations; consequently the TID amplitude is small. Near the poles the geomagnetic field is nearly perpendicular to the neutral fluid velocity so that the electrons are severely constrained and again the TID amplitude is small. At midlatitudes neither is the case and the amplitude is relatively large.
Fig. 3.5  Schematic Representation of the Coupling Between Gravity Waves and Electrons

The neutral wave motion (broken arrows) is nearly parallel to the phase fronts and is therefore nearly noncompressive. By ion-neutral collisions, the noncompressive neutral motion induces compressive electron motion (dotted arrows) along the magnetic field.  

(after Francis 1973b)

Fig. 3.6  Illustrating the competition between the requirements that both \( \langle U, \mathbf{l}_b \rangle \) and \( \langle k, \mathbf{l}_b \rangle \) be large if internal gravity wave is to produce a large amplitude ionospheric irregularity. For this diagram, \( \langle U, \mathbf{l}_b \rangle \) is largest when the wave propagates equatorward, but in this case \( \langle k, \mathbf{l}_b \rangle \) is so small that the wave produces relatively small ionospheric response. When the wave propagates poleward \( \langle U, \mathbf{l}_b \rangle \) small, but \( \langle k, \mathbf{l}_b \rangle \) is so much greater that the wave produces a significantly large response (after Hooke, 1970).
b. Altitude

The height dependence of the amplitude of a 90 minute large-scale TID was determined by Francis (1973a). A pronounced height variation of the TID amplitude is obtained, induced by the vertical structure of the ambient ionosphere through the term $1/\text{Ne}_0(d\text{Ne}_0/dz)$ of equation 3.5. This term is large and positive below the F region peak, vanishes at the peak and then approaches a constant negative value above the peak. When this $z$ (height) dependence is multiplied by the $z$ dependence of the neutral gravity wave obtained from equation 3.6, the result is a total $z$ dependence similar to the envelope function reproduced in Fig. 3.8. The physical mechanism causing this behaviour has been discussed by Thome (1968). The vertical wavelengths of large-scale AGWs are large enough to cause the entire F layer to oscillate up and down more or less in phase. The electron density at any given altitude is simply replaced alternately by the electron density slightly above and slightly below that altitude; consequently, the amplitude of the electron density fluctuation at any given altitude is roughly proportional to the vertical gradient of the ambient electron density at that altitude.

The above behaviour has been confirmed experimentally by Morgan and Ballard (1978) who found a sharp maximum in the TID amplitude that occurred within a height range of 20 to 60 km in the inflection region of the electron density profile below the peak. In the case of four events, the peak in the TID amplitude occurred just at the peak in the quantity $1/\text{Ne}_0(d\text{Ne}_0/dz)$. This suggests that the exponential growth of these TIDs has been
Fig. 3.7 TID amplitude versus magnetic dip
(after Francis, 1973a)

Fig. 3.8 Relative electron density perturbation versus altitude,
solid curve - instantaneous TID amplitude
broken curve - envelope within which the TID amplitude
oscillates as a function of time.
(after Francis, 1973a)
3.6.2 Characteristics of large-scale waves

a. Association with auroral and geomagnetic conditions

The association of large-scale TID observations with magnetic storms was noted by Wright (1961) and Chan and Villard (1962) and was confirmed by Georges (1968) and Titheridge (1971) who found a correlation between large-scale TIDs and large magnetic substorms (planetary geomagnetic index, $K_p > 5$). One-to-one correlation between specific magnetic substorms or geomagnetic sudden commencements and specific cycles of large-scale TID events has been reported by a number of authors, e.g., Davis and da Rosa, 1969; Davis, 1971; Harper, 1972; Testud, 1973.

b. TID velocities

Large-scale TIDs travel equatorward with speeds 300-1000 m/s (Francis, 1975; Richmond and Roble, 1979 and all references therein). Hernandez and Roble (1978) observed a northward propagating disturbance several hours after noting a disturbance travelling equatorward during a very 'active geomagnetic period ($K_p > 8$). However, such observations are very infrequent since large-scale disturbances are generally fully dissipated before they reach the opposite hemisphere. Davis and da Rosa (1969) determined a diurnal variation in the azimuth of propagation of ±20° about due south, suggesting that the source is located in the evening sector of the auroral oval.

Maeda and Handa (1980) obtained a spread in directions about 197°E and suggested that TIDs are only observed in the same longitude sector of their excitation. The 17° deviation of the mean propagation direction from due south
counteracted by dissipation at these heights and the peak in the TID amplitude arises mainly from the peak in the value of \( \frac{1}{Neo}(dNeo/dz) \).

3.6 OBSERVATIONS OF LARGE-SCALE TIDS

3.6.1 Experimental techniques

Many experimental techniques have been employed to observe large-scale disturbances either in the neutral or the charged species. These include:


b. Networks of ionosondes (Wright, 1961; Chan and Villard, 1962; Bowman, 1965; Thome, 1968; Arendt et al., 1971; Testud, 1973; Shashunkina, 1975; Chandra and Spencer, 1976; Morgan et al., 1978; Maeda and Handa, 1980)

c. Doppler (Chan and Villard, 1962; George, 1968; Jones and Reynolds, 1975)

d. Faraday rotation (Davis and da Rosa, 1969; Davis, 1971)

e. HF backscatter (Tveten, 1961; Hunsucker and Tveten, 1967)

f. Spacecraft neutral composition and temperature measurements (Chandra and Spencer, 1976; Trinks and Mayr, 1976; Potter et al., 1976; Chandra et al., 1979)

g. Fabry-Perot spectrometer (Hernandez and Roble, 1978)
indicates that the Coriolis force and possibly zonal winds may influence the propagation of large-scale waves. Maeda and Handa (1980) also reported that the TID directions were predominantly equatorward when the mean auroral electrojet (AE) index was high, implying that these TIDs were generated by a source of great longitudinal extent. For low mean values of the AE index the spread in the wave directions was large, implying that these waves were excited by less extended sources. These authors obtained an increase in TID speed with increasing wave period and with higher values of the Kp index.

c. Oscillatory nature

Large-scale TIDs have quasi-periodic features of 30 minutes to four hours, e.g. Georges, 1968; Thome, 1968; Francis, 1975; Richmond and Roble, 1979, but there has only been limited confirmation of the theoretical result of the increase in wave period with decreasing latitude, e.g. Shashun'kina (1975). Davis (1971) observed two pulse-like waves which were apparently excited by the same geomagnetic substorm. This author presented an interpretation in terms of two ducted modes with different speeds, but stated that the generation of multiple oscillations by individual substorms is not generally observed. The observed periodicity was therefore believed to be the consequence of a similar periodicity in the current source. Geomagnetic substorms often occur repeatedly during active conditions with a time-scale of one every one to three hours or so. This conclusion was supported by Testud (1973) who observed one event that consisted of two cycles whose separation (2.5 hours) remained constant over a distance of 2500 km. The
apparent constant period of the wave was in fact roughly equal to the interval between two polar substorms which Testud took to be the wave source.

d. Wavelength
In accordance with theory, Thome (1968) observed very little phase variation of the disturbance with height at low altitudes, but noted a phase advance with increasing height at greater altitudes. At these higher altitudes, tilted wavefronts (55 to 80° to the vertical) indicative of freely propagating waves, have been reported, e.g. Georges, 1968; Morgan and Ballard (1978). Large-scale waves have horizontal wavelengths \( \lambda_x \), greater than 1000 km but can be larger, and vertical wavelengths \( \lambda_z \), less than 1000 km but are generally much less. Waves of dimensions \( \lambda_x = 4000 \) km and \( \lambda_z = 500 \) to 800 km have been monitored by Harper (1972) and Testud and Vasseur (1969).

e. Amplitudes
Waves induce electron density variations at F region heights of up to 100% of the ambient electron densities (Thome, 1968; Harper, 1972; Roble et al., 1978). Large amplitudes have been observed in other parameters such as electron temperature variations of 100°K (Roble et al., 1978) and neutral temperature variations of 10°K (Trinks and Mayr, 1976).

f. Observations at spaced geographic locations
Potter et al. (1976) observed AGWs on a global scale finding an occurrence maximum in the midnight sector indicating a nightside source. Secondary maxima were noted at 1000 and 1700 LT which suggests that waves can propagate across the poles. Chandra et al. (1979) also observed waves on a global
scale reporting that the characteristics of the waves can differ significantly in different LT sectors. The wavefront extents of large-scale TIDs (in the E-W direction) has often been observed to be greater than 3000 km (e.g. Thome, 1968) but of a finite width (Chan and Villard, 1962). Davis and da Rosa (1969) observed strong similarities in the TID waveforms across 1500 km of their wavefronts but found poor correlation over E-W separations of 3000 km.

3.7 CONCLUSIONS

The properties of AGW/TIDs have been described with particular emphasis on large-scale waves. The midlatitude AGW response to an auroral disturbance is determined by, i) the nature of the source, and ii) the inhomogeneous and dissipative properties of the thermosphere. The ionospheric response to AGWs is highly anisotropic depending on, i) the wave parameters, ii) the ambient ionospheric conditions, and iii) the geomagnetic latitude. Large-scale AGW/TIDs have been monitored with many experimental techniques and their observed characteristics are generally in agreement with those predicted by theory. In the following chapter, a series of large-scale TIDs are observed by means of the HF direction finding technique.
CHAPTER FOUR

EXPERIMENTAL OBSERVATIONS AND HF RAYTRACING STUDIES OF LARGE-SCALE TIDS

4.1 INTRODUCTION

A number of large-scale TID events are present as quasi-periodic variations in the bearings of HF radio transmissions. The degree of correlation of the bearing error time signatures measured for geographically spaced paths is discussed in relation to, i) the response characteristics of the HF direction finding technique, ii) the present theories of atmospheric gravity wave (AGW) generation and their propagation from high to low latitudes, and iii) the anisotropic ionospheric response to AGWs. HF raytracing analyses are performed to determine the influence of the transmission path geometry on the TID-induced bearing errors.

4.2 TID OBSERVATIONS

4.2.1 Data analysis

From a visual inspection of the bearing records (bearing error versus time) sixteen large-scale TIDs (quasi-periods greater than 40 minutes) were identified. Their characteristics are listed in Table 4.1. A typical example of a large-scale TID event is reproduced in Fig. 4.1. Oscillations in bearing with a quasi-period of about two hours are clearly evident in the
<table>
<thead>
<tr>
<th>EVENT NUMBER</th>
<th>DATE</th>
<th>TIME [UT]</th>
<th>( \sum K_p ) 24hrs</th>
<th>TID CHARACTERISTICS</th>
<th>SPEED (m/s)</th>
<th>DIRECTION (deg)</th>
<th>PERIOD [mins]</th>
<th>AMPLITUDE [degrees]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>19/10/79</td>
<td>2100-2400</td>
<td>10</td>
<td>I</td>
<td>N-S( ^+ )</td>
<td>150±15</td>
<td>7-9</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>22/10/79</td>
<td>1900-2400</td>
<td>22</td>
<td>I</td>
<td>232±30</td>
<td>135±15</td>
<td>6-9</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>7/5/80</td>
<td>1900-2300</td>
<td>15-</td>
<td>I</td>
<td>22</td>
<td>78±15</td>
<td>5-8</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>11/5/80</td>
<td>1900-2300</td>
<td>34</td>
<td>I</td>
<td>I</td>
<td>69±15</td>
<td>6-7</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>12/5/80</td>
<td>1100-1600</td>
<td>28-</td>
<td>I</td>
<td>S-N( ^* )</td>
<td>&gt;180</td>
<td>7-9</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>13/5/80</td>
<td>2000-0200</td>
<td>20-</td>
<td>I</td>
<td>I</td>
<td>40±20</td>
<td>13-17</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>17/5/80</td>
<td>0800-1400</td>
<td>5</td>
<td>I</td>
<td>N-S( ^* )</td>
<td>100±15</td>
<td>12-15</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>20/6/80</td>
<td>0000-0400</td>
<td>12+</td>
<td>I</td>
<td>S-N( ^+ )</td>
<td>66±6</td>
<td>8-10</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>14/8/80</td>
<td>1900-0200</td>
<td>12+</td>
<td>I</td>
<td>S-N( ^+ )</td>
<td>45±6</td>
<td>7-10</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>18/8/80</td>
<td>1700-2000</td>
<td>24+</td>
<td>I</td>
<td>I</td>
<td>48±9</td>
<td>5-6</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>15/1/81</td>
<td>2200-0200</td>
<td>15+</td>
<td>I</td>
<td>I</td>
<td>&gt;150</td>
<td>6-7</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>26/3/81</td>
<td>1800-2200</td>
<td>31</td>
<td>I</td>
<td>N-S( ^* )</td>
<td>525±123</td>
<td>230±26</td>
<td>87±21</td>
</tr>
<tr>
<td>13</td>
<td>30/3/81</td>
<td>1800-0000</td>
<td>23+</td>
<td>I</td>
<td>241±50</td>
<td>231±10</td>
<td>85±10</td>
<td>5-6</td>
</tr>
<tr>
<td>14</td>
<td>31/3/81</td>
<td>0200-0600</td>
<td>30+</td>
<td>I</td>
<td>I</td>
<td>48±6</td>
<td>6-7</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>31/3/81</td>
<td>2000-0100</td>
<td>30+</td>
<td>I</td>
<td>S-N( ^* )</td>
<td>108±12</td>
<td>5-6</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>1/4/81</td>
<td>1800-2200</td>
<td>29+</td>
<td>I</td>
<td>294±46</td>
<td>197±10</td>
<td>50±10</td>
<td>5-7</td>
</tr>
</tbody>
</table>

\( ^+ \) see section 4.3.1

\( ^* \) see section 4.3.2

peak to peak amplitudes quoted

Table 4.1 The large-scale TID events

-66-
records for several of the transmission paths. The rapid fluctuation in the bearings is produced by wave-interference effects and can be eliminated by means of a low-pass filter. A linear interpolation technique is first applied to the unfiltered bearing records to obtain an equi-spaced time series (sampling interval = 1.5 minutes), the time series is then detrended (linear) and finally filtered. The filter is based on the Butterworth filter (Waldock, 1981) and its frequency response is given in Fig. 4.2. The application of the filter to the Luxembourg record of event 2 is illustrated in Fig. 4.3.

Smaller amplitude, shorter quasi-period (10 to 30 minutes) disturbances (medium-scale TIDs) occur frequently in the bearing records and are sometimes superimposed on the large-scale TIDs, however, these can also be eliminated from the data by choosing an appropriate cut-off frequency for the filter. The large-scale TIDs are themselves occasionally superimposed on large systematic ionospheric tilt (SIT) induced bearing error variations, which can be removed under favourable circumstances, as illustrated in Fig. 4.4 for the Luxembourg record of event 2. The TID velocities and quasi-periods listed in Table 4.1, were determined as described below.

a. TID velocity

TID velocities were calculated by means of a standard triangulation technique (Munro, 1958) which requires, i) a knowledge of the reflection point separations of three or more signals, and ii) the measured timelags between the TID-induced bearing errors on the spaced transmission paths. The timelags were determined by means of a cross-correlation technique.
Illustrating the frequency response of the digital filter for orders 3 to 8. The cut-off period has been set to 5 minutes and even a 3rd order filter is sufficient to produce greater than 10 dB attenuation at \( \omega_c \) (3 mins) (after Waldock, 1981)
Fig. 4.4 Removal of SIT error component from the bearings of the Luxembourg signal

EVENT 2

22 October 1979
running mean applied to unfiltered data (temporal window = 30 min)

20 October 1979
running mean (two hours)

SIT component removed
(a) - (b)
It was only possible to calculate the velocities of five of the sixteen TID observations, due either to the poor correlation of the TID-induced bearing errors on the spaced paths, or because the TIDs were observed in the records of fewer than three transmitters of known geographic location. The velocity errors quoted in Table 4.1 represent standard deviations: three unique velocity solutions are obtained from the timelag information of three paths and further solutions are obtained by assuming an appropriate error (1.5 minutes) in the timelag. The potential errors in the velocity values are quite large due to the assumptions employed in the calculations. The major assumptions are:

i. The signal reflection points are located above the path midpoints. This assumption can be seriously violated when large SITs are present in the ionosphere or the dominant HF transmission mode is a multihop path.

ii. The signals are reflected at equal heights. The transmissions monitored generally have unequal frequencies and path lengths resulting in reflection heights that can differ by tens of kilometers. The velocities calculated will be influenced by the height dependence of TID velocity (Morgan et al., 1978) and the downward phase propagating characteristics of TIDs (Reynolds and Morgan, 1975).

iii. The TID is non-dispersive over the zone of observation. This assumption is clearly not valid in many of the events since the waveform (shape) and the quasi-period are not preserved as the TID travels between the spaced transmission path reflection points.

It is very difficult to calculate the velocity errors
produced by the violation of these assumptions even with the use of sophisticated HF raytracing analysis, however, the errors can vary from a few percent to the order of magnitude of the calculated velocities, depending on the TID event considered. There is considerable debate concerning the nature of the calculated velocities, e.g. whether they are horizontal phase trace or horizontal group velocities (Hines, 1974; Morgan et al., 1978). It is therefore appropriate to simply refer to them as horizontal speeds, since there is no significant difference between phase or group speeds for large-scale waves with periods greater than 50 minutes (Francis, 1974).

b. TID quasi-period

The TID quasi-periods were determined either by visual inspection of the bearing records or by noting the time lag of the first maximum of the auto-correlation function (Davis, 1973). The power spectra of several of the TID events were determined by a modified fast fourier transform technique (FFT) based on the algorithm of Cooley and Tukey (1965). The power spectra for four of the transmission paths of event 2 are reproduced in Fig. 4.5. A given bearing record is prefiltered to remove frequencies greater than the Nyquist frequency to overcome the problems of aliasing and then multiplied by a cosine bell function to compensate for truncation errors caused by the finite nature of the bearing error time series. Performing an FFT on a time series containing a small number of wave cycles results in poor low frequency resolution of the power spectrum. Better resolution can be achieved by increasing

-74-
Fig 4.5 Power Spectra of Event 2

For each signal: Nyquist frequency = 0.3 cycles/min
256 data points (17.6 - 24.0 UT)
data prefiltered $w_1 = 16\text{min}$ $w_2 = 10\text{min}$
the length of the time series so that only a fraction comprises the TID, however, the reliability of the spectra is reduced, due to the poor low signal to noise ratio.

4.2.2 Geophysical data

The three-hour planetary geomagnetic activity index (Kp) for the data acquisition periods is displayed in Fig. 2.5 of chapter 2. Most of the large-scale TID events occur at times when the Kp index is relatively high (Kp > 4) and rapidly changing, although a few events coincide with periods of low Kp. Since this index may not be a useful indicator of auroral activity (Chandra et al., 1979), magnetometer (Tromso, Norway) and riometer (Kiruna, Sweden) data were also inspected for the data acquisition periods. Generally, TID occurrence corresponds well with disturbed auroral conditions but it is difficult to correlate specific cycles of the TID events with specific polar substorms, due to the complicated nature of the auroral activity and the dispersive nature of TIDs as they travel from high to low latitudes.

4.2.3 Vertical-sounder data

Hourly ionospheric data for a number of European ionosondes were examined for the data acquisition periods. Several of the longer period, larger amplitude TID events are clearly evident as quasi-periodic variations of the F-layer critical frequency (fc) and of the virtual height of the base of the F layer (h'F). The respective variations of fc and h'F during three twenty-four hour periods encompassing events 1 and 2 and a period of no TID activity, are presented in
Figs. 4.6 and 4.7. On Regular World Days (RWDs) soundings are generally made every 15 minutes, however, for the two RWDs which coincide with TID events (events 6 and 9), ionograms were available for only a few stations and hence no further study of the events was possible with ionosonde data.

In order to represent the ambient ionospheric conditions during a TID event, vertical profiles of electron density can be obtained from ionosonde data (Dudeney, 1978). When required, approximate signal reflection heights (assuming a single-hop path) were then calculated by means of an HF raytracing technique (Jones, 1966).

4.2.4 TID characteristics observed in the bearings of geographically spaced transmission paths

Due to dispersion, medium-scale TIDs continuously change their waveforms as they propagate and Georges (1968) reported that a wave can rarely be unambiguously identified at two stations, even if their separations are less than 200 km. In the present investigation, the horizontal separations of the signal reflection points are generally greater than a few 100 km and therefore problems can arise in tracking medium-scale waves from one transmission path to another. However, the very long horizontal wavelength (>1000 km) of large-scale TIDs suggests that their waveforms should be preserved over great horizontal distances and the induced bearing errors on spaced paths should be highly correlated.

The bearing measurements indicate that the degree of correlation of large-scale TID-induced bearing errors on spaced
Fig. 4.6. Diurnal variation of Tc for Stough (51°29'N, 0°57'W).

TIME (UT)

12 1 1 1 1 2 3 4 5 6 7 8 9 10 11 12

F layer Critical Frequency, fc(MHz)

Frequency Resolution = 0.1 MHz

- Event 1
- Event 2
- 17-18 October Quiet Day
- 19-20
- 21-22
- 22-23

17-18 October Quiet Day
Fig. 4.7: Diurnal variation of h_F for Slough

Virtual Height of Base of F layer, h_F (km)

Event 1: 19-20
Event 2: 22-23
17-18 October Quiet Day

Height Resolution = 5 Km
paths varies considerably from path to path and from event to event. Intuitively, it is expected that the degree of correlation should decrease as the reflection point separation is increased. Although this is generally the case, poor correlation is occasionally observed even between the errors of very similar paths. The characteristics of TIDs observed in the bearings of spaced paths are now related to:

a. The response characteristics of the HF direction finding technique (since the particular instrument employed to detect TIDs will influence their apparent characteristics) and

b. The temporal and spatial characteristics of AGW/TIDs.

4.3 CHARACTERISTICS OF THE HF DIRECTION FINDING TECHNIQUE

4.3.1 Transmission path geometry

Munro (1953) and Davis and Baker (1966) simulated the passage of a TID above a point on the earth's surface by means of a curved reflector model. Bramley (1953) studied the influence of a TID on the bearing of an oblique HF transmission by assuming a corrugated reflector model. Due to the difficulty in solving the transcendental equation that arises in the oblique case, extensive studies have only become feasible with the advent of computer iterative techniques (Walton and Bailey, 1976; Lyon, 1979).

In the present investigation the influence of the transmission path geometry on the observed large-scale TID-induced bearing errors is investigated using the corrugated reflector model. Lyon (1979) provided a comprehensive
description of the model so only a brief outline is presented here. The reflection surface is represented by

$$z = h + A \cos \left( \frac{2\pi x}{\lambda} \right)$$

where \((x, y, z)\) is a rectangular cartesian coordinate system, \(z\) is the height above the earth's surface, \(x\) represents horizontal distance in the wave propagation direction, and \(h, \lambda\) and \(A\) are the height, (horizontal) wavelength and amplitude of the corrugated surface respectively, as illustrated in Fig. 4.8. The horizontal motion of the reflection surface is taken into account by a phase factor. The specular reflection of a 'one-hop' ray from this surface is taken to represent the propagation of a HF radiowave through a TID. The location of the reflection point and the corresponding value for bearing error is calculated from the path geometry and the reflection conditions by means of an iterative technique. A number of reflection points often satisfy the specular reflection conditions, however, only one solution is generally physically realistic. The values of the model parameters in Table 4.2 are representative of typical transmission path geometries and large-scale TID characteristics. From these parameters, the dependence of the bearing error magnitude on, i) the path length \(R\), ii) the height of the reflection surface \(h\), and iii) the azimuthal orientation of the transmission path relative to the TID direction \(\theta\), has been evaluated (see Fig. 4.9 a,b,c).

The modelling studies indicate that the bearing error magnitude increases with decreasing \(R\), with the increase being most rapid at the shorter path lengths. Few clear examples of this path length dependence are evident in the bearing
Table 4.2

Corrugated Reflector Model Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Path Length</td>
<td>$R = 600\ km$</td>
</tr>
<tr>
<td>Angle between TID and ray direction</td>
<td>$\theta' = 90^\circ$</td>
</tr>
<tr>
<td>Height of Reflection Surface</td>
<td>$h = 300\ km$</td>
</tr>
<tr>
<td>Amplitude of TID</td>
<td>$A = 40\ km$</td>
</tr>
<tr>
<td>Wavelength of TID</td>
<td>$\lambda = 3000\ km$</td>
</tr>
</tbody>
</table>

Fig. 4.8 The Corrugated Reflector Model
Fig. 4.9 Dependence of bearing error magnitude on the path geometry.

(b) Height of reflection from surface (km)

Peak to peak bearing error magnitude (degrees)

Path length, R (km)

- R = 400Km
- R = 300Km
- R = 200Km

Peak to peak bearing error magnitude (degrees)
Fig. 4.9 Dependence of bearing error magnitude on the path geometry.

- Peak to peak bearing error magnitude (degrees)
- Azimuth angle between TID direction and HF path, \( \theta' \) (degrees)

Graph showing the dependence of bearing error magnitude on the path geometry with different R values (400 km, 800 km, 1200 km).
observations, since a number of other factors also influence the magnitudes of the TID-induced bearing errors (discussed later). One such example is found in event 15 (Fig. 4.10) for which the bearing error time signatures of the Luxembourg (R=633 km) and the Wertachtal, Germany (R=1050 km) paths are similar in shape but differ in their peak to peak magnitude (approximately 6 and 3.5 degrees respectively). The unequal bearing error magnitudes can be explained in terms of a path length dependence since, i) the true great-circle bearings of these two transmission paths are almost identical, ii) the transmission reflection heights differ by only a few tens of kilometers (assuming one hop paths), and iii) the TID amplitude can be assumed to change little as the disturbance travels across the two paths (as discussed later).

Large bearing error magnitudes are sometimes observed for the Moscow transmission, whereas small errors are calculated for such a long path (one hop). This behaviour can be explained if the dominant signal is a multihop mode, since large errors have previously been reported for such paths (Beckwith et al., 1972).

Varying the signal reflection height, h is generally equivalent to changing the signal frequency for a given path length. The bearing error dependence on the value of h is illustrated in Fig. 4.9b, however, these results will not be representative of the true situation since the TID amplitude can also vary significantly with height.

The dependence of bearing error magnitude on the azimuthal orientation of the signal path relative to the TID direction, θ'
Fig. 4.10 Unfiltered Bearing Data (EVENT 15) 31 March 1981

Figure 4.10 shows the unfiltered bearing data for EVENT 15 recorded on 31 March 1981. The data is plotted over time (X-axis) and bearing (Y-axis). Two sets of data are plotted: one for R = 1050 km and another for R = 633 km. The data is marked with various symbols and lines to indicate different measurements or events. The graph indicates a significant event at 2200 UT with no data recorded after that time.
is reproduced in Fig. 4.9c. The largest and smallest (zero) magnitudes are obtained for $\theta' = 90$ and $0$ degrees respectively. Consider event 12 (see Fig. 2.8 of chapter 2) in which the TID-induced bearing errors are much larger on the east to west (E-W) Prague path during the period 1830 to 2030 UT, than observed on the S-N 9.57 MHz Madrid path between 1900 and 2100 UT, both transmissions having similar frequencies and path lengths. Assuming the TID amplitude changes little as the disturbance travels across the two transmission paths, then these observations confirm the validity of the model results described above.

The model bearing error time signatures exhibit a characteristic sawtooth shape, which also depends on the azimuthal orientation of the signal path relative to the TID direction, $\theta'$, as illustrated in Fig. 4.11. The model parameters were chosen to highlight the sawtooth shape dependence on the value of $\theta'$. Similar results are obtained for the model parameters of Table 4.2, but the sawtooth shape is less pronounced. Fig. 4.12 illustrates schematically the predicted bearing error signatures of N-S and S-N propagating waves for an E-W transmission path. Reciprocity considerations dictate that the two sawtooth shapes are laterally reversed (mirror image in the ordinate). A qualitative assessment of TID direction should therefore be possible by comparing these model results with observed bearing records for a number of events. Several examples are now described, i) an equatorward propagating disturbance is indicated from the Moscow record of event 7 (Fig. 4.13), ii) the measured timelag between the Luxembourg and Wertachtal signatures of event 15 (Fig. 4.10)
Fig. 4.11 Influence of the TID direction on the bearing error

Model parameters

$H = 250\, \text{km} \quad \lambda = 400\, \text{km} \quad R = 600\, \text{km} \quad A = 6\, \text{km}$

$\theta'$ - Angle between the transmission path and TID directions (measured clockwise)
northward propagating TID

southward propagating TID

Schematic representation of the influence of the TID direction on the bearing error time signature (E-W transmission path)

Fig. 4.12
FIG. 4.13 Unfiltered Bearing Data

[Diagram showing time in hours, bearing in degrees, and distance in kilometers. Data points are marked with circles.]
agrees with the poleward direction indicated by the waveform shapes. This disturbance could be interpreted as a poleward medium-scale wave with a larger than average period. iii) the sawtooth signature observed in the Moscow record of event 5 (Fig. 4.14) suggests a poleward component in the TID direction unless the TID propagates in the SWW direction.

These studies confirm that a simple model can qualitatively simulate several characteristics of TIDs as observed in the variation of bearing with time. Such a model has the advantages over more sophisticated models that, i) only a small amount of computing resources is required, and ii) the model results are easy to interpret.

4.3.2 Realistic radio wave propagation effects

A more realistic technique which accounts for both the refraction and reflection of a radio wave through a disturbed ionosphere has been described by George and Stephenson (1969). This 'WAVE' model has previously been employed in a number of analyses, e.g. direction of arrival studies of HF propagation through medium-scale TIDs (George, 1972; Treharne, 1972). The major disadvantages of this model are the large computing resources required and the cumbersome nature of the homing-in algorithm that traces a ray between two fixed points on the earth's surface.

Ray paths are calculated by means of a 3-D raytracing program (Jones, 1966). The program performs the numerical integration of a set of differential equations (Haselgrove, 1954) which define the locus of the ray path and the components of the wave normal direction as the ray
progresses through the ionosphere. In the present investigation, the phase and group-refractive indices required for evaluating the differential equations are calculated by means of the Appleton-Hartree equation, an earth-centred dipole model of the earth's magnetic field and an ω-Chapman profile of the ambient electron density distribution. Rays are traced through a disturbed ionosphere given by

\[ N = N_0(r, \theta, \phi)(1 + \beta) \]  

where \( N_0(r, \theta, \phi) \) is the ambient electron density in earth-centred spherical polar coordinates and \( \beta \) is the perturbation produced by the TID

\[ \beta = \delta \exp - \left[ \frac{(r - R_0) - z_0}{H} \right]^2 \cos 2\pi \left\{ t' + \left( \frac{\pi}{2} - \theta \right) \frac{R_0}{\lambda_x} + \frac{(r - R_0)}{\lambda_z} \right\} \]

where \( R_0 \) is the earth's radius, \( z_0 \) is the height of the maximum wave amplitude, \( H \) is the wave-amplitude 'scale-height', \( \delta \) is the wave perturbation amplitude, and \( \lambda_x \) and \( \lambda_z \) are the horizontal and vertical wavelengths respectively. The characteristics of the perturbation are as follows, i) a horizontal sinusoidal time dependence, ii) downward sense of phase propagation, iii) an exponential growth of the wave amplitude with height and a subsequent attenuation above \( z_0 \), and iv) a north to south propagation direction.

The model results described below have been calculated for the ordinary (magnetoionic) component, although very similar results are obtained for the extraordinary mode. The elevation angle of the ray path at the transmitter is varied until the ground path length falls within a stipulated error (5 km) of the path length required. The initial ray azimuth lies along
the great-circle path between the transmitter and receiver. For the raypath endpoint to coincide exactly with the receiver location, the homing-in procedure must also vary the initial ray azimuth, however, this is a very time-consuming process. In the present investigation, only the elevation angle is varied since the model results are qualitatively similar to those calculated when both the elevation and the azimuth of the ray at the transmitter are varied.

The transmission path and 'WAVE' parameters employed in the modelling studies described below are listed in Table 4.3. Ray paths have been calculated through an ionosphere disturbed by a large-scale TID of quasi-period of about one hour. The bearing error signatures reproduced in Fig. 4.15 depart significantly from a purely sinusoidal or sawtooth shape. They comprise a sharp maximum (peak) and a broad minimum (trough) or vice-versa depending on the azimuthal orientation of the transmission path relative to the TID direction. Nevertheless, this peak/trough nature is superimposed on an underlying sawtooth behaviour which has a form consistent with the results obtained from the corrugated reflector model. The peak/trough variation in the bearing error reflects a similar characteristic form of the isoionic contours of the disturbed ionosphere (Fig. 4.16). The signal reflection height and elevation angle at the transmitter corresponding to the first simulation are reproduced in Fig. 4.17. The low reflection heights indicate that the signal frequency is very much smaller than the path MUF.

Peak/trough bearing error signatures are frequently observed in the bearing records. Several examples are now
I

J

J

J

OD ýl

p..

ýý

J

U'7

.ý

J

J

ýý

LU

=ý

J

N

O

OD J

'O

O.

u,

4.- uj

N

C

-ý

fD 0

--I
ni
orn

-

_

_

C)

=

O

O

ýO

--ý

N

W

ui

N

ii

=

=

O

_

_

W

C)

=

00

c.

_

O
O

S

O
O

_

O

v

rD

ý
CC)
Ul

O
O

p

_

O

y

y
x

o

D

N

N

y

-ý

N

o

D

-10
ni
-,

O

O

O

_

._

O

m

0
Cl.
ro

7v
g

O
O

h

--h

O

C) C7%
_
0
_
0 C>

N
=
_

O

O

WN

EA
_

f

C

C

ýJ

pJ
"

N

(D

-t-

-+

Oý

0
=

N

J

J

p

LAJ

Ö

J
ýa

2
rD

:wrr_
CD

1

ro
rn

11

A

--r
ßlß
N_
Bý
C
O

n
11

N

=-

-fi

z-

II

11

W
N

O

EI
OJ

3

--r
-I
0

cL
r0
N
-1-

"-

(A.

0
N

rD

-95-

n
--r
0

Z1
N

o


Fig. 4.15. WAVE Model Raytracing Results

(see Table 4.3)

-96-
Fig. 4.17
Ray Path Parameters corresponding to Simulation 1

Height of Reflection (km)

0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
120 150 200 250

Elevation Angle (degrees)

hm = 225 km (ambient)
described, i) equatorward propagating disturbances are indicated in the Luxembourg record of event 1 (Fig. 4.18) and the Prague record of event 12 (Fig. 2.8 of chapter 2), ii) poleward propagating disturbances are indicated in the 9.54 and 9.74 MHz records of event 8 (Fig. 4.19) and the Luxembourg record of event 9 (Fig. 4.20). These qualitative assessments of TID propagation direction are confirmed for events 8 and 12 from measured timelag information. Events 8 and 9 occurred during relatively quiet auroral conditions suggesting that they are poleward propagating medium-scale TIDs rather than equatorward large-scale waves.

Further simulations have been performed to determine the effect of changing the model parameters on the shape and magnitude of the predicted bearing error signature.

a. Changing the parameters $\phi$, $z_0$, and $H$ tends only to influence the bearing error magnitude

b. Increasing $\lambda_T$ while keeping $\lambda_X$ constant (effectively varying the period of the disturbance) tends to reduce the difference between the shape of the bearing error variation above and below the abscissa (see Fig. 4.21).

c. The effects of changing the signal path length $R$, the transmission path azimuth $\beta_T$ and the signal frequency $f$ are illustrated in Figs. 4.22, 4.23 and 4.24 respectively. These results are qualitatively similar to those calculated from the corrugated reflector model, quantitative differences exist due to the vertical amplitude dependence of the 'WAVE' model.

The 'WAVE' model can adequately reproduce a number of the TID effects evident in the experimental bearing records despite the inherent simplicity of the model. Further understanding of
Fig. 4.18: Untitled Bearing Data (EVENT 1) 19 October 1979

- R = 603 Km
- R = 931 Km
- R = 633 Km
Influence of Path Length (see Table 4.22)

Fig. 4.22 Wave Model Raytracing Results
Fig 4.23 WAVE Model Raytracing Results

I nfluence of transmission path azimuth

simulation $\frac{1 + 270}{9 \times 285}$

10 • 300
11 • 330
12 • 360
Fig. 4.24 WAVE Model Raytracing Results

Influence of Signal Frequency

(see Table 4.3)

(BEARING ERROR [DEGREES])

TIME (one cycle)

0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

0
1
2
3
the response characteristics of the direction finding technique to TIDs could only be achieved if the following model assumptions are modified, i) the horizontal sinusoidal time dependence, ii) the simple vertical amplitude profile, iii) the single spectral AGW component, and iv) the time-independent ambient ionosphere. A more realistic vertical amplitude profile of the perturbation can be obtained by solving the coupled Navier-Stokes and electron continuity equations (Francis, 1973b; Essex, 1976), however, this added complexity does not provide a significantly better understanding of the influence of TIDs on the radio propagation characteristics.

4.3.3 Proximity of the signal frequency to the path MUF

The lateral path deviation effects experienced by an HF radiowave travelling through the ionosphere depend on the horizontal gradients of electron density transverse to the signal path and the proximity of the signal frequency to the path MUF (Rawer, 1951; Waldo-Lewis, 1953; Titheridge, 1958; Beckwith et al., 1972). Very large bearing errors can be produced as the penetration condition is approached, and an example of this behaviour is found in the bearing measurements of the Luxembourg signal for event 6 (Figs. 4.25 and 4.26). The large scatter in the bearings between 2245 and 0030 UT indicates that the signal frequency is very close to the path MUF. The resulting TID-induced bearing errors are considerably greater than those observed for the Prague path, whose frequency is well below the respective path MUF.
4.4 Bearing Measurements Related to AGW/TID Theory

The characteristics of large-scale TIDs present in the midlatitude ionosphere can be related to, i) the AGW response to the auroral source, ii) the propagation characteristics of the waves as they travel from high to low latitudes, and iii) the ionospheric response to the AGW disturbance. These processes have complicated temporal and spatial dependencies and are now discussed in relation to the nature of TIDs at spaced geographic locations.

4.4.1 Bearing error magnitude

The magnitude of the observed TID-induced bearing errors depends on the transmission path geometry but also on the TID amplitude. In the present investigation, TIDs are observed at geographic locations with N-S separations of hundreds of kilometers, but generally less than 1000 km. Over such distances TID amplitudes do not change significantly (say, less than 10%), as indicated from the theoretical works of Richmond and Matsushita (1975) and Sterling et al. (1971) which considered the latitudinal dependences of the AGW and ionospheric responses respectively. This behaviour is supported by experimental observations, e.g. Davis and da Rosa (1969) reported TIDs that propagate over 800 km without appreciable attenuation.

The transmissions monitored during a particular TID event can be reflected at very different heights (E region heights to the F layer peak). Over this height range the TID amplitude can vary markedly (Francis, 1973a; Morgan and Ballard, 1978). A sharp
maximum in the TID amplitude occurs within a height range of 20-60 km in the inflection region of the electron density profile below the F layer peak and small TID amplitudes are produced at heights both close to the F layer peak and well below the inflection region.

Large-scale disturbances have previously been observed over E-W distances of greater than 1500 km (e.g. Thome, 1968; Davis and da Rosa, 1969) indicating that a large-scale TID can be present in all of the bearing records of a given event. Although this is generally the case, event 7 (Fig. 4.27) is a striking counter example. A large-amplitude, long period wave is evident in the bearings of the Moscow path but absent from the Luxembourg record. This may be due to the very large dissipative forces that act during the day, and/or the disturbance propagates largely to the east of the observation area.

4.4.2 Quasi-period of oscillations

AGW theory (e.g. Row, 1965; Testud, 1973; Richmond and Matsumita, 1975) and experimental observations (Shashun'kina, 1975) indicate that AGW quasi-period is a function of both altitude and the distance from the auroral source (see chapter three). Unequal quasi-periods are often observed in the bearings of the different transmission paths, but it is extremely difficult to quantify the dispersive AGW/TID behaviour for the following reasons,

a. The signal reflection points are often relatively close, therefore only small changes in quasi-period are produced.
Fig. 4.27 Filtered Bearing Data (EVENT 7)

W = 16 min W = 10 min

Luxembourg, 6.09000 M3, 17-4. 09.73

Great Circle Bearing 109.3

Moscow 10.92000 M3, 64.8

Great Circle Bearing 64.8 17-4. 09.73
b. The signal reflection points differ in both height and latitude and the changes in quasi-period due to their differences are combined in the final waveform.

c. Quasi-periods are difficult to measure (as discussed previously).

In addition, caution needs to be exercised in interpreting the oscillatory nature of a large-scale disturbance since the time interval between two consecutive 'cycles' may only be related to the time interval between two auroral substorms (Davis, 1971; Francis, 1975).

Three examples of the dispersive nature of TIDs are now presented:

a. Very long quasi-periods are evident in the Moscow and Ankara bearing records of event 5 (Fig. 4.14) and a much smaller period disturbance is possibly observed in the case of the SRI path. Predictions indicate that the Moscow and Ankara signals (probably multihop) are reflected in a region low in the ionosphere, where the dominant TID quasi-period is very long. The long periods are also indicative of a distant source location suggesting that this daytime event was generated on the nightside of the auroral oval.

b. The quasi-period of the disturbance observed in the Luxembourg record of event 6 (Fig. 4.26) is significantly shorter than the period evident in the Prague record. This dispersive behaviour can be explained by the higher reflection heights for the Luxembourg signal.

c. Close to the F layer peak, the ionospheric response to a large-scale AGW is small, and shorter period waves for which
compressional forces are more important can dominate the response (van Eyken, 1981). This can explain the very different quasi-periods observed in the Luxembourg and SRI bearing records of event 1 (Fig. 4.18) resulting from the frequency of the SRI path being close to the path MUF. However, this event occurred during quiet auroral conditions implying that the disturbance on the Luxembourg path may be localised rather than a TID. Unfortunately, no conclusive evidence for either interpretation can be gained from the bearing data or from the vertical-sounder ionospheric data available for this period.

4.4.3 Waveform (shape) of the TID-induced bearing errors

Theory predicts that the midlatitude AGW response to an isolated auroral substorm consists of a single impulse or sinusoid (Testud, 1973; Richmond and Matsushita, 1975). The ionospheric response, however, can depart significantly from a sinusoidal time dependence of the ambient F layer (Porter and Tuan, 1974). Many of the reported events occurred during the dusk transition period when the height and composition of the F layer change rapidly over a period of a few hours. Consequently, significant changes in the path geometry can occur during the time it takes a disturbance to travel across the signal reflection point. These considerations suggest that it is not surprising that the bearing measurements often exhibit waveforms quite different from a sawtooth or a peak/trough behaviour, as predicted in section 4.3. The ambient ionosphere changes much less rapidly near noon and this may explain the regular sawtooth waveforms evident in the bearing
measurements of the two daytime events (events 5 and 7).

TIDs are often observed superimposed on large SIT-induced bearing errors. These SIT effects are caused by the spatial and temporal variations of the regular ionosphere and are very path dependent (see chapter six), causing the correlation of TID-induced bearing errors on spaced paths to be reduced. Examples of this behaviour are evident in events 13 and 16 (Figs. 4.28 and 4.29 respectively) in which the SIT effects are much larger in the bearings of the Luxembourg path.

During the early stages of a large geomagnetic storm large decreases in the critical frequency and large increases in the height (up to 100 km) of the F layer can be produced by neutral composition effects of enhanced equatorward propagating neutral winds (Kane, 1973; Jung and Prölss, 1978; Essex, 1981). These ionospheric changes occur over only a few hours and can therefore significantly influence the time dependence of changes in, i) the ionospheric response to AGWs, and ii) the transmission path geometries. Event 4 (Fig. 4.30) occurred during such ionospheric conditions, as indicated from ionosonde data available for this period. The bearing errors recorded for the Prague and Vienna transmission paths after 2000 UT have quite different quasi-periods and waveform shapes. The poor correlation between the respective bearing error signatures is difficult to explain, even during these disturbed ionospheric conditions, since their reflection points (assuming single hop propagation) differ only by about 150 and 15 km in the horizontal and vertical directions respectively. A tentative explanation can be proposed in which the antenna polar diagrams
Fig. 4.28 Unfiltered Bearing Data [EVENT 13] 30 MARCH 1981

BEARING IN DEGS

BEARING IN DEGS

BEARING IN DEGS

BEARING IN DEGS

R = 1050 Km

R = 931 Km

R = 633 Km
Fig. 4.30 Unfiltered Bearing Data [EVENT 4] 11 May 1980

BEARING IN DEG.

R = 2,586 km

R = 12,20 km

R = 19,93 km
for the two transmitters are such to produce different mode components in the received signals. The respective dominant modes would then travel through the TID along different paths and be affected unequally. The bearing errors for these two paths are dissimilar on other occasions, perhaps supporting the suggestion of different propagation modes. A similar explanation can be proposed for event 10 (Fig. 4.31) in which a disturbance of long quasi-period is observed in the bearings of the 15.89 MHz transmission, but is apparently absent from the 15.69 MHz Moscow record. Unlike event 4, this disturbance occurred at a time of low auroral activity. The validity of this interpretation for event 10 depends to some extent on whether the two Moscow transmitters are colocated.

The poor correlation observed in event 4 could alternatively be explained as follows. The Prague disturbance may be the interference in some fashion of two waves of unequal quasi-periods, i) a long period wave which is also observed in the bearings of the Vienna path, and ii) a shorter period, localized disturbance that does not affect the Vienna bearings. A similar explanation can be proposed for the observation of two peaks in the bearings of the Prague record in event 2 (Fig. 4.1) between 2030 and 2230 UT, whereas in the Luxembourg bearings only a single peak is evident.

4.5 CONCLUSIONS

Previous HF direction-finding studies of TIDs have been
Fig. 4.31 Unfiltered Bearing Data [EVENT 10]. 18 August 1980

R = 1332 km

R = 1277 km

R = 2586 km
confined to medium-scale events. In the present investigation, the technique is employed for the investigation of large-scale TIDs by simultaneously monitoring the bearings of a number of geographically spaced transmitters. The characteristics of sixteen events observed in the bearing measurements have been related to the response characteristics of the HF direction finding technique and to the present theories of the generation and propagation of large-scale AGW/TIDs.

Despite the very long wavelength of large-scale TIDs poor correlation is often observed between the bearings of spaced transmission paths and this behaviour allows the velocity of only five of the TID events to be determined. Extensive HF raytracing studies of radio propagation through 3-D models of the disturbed ionosphere, have shown that the transmission path geometry has an important influence on the magnitude and time variation of the TID-induced bearing errors. A large number of other factors reduce the correlation on spaced paths, these include:

a. The proximity of the signal frequency to the path MUF.

b. The presence of systematic ionospheric tilts.

c. The relative mode strengths are affected by the radiation pattern of the transmitting array.

d. The TID amplitude and quasi-period are functions of the height of the reflection point and its distance from the auroral source.

e. The interference between more than one AGW/TID that may be present.

f. The height and composition of the ambient F layer can
change significantly during a TID event producing changes in
the TID characteristics and the path geometry.

g. Several of the events may be medium-scale TIDs.

Unfortunately, it has only been possible to propose
tentative explanations for several instances of poor
correlation due to the lack of supportive data.
5.1 INTRODUCTION

The large-scale temporal and geographical variations of electron density in the quiet ionosphere (SITs) cause the deviation of HF signals from the great-circle plane containing the transmitter and the receiver. Both elevation and azimuthal angles-of-arrival are affected, the latter being of major interest to the present investigation. This chapter reviews the previous observations of SIT-induced bearing errors and describes how the tilt behaviour is dependent on the solar-zenith angle. A number of theoretical investigations and raytracing simulations are also described; these studies provide a means of understanding the nature of these path deviation effects.

5.2 PREVIOUS OBSERVATIONS OF SITs

Ross and Bramley(1947) observed bearing errors of 10-20 degrees near to local sunrise in the records of an HF transmission which propagated from north to south (N-S) over a 400 km path. This 'sunrise effect' was also reported by Crone(1944) for a N-S 1800 km transmission path. This author found that the peak in the bearing error variation correlated well with the time of local sunrise throughout a nine month observation period.
path and found a diurnal rate of change of $0.15$ degrees per hour and $0.6$ degrees per hour for one hop F and two hop F signals respectively. Similarly, Titheridge (1958) noted a uniform change of bearing for N-S transmission paths (2000 to 10,000 km) of $n/f$ degrees per hour, where $n$ is the number of hops and $f$ is the signal frequency.

The SIT-observations described above refer to mid-latitude transmission paths and are therefore of particular relevance to the measurements made in the present investigation. However, much larger and more complicated SIT effects occur both at high latitudes (e.g. due to the midlatitude trough, Halcrow and Nisbet, 1977) and for trans-equatorial transmission paths (Davies and Jones, 1968).

5.3 INTERPRETATION OF IONOSPHERIC TILT EFFECTS

The path deviation effects experienced by an HF radiowave due to SITs can be represented by an effective reflection surface. This surface is located in space along the signal path (at the path midpoint for a one hop signal) and has an effect on the ray geometry (specular reflection) equivalent to the net effects of ionospheric reflection and refraction. Croft and Fenwick (1953) defined, i) an effective tilt magnitude as the angle between the effective reflection surface and a local horizontal plane, ii) an effective tilt azimuth as the azimuthal direction of the maximum slope of the reflection surface. The bearing error $\delta \beta$ corresponding to an effective reflection surface for which the tilt azimuth is normal to the
Before the development of wide aperture DFs and computer averaging techniques, SITs were often difficult to observe due to the large variances associated with the bearing measurements. Nevertheless, large SIT-induced bearing errors were occasionally noted at times when the signal frequency was close to the path MUF. Ross (1949) observed large errors in the bearing records made over a 90 km path at a time when the path MUF for the extra-ordinary mode was very close to the signal frequency; a similar situation also occurred a little later for the ordinary mode. Bramley (1953) correlated a post-sunrise decrease in the bearing error with a decrease in the ratio of the signal frequency to the path MUF.

Observations of SIT-induced bearing errors are not confined to S-N paths or to times close to local sunrise; for example, Kasuya (1954) and Burtnyk et al. (1962) noted significant errors (several degrees) near to local sunset. These evening tilts were found to extend over longer periods and to be less closely correlated with local sunset than were the morning tilts with local sunrise. Morgan (1971, 1974) reported bearing measurements over both 1200 km E-W and 1300 km S-N paths. For the S-N signal, the bearing errors observed at sunset were slightly smaller and of the opposite sign to those noted at sunrise. For the E-W signal, the mean bearing errors observed at both sunrise and sunset were close to zero, whilst during the day a positive bias in the bearing error was noted.

The occurrence of significant bearing errors during the day have also been reported by a number of authors. Bramley (1955) monitored the bearings of a 700 km transmission
ray direction and $\alpha_T$ is the tilt magnitude is given by (approximately)

$$\delta \beta = \frac{2 h \alpha_T}{R} = \alpha_T \tan \Delta$$

where $R$ is the ground range, $h$ is the height of the reflection surface and $\Delta$ is the elevation angle (see Fig. 5.1).

Kasuya (1954) explained the diurnal variation in the SIT-induced bearing errors of a S-N transmission path by means of a simple 'mirror analogy' (see Fig. 5.2). The negative bearing errors observed between midnight and midday are due to the sub-solar point lying to the east of the undeviated signal path. This implies that the isoionic contours are lower to the east producing an eastward deviation of the signal and therefore negative bearing errors. Similarly, the positive SIT-errors observed between midday and midnight are the result of the sub-solar point lying to the west of the undeviated signal path. At midday and midnight the SIT-errors are close to zero since at these times the ionosphere is symmetrical with respect to the undeviated signal path.

The diurnal variation of the SIT-errors of an E-W path have been explained by Morgan (1971, 1974). The bearings are affected by the N-S component of the effective ionospheric tilt which is governed by the latitude of the reflection point relative to the sub-solar point. As the latitude of the sub-solar point never exceeds 23 degrees north, then the sun lies to the south of a midlatitude signal path (northern hemisphere) throughout the day. This implies that the isoionic contours are lower to the south resulting in the southward
Fig. 5.1 Geometry of the specular reflection from a tilted plane surface.

Fig. 5.2 Explanation of diffraction deviation caused by tilted reflection layer.
deviation of the signal path and giving rise to positive bearing errors. At night these tilts may be expected to be small due to the absence of the solar production of ionisation.

The diurnal behaviour of SIT-induced bearing errors has been explained above solely in terms of a solar-zenith angle dependence of the SITs. The adequacy of such a simplistic treatment is discussed in chapter six.

5.4 THEORETICAL INVESTIGATIONS OF PATH DEVIATION EFFECTS

5.4.1 Snell's Law treatment

Titheridge (1956, 1958) calculated the path of a ray through a non-horizontally stratified ionosphere using an analytical treatment based on Snell's Law. At a given height, a ray is incident at an angle i to a plane boundary, tilted at an angle \( \alpha \) to the horizontal, between media of refractive index \( n \) and \( n+dn \) (see Fig. 5.3). Snell's Law is applied in the directions Ox and Oy. The two relationships so formed are combined to obtain the mean change in the horizontal direction of propagation over the two corresponding sections of the ray path. The total deviation in the bearing of the received signal is then given by

\[
\delta \beta = \cot i_0 \int_0^1 \alpha_2 \, du \quad 5.2
\]

where \( i_0 \) is the mean of the angles of incidence and emergence of the ray at the base of the layer, \( \alpha_2 \) is the lateral component of the tilt \( \alpha \) and \( u \) is the refractive index for the equivalent vertical ray of frequency \( f \cos(i_0) \), i.e.
\[ u = \cot(i)/\cot(i_0); \quad u=1 \text{ at the base of the layer, } u=0 \text{ at reflection.} \]

From equations 5.1 and 5.2 it follows that an ionosphere, for which the tilt \( \alpha \) of the planes of constant ionisation varies with height, causes the same deviation in the bearing of a reflected radiowave as a constant lateral (transverse) tilt \( \alpha_T \) given by

\[
\alpha_T = (\delta \beta) \tan i_0 = \int_0^1 \alpha_2 \, du
\]

Equation 5.3 is applicable to any type of layer, any variation of tilt \( \alpha \) with height and any angle of incidence. It is exact when the tilt is entirely perpendicular to the direction of propagation of the radiowave.

If the angle \( i_0 \) is measured from the horizontal, then any tilt \( \alpha_0 \) of the layer base can be allowed for by supposing the reflecting plane to have an effective tilt of \( \alpha_0 + \alpha_T \). Similarly, the components of the tilt from any other cause can be separately evaluated; the overall tilt and the resulting deviations being simply the linear sum of such components.

Equation 5.3 indicates that the effective ionospheric tilt is the integrated effect of the tilts in the isoionic contours along the entire ray path and is therefore dependent on the depth of penetration of the ray into the layer. Hence, the effective tilt is a function of both the signal frequency and elevation angle as well as the gradient of electron density. Theoretical treatments presented by Waldo-Lewis(1953) and Rawer(1951) predict the presence of very large effective tilts at signal frequencies close to the path MUF.
The application of equation 5.3 can be illustrated for a model ionosphere for which the tilt of the isoionic contours increases linearly with electron density (Titheridge, 1958). The dependence of the tilt angle $\alpha$ on electron density $N$ is given by $\alpha = K_1 N$ where $K_1$ is a constant. Since

$$\cos^2 i_0 (1 - u^2) = 1 - \frac{\sin^2 i_0}{\sin^2 i} = 1 - n^2 = \frac{Ne^2}{\pi m f^2}$$ 5.4

where $f$ is the signal frequency, $e$ is the electronic charge, and $m$ is the electron mass, then

$$\alpha(u) = K_1 N = K \cos^2 i_0 (1 - u^2)$$ 5.5

where $K$ is a constant.

From equations 5.3 and 5.5, the effective ionospheric tilt is given by

$$a_T = \frac{2}{3} K \cos^2 i_0$$ 5.6

which is two-thirds the tilt of the isoionic contour at the reflection height ($u=0$). This result is applicable to any shape of the reflecting layer and any angle of incidence.

Titheridge (1958) found that for a linear layer, the effective ionospheric tilt is equal to one-third of the tilt of the virtual heights. The relationship between the gradient in the virtual and effective heights, and the tilt of the isoionic contour at the reflection height for the case of a linear layer is illustrated in Fig. 5.4.

Titheridge (1958) also determined the effective tilt due to a gradient in the critical frequency, $f_c$ and a peak height, $h_m$ of a parabolic layer of semi-thickness, $y_m$. If $f_c$ varies by $df_c$
Fig. 5.3 Illustration of Snells Law Treatment (after Titheridge, 1958)

Fig. 5.4 Schematic representation of ionospheric tilts (after Titheridge, 1956)
in a horizontal distance $ds$, the overall tilt in the direction of $ds$ is given by (in the transverse case)

$$\alpha_T = y_m \frac{d f_c}{f_c ds} B(z)$$  \hspace{1cm} 5.7

where $B(z) = \frac{1}{4} (z + \frac{1}{z}) \ln \left| \frac{1 + z}{1 - z} \right| - \frac{1}{2}$

and $z = \frac{f \cos \phi}{f_c}$

The function $B(z)$ is infinite at $z=1$ but drops rapidly; for example, $B(z)$ varies from 0.5 to 1.0 as $z$ changes from 0.8 and 0.9. Consequently, the effective tilt and bearing error become very large as the penetration condition ($z=1$) is approached.

Titheridge (1958) also considered the case of a signal which penetrates a lower ionospheric layer and is reflected at an upper layer. This author described the effect of a penetrated layer in terms of an optical analogy. If a layer is equivalent to a parallel-sided slab of glass then no path deviation effects will be produced even if the slab is tilted. However, if the layer is equivalent (in cross-section) to a prism, then it produces a deflection of the ray path; the deflections on the upward and downward paths being additive. The overall effective tilt is given by

$$\alpha_T = \alpha_1 + 2 \alpha_2 + \ldots \ldots + \alpha_r + \alpha_{or}$$  \hspace{1cm} 5.8

where $\alpha_1, \alpha_2, \ldots$ are the effective tilts corresponding to the penetrated layers 1,2,$\ldots$; $\alpha_r$ is the effective tilt of the reflecting layer and $\alpha_{or}$ is the tilt of the base of the reflecting layer.
The path deviation due to a gradient in critical frequency can be obtained from equations 5.7 and 5.8. Titheridge (1958) employed them in conjunction with experimentally observed values of $i_0$ and the gradients of the monthly median critical frequencies of the F1 and F2 layers measured at the centre of a S-N transmission path. This author found that the major contribution to the daytime variation of the observed bearings of certain transmission paths (a rate of change of $1/f$ degrees per hour per hop) of the observed bearings was caused by the refractive effects of the signal in the F1 layer. Similar conclusions were reported by Whale (1959).

5.4.2 Phase path treatment

Most direction finders are basically phase-sensitive instruments which determine the orientation of lines of constant phase in the wavefield of an HF signal at the reception point. These lines are intersections of surfaces of constant phase with the horizontal plane. A surface of constant phase links points on the various ray paths with the same value of the phase path parameter $P$. Hence, $P$ is an important parameter in DF theory.

Gething (1965, 1974, 1978) considered two antennae mounted a distance 'a' apart on a line perpendicular to the great-circle path to the transmitter (see Fig. 5.5). In the absence of any lateral tilt there is no phase difference between the signals at the two antennae, assuming a plane wavefront for the radiowave. When a lateral tilt $\alpha$ is present, the respective phase paths from the transmitter to the two elements differ by
At a given moment in time, the difference in phase of the signals \( \Phi \) (arriving at an elevation angle \( \Delta \)) at the two antennae is equal to

\[
\Phi = \frac{2\pi}{\lambda} d' = \frac{2\pi}{\lambda} a \sin(\delta \beta) \cos \Delta
\]

where \( \lambda \) is the wavelength of the signal and \( d' \) is defined in Fig. 5.5.

Phase and phase path are simply related, for incremental changes

\[
\delta \Phi = \frac{2\pi \delta P}{\lambda}
\]

The bearing of a signal is defined as the direction of the phase propagation vector as measured at the receiver, hence, the bearing error is given by

\[
\delta \beta = \frac{\delta P}{a \cos \Delta}
\]

where \( \sin(\delta \beta) = \delta \beta \) for small \( \delta \beta \) (approximately).

The difference in the phase path at the two antennae is assumed to be due to a transverse gradient of some ionospheric parameter \( I \), represented by \( dI/ds \) where \( ds \) is an element of distance transverse to the ray path. Since for a one-hop path the signal reflection points are separated by approximately \( a/2 \), then the difference in the value of \( I \) at the two reflection points is given by \( a/2 (dI/ds) \). Hence,

\[
\delta P = \frac{a}{2} \left( \frac{d I}{d s} \right) \left( \frac{d P}{d I} \right)
\]
and from equations 5.11 and 5.12, the bearing error is given by

\[ \delta \beta = \frac{1}{2 \cos \Delta} \left( \frac{d I}{d s} \right) \left( \frac{d P}{d I} \right) \]  

Equation 5.13 can be determined analytically if a suitable expression is found for \( P \) and \( dP/dI \). The gradient \( dI/ds \) can be estimated from vertical-sounder data.

The theory can be verified for the case of a tilted plane mirror (Gething, 1978). The transverse gradient is then \( dh/ds \), where \( h \) is the height of the mirror. Geometrical considerations give

\[ \sin \Delta = \frac{2 h}{p} \] \hspace{1cm} 5.14
\[ \tan \Delta = \frac{2 h}{R} \]

where \( R \) is the ground range. The total differential \( dP/dI \) for a constant \( R \) is given by

\[ \frac{d P}{d I} = \frac{\partial P}{\partial I} + \frac{\partial P}{\partial \Delta} \frac{d \Delta}{d I} \] \hspace{1cm} 5.15

where \( \frac{d R}{d I} = 0 \)

From equations 5.12 to 5.15 the bearing error is then found to be

\[ \delta \beta = \frac{d h}{d s} \tan \Delta \] \hspace{1cm} 5.16

This relation is equivalent to equation 5.1 which was obtained from purely geometrical considerations.

Gething (1978) also derived expressions for the bearing
errors produced by transverse gradients of various ionospheric parameters. The ionosphere is assumed to have a parabolic distribution of electron density and an analytical expression for the phase path can be obtained from

$$ P = \int n \, d\ell $$

where \( n \) is the phase refractive index and \( d\ell \) is an increment of distance along the raypath. If earth-curvature effects are ignored, then the following relationships can be obtained for the bearing error

**a. transverse gradient in critical frequency**

$$ \delta \beta = \frac{y_m \cot i_0}{2 \, f_c} \left[ \frac{d f_c}{d\ell} \right] \left[ 1 - \frac{1}{2} \left( z + \frac{1}{z} \right) \ln \left( \frac{1 + z}{1 - z} \right) \right] $$

This expression agrees with the relationship derived by Titheridge (1958), that is equation 5.7.

**b. transverse gradient of the F layer peak height, \( h_m, h_b \)**

\( F \) layer base height is held constant and \( y_m \) is allowed to vary as

$$ \frac{dh_m}{d\ell} = \frac{dy_m}{d\ell} $$

giving

$$ \delta \beta = \frac{\cot i_0}{2} \left[ \frac{dh_m}{d\ell} \right] \left[ 1 + \frac{1}{2} \left( z - \frac{1}{z} \right) \ln \left( \frac{1 + z}{1 - z} \right) \right] $$

**c. transverse gradient of \( h_b \) and \( h_m \) with \( T = \frac{y_m}{h_b} \) constant**

$$ \delta \beta = \frac{\cot i_0}{2} \left[ \frac{dh_m}{d\ell} \right] \left[ \frac{2 + T}{1 + T} \right] + \frac{T(z - \frac{1}{z})}{2(1 + T)} \ln \left( \frac{1 + z}{1 - z} \right) $$

**d. penetration through one layer, denoted by subscript 1, and reflected at a higher layer, denoted by subscript 2. Both**
critical frequencies vary with position and all other parameters are constant

\[
\delta \beta = \frac{y_m \cot i_0}{\frac{df_1}{ds}} \left[ 1 - \frac{1}{2} \left( \frac{z_1 + 1}{z_1} \right) \ln \left( \frac{z_1 + 1}{z_1 - 1} \right) \right] + \frac{y_m \cot i_0}{\frac{df_2}{ds}} \left[ 1 - \frac{1}{2} \left( \frac{z_2 + 1}{z_2} \right) \ln \left( \frac{1 + z_2}{1 - z_2} \right) \right]
\]

Morgan (1971) employed these relationships to try to explain the zero bearing error observed near sunrise and sunset for an E-W path. Ionospheric data from four ionosondes, situated in the vicinity of the path midpoint, were employed to calculate the transverse gradients. Between 2200 and 0600 UT the signal frequency was close to the path MUF and Morgan (1971) assumed that the bearing error was governed only by gradients of the critical frequency and the F layer peak height. The bearing errors produced by these two parameters were calculated to have equal magnitudes but opposite signs. It was therefore concluded that this was consistent with the experimental observations.

5.5 MODELLING SITS

5.5.1 Using vertical incidence (V.I.) ionospheric data

Ross and Bramley (1947) estimated the effective ionospheric tilts near to local sunrise on the basis of changes in the virtual height of a signal having the equivalent vertical incidence frequency corresponding to an 400 km N-S path. This information was determined from ionograms for a single station
located at the transmission path midpoint for which time and longitude were suitably interchanged. Effects were found to be most marked when the signal frequency was close to the path MUF and good qualitative agreement with the measured angle-of-arrival information was achieved.

Rao (1968a) employed the 3-D raytracing program of Jones (1966) to simulate bearing data of a 1330 km S-N transmission path. Rays were traced through an electron density model obtained from electron density versus height profiles deduced from ionograms collected at a network of five V.I. ionosondes located near the oblique path midpoint. In the case of a one hop F transmission during the night hours, excellent agreement was found between the measured and simulated bearing errors. This demonstrates that vertical-sounder data employed in conjunction with raytracing can provide, under favourable circumstances a realistic representation of the most significant path deviation effects.

In chapter 6, vertical-sounder data is employed in conjunction with both the Jones (1966) program and a simple specular raytracing technique in order to simulate the path deviation effects experienced by several of the transmission paths for which bearing data is available.

5.5.2 Using numerical prediction maps of ionospheric data

Beckwith et al. (1972, 1974), Beckwith (1971, 1973), and Rao and Beckwith (1974) employed a specular raytracing technique together with a 3-D ionospheric model based on predictions to simulate the path deviation effects of a medium range
midlatitude path within the U.S.A.

a. The technique

The 3-D reflective and refractive properties of the ionosphere are represented by a spherical reflection surface eccentric with the earth's surface. The centre and radius of curvature of the effective reflection sphere are determined by means of an equivalent-frequency virtual-height calculation.

The geographic location of the reflection point is first estimated for a specified mode configuration, for example, in the one hop F case this is taken to be the geometric midpoint of the transmission path (see Fig. 5.6). Four geographic locations in the vicinity of the estimated reflection point are then chosen so that the three outer points are symmetrically spaced on a circle of radius 50 km centred at the inner point. For each geographic location the equivalent vertical-incidence virtual-height (hv') is calculated using

i. the 3-D ionospheric model

ii. the geometry pertinent to the transmitter-receiver configuration

iii. the frequency of the transmission

The four virtual heights determined in this manner define the equivalent spherical reflection surface.

The 3-D ionospheric model assumes a vertical distribution of ionisation consisting of two distinct parabolic (E- and F2-) layers (Barghausen et al., 1969), whose parameters (fc, hm, ym) are obtained directly or indirectly from sets of numerical coefficients (CCIR, 1967). These are specified for high and low Zurich sunspot numbers and a linear interpolation is applied.
Fig. 5.5 Wavefield at antennae.

Fig. 5.6 Geometry of the equivalent ionospheric reflection surface.
(after Beckwith et al., 1972)
for intermediate values of solar activity. The coefficients represent the predicted monthly median values and are in the form of numerical maps which describe, on a world-wide basis, the variation of various ionospheric parameters as a function of geographic location and time. The numerical mapping method is based on the early work of Jones and Gallet (1962a, b).

Ray path parameters are determined purely from geometrical considerations, namely the specular reflection from the spherical reflection surface in the plane containing the transmitter, the receiver and the centre of curvature of the sphere.

b. Comparison of the model results with experimental data
Rao and Beckwith (1974) compared their model results with experimental data acquired for a 1330 km S-N transmission path located within the U.S.A. The bearing error variation near to local sunrise over a five-day period each month for twelve months is reproduced in Fig. 5.7. A significant overall agreement is evident despite the day-to-day variability of the path deviation effects and the limitations inherent in the technique. These limitations are discussed in chapter six where the technique is employed to model the path deviation effects of a number of European transmissions.

5.6 CONCLUSIONS

SIT-induced bearing errors have been reported for a wide range of transmission paths and propagation conditions. The path deviation effects experienced by an HF signal can be
Fig. 5.7 Experimentally observed (solid curves) and monthly mean predicted (dashed curves) azimuthal angle-of-deviations from the great-circle azimuth during the sunrise period. (after Rao & Beckwith, 1974)
conveniently represented by an effective reflection surface. Effective ionospheric tilts are related both to the gradients of electron density produced by the zenith-angle variation of the solar-ionising radiation and to the proximity of the signal frequency to the path MUF. The results of a range of theoretical investigations have been described which are in general agreement with the experimental observations.
6.1 INTRODUCTION

The DF observations undertaken during the present investigation exhibit many clear examples of the effect of SITs on the measured bearing error. These are studied by means of raytracing analyses which incorporate a range of ionospheric models. The physical mechanisms which produce the most significant SIT-induced bearing errors are also examined.

6.2 SIT-INDUCED BEARING ERROR OBSERVATIONS

6.2.1 Measurements for E-W paths

The diurnal variation of the averaged bearing error obtained for the Luxembourg transmission is presented in Fig. 6.1. A two-hour running mean was applied to the unfiltered bearing measurements to eliminate the major effects of wave interference and TIDs. Averaging over such a long period tends to reduce the apparent magnitude of the SIT effects, particularly at times when the bearing changes most rapidly. Consequently, shorter averages are employed for many of the examples presented later. The averaged bearings are plotted about the true great-circle bearing of the transmitter. The thick curve of Fig. 6.1 represents the mean of the six thin
Fig. 6.1: Averaged bearing error for the Luxembourg signal.

Temporal window = 2 hours

Azimuth $\theta = 109.3^\circ$
Range $R = 633$ km
Frequency $f = 6.09$ MHz
curves (the measured bearing error variations on different days during October 1979). The day-to-day variability about the mean curve is represented as a standard deviation in Fig. 6.1b. The times of local sunrise and sunset at the path midpoint (15th day of the month, altitude of 250 km) are indicated by vertical lines. These times differ from those shown by up to an hour, depending on the day of the month and altitude considered.

The averaged bearing error curves exhibit several well-defined features, i) a positive bias during the day, ii) a two to three hour period of positive errors after midnight, and iii) periods of large negative errors close to local sunrise and between local sunset and midnight. A similar variation is evident in the averaged curves for the same path, but for measurements made during May 1980 (see Fig. 6.2), August 1980 and March 1981. The negative errors for these months, however, are displaced with respect to the local sunrise and sunset times. The interpretation of these and other observed SIT effects are presented in section 6.4. Similar observations for the Prague and Vienna transmissions are reproduced in Fig. 6.3 and 6.4 respectively, although the negative errors which occur during the evening/night period are very much smaller on these paths.

The day-to-day variability in the SIT errors for the Luxembourg path is significantly larger close to local sunrise and during the night than during the day. This behaviour is caused by, i) the signal frequency is closer to the path MUF during the night, therefore small day-to-day differences in the horizontal electron density gradients will produce large
Fig. 6.2. Averaged bearing error for the Luxembourg signal.
FIG. 6.4 Averaged bearing error for the Vienna signal
differences in the effective ionospheric tilts, ii) large-scale TIDs are often present (at night) during geomagnetically disturbed conditions, and iii) the reliability of the bearing measurements deteriorates during the nighttime.

6.2.2 Measurements for SE-NW paths

The averaged bearing error data for the Issoudun transmission are presented in Fig. 6.5. The features which are consistently reproduced from day-to-day for this path are, i) large negative errors after local sunrise, ii) small errors during the day, iii) significant negative errors during the early evening hours, and iv) large positive errors after about 2100 UT. The difference in the times for which data are available for the Issoudun and Luxembourg transmissions arises from the different schedules of their respective broadcast services. Only a small amount of data was recorded for the Issoudun path during May 1980, however, the existence of a similar bearing error variation during the evening hours is evident from Fig. 6.6. This evening behaviour is also present in the bearing measurements for the Geneva transmission (see Fig. 6.6b).

6.2.3 Measurements for other paths

The bearing measurements for the 9.57 and 11.92 MHz transmissions are presented in Figs. 6.7 and 6.8 respectively. Although both these frequencies are labelled as Madrid signals, the respective transmitters have a ground separation of some 40 km. During the period of observation (1400 to 2100 UT) the
Figure 65. Averaged bearing error for the Issoudun signal during October 1979.

Temporal window = 1 hour

Azimuth $\phi = 148.6^\circ$

Range $R = 603$ Kms

Frequency $f = 6.175$ MHz
Figure 6.6 Averaged bearing error data for two SE-NW paths.

Geneva
October 1979

Issoudun
May 1980

\[ \theta = 131.8^\circ \]
\[ R = 860 \text{ km} \]
\[ f = 9.943 \text{ MHz} \]

Fen mean curve
Temporal window = 1 hour
Fig. 6.7: Averaged bearing error data for the 957 MHz signal.
FIG. 68. Averaged bearing data for the 11.92 MHz Madrid signal.

- No mean curves.
- Temporal window = 30 minutes.

- For the month of March 1981.
- For the month of January 1981.
- For the month of May 1980.
- For the month of October 1979.

- Mean bearing error (degrees).
measured bearing errors are generally positive and some dependence of the error variation on season is evident.

Further examples of the dependence of SIT effects on the transmission path and season are evident in the bearing measurements for the Copenhagen and the unidentified 11.705 MHz signals (Fig. 6.9), the Moscow signal (Figs. 6.10 and 6.11) and the unidentified 13.725 MHz transmission (Fig. 6.12). The bearing variation for this last station is the lateral inversion about the abscissa of that observed for the Luxembourg path. Such a relationship is expected on purely geometrical grounds if the true great-circle bearing of the 13.725 MHz transmission is assumed to be about 243 degrees.

### 6.3 Raytracing Analyses

In order to interpret the SIT observations some knowledge of the ionospheric gradients of electron density is required. The vertical distribution of electron density above a point on the earth's surface is generally determined from an ionospheric model (e.g. Dudeney, 1978) which requires as input various ionospheric parameters, e.g. $f_c, M(3000)F_2, f_0E, h'F, (h'F,F_2)$. These parameters can be obtained either from, i) global maps of numerical coefficients which describe their monthly median behaviour as a function of geographic location and time (CCIR, 1967), or ii) tables of routinely scaled parameters for ionograms taken at specific locations and times. Electron density profiles above two or more spaced locations provide horizontal gradient information. The lateral path deviation effects experienced by a wide range of transmission paths can
Fig. 6.9 Average bearing error data for two NE-SW paths.

August 1980

Station unknown

Temporal window = 30 minutes

Temporal window = 1 hour

0.0 200 400 600 800 1000

Time (UT) [hours]

0.0 0.5 1.0 1.5

BEARING ERROR [DEGREES]

No mean curve

R = 11.705 MHz

T = 71.705 MHz

R = 10.34 km

T = 75.165 MHz

β = 53.0° (estimated)

January 1981

Copenhagen
Fig. 6.10 Average bearing error data for the Moscow signal.
Fig 6.1: Averaged bearing error for the Moscow signal.

\[
\begin{align*}
\text{May 1980} & \quad \text{Phase} = 64.8^\circ \\
R & = 2586 \text{ km} \\
\nu & = 1.92 \text{ MHz}
\end{align*}
\]

- Temporal window = 2 hours
- No sunrise or sunset

TIME [UT]
Fig. 6:12 Averaged bearing error data for an W-E path.

Temporal window = 2 hours
R = 2.430 (estimated)

June 1980
be calculated from this gradient information by means of raytracing analyses. Three raytracing studies are now presented in an attempt to account for the experimental observations.

6.3.1 The 3-D ionospheric model based on predictions (model 1)  
a. Model results  
The computer program (Beckwith, 1973) incorporating the 3-D ionospheric model developed by Beckwith et al. (1972), described in chapter five, has been adapted for use on the University’s mainframe computer (CDC Cyber 73). The numerical prediction coefficients (CCIR, 1967) required by the program were provided by the Rutherford-Appleton Laboratory. This model has previously only been employed to calculate the lateral path deviation effects for a N-S path located in North America. The analysis technique is now applied to paths located within Europe:

I) Predictions for the Luxembourg path  
The predicted SIT-induced bearing errors (curve I) for the Luxembourg path are compared with the corresponding experimental observations (curve II) in Fig. 6.13 (one hop paths have been calculated through an ionosphere representative of October median, sunspot number = 200 conditions; no E layer is incorporated in the ionospheric model). Very poor agreement exists between the two curves. The diurnal behaviour of the tilt azimuth and tilt magnitude of the effective reflection surface (at the reflection point) are reproduced in Fig. 6.14. The value of the tilt azimuth relative to the azimuth of the transmission path determines the sign of the bearing error.
Bearing Error (degrees)

Receiver Location: 52°N, 2°E
One Hop Path: δ = 1 09.3°
Spot Number = 200
October Median data: r = 633 Km, f = 6.09 MHz

MODEL I

Time (UT)
18 19 20 21 22 23

Curve II: Experimental (mean curve)
Curve I: Model Results

Model I comparison of the predicted and observed S11 errors for the Luxembourg path.

Fig. 4.6.13
Fig. 6.14 Diurnal behavior of the effective ionospheric tilt (corresponding to Fig. 6.13)

TIME [UT]

Tilt Magnitude (degrees)

Tilt Azimuth (degrees)

Positive errors

Negative errors

Gaps in curves generally denote penetration
Large bearing errors are produced when a relatively large tilt magnitude coincides with a tilt azimuth significantly different from the path azimuth (Beckwith et al., 1972).

The diurnal variation of, i) the ratio of the signal frequency to the path MUF, and ii) the virtual heights of the base of the assumed parabolic layer and the effective reflection surface at the path midpoint are presented in Fig. 6.15. The small bearing errors predicted for the period 0800 to 1800 UT occur when the signal frequency is much lower than the path MUF. The simulation was repeated, but with the inclusion of an E layer in the ionospheric model. Between 0900 and 1400 UT, reliable model results are not obtained since the signal frequency is close to the E layer lowest usable frequency (the model calculates signal strength loss calculations for each mode). At other times, the predicted bearing errors are almost identical to those obtained when the E layer is neglected.

The program has the facility to determine the influence of the parameter gradients (of fc etc.) on the effective ionospheric tilt. Each curve of Fig. 6.16 was determined by assuming that the path deviation effects were due to the gradient of a single ionospheric parameter, the other parameter gradients being kept constant. The gradient of the F layer base height, hb provides the major contribution to the bearing error when the signal frequency is significantly lower than the path MUF. When the signal frequency is close to the path MUF, however, the F layer critical frequency gradient makes the dominant contribution.
Figure 6.16 Parameter gradient contributions to the predicted bearing error (see Fig. 6.13)

Gradient contribution only
The model results of four further simulations are presented in Fig. 6.17. The model input parameters are altered to determine whether better agreement with the experimental observations for the Luxembourg path could be achieved. These results differ from those obtained above only in the magnitude of the predicted errors and the times at which the rays penetrate the model ionosphere. These differences arise from the different times at which MUF proximity effects occur for the various simulations.

II. Predictions for the Issoudun path

One and two hop F calculations for the Issoudun path are compared with experimental observations in Fig. 6.18. Much better agreement is obtained for this path, although significant differences still exist between the observed and predicted curves. The large errors predicted before 0700 UT and after 1900 UT in the one hop case are caused primarily by the gradient of the F layer critical frequency. The smaller negative errors around 1600 UT are produced by the gradient of the F layer base height.

III. Predictions for other transmission paths

Model calculations for a number of other paths are reproduced in Fig. 6.19. Generally, there is disappointing agreement between these predictions and the experimental observations, although qualitative agreement is obtained for N-S paths.

As an experiment the latitude of the Luxembourg-DF path was varied and the results are presented in Fig. 6.20. Better agreement with the measured results is achieved as the path is displaced to lower latitudes. Particularly good agreement is
Fig. 6.17 Predicted errors corresponding to the model parameters of Fig. 6.13 except for a single parameter which has been altered.

- I: May conditions
- II: Support number = 200
- III: Support number = 100
- IV: HOP F
- V: HOP F

Bearing Error (degrees) vs. Time (UT)
Figure 6.18 Comparison of the predicted and observed ST errors for the Issoudun path.

Curve I: Experimental (mean curve)
Curve II: Model results (1 hop) F
Curve III: Model results (2 hop) F

Receiver location: 52°N, 2°E
\(\phi = 148.6°\)
R = 60.3 km
\(f = 6.175\) MHz

Sunspot number = 200
October Median data
Fig. 6.20 Predicted SLR errors for the Luxembourg path displaced to lower latitudes except receiver latitude $\theta$. See Fig. 6.13.

Model parameters - see Fig. 6.13.
obtained for the receiver (DF) located at a latitude of $30^0$ N. At these latitudes, the base height gradient makes the most important contribution to the path deviation effects, since higher F layer critical frequencies cause the transmission to be reflected lower in the ionosphere. A similar improvement between the predicted and observed errors is also found for other transmission paths displaced to lower latitudes. An explanation for the improved results is suggested in the next section.

b. Limitations of model 1

The major limitations of model 1 are now discussed in order to explain the poor agreement between the predicted and measured SIT errors:

I. Although no provision is made in the model for the high-angle (Pedersen ray), this mode should not be important for the transmission paths considered.

II. The earth's steady magnetic field is neglected in the model calculations. Rao (1969) has established that 'magneto-ionic' deviations are small for paths similar to those considered here.

III. Lateral path deviation effects due to underlying layer refraction are also neglected. However, these should be important for only a short time when the signal is reflected very close to the base of the F layer. A separate F1 layer is not incorporated into the ionospheric model, but the model results described above should not be affected since the F1 layer is primarily a summer low solar activity phenomenon.

IV. The poor agreement between the predicted and observed
October conditions

Figure 6.21 Comparison of the N-S gradients (differences) in the ionospheric parameter. H'F obtained from the vertical sounder data (median 1979, spot number = 200).
bearing errors is undoubtedly due to the inadequacy of the numerical maps of ionospheric characteristics to accurately represent the large-scale geographic and temporal variations of the quiet ionosphere. This inadequacy is now demonstrated.

The N-S gradients (differences) in the ionospheric parameter $h'F$, obtained from published ionosonde data, are compared with those derived from numerical prediction maps (CCIR, 1967) in Fig. 6.21. The differences in the values of $h'F$ (October 1979 conditions) have been calculated for three pairs of stations which are separated by different latitudinal distances. The locations of these stations are indicated in Fig. 6.22. The inadequacy of the prediction maps to yield accurate ionospheric gradient information is obvious, e.g. the evening/night negative tilts (a positive tilt is defined as ' $h'F$ lower to the south') evident in the measured ionosonde data are predicted several hours earlier by the prediction maps. Similar comparisons are presented in Figs. 6.23 and 6.24 for N-S gradients in $f_c$ and $h_m$ respectively. Marked differences between the true and predicted gradients of these parameters are also obtained.

The prediction maps (CCIR, 1967) assume a linear relationship between $f_c$ and solar activity, whereas for high solar activity the actual relationship departs significantly from this. Gradient information has been obtained from a second set of coefficients (CCIR, 1970) which represents this relationship as a second degree polynomial, but without improvement.

Bradley (1982, private communication) suggested that the
Fig. 6.22. The locations of the ionosonde stations referred to in chapter six.

- Bearing of station 2 from 1 and 2
- Distance between stations

<table>
<thead>
<tr>
<th>Depth (km)</th>
<th>Station 1</th>
<th>Station 2</th>
<th>M</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>DB</td>
<td>DB</td>
<td>M</td>
<td>S</td>
</tr>
<tr>
<td>1528</td>
<td>DB</td>
<td>DB</td>
<td>M</td>
<td>S</td>
</tr>
<tr>
<td>191</td>
<td>DB</td>
<td>DB</td>
<td>M</td>
<td>S</td>
</tr>
<tr>
<td>235</td>
<td>DB</td>
<td>DB</td>
<td>M</td>
<td>S</td>
</tr>
<tr>
<td>1796</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>S</td>
</tr>
<tr>
<td>191</td>
<td>DB</td>
<td>DB</td>
<td>M</td>
<td>S</td>
</tr>
<tr>
<td>235</td>
<td>DB</td>
<td>DB</td>
<td>M</td>
<td>S</td>
</tr>
<tr>
<td>340</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>S</td>
</tr>
<tr>
<td>375</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>S</td>
</tr>
<tr>
<td>385</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>S</td>
</tr>
</tbody>
</table>
\( \Delta f_0, F2 \text{ (MHz)} \)

- If \( \Delta fc = 0 \), the value of \( F_2 \) is greater to the north.

- \( f_0, F2 \text{ (MHz)} \)

See Fig. 6.21

October conditions

Vertical Sonde data (median 1979)

Obtained from: Project Prediction Maps (sunspot number = 200)

Fig. 6.23 Comparison of the N-S gradients differences in the ionospheric parameter \( f_c \)
$\Delta h_m = 0$, if the value of $h_m$ is greater to the north

October conditions

From Fig. 6.21: Vertical sounder data (median 1979)

Comparison of the N-S gradients (differences) in the ionospheric parameter $h_m$
prediction maps do not yield accurate gradient information because the CCIR coefficients include harmonics up to an order determined by the fit to the vertical-sounder data (employed in generating the maps). The highest order is chosen to minimize residuals and effectively specifies the minimum gradient size that is represented. This scale size must be considerably greater than the size responsible for lateral deviations which produce the observed bearing errors.

The E-W gradients of $h'F$, $fc$ and $hm$ are similarly reproduced in Figs. 6.25, 6.26 and 6.27 respectively. Clearly the prediction maps yield more accurate gradient information in the E-W than N-S directions, particularly for the $fc$ gradient, and during the dawn and dusk periods when very large gradients are maintained over a wide spatial area. This behaviour explains the better qualitative agreement between the predicted and observed SIT-errors for paths affected chiefly by the E-W ionospheric gradients.

At lower latitudes, the large-scale regular ionospheric tilts are generally much larger, indicating that they are more likely to be preserved in the analyses which generate the maps. This explains the improvement in the model results for transmission paths displaced to lower latitudes, as reported earlier.

A technique for generating more realistic prediction maps, based on their updating with hourly vertical-sounder data, has been reported by Rush and Miller(1973) and Rush and Edwards(1976). This updating procedure is accomplished by computing differences between the predicted median and the
\( \nabla \Delta h'F = 0 \) if the value of \( h'F \) is greater to the east.
Figure 6.26. Comparison of the F2-equivalent (differences) in the ionospheric parameter Fc0 obtained from October conditions B. Vertical-sounder data from Figure 6.26. [vertical lines and data points on graph]
Fig. 6.27 Comparison of the E W gradients (differences) in the ionospheric parameter hm

Obtained from a. Prediction maps b. Vertical-sounder data

October conditions

TIME (UT)
hourly values of a given ionospheric parameter at specific locations where hourly data are available. These differences are extended to areas where vertical-sounder data are not available by employing weighting factors which are derived from correlation coefficients determined from past observations. Maps can be generated which produce daily values that have accuracies similar to the observations with which the maps are updated. Unfortunately, this updating technique was not available during the present investigation.

6.3.2 A simple model based on a plane effective reflection surface (model 2)

a. The technique

Ionospheric data for a pair of geographically spaced locations (say $I_1$ and $I_2$) provides ionospheric gradient information in the direction $I_1I_2$. This information can be employed to calculate the lateral path deviation effects experienced by an HF transmission orthogonal to the direction $I_1I_2$. Fortunately, a large number of ionosondes exist in Europe so that SIT-induced bearing errors can be calculated for transmissions having a wide range of azimuths. Alternatively, numerical prediction maps can provide the required ionospheric information. This model is now described:

i. The vertical distributions of electron density at $I_1$ and $I_2$ are assumed to be of a parabolic form,

$$\frac{f_N^2}{f_c^2} = 1 - \left( \frac{h_N - h_m}{y_m} \right)^2 \quad 6.1$$

where $f_N$ is the plasma frequency at a height $h_N$. The
Ionospheric characteristics $h_m$ and $y_m$ are derived either from the empirical expressions given by Bradley and Dudeney (1978) which fully take into account the effects of underlying ionisation, or from the Lucus and Hayden (1966) model which grossly underestimates these retardation effects.

ii) Equivalent-frequency virtual heights are calculated for locations $I_1$ and $I_2$ ($h_{v_1}'$ and $h_{v_2}'$ respectively).

Assuming a flat earth/ionosphere the secant law gives

$$f_v = f_{ob} \cos i_o$$  \hspace{1cm} (6.2)

where $f_v$ is the equivalent vertical-incidence frequency (of the wave that propagates vertically and is reflected from the same real height as an obliquely propagating wave of frequency $f_{ob}$); $i_o$ is the angle between the ray and the vertical at the bottom of the ionospheric layer.

Martyn's equivalent path theorem states that the virtual heights of reflection of the obliquely propagating wave, $h_{ob}'$ and the equivalent vertical wave, $h_v'$ are equal. This allows the actual ionosphere to be simulated by a fictitious mirror-like reflector located at the virtual height of reflection $h_v'$ (as illustrated in Fig. 6.28). From geometrical considerations

$$\tan i_o = \frac{R}{2 h_v'}$$  \hspace{1cm} (6.3)

where $R$ is the path length, and

$$h_v' = h_b + \frac{1}{2} y_m \frac{f_v}{f_c} \ln \left( \frac{f_c + f_v}{f_c - f_v} \right)$$  \hspace{1cm} (6.4)

which is derived in standard texts (e.g. Davies, 1969), and
Fig. 6.28 Equivalent Waves of frequencies $f_{ob}$ and $f_v$
assumes a parabolic layer, no geomagnetic field and a flat ionosphere.

Equations 6.2, 6.3 and 6.4 can be solved by iterative means to obtain the unknown $h'v$.

iii) the (effective) tilt, $\alpha_T$ in the direction $I_1I_2$ is given by

$$\alpha_T = \frac{h'v_2 - h'v_1}{d} \quad 6.5$$

Assuming the radio wave is reflected at the midpoint of the line joining the two locations $I_1$ and $I_2$, then the height of reflection $h'_m$ is given by

$$h'_m = \frac{h'v_1 + h'v_2}{2} \quad 6.6$$

and the bearing error $\delta\beta$ is given by

$$\delta\beta = \frac{2h'_m \cdot \alpha_T}{R} \quad 6.7$$

b. Model results

I. simulations employing prediction data

In order to compare the bearing errors calculated by the above procedure (model 2) with those obtained from model 1, ionospheric gradient information was obtained from numerical prediction maps (CCIR, 1967). The comparison of the errors predicted for a 6.09 MHz transmission are presented in Fig. 6.29. The same transmission path geometry ($R$ and $\beta$) were employed and the ionospheric layer parameters $h'_m$ and $y'_m$ were derived from the model of Lucus and Hayden (1966). A similar comparison is presented in Fig. 6.30 for a transmission path displaced to the south. Qualitative agreement is obtained
Fig. 6.29 Comparison of the predicted bearing errors obtained from model 1 (curve I) with those from model 2 (curve II).

Model 1: Receiver location = 52°N, 2°E

Model 2: I', 50.87°N, 2.12°E

Penetration occurred 0.400 UT (both curves) calculations performed every hour (both models)

CICR (1967) coefficients employed

October, snapshot number = 200

\[ \phi = 90^\circ \]
\[ R = 633 \text{ Km} \]
\[ f = 6.09 \text{ MHz} \]
Fig. 6.30 Comparison of the bearing errors predicted by model 1 (curve I) with those of model 2 (curve II).

CCIR coefficients, October, sunspot number=200

C1 = 30.8 deg 0.97 E

model 1: receiver location = 30°N, -2°E

1, 29.1, 0.97 E

R = 633 km
\beta = 90°

f = 6.09 MHz

one hop
between the bearing errors predicted by the two models, although poorer agreement is evident for the more northerly transmission path, particularly at times when the signal frequency is close to the path MUF.

II. Simulations employing vertical-sounder data

The lateral path deviation effects (calculated from model 2) experienced by three transmissions of equal frequencies and ranges, but of different path azimuths (124°, 152° and 168°), are reproduced in Fig. 6.31. Data from three pairs of ionosondes were employed in these calculations, as indicated in Fig. 6.31. In these calculations, the layer parameters hm and ym were derived from the Bradley and Dudeney (1973) model.

The transmission path azimuth corresponding to curve III differs from the bearing of the Luxembourg transmission by only about 15° (124° and 109.3° respectively), therefore, the calculated results (curve III) can be reasonably compared with the observed errors for the Luxembourg path (Fig. 6.1). Similar features are evident, e.g. the negative errors during the early morning and evening/night periods and the positive errors during the middle of the day. These model results therefore represent a considerable improvement over the predictions of model 1, although significant differences still exist.

The diurnal variation of the gradients of the ionospheric parameters f_c, h'_F, h_b and h_m, corresponding to simulation III are reproduced in Fig. 6.32. A similar time dependence to the bearing error is present in the variation of h'_F and h_b gradients, due to the low signal reflection heights for a large
FIG. 6.31 Bearing errors calculated from model 2 employing vertical sounder data.

October 1979 median data
R = 6.33 km
f = 6.09 MHz

Curve III $\beta = 124.0^\circ$
Stough and Laninon

Curve II $\beta = 152.0^\circ$
De Bell and Laninon

Curve I $\beta = 168.5^\circ$
Stough and De Bell transducers

BEARING ERROR (DEGREES)

TIME (UT)
7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
Fig. 6.32. Diurnal variations of the ionospheric parameter gradients (differences) between Slough and Lannion (corresponding to curve II of Fig. 6.32).

\[ \Delta f_0 F_2 (MHz) \]

\[ \Delta \text{HEIGHT (KM)} \]

\[ + \Delta h \text{F} \]

\[ \Theta \Delta h \text{F} \]

\[ \text{TIME (UT)} 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 \]

-189-
part of the day. During the period 1000 to 1600 UT, the hb gradient is positive, whereas the h'F gradient has zero or negative values. This behaviour suggests that the assumption hb = h'F (Lucus and Hayden, 1966 model) which is employed in model 1 is inadequate, and E layer retardation effects need to be taken into account when calculating hb (as performed in model 2 by means of the Bradley and Dudeney, 1973) model.

The bearing errors calculated for a 6.175 MHz transmission having a path azimuth of 152°, and the observed errors for the Issoudun path are compared in Fig. 6.33. Good qualitative agreement is obtained despite the following, i) only hourly vertical-sounder data were employed, ii) the observed error curve is the mean of the bearing measurements made on only five days, and iii) the midpoint of the line joining the two ionosonde stations is displaced from the midpoint of the Issoudun path.

c. Limitations of model 2

The limitations of model 2 arise from the following considerations:

I. Large errors in the path geometry may be obtained, particularly at signal frequencies close to the path MUF since, i) only the transverse ionospheric tilt component is considered, and ii) the ionospheric gradient is assumed constant over a wide area (plane mirror approach).

II. Path deviation effects for multihop signals cannot be calculated since ionospheric information is only available for one reflection point. This is a major consideration since for most of the paths under consideration the two hop F mode is
Fig. 6.33 Comparison of the calculated (model 2) and observed bearing errors

Curve I: Experimental data (mean curve), \( \phi = 14.8^\circ \)

Curve II: Model 2

Both curves - October 1979, \( T = 6.775 \text{ MHz} \), \( R = 603 \text{ km} \)

\( \phi = 152.0^\circ \)

Median data: De Bilt and Larnion sounders

for the Issoudun path
active during the daytime (Appleton Laboratory HF field strength program).

III. A parabolic electron density distribution is representative of the nighttime ionosphere, but is inadequate to represent the ionosphere at low altitudes during the day. The empirical model described by Dudeney (1978) overcomes many of the limitations found in practice with other ionospheric models. The 'Dudeney' model divides the ionospheric profile into three segments: a parabolic E layer, a secant expression for the E-F 'transition region and a cosine function for the F2 region. As for the parabolic model, it can be specified in terms of readily available (or derivable) ionospheric parameters. The electron density profiles calculated from vertical-sounder data for the Slough and Lannion stations by the 'Dudeney' model are reproduced in Fig. 6.34a. The corresponding profiles for a parabolic model are displayed in Fig. 6.34b. The height difference between the profiles at a given plasma frequency provides a measure of the magnitude and sign of the gradient of electron density between the two stations. Similar gradients are indicated from the two pairs of profiles, however, the effective ionospheric tilts experienced by signals reflected at low altitudes in the respective models may be quite different, due to the unequal reflection heights.

These considerations suggest that some improvement in model 1 may be achieved by employing the 'Dudeney' model in place of a parabolic distribution. Unfortunately, a simple analytic expression for virtual height does not exist for this model, unless it is modified by, i) replacing the cosine term
Comparison of the electron density profiles having

a. a form given by Dudeney (1978)

b. a parabolic distribution

Ionospheric data for 1200 UT October 1979

I Slough
II Lannion
by a sech term, and ii) not including an F1 layer. Virtual heights can be obtained for the unmodified 'Dudeney' model by means of numerical techniques, however, to incorporate these into model 1 would require extensive modification to the 3-D technique.

IV. A major limitation of employing median ionospheric data (predicted or measured) is the day-to-day variation of SIT effects. Unfortunately, when employing vertical-sounder data for individual days irregularities very often obscure the background ionospheric tilts. The considerable day-to-day variability of the ionospheric tilts is evident from the h'F parameter gradient between the Slough and Tortosa stations (see Fig. 6.35).

6.3.3 Realistic raytracing investigation

a. The technique

A realistic raytracing program (Jones, 1966) is now employed with the 3-D ionospheric model developed by Rao (1968a). The electron density and its gradients along the ray path are determined from the electron density profiles above four geographic locations (say stations A, B, C, D).

Consider a particular spherical shell passing through the ray point at which gradients are required. Let \( (\Delta N)_1 \) be the difference in the electron densities above A and B, and let \( (\Delta N)_2 \) be the difference above C and D. If \( (\Delta \theta)_1 \) and \( (\Delta \theta)_2 \) are the differences in colatitudes for the two pairs of stations, and \( (\Delta \phi)_1 \) and \( (\Delta \phi)_2 \) are the corresponding differences in longitudes, two equations can be written as
Fig. 6.35 Illustration of the day-to-day variability of the ionospheric parameter gradients.

Five days data (19th to 23rd October 1979)

\[
\Delta h'F = h'F (Soft) - h'F (Tortosa)
\]
where $\frac{\partial N}{\partial \Theta}$ and $\frac{\partial N}{\partial \Phi}$ are the partial derivatives of electron density with respect to colatitude and longitude respectively. These two equations are solved for $\frac{\partial N}{\partial \Theta}$ and $\frac{\partial N}{\partial \Phi}$ and the horizontal gradient is determined. The vertical component of the gradient is obtained from the electron density profile of station A, and the electron density value at station A with the computed horizontal gradients.

### b. Model Results

One hop, ordinary mode raypaths were calculated in the simulations described below. The transmitter and receiver locations were chosen to ensure that the path midpoint is close to the centre of the network of ionosonde stations. Electron density profiles were determined from the 'Dudeney' model together with:

1) data input from the numerical prediction maps

The model results obtained for E-W and S-N transmission paths are presented in Figs. 6.36 and 6.37 respectively. Results obtained for identical paths by model 1 are also displayed. The two models predict a similar diurnal variation of bearing error, however, several differences are evident in the sign and magnitude of the predicted errors. These are probably caused by the unrealistic assumptions of model 1 discussed earlier.
Fig. 6.36 Comparison of the bearing errors predicted by the models 1 and 3

Ionospheric data

CCIR (1967)

October sunspot number = 150

curve I: model 1  
receiver location $52^\circ N$, $2^\circ E$

curve II: model 3
A: $5^\circ N$, $5^\circ E$
B: $4^\circ N$, $5^\circ E$
C: $5^\circ N$, $0^\circ E$
D: $5^\circ N$, $10^\circ E$
BEARING ERROR (DEGREES)

TIME (UT)

-198

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

I

II

Curve I: Model 1

Receiver location 52°N, 2°E

October sunset number = 150

Ionospheric data (CIR, 1967)

Comparison of the bearing errors predicted by the models 1 and 3

\[ \theta = 0.0 \]

\[ R = 600 \text{ km} \]

\[ f = 6.09 \text{ MHz} \]
II) input from vertical-sounder data

Model results are presented for the Issoudun and Luxembourg transmissions in Fig. 6.38. Reasonable agreement is obtained between predictions and observations for the Issoudun path, but much poorer agreement is found for the Luxembourg path. In this case, the (one hop) errors predicted for the daytime will underestimate the true errors, if the propagation mode is a multihop signal. The lack of agreement for the evening/nighttime period is more difficult to explain, but is probably due to the unreliability of the vertical-sounder data for one of the stations.

The major drawback of this model is the use of the Jones 3-D raytracing program which requires a very large amount of computing resources.

6.4 INTERPRETATION OF EXPERIMENTAL OBSERVATIONS IN TERMS OF IONOSPHERIC PROCESSES

The time variation of electron density $N(h,t)$ at any geographical location is given by the continuity equation

$$\frac{\partial N}{\partial t} = q - L(N) - \text{div} (Nv)$$

where $q$ and $L$ represent the rates of production and loss of ionization respectively, and $\text{div}(Nv)$ is the rate at which ionization diverges due to dynamic processes. The ionic drift velocity $v$ is determined from separate equations of motion for the ionospheric plasma. A photochemical regime exists where the ion production and loss terms dominate the transport term in the continuity equation. This occurs in the lower ionosphere where the ion lifetimes are short. In the F2 region,
Fig. 6.38 Bearing errors calculated by model 3 with vertical sounder data.
ion lifetimes are much longer and dynamic processes play a very important role.

6.4.1 The photochemical regime

A useful approximation to the F region production function is provided by the Chapman theory (Chapman, 1931) which assumes, i) a single monochromatic radiation and, ii) a plane stratified exponential atmosphere of scale height $H$, having a single gaseous component. Details of the theory are given in numerous reference works, e.g. Rishbeth, 1967; Rishbeth and Garriott, 1968; Hargreaves, 1979.

The Chapman production function, $q$ is given by

$$q(z, \chi) = q_0 \exp(1 - z - e^{-z} \sec \chi)$$  \hspace{1cm} 6.11

where $z$ is equal to $(h-h_0)/H$, $h$ is the height above the earth's surface, $\chi$ is the solar-zenith angle and $q_0$ is the peak rate of production for an overhead sun ($\chi=0$) which occurs at a height $h_0$. A similar function is obtained for electron density when electron loss processes are also taken into account. The ion production and electron density variations are represented as a function of height and solar-zenith angle in Fig. 6.39. Below the F layer peak, this simple treatment indicates that contours of constant ionization will be lower at geographic locations where $\chi$ is smaller. The diurnal behaviour of SIT-induced bearing errors has previously been interpreted in terms of this solar-zenith angle control of the ionosphere (e.g. Morgan, 1974) as described in chapter five. Many features of the experimental observations presented in section 6.1 can also be explained in

(Fig. 6.39)
this way, as discussed below.

The bearings of the SE-NW Issoudun transmission (see Fig. 6.5) will be greatly affected by the very large E-W ionospheric gradients, particularly during the dawn and dusk transition periods. The following features can be explained by the solar-zenith angle control of the E-W gradients, i) the large negative bearing errors in the morning, ii) the small errors around midday, and iii) the large positive errors in the evening. Positive errors are similarly present between 1500 and 2100 UT for the two S-N Madrid transmissions (see Figs. 6.7 and 6.8).

The solar control of the N-S ionospheric gradients is evident in the daytime bearings of the E-W Luxembourg, Prague and Vienna paths (see Figs. 6.1, 6.3 and 6.4 respectively). Positive errors are caused by lower isionic contours to the south, which are in turn produced by the subsolar point being to the south of the respective undeviated paths.

The simple Chapman theory cannot explain a number of features evident in the experimental observations. The anomalous features present in the bearing variation of the Luxembourg transmission (see Fig. 6.1) are, i) the large positive errors in a two or three hour period after midnight. At night, the horizontal gradients of electron density should be small due to the absence of electron production, although the behaviour of the effective ionospheric tilt is more complicated due to the closer proximity of the signal frequency to the path MUF, ii) the large negative errors in the early morning, although these may possibly be produced by the large
solar-induced E-W ionospheric tilts present at this time. The Luxembourg station has a true great-circle bearing of 109.3°, therefore, the E-W gradients of electron density can affect the bearings of this path, and iii) the large negative errors which occur during the four hour period before midnight. These nighttime errors are also present in the Prague and Vienna bearing observations, but are of a smaller magnitude. The bearings for the SE-NW Issoudun (Figs. 6.5 and 6.6a) and Geneva (Fig. 6.6b) transmissions also exhibit negative errors in the evening, over a period of a few hours. The variation of bearing error during the evening/night periods tends not to be correlated with the time of local sunset, therefore supporting the suggestion that these effects are not solar controlled. The major limitation of the Chapman theory is the neglect of plasma transport processes which are now discussed.

6.4.2 Transport of ionization

In the midlatitude F2 region, the vertical gradient of electron density exceeds the horizontal gradients (except near to the F layer peak). Consequently, vertical ion drift is generally more effective than horizontal ion drift in causing changes in electron density (Rishbeth, 1967). The major processes contributing to the transport of ionization are:

a. Ionic drifts driven by electric fields. These fields are generated by the dynamo action of the E layer and are mapped into the F region along highly conducting geomagnetic field lines. The direction of the ionic drift is perpendicular to both the electric field and the geomagnetic field.
directions. Due to the dip of the geomagnetic field, the effects of the drifts on the ionosphere are small compared with neutral wind effects (Bramley, 1969; Bramley and Ruster, 1971).

b. Thermal expansion and contraction of the neutral atmosphere and plasma. These processes cause only minor effects at the peak of the F layer, but can perturb the shape of the layer below the F layer peak (Rishbeth, 1967).

c. Flow of plasma between the ionosphere and outer atmosphere. This plays an important role in maintaining the F layer at night (e.g. Stubbe, 1968).

d. Ambipolar (plasma) diffusion. In the F2 region, ions and electrons diffuse under the action of gravity and partial pressure gradients, hindered only by collisions with the neutral air. The F2 layer peak cannot be accounted for by a peak in electron production, but forms (in the absence of vertical plasma drift) at a height where the diffusion rate, \( \beta \), is comparable to the electron loss coefficient (attachment), \( \beta \), given by

\[
\beta \sim d
\]

where

\[
d = \frac{D_p}{H_i^2} \sin^2 I
\]

where \( D_p \) is the plasma diffusion coefficient, \( H_i \) is the scale height of the ionizable gas and \( I \) is the magnetic dip angle.

e. Ionic motion caused by momentum transfer from neutral winds (King and Kohl, 1965; Rishbeth, 1972). The ion velocity \( \mathbf{v} \) equals the geomagnetic field-aligned component of the wind velocity \( \mathbf{u} \); in vector terms
The vertical ion drift velocity, \( W \), produced by a horizontal wind, \( U \) blowing at an azimuth \( \theta \) is given by
\[
W = -U \cos(\theta - D) \cos I \sin I
\]  

where \( D \) is the declination of the geomagnetic field.

A drift \( W \) shifts the F layer peak by a vertical distance \( \Delta h_m \) (Rishbeth and Barron, 1960; Rishbeth, 1966) given by
\[
\Delta h_m \approx 0.9 \frac{W}{d(h_m)} \approx \begin{cases} 
0.55 \frac{W}{\beta(h_m)} & \text{day} \\
0.12 \frac{W}{\beta(h_m)} & \text{night}
\end{cases}
\]

where \( \beta(h_m) \) and \( d(h_m) \) are the loss coefficient and the diffusion rate at the peak. By day, the effect of a displacement \( \Delta h_m \) on the peak electron concentration, \( N_m \) is given by (very approximately)
\[
\Delta (\ln N_m) \approx 0.4 \frac{W}{H_i \beta(h_m)}
\]

All these processes have been omitted in the simple Chapman approach.

### 6.4.3 Anomalous SIT effects

The negative bearing errors obtained during the evening/night period, for a number of transmission paths, implies that the isoionic contours are lower to the north or north-east. This behaviour is also suggested from the lower values of the parameters \( h_F^* \) and \( h_m \) to the north as indicated in the vertical-sounder data presented in Figs. 6.22, 6.23 and
Lines of constant h'F and hm are not isoionic contours, however, and at heights near the F layer peak, the isoionic contours are lower to the south, as implied by the lower values of fc to the north. The height dependence of the sign of the N-S tilt is evident in the profiles of Fig. 6.36.

The modelling studies indicate that signal reflection heights are quite low during the evening, implying negative bearing errors, while later in the night signals are reflected higher (signal frequency is closer to the path MUF), and positive errors should occur.

The lower values of hm to the north can be explained in terms of the latitudinal variation of diffusion and neutral wind effects, arising from the variation of magnetic dip with latitude. The \( \sin^2 I \) term of equation 6.13 has values 0.82 and 0.68 at the two geographic locations A(50° N, 0° E) and B(40° N, 0° E). The diffusion rate is therefore smaller to the south, implying that equation 6.12 is satisfied at a greater altitude, and the F layer peak (to the south) will form at a lower height. This assumes correctly that the loss coefficient \( \beta \) varies little over the latitudinal distances considered (\( \beta \) is proportional to the concentration of molecular nitrogen and molecular oxygen, which differ by only a small amount at locations A and B; MSIS neutral thermospheric model, Hedin et al., 1977).

The F layer peak is shifted by a vertical ion drift \( W \) produced by neutral wind effects. At middle latitudes, \( W \) is typically 30 m/s by day and 100 m/s at night (Rishbeth, 1972). The values of \( \Delta hm \) (see Fig. 6.16) will be of the order of tens
of kilometers, depending on the time of day, the atmospheric conditions and the dip of the geomagnetic field. At night, the thermospheric winds travel equatorward producing upward ion drifts which raise the layer. Since wind effects are larger at low latitudes (the sinIcosI term of equation 6.15 takes on values 0.38 and 0.48 at locations A and B respectively) then the F-layer (or hm) will be lower to the north. The lower values of fc to the north are also consistent with the above behaviour (the raising of the F layer peak transports ionization into a region of smaller loss rate and the equilibrium electron density increases).

It follows from the above discussion that important N-S (latitudinal) differences in the height and composition of the F layer are produced by the magnetic dip dependence of diffusion and wind effects. Transport effects are very large at the F-layer peak, but will be smaller at lower F region heights. Due to the complexity of the ionospheric processes governing the F layer height and composition, model calculations are required in which the electron continuity equation and the equation of motion of the neutral atmosphere are solved simultaneously (Bailey et al., 1969; Torr and Torr, 1973). Such studies may allow firmer conclusions to be drawn concerning the nature of the ionospheric processes that cause the anomalous SIT effects.

6.5 CONCLUSIONS

Three raytracing analyses have been performed to simulate
the effect of SITs on bearing measurements:

a. A 3-D ionospheric model based on predictions (model 1) allows the rapid calculation of lateral path deviation effects, but the errors obtained are in poor agreement with experimental observations, particularly for E-W paths. The major cause of this behaviour is the inadequacy of the numerical prediction maps to yield accurate gradient information.

b. A simple model that allows vertical-sounder data to be employed (model 2), produces results in better agreement with the measured bearing errors. This suggests that the results for model 1 could be improved if the prediction maps were updated by vertical-sounder data.

c. The Jones 3-D raytracing technique together with a 3-D ionospheric model (model 3) predicts errors similar to those obtained from model 1, although some differences exist. Nevertheless, this agreement demonstrates the efficacy of employing vertical-incidence virtual-height information (as in model 1) in the determination of lateral path deviation effects. The disadvantage of employing the Jones raytracing program is the great amount of computing resources required. Disappointing agreement between calculated and measured bearing errors is found when employing vertical-sounder data in the 3-D model. Better agreement could probably be achieved by employing data for a different network of sounders and by calculating multihop paths.

SIT-induced bearing errors have previously been explained solely in terms of a solar-zenith angle dependence of the ionospheric tilts. In the present investigation, anomalous
features observed in the nighttime bearing errors cannot be explained in this way. Dynamic processes are also important in determining the F layer electron density distribution and the influence of plasma diffusion and neutral wind effects on the N-S horizontal ionization gradients is discussed.
CHAPTER SEVEN

THE REDUCTION OF BEARING ERRORS DUE TO PROPAGATION EFFECTS
IN HF DIRECTION FINDING

7.1 INTRODUCTION

The intrinsic high accuracy of modern HF direction finders (accuracies better than 0.5 degrees standard deviation) cannot be realized in practice due to bearing errors which are produced by perturbations and tilts present in the ionosphere. It is therefore of great importance, either to apply corrections for these propagation bearing errors on an instantaneous basis or to assign confidence limits on their likely magnitudes. In the present investigation, the bearing errors arising from magneto-ionic deviations (Davies, 1968; Rao, 1969; Gething, 1978) are not considered since they are generally very small (< 0.1°) for the transmission paths under consideration.

7.2 METHODS OF CORRECTING FOR TID-INDUCED BEARING ERRORS

7.2.1 Doppler techniques

Since the occurrence and form of a TID cannot be predicted, any correction technique must be based on the observation of a TID at or near the time at which the bearings are measured. The Doppler technique has been employed extensively for this purpose. The passage of a TID displaces
the reflection point of an HF wave reflected from the ionosphere. The frequency of the reflected wave will thus be slightly changed by the Doppler effect associated with the moving reflector. The magnitude of the Doppler shift will depend on the rate of movement, i.e. the rate of change of phase path of the radio wave. By employing at least three spaced Doppler transmitters (very highly stable in frequency) together with a central receiver, the TID velocity can be determined from the time displacements between the Doppler records and the transmitter-receiver path geometries (Davies et al., 1962; Georges, 1967).

Jones et al. (1975) employed four Doppler sounders (the additional sounder provides the redundancy to allow a cross-check on the TID velocity measurements) spaced by approximately 60 km and situated near the midpoint of an oblique 260 km HF path. The frequency of this transmission was chosen to give approximately the same reflection height in the ionosphere as the transmissions of the Doppler system. The fluctuations in the bearings recorded by the DF correlated well with individual TIDs (generally medium-scale) observed in the Doppler records. Bearing errors produced by the TID were simply calculated from the values of the ionospheric tilt at right-angles to the HF path and the approximate reflection heights determined from ionograms obtained for the path midpoint. Good agreement between the calculated and measured bearing errors was achieved during single-mode propagation conditions.
An extension of the Doppler technique to longer propagation paths has been reported by Jones and Spracklen (1976). A highly stable HF transmission received over an 850 km path is monitored by a DF. The signal frequency is also recorded at three receiving sites spaced symmetrically around the DF and located at the vertices of an equilateral triangle of side approximately 122 km. A fourth Doppler receiver is colocated at the DF. These authors have described in detail how the displacement of the reflection point can be calculated and an accurate correction made for the error in the observed bearing. A major drawback of this technique is that integration of the frequency record over a time interval greater than half the TID period is necessary to obtain a bearing correction. Operationally, the signal may be of short duration, say of the order of a few minutes, and such an integration procedure could not therefore be applied. Another major problem in implementing this system is that, in general, the target transmitter will not be frequency stable, moreover it will probably be modulated in some way. Thus, it is difficult to distinguish frequency shifts due to ionospheric changes from those originating at the transmitter.

Jones and Spracklen (1978) described a 'difference Doppler' technique in which the Doppler frequency changes at each pair of receivers (employing the same experimental configuration as Jones and Spracklen, 1976) are compared to produce a 'difference Doppler' frequency. If the ionosphere is stationary then all three difference frequencies are zero, however, if ionospheric tilts exist then the difference frequencies are non zero.
Hence, by using this technique, frequency instabilities of the transmitter are overcome since these will be the same at each receiver and therefore nulled by the difference process. The magnitude of the difference frequencies provides a quantitative index of bearing accuracy. Those bearings classified as accurate have considerably less error and smaller variances than those of the complete data set. When more than one propagation mode is present multiple frequency shifts are observed in the received signals, however, the dominant modes can be identified by spectral analysis (Jones and Spracklen, 1977). When an E-mode is dominant, bearings are characterized by low variances and therefore accurate bearings can be identified at these times by the 'difference Doppler' technique.

7.2.2 Reference to bearings of known transmitters

The major disadvantage of the 'difference Doppler' system is the need for Doppler receivers at three locations other than at the DF site, whereas operationally a minimum of auxiliary equipment is desired. An error correction technique based on monitoring the bearings of reference transmitters (of known location) close to the target transmitter (of approximately known location) was suggested by Jones and Reynolds (1975). These authors reported good correlation during single-mode propagation conditions between the bearings of pairs of transmitters closely spaced in both frequency and geographic location. Reynolds and Morgan (1975) studied the effect of frequency separation on the correlation of bearing errors.
measured on HF signals from colocated transmitters. During one TID active period, excellent correlation was obtained between the bearing fluctuations (quasi-periods of about 20 minutes) of two transmissions having a frequency difference of 4.5 MHz (when a suitable timeshift was applied to one of the bearing records). Poorer correlation was obtained for TIDs having shorter quasi-periods, e.g. 12 minute waves were well correlated only on transmissions having frequencies that differed by less than 2 MHz.

In the present investigation, the bearings of a matrix of reference transmitters are monitored continuously to provide real-time TID velocities and magnitudes. Initially, it was proposed that a suitable choice of reference transmitters would allow some form of bearing correction to be undertaken. The feasibility of this approach is now discussed:

a. Medium-scale events

Medium-scale TIDs have quasi-periods generally shorter than thirty minutes and the TID-induced bearing error magnitudes are usually less than ±3 degrees for signal path lengths of 600 to 3000 km (Ross and Bramley, 1949; Bramley and Ross, 1951; Bramley, 1953; Morgan, 1972). They occur as single cycles or as wavetrains and are observed predominantly during the day. These waves are dispersive and change their shape as they propagate over distances less than 100 km (Georges, 1968). Therefore, to measure their characteristics, transmissions must be monitored which have reflection point separations much less than 100 km in the horizontal direction. Although a library of many hundreds of transmitters of known location can be compiled,
their ground separations are generally greater than 200 km (reflection point separations greater than 100 km). Consequently, the correction of bearing errors induced by medium-scale TIDs is not generally possible by reference to bearings measured on known transmitters.

b. Large-scale TID events

Large-scale TIDs produce bearing error variations with quasi-periods from about thirty minutes to four hours and magnitudes generally less than five degrees for signal path lengths 600 to 3000 km. These events seldom have more than two or three cycles, their sources lie in the auroral zone and they propagate in an equatorward direction. Due to their very long wavelengths (> 1000 km) dispersive effects occur only over great horizontal distances of the order of 500 km. Transmitters having ground separations of 200 to 500 km exist in large numbers suggesting that good correlation between the bearings of such spaced transmissions should be achieved.

An extensive study of the correlation of large-scale TID-induced bearing errors on spaced paths (see chapter four) has demonstrated that poor correlation is often obtained even in the case of transmissions with closely spaced reflection points. This behaviour has been explained in terms of the response characteristics of the HF direction finding technique and the complicated temporal and spatial characteristics of AGWs. In practice, it is also extremely difficult to find a matrix of reference transmitters (three or more) that satisfy the following criteria:

1. The true position of each of the reference
transmitters is accurately known.

ii. The signals are reliably received for a reasonable length of time (of the order of hours).

iii. The signals have similar path geometries so they are reflected near the same height (and near to the height at which the target transmitter is reflected).

iv. The signal frequencies are high enough to ensure single-mode propagation, or at least to ensure that wave-interference effects are small. 'Broadcast' transmissions are particularly inadequate in this respect since they generally have frequencies well below the path MUFs' for the intended reception areas. Unfortunately, to satisfy this last criterion the reflection heights are often required to differ significantly from the reflection height of the target transmitter.

The above considerations indicate that even in the case of large-scale TIDs, bearing correction is generally not possible by the reference transmitter technique.

7.3 METHODS OF CORRECTING FOR SIT-INDUCED BEARING ERRORS

7.3.1 Regression analysis

SIT-induced errors can have similar magnitudes to those caused by TIDs, therefore, a technique for reducing these errors is also required. Regression analysis is based on the assumption that the bearing correction required on a specified path is a function of various path parameters and ionospheric conditions. From an analysis of many observations on a wide
range of paths it might be possible to determine a suitable relationship between bearing correction and the most important factors. For example, the observations presented in chapter six indicate that bearing errors of one to three degrees are observed for most E-W paths during the day and significant negative errors are sometimes evident during the night, moreover on S-N paths large negative errors after local sunrise, small errors around noon and large positive errors in the evening are generally observed. However, SITs are highly dependent on the ionospheric conditions as dictated by solar activity, season, geomagnetic activity etc. Consequently a very large data set would have to be analysed to achieve corrections on a regular basis.

7.3.2 HF raytracing

HF raytracing through a 3-D representation of the ionosphere enables the rapid calculation of SIT-induced bearing errors for various path geometries and ionospheric conditions. The use of monthly median prediction maps of ionospheric characteristics in these calculations is described in chapter six. Unfortunately, the considerable smoothing of the ionosonde measurements involved in the preparation of such maps leads to inaccuracies in the predicted horizontal electron density gradients. This technique in its present form can therefore only predict the qualitative behaviour of SIT effects and then only for S-N paths which are influenced by the very large dawn/dusk E-W ionospheric gradients. Nevertheless, it is envisaged that vertical-sounder data can be employed to update
the numerical maps of coefficients so yielding more accurate ionospheric gradient information. These updated maps should enable more accurate predictions of SIT errors to be obtained for both S-N and E-W transmission paths. The vertical-sounder data employed can be taken from a previous year for which similar ionospheric conditions were prevalent, hence, this technique is suitable in an operational sense since it does not require real-time ionosonde measurements.

In addition to accurate gradient information, the successful prediction of SIT-induced bearing errors requires a knowledge of the dominant propagation mode for the particular path. A suitable HF field-strength prediction program can be employed for this purpose, e.g. the 'APPLAB' program developed at the Appleton laboratory. The day-to-day variability of the ionosphere will set a limit on the accuracy of the predicted SIT-induced bearing errors, however, some knowledge of the ambient ionospheric conditions can be gained from the reference transmitter technique.

7.4 WAVE-INTERFERENCE EFFECTS

Due to the effects of wave-interference, arising from the presence of a number of propagating HF waves at the DF, bearing measurements often exhibit rapid fluctuations with periods of the order of fractions of a second to minutes and errors of a few tenths to many degrees (Bain, 1953, 1956; Hayden, 1961). Each ray path (mode) involves some combination of reflections from the major ionospheric layers and from the earth's surface.
Furthermore, each mode cannot, in general, be represented as a single plane wave but rather consists of a narrow angular distribution of energy (Booker et al., 1950). Each of these modes consists of a specular component surrounded by a cone of many diffracted rays. In the case of intermediate-range temperate latitude paths, however, the distribution of diffracted energy is often very narrow resulting in a net effect closely approximating that of a single ray. The received angular energy spectrum is further complicated by the possible simultaneous existence of high and low ray propagation for certain modes (for frequencies close to the path MUF), and magneto-ionic splitting of the rays into O and X components.

The occurrence of a multiplicity of rays, each arriving at the DF site via a somewhat different ionospheric path, results in a very complicated spatial field configuration at any instant in time. Since each ray is normally subject to phase and amplitude variations due to time variations in the ionosphere along each path, this field configuration also varies with time in an extremely complicated manner. Wave-interference effects in the measured bearings are caused by the inability of the DF to resolve or separate the various rays producing the field configuration.

In the simple two ray case where one ray is dominant, the indicated bearing fluctuates (typically about ± 0.5°) about a mean value very close to the dominant ray bearing even when the ray elevation angles differ considerably (capture effect). However, if the ray amplitudes have nearly equal amplitudes the situation is more complicated and much larger swings in the
bearing can occur. The mean bearing in this case depends on the proportion of the observation time for which each mode component is stronger.

Time averaging techniques can be employed to achieve significant reductions in wave-interference errors (Bailey, 1954; Bain, 1956). However, the two major limitations of time averaging are, i) the incident signal may be of a very short duration, and ii) a stationary random process is assumed whereas, in practice, this is often not true even over periods of a few minutes. Consequently, other methods for reducing these errors would be useful, e.g. by mode resolution.

Modes may be resolved in several ways as indicated below:

a. The conventional DF does not usually have a large enough aperture to resolve modes in azimuth. Very large linear arrays designed primarily for research are capable of separating the azimuth of arrivals for two modes separated by one degree with an accuracy of 0.1 degrees, but only for signals arriving from broadside directions. The implementation in a circular form to provide the same resolution in any azimuth sector would be very expensive.

b. Modes can generally only be separated in the time domain in the case of pulsed signals, e.g. Bramley and Ross, 1951).

c. Resolution techniques that employ the relative Doppler shifts between modes (e.g. Davies and Baker, 1966) generally require transmitters with carriers that are very stable in frequency. In addition, there is the difficulty that the ionosphere does not always produce significant differences.
in the Doppler shifts between modes.

d. Some success in mode resolution with respect to elevation angle has been achieved by a number of authors, e.g. for a Wullenweber (Jones et al., 1966); by means of a separate vertical array (Gething et al., 1969; Kelso, 1972).

Possible techniques for the calculation of ray parameters from the resolution of multicomponent wavefields in the case of circular arrays (Wullenweber) have been discussed extensively by Gething (1978). These involve the comparison with the aid of iterative methods of a single waveform (radiation pattern) or a sequence of waveforms from a spinning goniometer with waveforms generated by computer simulation. When elevation angles are determined separately, mode identification is assisted and unknowns are removed from the numerical analyses. However, these techniques are not sufficiently powerful to resolve the many complex types of wavefield possible in a limited sample of noisy data.

The development of multichannel receivers in recent years and the availability of digital computers allow the digital processing of signals (amplitude and phase) measured at the elements of a DF array. These signals are measured and recorded digitally and processed by a computer program in a procedure referred to as wavefront analysis (WFA). Some success has been achieved in resolving multiray wavefields by means of WFA but many problems remain to be overcome before the technique is widely employed in operational DF (Gething, 1978).

In the present investigation, signal processing is basically of an analogue nature (spinning goniometer) and no
elevation angle measurements are available. Consequently, the reduction of wave-interference effects is achieved simply by time averaging (over a period of eight seconds for each bearing measurement; eight or nine transmitters are generally monitored so that one bearing is obtained for each signal every 80 to 90 seconds). A significant reduction in the wave-interference error can be achieved by averaging the bearing over a longer time period (monitoring fewer than eight transmissions).

7.5 ASSIGNING VARIANCES TO BEARING MEASUREMENTS

7.5.1 Reference transmitter technique

A variance can be associated with a single bearing measurement, although it must be calculated from a large sample of bearings taken on the same transmitter under identical conditions. This variance value places confidence limits on the possible errors present in a bearing measurement of a given transmission. An important parameter in HF direction finding is the 'DF variance', \( V \) defined as

\[
V = \sigma^2 + (SE)^2
\]

where \( \sigma \) is the standard deviation about the sample mean, with wilds excluded (i.e. bearings outside some arbitrary limit such as \( \pm 10^\circ \)); \( SE \) is the systematic error, i.e. the mean value of the bearing errors excluding wilds. 'DF Variance' is the parameter employed in weighting bearings and in computing probability areas (Gething, 1978).
Variance depends on a number of factors such as the DF site, the instrument, the time period over which the bearings are averaged, the 'nature' of the transmission and the ionospheric conditions. The dependence of the measured variances on the transmitter and the transmission path is illustrated by the following examples:

a. The bearings recorded for four paths on 23 March 1981 are reproduced in Fig. 7.1. The calculated standard deviations of the bearings for the Madrid and S.R.I. signals (0.61° and 1.52° respectively) are significantly smaller than those for the Gibraltar and Issoudun signals (2.13° and 2.96° respectively). Smaller values of standard deviation correspond to higher values of the signal frequency to the path MUF ratio (f/MUF) for which single-mode conditions are more likely (see Fig. 7.2). This is a typical result, although exceptions do occur as illustrated by the following example.

b. The bearings recorded for three Berlin frequencies on 15 August 1980 are presented in Fig. 7.3. The scatter in the bearings for the 14.46 MHz signal is significantly less than for the 11.459 MHz signal, however a corresponding decrease in wave-interference effects is not observed for the 16.358 MHz signal and periods of large scatter are evident. The 16.358 MHz frequency is close to the path MUF implying that wave interference arises from the presence of both the 0 and X magneto-ionic components and the high and low angle rays.

c. The severity of wave-interference effects can change dramatically when propagation switches from single to multimode conditions (or vice-versa). This result is evident in the
Fig. 7.2 Calculated values of $f/MUF$

- **S** - S.R.I
- **M** - Madrid
- **I** - Issoudun
- **G** - Gibraltar

**MARCH median**

Sunspot number = 110

**Time (UT)**

14 15 16 17 18 19 20 21 22 23

signal frequency / MUF
FIG. 3: Unfiltered Bearing Data for 15 August 1980
Vienna records reproduced in Fig. 7.4. The smaller variances before 0600 UT in (a) and after 2100 UT in (b) correlate with times when the one hop F mode is dominant, as indicated by predictions (see Fig. 7.5). This behaviour occurs regularly for the Vienna signal as it does for other transmission paths.

d. The radiation pattern of a transmitter array can influence the relative mode strengths of an HF transmission (e.g. Ross, 1975). The wave-interference errors present in the records of the 10.125 and 10.308 MHz Prague transmissions often have quite different magnitudes (as illustrated in Fig. 7.6). This behaviour can possibly be explained if the respective radiation patterns for the two transmissions are appreciably different. The transmitter power and type (CW, FSK, etc.) will also influence the nature of the received signal, and possibly the measured variances. Signal-to-noise considerations suggest that a wider receiver bandwidth should be employed for CW than FSK transmissions.

Previous statistical analyses of samples of bearings taken on large numbers of reference transmitters under a wide variety of propagation conditions (e.g. Morgan and Sweet, 1971) have produced similar results to those presented above. Such analyses 'sort' bearing errors against various path parameters and lower variances are obtained for, i) single hop modes compared with multihop modes, ii) low-angle rays compared with high-angle rays, iii) E modes compared with F modes, and iv) single-mode rather than multimode conditions. These variances are for 'average' propagation conditions whereas, in practice, variances can differ markedly from day-to-day for a given transmission path. This behaviour arises from solar and
Fig. 24 Example of the change from single to multimoode conditons (and vice versa)
Signal frequency / MUF

numbers=signal component signal-to-noise ratio (dB)

Support number = 180

MUF median

Resolution = 2 hours

RX and TX - isotropic vertical polarization

noise power bandwidth = 750 Hz

Transmitter power = 100 kW

Fig. 15 Application HF Field Strength Predictions for the Vienna-DF path
seasonal variations, but also from the effects of TIDs, spread F, sporadic E, etc. Degradation of fix accuracy has been detected during periods of ionospheric storms (Gething, 1961; Morgan, 1972). The above processes will affect the nature of the ionospheric tilts in addition to influencing the relative strengths of the propagation modes present in a given signal.

The bearings measured for the eight transmitters monitored on 26 March 1981 have been presented in chapter two (Fig. 2.7). The bearing accuracies for the various signals are quite different due to the path dependence of the SIT, TID, and wave-interference effects, as discussed previously. In addition, larger variances occurred on this day compared to other days for which no TID activity is present. The day-to-day change in the bearing accuracy for the Luxembourg signal is illustrated in Fig. 7.7. These examples suggest that some form of real-time monitoring of the ionospheric conditions is required in a variance-estimation method.

By continuously monitoring a matrix of reference transmitters the variances corresponding to the different propagation conditions can be determined experimentally. It is therefore possible to immediately assign a variance to any target transmission that may be monitored, even if it is only present for a very short time.

In practice, variances for a large number of reference transmitters can be obtained by, i) monitoring more than the eight or nine transmitters of the present study, ii) periodically changing the transmitter set, and iii) employing more than one DF6/receiver. Consequently, the variances for a
Fig. 7.7 Unfiltered Bearing records for the Luxembourg signal
wide range of propagation conditions can be determined. The precise nature of the variance-estimation technique will depend on the operational circumstances, but the real-time variance information should be supplemented by, i) HF field-strength predictions, and ii) knowledge of variances measured previously during similar ionospheric conditions.

7.5.2 Quality Factor

The DF associates a quality factor (QF) to each bearing determination and this is intended to provide a measure of bearing reliability. The parameter has values from zero (highest reliability) to 99 (lowest reliability). The associated QFs corresponding to the bearing measurements of Fig. 7.1 are presented in Fig. 7.8. The relationship between bearing error and quality factor is considered below:

a. The QF is not simply related to the presence of TIDs or SITs.

b. Low values of QF are predominant during periods of little wave-interference, while a large spread in values of QF are produced during severe wave-interference conditions.

The standard deviation of the bearings of eight reference transmitters monitored during a sixteen hour time period has been determined for different values of the QF cut off (defined as the QF above which associated bearings are discarded before the standard deviation is determined). The variation of standard deviation with QF cut off from 0 to 99 together with the percentage of the total data available at each cut-off value is presented in Fig. 7.9. Standard deviation is found to
Fig. 7, 8 Quality Factor values (23 March 1981)

Issoudun

Gibraltar

S.R.I

Madrid
Percentage of data available 

6400 bearing measurements

October 23, 1979

FIG 7.9 An Investigation of Quality Factor

Standard deviation (degrees) +

Quality Factor cut-off

Percentage of data available %
decrease as QF cut off is reduced.

c. One-to-one correlation between low values of QF and small bearing errors is not always obtained, even when TIDs and SITs are absent. This suggests that the QF in its present form has limited practical use, since information gained from the QF can be obtained directly from the scatter in the measured bearings. Further work is being undertaken at Leicester to improve the algorithm which produces the QF.

7.6 CONCLUSIONS

The reduction of the bearing errors produced by TIDs is generally not possible by means of the reference transmitter technique due to the following factors, i) the complicated spatial and temporal characteristics of TIDs, ii) insufficient numbers of suitable reference transmitters, and iii) TID-induced bearing errors are often obscured by wave-interference effects, although these can be significantly reduced by averaging the bearings over a suitably long period. In this case, the bearings take on the variation of the dominant mode. The lack of mode identification/resolution prevents the geometrical factor linking the bearing error with the ionospheric tilts to be ascertained. Nevertheless, HF field-strength programs can give useful information concerning the relative strengths of the modes present in a given signal.

Lateral path deviation effects due to SITs can be predicted by employing a suitable HF raytracing technique. The limitations of analyses based on numerical maps of prediction
coefficients are, i) the prediction maps do not yield accurate gradient information, and ii) the day-to-day variability of SIT effects can be large for some paths, particularly for the night period. Nevertheless, some improvement in the accuracy of the predictions can probably be achieved by updating the prediction maps with readily available vertical-sounder data.

A variance-estimation scheme is proposed whereby confidence limits can be readily assigned to the bearing of a target transmission. Variances for a wide range of propagation conditions can be determined experimentally by monitoring the bearings of transmitters of known bearing. This approach probably has more immediate practical application than some of the error correcting methods discussed.
The initial objectives of this thesis were to:

a. correct for TID-induced bearing errors by reference to bearings of known transmitters,

b. predict the bearing errors produced by SITs, and

c. determine bearing reliability such as the expected bearing error limits.

In order to quantify the bearing errors produced by propagation effects, measurements of the bearings of European transmitters of known location were made during the period 1979 to 1981, and some 700 hours of data examined. These experimental observations confirm that the major types of propagation bearing error are produced by wave interference, TIDs and SITs.

Medium-scale TIDs are dispersive and change their shape over distances of a 100 km or less, therefore to measure their characteristics, transmissions having reflection point separations of much less than 100 km are required. Unfortunately, insufficient numbers of transmitters exist to allow such measurements to be made on a regular basis. Consequently, attention has been primarily confined to the study of large-scale TIDs, for which dispersive effects are significant over considerably greater horizontal distances.

The bearing measurements indicate that the degree of correlation of large-scale TID-induced bearing errors on spaced
paths varies considerably from path to path and from event to event. Correlation decreases as the reflection point separation increases, but poor correlation can also occur between the bearing fluctuations of quite similar paths.

Extensive ray-tracing calculations have demonstrated that even simple TID models can qualitatively reproduce several characteristics of TIDs, observed as the variation of bearing with time. These studies indicate that the transmission path geometry has an important influence on the magnitude and time variation of TID-induced bearing errors.

Many other mechanisms have been identified which reduce the correlation of bearing errors on spaced paths. These can be related either to the response characteristics of the HF direction finding technique or to the temporal and spatial characteristics of AGW/TIDs. Because of the complicated nature of these mechanisms it is generally not possible to correct for TID-induced bearing errors by reference to the bearing fluctuations of known transmitters.

The data collected are particularly suitable for the study of lateral path deviation effects produced by the large-scale geographic and temporal variations of the ionosphere. These SIT-induced bearing errors exhibit marked transmission path and seasonal dependencies. In accordance with previous studies, the diurnal variation of these errors can be largely interpreted in terms of a solar-zenith angle dependence of the ionospheric tilts.

Anomalous features observed in the nighttime bearing errors, however, cannot be explained in this way. This
behaviour has been accounted for in terms of the effects of transport processes on the F layer electron density distribution. In particular, the influences of plasma diffusion and neutral winds on the latitudinal gradients of electron density have been considered.

Examination of the experimental observations indicates that SIT-induced bearing errors are reproduced from day-to-day in a consistent manner. Raytracing analyses have been performed to reproduce these errors for a wide range of transmission paths and ionospheric conditions.

A previously developed model based on ionospheric predictions has allowed the rapid calculation of lateral path deviation effects, but the results obtained exhibit poor agreement with the experimental observations, particularly for E-W paths. The major cause of the poor results is identified as the inadequacy of the ionospheric predictions to yield accurate gradient information.

A simple model has been developed in which vertical-sounder data is employed to provide the ionospheric gradient information. Despite the simplicity of the model, better agreement between the calculated and measured bearing errors is obtained than is the case using ionospheric prediction information. The efficacy of employing equivalent vertical-incidence virtual-height information to define an effective reflection surface has been confirmed by raytracing analyses using the Jones 3-D program.

A variance-estimation scheme is proposed whereby confidence limits are assigned to a target bearing. The
combined effects of wave interference, TIDs and SITs present in the bearing measurements are strongly dependent on the transmission path and the ionospheric conditions. By continuously monitoring a number of reference transmissions the variances corresponding to the propagation conditions can be determined experimentally. It is therefore possible to immediately assign a variance to a target transmission, even if it is present only for a very-short time.

Suggestions for future work

a. The mechanisms that produce poor correlation of the TIDs observed in the bearings of spaced transmission paths need further study. In future investigations, the bearing data should be supplemented by additional ionospheric information from another experimental technique. Ideally, some form of identification of the mode content of the received signal should also be available. This additional information would enable many of the uncertainties involving the nature of the mechanisms to be resolved.

b. The accuracy of the SIT-induced bearing errors predicted by raytracing analyses could probably be improved by employing vertical-sounder data to update the prediction maps.

c. The interpretation of the anomalous SIT errors observed during the night would be aided by theoretical studies involving an ionospheric model in which the electron continuity equation is solved simultaneously with the equations of motion of the neutral atmosphere.

d. The precise nature of the proposed variance-estimation technique will depend on the operational circumstances and
limitations. These should be identified and a practical system developed to allow variance estimations to be made in a routine manner.

e. The 'quality factor' associated with each bearing determination has been found to have limited practical use. Some improvement in the DF algorithm which produces this parameter may be possible.

The accuracy of HF direction finding is currently limited by the ionospheric conditions. The conclusions and suggestions for further studies outlined above should lead to a marked improvement in the operational accuracy of DF systems.
REFERENCES

Adcock, F. (1919), British Patent 130490


Bain, W.C. (1953), Proc. IEE, 100(III), p.253

Bain, W.C. (1956), Proc. IEE, 103B, p.113


Beckwith, R.I., Bailey, A.D., Rao, N.N. (1972), Univ. Illinois, RRL Tech. Rept. 32


Bradley, P.A. (1982), Private Communication


Bramley, E.N. (1955), Proc. IEE, 102B, p.544


CCIR(1970), Rept. 252-2, Documents of the 12th Plenary Assembly, CCIR, New Delhi, (ITU Geneva)
Cooley, J.W., Tukey, J. (1965), Math. of Computation, 19, p. 297
Croft, T.A., Fenwick, R.B. (1963), J. Res. NBS 67D, p. 735
Crone, W. (1944), Dtsche Luftfahrtforsch, Untersuchungen und Mitteilungen, No. 802/1 (see Landmark, B. 1968 Agard XXIX Lecture Series III-56)
Davies, K. (1968), AGARD Lecture Series XXIX, 'Radio Wave Propagation'
Davies, K., Jones, R.M. (1968), Proc. 11th AGARD-EDC Symposium, Technivision, Maidenhead, England
Earp, C.W., Godfrey, R.M. (1947), J. IEEE, 94, IIA, p.705
Gething, P.J.D. (1966), Proc. IEE, 113, p.49
Gething, P.J.D. (1978), 'Radio direction-finding and the resolution of multicomponent wavefields', Peter Peregrinus Limited
Halcrow, B.W., Nisbet, J.S. (1977), Radio Sci., 12, p.815


Hockley, J. (1973), Rad. and Elect. Eng., 43, p475


Jones, T.B., Kantarizis, E., Morgan, A.D. (1975), Proc. IEE, 122, p.241


Jones, W.B., Gallet, R.M. (1962), J. Res. NBS, 66D, p.419

Kelso, J.M. (1972), Radio Sci., 7, p.245
Kennelly, A.E. (1902), Electrical World and Engineering, 15, p.473
Maxwell, J.C. (1873), 'A Treatise on Electricity and Magnetism', Vols. I and II, OUP
Morgan, A.D. (1971), Canukus Target Location Conference, Paper No. UK 3
Morgan, A.D. (1972), AGARD Conf. Proc., No. 115, Wiesbaden, p. 28-1
Munro, G.H. (1958), Aust. J. Phys., 11, p.91

Rao, N.N. (1968a), Radio Sci., 3, p. 796
Rao, N.N. (1968b), Radio Sci., 3, p. 1113
Rao, N.N. (1968c), Radio Sci., 3, p. 1119


Ross, W. (1947), J. IEE, IIIA, p. 147
Ross, W. (1949), DSIR Special Report No. 19
Ross, W., Bramley, E.N. (1947), Nature, 159, p.132
Smith-Rose R.L., Barfield, R.H. (1926), J. IEE, 64, p.831
Strangeways, H. (1979), 'Ionospheric Topography Progress Report', No. 22, Leicester University
Walton, E.K., Bailey, A.D. (1976), Radio Sci., 11, p. 175
Wright, M.D. (1961), Nature, 190, p. 898

Text cut off in original
BEST COPY

AVAILABLE

Variable print quality