Observations of HF Radiowaves Propagated Over High Latitude Paths

Thesis submitted for the degree of Doctor of Philosophy at the University of Leicester

Stephen Milan
Department of Physics and Astronomy
University of Leicester

October 1994
Observations of HF Radiowaves Propagated Over High Latitude Paths

Stephen Milan

ABSTRACT

This thesis presents results from the NONCENTRIC experiment, conducted by the University of Leicester, in which five trans-auroral and polar cap HF propagation paths were monitored during two one-month campaigns. Signal strength, noise level, Doppler spreading and signal recognition were determined each hour on fourteen frequencies in the range 3 to 23 MHz.

The diurnal variations in signal recognition and signal strength are consistent with the ionospheric changes produced by solar illumination. This behaviour is modified by processes occurring within the auroral zone and polar cap during disturbed geomagnetic conditions.

Electron densities within the auroral ionosphere increase as a consequence of particle precipitation during disturbed conditions. Auroral enhancement of the D region attenuates signals on trans-auroral paths. This occurs in bursts with durations of tens of minutes, simultaneous with the occurrence of substorms. The magnitude of the absorption can be correlated with changes in the geomagnetic field at geosynchronous orbit. In contrast, enhanced electron densities within the auroral zone E and F regions increase the maximum frequency of propagation, especially at night, thus extending the propagation bandwidth. Within the winter polar cap ionosphere, sporadic high frequency propagation becomes possible from ionization patches convected from the dayside ionosphere.

Prolonged periods of geomagnetic disturbance result in global changes in F region electron density, known as ionospheric storms. Four storms studied produce a decrease in the maximum usable frequency, a degradation of propagation reliability and, within the polar cap, an increase in the lowest usable frequency. Such periods of propagation degradation typically have a duration of five days, even during mild storm conditions.

This comprehensive study of high latitude HF propagation has produced an improved understanding of the changes in the electron density of the polar ionosphere during both quiet and geomagnetically disturbed conditions.
## Contents

### Chapter 1 Introduction

1.1 Introduction .......................... 1  
1.2 The NONCENTRIC experiment ....... 2  
1.3 Arrangement of the present investigation .......... 2

### Chapter 2 Review of HF propagation in the high latitude ionosphere

2.1 Introduction .......................... 4  
2.2 The ionosphere ......................... 4

2.3 The high latitude ionosphere: morphology and coupling with the magnetosphere .......... 5  
2.3.1 The substorm ....................... 6  
2.3.2 Geomagnetic activity and indices .......... 6  
2.3.3 The auroral oval .................... 6  
2.3.4 Precipitation regions within the auroral zones ........ 7  
2.3.5 The polar cap ....................... 9  
2.3.6 Sporadic E and auroral E ............... 10  
2.3.7 The mid latitude trough ............... 11  
2.3.8 Geomagnetic and ionospheric storms .......... 12

2.4 HF propagation ........................ 14  
2.4.1 Reflection of radio waves ............. 14  
2.4.2 Radio wave paths in the ionosphere ........ 16  
2.4.3 Propagation in the presence of horizontal plasma gradients and sporadic E ........ 18  
2.4.4 Scatter propagation ................. 19  
2.4.5 Attenuation of radio waves ............. 20  
2.4.6 Signal fading ........................ 21

2.5 Summary ................................ 22

### Chapter 3 Data collection: the NONCENTRIC experiment

3.1 Experimental arrangement .............. 23  
3.1.1 Propagation paths ................... 23  
3.1.2 Transmitter duty cycle ............... 24  
3.1.3 Transmission sequence ............... 24  
3.1.4 Transmitter and receiver antennas ....... 24  
3.1.5 Receivers .......................... 25  
3.1.6 Scaling of signal characteristics ........ 26  
3.1.7 Signal recognition ................... 27

3.2 Limitations of the data .................. 28  
3.2.1 Equipment limitations ............... 28  
3.2.2 Time resolution ..................... 29  
3.2.3 Sampling interval .................... 29

3.3 Interference .......................... 30  
3.3.1 Local interference .................... 30  
3.3.2 Global interference ................... 31

3.4 Summary ................................ 31
Chapter 4 Preliminary analysis

4.1 Introduction 32

4.2 Recognition of signals 32
  4.2.1 Diurnal variation of signal recognition 32
  4.2.2 Seasonal variation of signal recognition 34
  4.2.3 Dependence of signal recognition on path length 35
  4.2.4 Dependence of signal recognition on geomagnetic activity 35

4.3 Spread index 36
  4.3.1 Path dependence of spread index behaviour 36
  4.3.2 Auroral oval dependence of spread index 38
  4.3.3 Solar-controlled diurnal and seasonal variation of spread index 39

4.4 Signal level 39
  4.4.1 Path dependence of signal level behaviour 40
  4.4.2 Solar and auroral control of absorption 41
  4.4.3 Dependence of signal level on fading 41

4.5 Summary 42

Chapter 5 Radiowave absorption in the high latitude ionosphere

5.1 Introduction 43

5.2 Daytime variation of HF signal level 43
  5.2.1 Solar control of D region electron density and HF signal level 43
  5.2.2 Comparison of low and high latitude signal level variation 45

5.3 Auroral absorption and the magnetospheric substorm 46
  5.3.1 Geomagnetic dependence of auroral absorption 46
  5.3.2 The spatial and temporal pattern of auroral absorption 46
  5.3.3 Comparison of $K_p$ and $D_s$ as indicators of absorption level 47

5.4 Model of energy flux of $>30$ keV electron precipitation in the auroral zone 49
  5.4.1 The HARDY model of auroral precipitation 49
  5.4.2 Extension to the HARDY model 49
  5.4.3 Model results 50

5.5 Relationship between substorm occurrence and HF absorption; substorm correlated absorption bursts (SCABs) 51
  5.5.1 Determination of substorm occurrence 52
  5.5.2 Correlation between substorm occurrence and auroral absorption 54
  5.5.3 Local time distribution of SCAB occurrence 54
  5.5.4 Substorm intensity and associated absorption level 56

5.6 Substorm occurrence, SCAB magnitude, and the relationship with $K_p$ 57

5.7 Summary 58

Chapter 6 Variations in the Maximum Usable Frequency

6.1 Introduction 60

6.2 Ionosonde observations of ionospheric parameters and comparison with NONCENTRIC observations 60
  6.2.1 The ionosonde stations 60
  6.2.2 Mean ionospheric parameters 61
  6.2.3 Comparison of mean ionospheric parameters with NONCENTRIC observations 62
6.3 Enhancements of the MUF on trans-auroral paths: anomalous propagation events (APEs) 64
6.3.1 Anomalous propagation events 64
6.3.2 Geomagnetic dependence and local time distribution of APE occurrence 65
6.3.3 Possible mechanisms for APE production 65
6.3.4 Auroral and sporadic E occurrence and relationship to APE occurrence 66
6.4 Enhancements of the MUF on polar cap paths 68
6.4.1 Alert, summer campaign 68
6.4.2 Fairbanks, summer campaign 70
6.4.3 Fairbanks, winter campaign 71
6.5 Influence of the mid latitude trough on the time of evening signal loss 72
6.6 Summary 74

Chapter 7 The response of HF propagation to geomagnetic storms
7.1 Introduction 75
7.2 Geomagnetic storms during the experimental campaigns 75
7.3 Ionosonde and oblique path observations during geomagnetic storms 77
7.3.1 Storm I 78
7.3.2 Storm II 79
7.3.3 Storm III 80
7.3.4 Storm IV 81
7.3.5 Comparison of observed storm effects 81
7.4 Storm-time absorption in the polar cap 84
7.4.1 Storm I 84
7.4.2 Storm II 84
7.4.3 Comparison and discussion of absorption observations 85
7.5 Summary 86

Chapter 8 Summary and suggestions for further work
8.1 Introduction 88
8.2 Summary and conclusions 88
8.2.1 General diurnal behaviour 88
8.2.2 Trans-auroral paths 89
8.2.3 Polar cap paths 90
8.2.4 Storm-time propagation 91
8.2.5 Summary 91
8.3 Further work 92

References 94
Acknowledgements

I am sincerely grateful to all the people who have made this thesis possible:

Professor Tudor Jones for combining the roles of supervisor and literary critic, and Dr Mike Warrington for producing the NONCENTRIC experiment and for always being round when I needed him.

Mark Orrell and Budgie Dhanda for collecting the original data.

Nigel Wade, Dr Chris Thomas and Andrew Stone for maintaining the Ionospheric Physics Group computers.

For providing data employed in this thesis I would like to thank: Matthew Wild at the World Data Centre C1 for ionosonde and IMP-8 data; Dr Geoff Reeves of the Los Alamos National Laboratory for LANL satellite CPA data; the Space Environment Laboratory at NOAA for GOES-7 satellite magnetometer data.

Professor K A Pounds and Professor T B Jones, who, as heads of the Department of Physics and Astronomy, gave me the opportunity to work at the University of Leicester.

The Ministry of Defence for providing financial support during the period of the study.

Special thanks are due to Dr Mark Lester for constant encouragement and support. Also, all the members of the Ionospheric Physics Group, past and present, including Tim Yeoman, Paul Eglitis, Alan Stocker, Farideh Honary, Eric Bradshaw, Darren Wright and Brian Shand, but especially Jackie Davies for never taking things too seriously.

My parents, who will now be able to stop asking me when I’m going to finish.

Finally, I would like to thank Judith for making the last five years happy ones. It is to her, and to the memory of Derek, that I would like to dedicate this thesis.
Chapter 1

Introduction to the NONCENTRIC experiment

1.1 Introduction

The ionized region of the earth’s upper atmosphere, the ionosphere, has the property of reflecting radiowaves in the HF frequency band (from 3-30 MHz). This propagation is possible to distances in excess of 4000 km, beyond the curvature of the earth’s surface, by means of these reflected waves, and in the past this has been employed as the primary mode of global communication. Problems arise, however, due to over-occupancy of the HF band and interference between users. Also, the reliance of HF propagation on the ionosphere means a susceptibility to disruption of communication during periods of ionospheric disturbance. The use of higher frequencies (GHz) and the implementation of radio repeaters aboard geostationary satellites allows a greater bandwidth for communication and less disruption by geophysical conditions. Consequently, HF radio communication has been largely superseded. Geostationary satellites, however, are expensive and subject to failure, and suffer line-of-sight problems in the extreme polar regions. The need for reliable communications at polar latitudes for scientific, military, and commercial purposes has resulted in a renewed interest in the use of the HF band.

The ionosphere is generally produced by photo-ionization of the ambient atmospheric gases (oxygen and nitrogen) by incoming solar radiation (e.g. UV, X-rays). At mid and low latitudes, electron production depends on the solar flux and therefore follows the solar zenith angle to give a marked diurnal variation. Loss of free electrons occurs by recombination, attachment and diffusion, and so the ionospheric layers tend to be reduced during night times as the loss processes dominate. The polar ionosphere is highly disturbed due to its interaction with the magnetosphere and the solar wind. Solar wind particles gain entry into the magnetosphere where they are constrained to move along the magnetic lines of force and are guided into the region of the high latitude ionosphere known as the auroral zone. Interaction of these precipitating particles with the atmosphere produce the visible aurora and enhanced ionospheric electron densities. Electric fields generated in the magnetosphere produce large-scale motions of the ionospheric plasma, which result in regions of ionospheric enhancement, for example in the night-time polar cap, and ionospheric depletion, as in the mid latitude trough. These regions in the polar ionosphere produce deviations in the paths of radiowaves and attenuation of radiowave signal levels. They also impose frequency and time distortion on radiowaves, which can produce degradation in some types of communication systems. Recent work in wideband HF communication, made possible by fast digital technology, allows real-time determination of the ionospheric transfer function and consequent reconstruction of the transmitted signal at
the receiver. On the other hand, narrow-band communication is possible when a knowledge of the high latitude ionosphere allows the frequency to be selected so as to minimise signal distortion. This works well when a network of propagation paths is available and signals can be routed around regions of disturbance. An understanding of the polar ionosphere is, then, essential for planners and operators of HF radio systems which propagate through these regions.

The distortion imposed on HF signals by the ionosphere means that they are a useful diagnostic tool for the ionospheric physicist. Modelling of the ionosphere and understanding of its interaction with the magnetosphere and interplanetary space is restricted by the lack of detailed knowledge of the morphology and behaviour of the polar ionosphere where strong interactions take place. Since the polar regions are disturbed and conditions can change over small intervals of space and time, an experiment (NONCENTRIC) was designed to observe these changes by means of their effects on the propagation characteristics of HF radio signals.

1.2 The NONCENTRIC experiment
A frequency agile transmitter was constructed at Leicester University and installed within the polar cap at Clyde River, in the Canadian Northwest Territories. Five receiving and data-logging systems, also built at Leicester University, were deployed at various high and mid latitude locations to provide four trans-auroral propagation paths and one path contained entirely within the polar cap. The lengths of the propagation paths ranged from approximately 1000 km to 4000 km. Fourteen frequencies in the HF band, between 3-23 MHz, were transmitted and received every hour during two 24 day campaigns undertaken during the summer of 1988 and the winter of 1989. Each received signal was analysed to determine signal level, noise level, Doppler frequency spreading and signal recognition.

1.3 The present investigation
This study has, for the first time, allowed simultaneous monitoring of the HF propagation medium on five separate paths within the high latitude ionosphere. The amplitude and frequency of the received signals have been measured, since these parameters are sensitive indicators of disturbed ionospheric conditions. The ability to determine the state of propagation over a large portion of the auroral oval and polar cap is essential for interpretation of the results within the context of large scale features in the polar ionosphere. The month-long experimental campaigns are sufficient to provide an understanding of the repeatability and duration of sporadic events occurring within the solar-terrestrial environment which perturb the propagation from its quiet time behaviour.

The diurnal variation of the frequencies that propagate on the five paths during geomagnetically quiet and disturbed periods is investigated. Regions within the high
Latitude ionosphere producing signal attenuation and signal distortion are identified, and their dependence on geomagnetic activity determined. Also, periods when the frequency range of propagating frequencies is anomalously enhanced or reduced are investigated and their relation to ionospheric phenomena determined. The experimental results are interpreted in terms of the basic physical processes which control the behaviour of the ionosphere and magnetosphere. A self-consistent explanation of the experimental data has thus been produced in terms of physical models of both the disturbed and undisturbed ionosphere during both the summer and winter periods.
Chapter 2

Review of HF propagation in the high latitude ionosphere

2.1 Introduction

Long distance, over the horizon radio communications are achieved due to the presence of free electrons in the atmosphere: the ionosphere. Refraction of HF radio waves within the ionosphere allows reflection of signals from altitudes between 100-450 km to locations beyond the line of sight of the transmitter. Spatial and temporal variations in the electron density of the ionosphere are of paramount importance in determining the frequencies available for propagation in point-to-point communications, and consequently the formation and morphology of the ionosphere are discussed in this chapter, with special regard to the high latitude ionosphere. The propagation of radiowaves in the ionosphere is also reviewed.

2.2 The ionosphere

The sun emits radiation and particles which can ionize the atoms and molecules of the Earth’s atmosphere. These electron production processes are balanced by a number of loss processes such as recombination and attachment. In the lower atmosphere, below 60 km altitude, the loss processes dominate due to high atmospheric densities. Above 60 km free electrons can exist for considerable periods of time. This ionized region, extending from 60 km to 1000 km is known as the ionosphere. At night the ionising photon flux is no longer present and loss processes dominate, resulting in a depletion of the ionospheric electron density.

As a consequence of the stratification of atmospheric constituents by gravity and the variation in ionization cross-section between these constituents, the electron density of the ionosphere varies with height and exhibits several distinct ionospheric regions: the D, E, and F regions. Typical mid latitude electron density profiles are illustrated in FIGURE 2.1.

The D region extends from 60 km to 90 km, with an approximate daytime electron density of $10^{10}$ m$^{-3}$. During the night the D region completely recombines.

The E region extends from 90 km to 120 km, and has a typical daytime electron density of $10^{11}$ m$^{-3}$. The electron density of the E region varies regularly with the level of solar illumination; in the lower ionosphere, especially the E region, where the recombination rate is sufficiently high that production and loss processes reach quasi-equilibrium, the electron density, $N_e$, is a function of the intensity of ionising flux falling per unit area (CHAPMAN, 1931).

$$N_e \propto \cos\chi,$$  \hspace{1cm} [2.1]

where $\chi$ is the solar zenith angle and $0.2 < n < 0.9$ (DAVIES, 1990).
FIGURE 2.1 A schematic diagram illustrating typical mid latitude electron density profiles for night and day at solar maximum and solar minimum. The altitudes of the D, E, F1 and F2 regions are also indicated.
The F region, in which the maximum electron density generally occurs, extends above 120 km. During the day the F region is sometimes composed of two parts, the F1 and F2 regions. The electron density in the F region increases rapidly after sunrise, but decreases slowly after sunset. In contrast to the expected solar control, the electron density of the F region is greater in winter than in summer, the \textit{seasonal anomaly}, as a consequence of a seasonal variation in chemical composition, and lower winter temperatures retarding recombination processes.

The solar flux, and consequently the electron density of the ionosphere, varies with an 11 year period, the \textit{solar cycle}. The peak ionospheric electron density can vary by over a factor of 2 between solar minimum and solar maximum.

The electron density profile of the ionosphere can be described by the maximum electron densities in the E and F regions, \( N^E \) and \( N^F_2 \), and the heights of these maxima, \( h^E \) and \( h^F_2 \). Associated with these electron densities are plasma frequencies or \textit{critical layer frequencies}, \( f^E_0 \) and \( f^F_2 \), the maximum frequencies of vertically propagating radio waves that are reflected by the ionospheric regions (see §2.4).

\subsection*{2.3 The high latitude ionosphere: morphology and coupling with the magnetosphere}

The morphology and dynamics of the high latitude ionosphere are controlled not only by solar illumination, but also by the mapping of the geomagnetic field into the \textit{magnetosphere}, illustrated in \textbf{FIGURE 2.2}. At low latitudes the field is closed and approximates to a dipole inclined at 12° to the Earth’s axis; the space enclosed within this region of the geomagnetic field, which corotates with the Earth, is known as the \textit{plasmasphere}. At high latitudes, however, interaction with the solar wind and interplanetary magnetic field (IMF) causes the terrestrial field to be swept back to form the \textit{magnetotail}. The centre of the magnetotail, the \textit{plasma sheet}, maps to the \textit{auroral oval}, the region of the high latitude ionosphere where aurora are observed (FELDSTEIN, 1963; FELDSTEIN and STARKOV, 1967), produced by precipitation of magnetospheric particles into the ionosphere along the magnetic field lines. Enclosed within the auroral zone is the \textit{polar cap}, that region of the ionosphere where magnetic field lines are open, mapping into the \textit{magnetospheric lobes} which are linked to the IMF. New open field lines are created by \textit{magnetic reconnection} between the IMF and the front of the magnetosphere, especially when the two fields are antiparallel (\textit{i.e.} the geocentric solar ecliptic (GSE) north-south component of the IMF is in a southward orientation, \( B_z < 0 \) nT), and are dragged back into the magnetospheric lobes by the flow of the solar wind (DUNGEY, 1961). Under steady state conditions these open field lines are subsequently closed by magnetic reconnection at the distant neutral line approximately 100 \( R_e \) down the magnetotail, and then propagate back towards the front of the magnetosphere due to magnetic tension. This cyclical motion of the field lines is known as \textit{magnetospheric convection} (see \textbf{FIGURE 2.3}).
Figure 2.2 Two schematic representations of the magnetosphere, illustrating the magnetic field configuration and the locations of the various magnetospheric plasma populations.
FIGURE 2.3 The interaction between the interplanetary and geomagnetic fields and the resulting motions of the plasma (open arrows). Two reconnection sites are present, at the front of the magnetosphere (left) and in the magnetotail. Increasing numbers indicate the motion and direction of individual field lines.
2.3.1 The substorm  
Coupling between the IMF and the magnetosphere results in energy transfer from the magnetosphere to the high latitude ionosphere in a process known as the substorm (Akasofu, 1964; 1968). Unequal rates of reconnection at the front of the magnetosphere and at the distant neutral line during the substorm growth phase can lead to an accumulation of open flux in the magnetospheric lobes (McPherron, 1970). The increase in open flux in the lobes increases the size of the polar cap and the auroral oval moves to lower latitudes. This situation cannot be sustained and an explosive release of the stored energy into the auroral zones occurs, the substorm expansion phase. This process is not fully understood but involves a reconfiguration of the near-earth magnetic field (Nagai, 1982) and a reduction in the amount of open flux in the lobes (reducing the size of the polar cap), possibly as a result of the formation of a near-earth neutral line (NENL) (e.g. Hornes, 1979). Simultaneous earthwards acceleration of plasma from the magnetotail, a particle injection (e.g. Sauvaud and Winckler, 1980), enhances the level of precipitation in the auroral oval producing the auroral breakup (Akasofu, 1968), and increases the number of particles trapped in the near-earth dipolar magnetic field, the ring current (see Figure 2.2). During the substorm recovery phase the auroral intensity decreases.

2.3.2 Geomagnetic activity and indices  
The level of energy transfer from the magnetosphere to the ionosphere does not remain constant, but is dependent on the efficiency of coupling between the IMF and the magnetosphere. Currents flowing in the magnetosphere and auroral ionosphere during periods of enhanced loading and unloading of the magnetosphere produce magnetic disturbances measured at ground magnetometers. Prolonged periods of geomagnetic disturbance are known as geomagnetic storms. The level of disturbance, or geomagnetic activity, is quantified by several indices, each of which emphasise different aspects of the disturbance. The three-hour resolution planetary indices $K_p$ and $a_p$ give an indication of the global level of geomagnetic activity, determined from mid latitude magnetometer stations. $AE$, $AU$ and $AL$, the one-minute resolution auroral indices, measure the level of disturbance produced by currents flowing in the auroral ionosphere, and are determined from a chain of magnetometers located at auroral latitudes. $D_s$, one-hour resolution, measures the change produced in the equatorial magnetic field by enhancement of the ring current. As the ring current is generally enhanced during geomagnetic storm periods, $D_s$ is known as the storm disturbance index.

2.3.3 The auroral oval  
The auroral zones are regions of electron and proton precipitation into the ionosphere from a variety of magnetospheric sources. They take the form of two annuli, centred on the north and south geomagnetic poles, at latitudes of approximately 70°N and 70°S (Fieldstein, 1963). The precipitation causes ionization of the atmosphere, in addition to the ionization produced by solar radiation, and excitation of atmospheric constituents to produce the luminous aurora. The poleward and equatorward boundaries of
the auroral (or Feldstein) oval, the region of most statistical likelihood of observation of luminous auroral features, were determined from photographic data (FELDSTEIN and STARKOV, 1967) for different levels of geomagnetic activity. The oval boundaries can be represented by Fourier series of latitude as a function of MLT, whose constants are determined by least-squares fits to the data (HOLZWORTH and MENG, 1975). FIGURE 2.4 illustrates the boundaries of the auroral oval for different levels of $K_p$.

The orientation of the oval is fixed with respect to the sun, the Earth rotating beneath it. The statistical auroral oval broadens and moves to lower latitudes with increasing geomagnetic activity.

2.3.4 Precipitation regions within the auroral zones Precipitating particles dissipate their energy by ionization of neutral atmospheric constituents, 35 eV being the mean energy expended for each ion-electron pair produced. The altitude distribution of energy deposition, and thus electron density enhancement, is highly dependent on the initial particle energy, higher energy particles penetrating to lower altitudes before ionization commences (REES, 1963; BERGER and SELTZER, 1970). Approximately, electrons with energies greater than 30 keV produce ionization in the D region, particles with energies between 1 keV and 30 keV ionize the E region, and particles with energies of less than 1 keV produce ionization in the F region. The energy flux of the precipitation within the oval increases with increasing geomagnetic activity, by a factor of 8 on the nightside and a factor of 2 on the dayside (SPIRO et al., 1982).

The major proportion of energy deposited in the auroral zone is in the diffuse subvisual continuous aurora as opposed to the bright discrete auroral forms from which the Feldstein oval is defined (FELDSTEIN and GALPERIN, 1985). The continuous aurora comprises a uniform band of lower energy (0.1-10 keV) precipitation which originates in the central plasma sheet of the magnetosphere (WHALEN, 1983). It is coincident with the auroral oval and displays the same dynamic behaviour under different geomagnetic activities, extending to lower latitudes and broadening as geomagnetic activity increases (FELDSTEIN and GALPERIN, 1985). The latitudinal variation in the energy flux, $E$, in the continuous aurora can be approximated by a Gaussoid, defined by $E_{\text{max}}$, the maximum flux, $\varphi_{\text{max}}$, the latitude of maximum flux, and $\sigma$, the half width at half maximum. The level of precipitation is uniform in local time (except for 2-3 hours near local noon) and changes in the level of precipitation occur simultaneously over the whole ring. Changes in $\varphi_{\text{max}}$ and $E_{\text{max}}$ are correlated with the occurrence of substorms (WHALEN, 1983). The energy distribution of precipitation in the continuous aurora is generally Maxwellian (SHARBER, 1981). The hardest electrons precipitating in the continuous aurora ionize the neutral atmosphere at E region altitudes, enhancing the E region electron density to produce auroral $E$ (WHALEN, 1983; ROBINSON and VONDRAK, 1985), the critical frequency, $f_0E_0$ (MHz), being related to the precipitating energy flux, $E$ (erg cm$^{-3}$s$^{-1}$), by $E \sim 0.01 f_0E_0^4$ (OMHOLT, 1955; WHALEN et al., 1971). The softer electrons produce enhanced ionization at F region altitudes and
Figure 2.4  The location of the Feldstein oval for $K_p = 0$, 2, and 4. The coordinates are geomagnetic latitude and magnetic local time.
Top-sounding from the Alouette spacecraft at altitudes of \(~1000\) km have shown electron density enhancements coincident with the auroral oval, known as the plasma ring (Thomas and Andrews, 1969). This region extends to F region altitudes where it comprises strong field aligned irregularities and has been named the F Layer Irregularity Zone (FLIZ) by Pike (1971) and Wagner and Pike (1972). This enhancement of the F region is spread in height and extends to low altitudes (Jones, 1980). The FLIZ is coincident with the spread F oval (Akasofu, 1968) and the poleward edge of the mid-latitude trough. Enhanced F region critical frequencies associated with the plasma ring have been observed in high latitude ionogram traces (Andrews and Thomas, 1969).

The equatorward boundary of the auroral oval is a region of high energy (>30 keV) particle precipitation from the ring current (McDiarmid and Burrows, 1964; Hartz and Brice, 1967; Craven, 1970), termed the diffuse equatorward luminosity region by Feldstein and Galperin (1985). After injection into the midnight sector ring current during substorms particles bounce between conjugate points in the northern and southern auroral zones, simultaneously drifting eastwards (electrons) or westwards (protons) under the influence of magnetic gradient and curvature drifts. Ring current particles are gradually precipitated into the ionosphere due to collisions with atmospheric particles. The electron precipitation maximises in the morning sector, as a consequence of the eastwards drift from the midnight sector injection region. Electrons of these energies cause enhancement of the D region electron density and consequently give rise to auroral absorption of HF radio waves, observed in the riometer studies of Hartz et al. (1963) (Figure 2.5). Hartz and Brice (1967) subdivided this electron precipitation into two broad regions, the splash precipitation region in the midnight sector and the drizzle precipitation region in the morning sector (Figure 2.6). Splash precipitation is characterised by short, intense bursts of electron flux, often associated with discrete auroral forms, and short bursts of auroral absorption; drizzle precipitation is associated with steady, slowly varying weak auroral forms, and slowly varying auroral absorption. The spatial patterns of precipitation and auroral absorption determined by various workers (e.g. Hartz et al., 1963; Hartz and Brice, 1967; Craven, 1970) are time-averaged and make no distinction between quiet periods and periods of auroral activity associated with substorms. That such a dependence exists is argued by Akasofu (1968) and is suggested by the observation of Jelly and Brice (1967) that absorption is enhanced in the morning sector after the occurrence of substorms. Proton precipitation from the ring current maximises in the pre-midnight sector due to its westward drift from the midnight injection point. This precipitation gives rise to auroral E with a maximum occurrence (80% for \(f_E > 3\) MHz and 60% for \(f_E > 6\) MHz) between 19-01 MLT (Pfittinger and Gassmann, 1971).

In the noon sector auroral oval the magnetic field lines map to the magnetospheric neutral points and allow direct precipitation of plasma which has diffused across the magnetosheath from the solar wind. This region of the auroral oval is known as the cusp region. The precipitation in the cusp comprises low energy electrons, less than 2 keV...
FIGURE 2.5  The percentage of times of occurrence of auroral radio absorption greater than or equal to 1 dB on 30 MHz (HARTZ et al., 1963).
FIGURE 2.6  The average location of the drizzle (dots) and splash (triangles) >30 keV electron precipitation regions. The coordinates are geomagnetic latitude and magnetic local time (HARTZ and BRICE, 1967).
2.3. The polar cap

The region of the ionosphere enclosed within the auroral oval is known as the polar cap. In summer the polar cap is under constant solar illumination and in winter the polar cap is in constant darkness; in the latter case the ionospheric electron density of the polar cap is maintained by sources other than solar photo-ionization and the structure of the polar cap is dependent on the northward or southward orientation of the interplanetary magnetic field (Carlson, 1994). Magnetospheric electric fields, produced by magnetospheric convection, map into the polar cap. Under $B_z < 0$ nT conditions, an $E \times B$ force then drives convection of plasma across the polar cap in a generally antisunward direction, with return flows at subauroral latitudes in a two cell pattern (e.g. Heppner, 1977; Heelis et al., 1982; Cowley and Lockwood, 1992) (Figure 2.7a). Depending on the orientation of $B_z$, the east-west component of the IMF, the flow exhibits a dawn-dusk asymmetry, changing the sense of the east-west component of flow across the polar cap. When the front of the convection pattern reaches into mid latitude sunlit ionosphere, high electron density plasma can be convected into the dark (and consequently low electron density) polar cap. At F region altitudes plasma recombination rates are sufficiently low that high electron densities can be maintained for several hours, comparable with the drift time of convection across the polar cap, leading Knudsen (1974) to predict a polar cap tongue of ionization. Modelling by Schunk et al. (1980) indicated a modulation of the electron density of the convected plasma by the rotation of the geomagnetic pole about the geographic pole, moving the convection pattern closer to the dayside, into regions of high electron density, at 17 UT and away from the dayside, into regions of lower electron density, at 05 UT. The F region electron density enhancement takes the form of antisunward convecting -1000 km diameter patches with 2-10 times the background density (Weber et al., 1984; Buchau et al., 1985), however, and not a spatially continuous tongue of ionization. The process that segments these patches is not fully understood, but suggestions include transient reconnection (Lockwood and Carlson, 1992), creation of patches by changes in the flow pattern due to changes in $B_z$ (Soika et al., 1993), and depletion of electron density between patches by fast east-west plasma flows near the cusp region (Valladares et al., 1994). These patches, once convected across the polar cap, can enter the return flow where they enhance the F region electron density at the equatorward boundary of the auroral oval (De la Beaujardière et al., 1985; Senior et al., 1987; Rodger et al., 1992). The convecting patches become elongated by the flow (Robinson et al., 1985) to form a wall of electron density enhancement, sometimes known as the boundary blob (Rino et al., 1983) or ionization channel (Senior et al., 1987).

Under $B_z > 0$ conditions the convection pattern becomes weaker and distorted from the $B_z < 0$ case, mainly by rotation about the pole (Heppner and Maynard, 1987), or becomes three- or four-celled (e.g. Heelis et al., 1986), so that regions of the convection flow turn
FIGURE 2.7 Conditions characterizing (a) the southward and (b) northward IMF state of the polar cap ionosphere. The coordinates are corrected geomagnetic latitude and local time. In panel a are indicated the boundaries of the auroral oval, typical polar cap convection flow lines, and two ionization patches approximately 1000 km in diameter. In panel b are indicated the auroral oval boundaries and two polar cap arcs. (CARLSON, 1994).
sunward. Under such conditions the convection pattern shrinks and the production of ionization patches ceases. In addition to the changes in the convection flow, polar cap arcs are formed (FIGURE 2.7b). These arcs are aligned in the noon-midnight meridian and can reach right across the polar cap, in which case they are known as theta aurora (FRANK et al., 1986). Precipitation in the arcs can produce significant ionization at F and E region altitudes.

A region of depleted F region electron density, the polar hole or cavity (PIKE, 1971), can form near local midnight in the polar cap. The polar hole is formed by slow convection of the F region plasma in the absence of electron production, so loss processes dominate and the F region electron density decreases. Under $B_z < 0$ nT conditions, when a two-cell convection pattern is present, the polar hole is shallow as high electron density plasma is constantly convected from the dayside. Under $B_z > 0$ nT conditions, however, when four-cell convection pattern is present, the two nightside cells are in constant darkness and the polar hole can become deep (CROWLEY et al., 1993).

After the mid-latitude region, the polar cap is regarded by many as the most benign for HF radio communications as propagation is usually undisturbed except for polar cap absorption (PCA) events. PCAs cause severe black-out over the whole HF spectrum across large regions of the polar cap and have durations from hours to days, but are generally quite rare, occurring approximately once per month at solar maximum, and less often at other times (COLLINS et al., 1961). PCA absorption is caused by intense enhancement of the D region electron density by solar proton events (SPEs), protons with energies greater than 1 MeV ejected from the sun; these particles are confined to the polar cap, geomagnetic latitudes greater than $-65^\circ$, by the terrestrial magnetic field (HOLT, 1968; DAVIES, 1990).

2.3.6 Sporadic E and auroral E
Sporadic E ($E_s$) and auroral E are enhancements of the E region electron density, usually in a thin (2-10 km) layer at an altitude of approximately 110 km. At mid-latitudes sporadic E can be produced by interaction of the E region plasma with neutral winds. Collisions couple the electrons to the horizontally moving neutrals, producing a vertical component in the motion due to the constraint that electrons move parallel to the inclined magnetic field. In the presence of wind shears this can concentrate E region electrons in a thin, dense sporadic layer (WHITEHEAD, 1961). Sporadic E is most significant in summer months (STEVENS, 1968).

Auroral E is produced by precipitation of 10-30 keV particles in the auroral oval (e.g. WHALEN, 1983). The E region electron density enhancement can persist longer than the source precipitation despite rapid recombination rates at E region altitudes due to the presence of metallic ions which have a long diffusion time, maintaining the high electron density for several hours (TURUNEN, 1976). The auroral E region moves to lower latitudes and intensifies under disturbed magnetic conditions, correlated with movements of the auroral oval (BESPROZVANNAYA et al., 1980; WHALEN, 1983). HF backscatter observations at Andoya, Norway, indicate that patches of sporadic E or auroral E can be up
Chapter 2

Theory and review

11

to 1500 km in extent, and occur most frequently near local midnight (FOLKESTAD, 1968).

Polar cap sporadic E, predominantly occurring during the summer and under quiet magnetic conditions, is a consequence of the redistribution of E region electron density under the influence of the convection electric field in a similar manner to neutral winds at mid-latitudes (BESPROZVANNAYA et al., 1980). Polar cap sporadic E occurs predominantly in the evening and night sectors (18-01 MLT), and the spatial extent of the region of greatest probability of polar cap E, extends to lower (auroral) latitudes in these time sectors. The occurrence of polar cap E, with \( f_E > 3 \) MHz shows a dependence on the orientation of the IMF, with the greatest occurrence (~50%) for \( B_y < 0 \) nT and the lowest (~20%) for \( B_y > 0 \) nT.

2.3.7 The mid latitude trough

The mid latitude trough is a depletion of the F region electron density at the equatorward edge of the nightside auroral oval. A schematic diagram of the location of the trough is illustrated in FIGURE 2.8. Several mechanisms have been proposed for the formation of the trough: convection of low density plasma from the nightside ionosphere (WATKINS, 1978); stagnation of plasma caught between the two regimes of high latitude convection and mid latitude corotation in the evening sector, leading to recombination of the plasma over time (KNUDESEN, 1974; SPIRO et al., 1978); an increased recombination rate of the F region plasma and upwelling of the plasma, both as a consequence of frictional heating between convecting ions and slower-moving neutrals (ST. MAURICE and TORR, 1978).

Many observations of the trough are topside soundings from satellites (e.g. Alouette I and II: MULDREW, 1965; HALCROW and NISBET, 1977). Bottomside soundings from ionosondes can be difficult to interpret due to off-vertical reflections from the trough walls, but with direction finding techniques trough motion can be deduced from the ground (e.g. RODGER and PINNOCK, 1980; RODGER et al., 1986; WHALEN, 1987). Computational models of the high latitude ionosphere have been formulated which can be used to investigate the behaviour of the trough (e.g. QUEGAN et al., 1982; SOJKA and SCHUNK, 1989).

The poleward wall of the trough is formed by enhancement of the ionospheric plasma density by low energy electron precipitation in the auroral oval (RODGER et al., 1986). Evidence also exists for contribution to the poleward edge by solar-produced plasma convected from the dayside ionosphere, across the polar cap, and along the convection return flow (RINO et al., 1983; DE LA BEAUJARDIÈRE et al., 1985; SENIOR et al., 1987; RODGER et al., 1992). The equatorward wall marks the boundary between the trough and the solar-produced mid-latitude ionosphere.

The trough is generally a night-time phenomenon as solar-illumination causes the trough to fill; the dawn trough wall is usually coincident with the solar terminator. The dusk wall, or leading edge, however, can extend into the daylit hemisphere, its local time position becoming earlier during periods of increased geomagnetic activity (WHALEN, 1987).
FIGURE 2.8 A schematic diagram of the location of the mid latitude trough and auroral oval. The location of the dawn wall of the trough is usually coincident with the solar terminator. The location of the dusk trough wall, or leading edge, is dependent on $K_p$. 
HALCROW and NISBET (1977) derived an empirical model of trough location and local time extent from Alouette I and II data. In this model, the trough is filled at dawn once the solar zenith angle $\chi = 87^\circ$. The modelled location of the dusk wall of the trough is more complicated: the trough opens at a local time 1.5 hours after the sun reaches a solar zenith angle dependent on geomagnetic activity. The local time at which the trough forms is given by

$$\chi(LT - 1.5) = 87^\circ - 3^\circ x (K_p - \frac{1}{2}).$$  \[2.2\]

However, SOIKA et al. (1981) show that a small change in the latitudinal position of the auroral oval with respect to the convection pattern can cause the extent of the trough to change by more than 2 hours of MLT. Such a behaviour is not adequately parameterised by $K_p$.

The trough moves equatorward as local time progresses (KÖHNLEIN and RAFTT, 1977; RODGER and PINNOCK, 1982). Superimposed on this is an equatorward motion of the trough during periods of enhanced geomagnetic activity. This is a consequence of the expansion of the polar cap convection pattern and the equatorward motion of the auroral oval under increased activity, both of which are responsible for the formation of the trough. Models of trough location include terms in $K_p$ to describe this geomagnetic dependence. It has been shown, however, that $K_p$ is a poor indicator of trough position: the local time at which the poleward edge of the trough crosses a specific invariant latitude has been shown to vary by up to 7 hours under similar geomagnetic conditions (RODGER et al., 1986).

The extent and position of the trough have a universal time dependence due to the rotation of the geomagnetic pole around the geographic pole. The trough position remains constant relative to the geomagnetic pole, but this moves with respect to the terminator of solar illumination throughout the day. In the HALCROW and NISBET model (1977) the dawn and dusk walls are positioned relative to the terminator, so the longitudinal extent of the trough varies with the orientation of the geographic and geomagnetic poles to the sun. At 05 UT when the magnetic pole is orientated away from the sun the trough is at its maximum extent and at 17 UT it is at its minimum. Seasonal changes in the position of the solar terminator also lead to changes in trough extent. In summer most of the polar region is constantly illuminated so the trough is of very limited extent, confined to a few hours in the midnight sector. In winter the sense is reversed and the trough becomes a major feature of the high-latitude ionosphere.

2.3.8 Geomagnetic and ionospheric storms  Geomagnetic storms are a consequence of shockwaves in the solar wind, produced by solar flares or coronal mass ejections (CMEs), or prolonged periods of coupling between the IMF and the magnetosphere. Geomagnetic storms can be identified by the characteristic variation they produce in $D_s$ (MATSUSHITA, 1967). The typical variation includes the main phase, in which $D_s$ falls rapidly below the prestorm value, and the recovery phase, a slow return to quiet-time levels. The main phase is often preceded by a small, short-lived increase in $D_s$, the initial phase, produced by
compression of the geomagnetic field by a shockwave in the solar wind. An abrupt change in magnetometer records at the start of a storm is known as a storm sudden commencement or SSC. During geomagnetic storms, enhanced energy deposition in the auroral ionosphere, by precipitation and joule heating, causes heating of the atmosphere and global modification of the neutral wind pattern. These changes have a direct influence on the ionospheric electron density, especially the F region, which becomes perturbed from quiet-time values for hours or days, an ionospheric storm. Two effects are produced in the F region: an increase in $f_{o}F_2$, the positive storm effect, and a decrease in $f_{o}F_2$, the negative storm effect.

Travelling atmospheric disturbances (TADs) are excited by the enhanced energy input in the auroral oval and propagate equatorwards; TADs propagating equatorwards cause $h_mF_2$ to increase as the ionospheric plasma is pushed up inclined magnetic field lines. At higher altitudes the recombination rate of the plasma is lower and so in the presence of electron production (insolation on the dayside) $N_aF_2$ increases (PROLSS, 1993): the positive storm effect. Consequently the positive storm is not expected on the nightside, and is most conspicuous in the afternoon sector where electron density is usually decreasing.

The quiet-time vertical distributions of the atmospheric neutral species approximate to diffusive equilibrium above ~100 km. The influx of energy in the auroral oval during storm periods, however, modifies the high latitude wind patterns, introducing a significant vertical component due to upwelling of the heated atmosphere and the flow necessary to balance the divergence of the horizontal storm-time wind patterns (Rishbeth et al., 1987). Mixing of the neutral atmosphere results and as a consequence the composition at F region altitudes becomes more molecular (an increase in the ratio of $N_2$ to O). This composition change causes a decrease in $N_aF_2$: the negative storm effect. This occurs throughout the auroral oval, but at mid latitudes it is usually first observed in the midnight sector as it is here that the composition disturbance is carried equatorward by the neutral atmosphere circulation pattern (generally winds travelling away from sub-solar point towards the nightside). Once at mid latitudes the disturbance is carried into the morning and noon sectors by the rotation of the earth (PROLSS, 1993).

The variation of $f_{o}F_2$ observed during storm times is dependent on both geographic location and the local time at the onset of the storm disturbance. At mid latitudes, positive storm effects are usually only observed if the disturbance occurs during the daytime (Appleton and Piggott, 1952) and negative storms commence most frequently in the early morning and infrequently in the noon and afternoon sectors (Jones, 1971).

Storm-time precipitation in the auroral oval produces enhanced ionization in the ionosphere, though the time-scale of the response is altitude dependent, predominantly as a consequence of the relative importances of photo-chemical and transport time constants. The electron density at lower altitudes, E region to lower F region, responds almost instantly to particle precipitation; at higher altitudes the response is delayed but can continue to be felt for hours or days after the initial disturbance (Sojka and Schunk, 1983). Auroral E or storm E, is a conspicuous feature of the night-time auroral E region
during geomagnetic storms (Reid, 1967).

Sojka and Schunk (1983) modelled the response of the high latitude F region to idealised magnetospheric storm inputs: an enhanced convection pattern and enhanced precipitation within the auroral oval. The model results indicated a near 10-fold increase in the F region concentration of O+ ions (indicating a similar increase in electron density) within the afternoon and evening sectors of the polar cap at 300 km altitude due to the increased convection and an increase in electron densities within the auroral oval due to enhanced precipitation. At F region altitudes these storm enhancements of electron density persisted for several hours after the precipitation and convection returned to quiet-time conditions. At sub-auroral latitudes electron densities decreased as a consequence of a deepening and enlarging of the mid latitude trough. Obviously there can be a difficulty in interpreting changes in electron densities in the night-time sub-auroral ionosphere as both a negative storm effect and an enlarged trough are expected. Competition between the storm-time processes results in a highly complex behaviour that is difficult to predict; the effect of a storm at a given location is dependent on local time, season, latitude and storm time (time elapsed since the storm commencement).

Shock waves in the solar wind can be accompanied by energetic (MeV) solar protons ejected from the sun during solar flares and CMEs. These protons, which produce polar cap absorption (PCA) events (see §2.3.5), can stream ahead of the solar wind shock front and so be incident on the Earth hours or days before the geomagnetic storm effects are felt.

On 31 December 1901 Marconi succeeded in transmitting a radio signal from Cornwall, England, to Newfoundland, Canada, proving that radiowaves could travel around the curvature of the Earth. Independently, Kennelly (1902) and Heaviside (1902) suggested that signals were reflecting from a conducting layer of atmospheric ions at an altitude of approximately 80 km. Appleton and Barnett (1925a and b) and Breit and Tuve (1925, 1926) demonstrated the existence of reflecting layers in the ionosphere and determined their approximate heights. After it was shown that the Earth's magnetic field influenced the return of radiowaves (Appleton, 1925; Nichols and Scheleng, 1925) the magneto-ionic theory was developed by Appleton and Builder (1932).

In the following sections the propagation of HF radiowaves and their reflection by the ionosphere are briefly discussed.

2.4 HF propagation

The passage of a radiowave through the ionosphere is influenced by many factors, but predominantly the electron density profile. The refractive index of an electromagnetic wave propagating through an isotropic ionized medium in the presence of a magnetic field is given by the Appleton-Hartree formula:
Chapter 2  Theory and review

\[ n^2 = (\mu - i\gamma)^2 = 1 - \frac{X}{1 - iZ - \frac{Y^2}{2(1 - X - iZ)^2} \pm \frac{Y^2}{4(1 - X - iZ)^2 + Y^2}}. \]  \[ 2.3a \]

where \( n \) is the complex refractive index, and

\[ X = \frac{\omega_p^2}{\omega^2} = \frac{N_e e^2}{\varepsilon_0 m_e \omega_p}, \quad Y = \frac{\omega_p}{\omega} = \frac{eB}{m_e \omega}, \quad Z = \frac{\nu}{\omega}. \]  \[ 2.3b, c, d \]

\( f \) and \( \omega \) are the radiowave frequency and angular frequency respectively \((\omega = 2\pi f)\). \( \omega_p \) and \( f_e \) are the plasma angular frequency and frequency respectively \((\omega_p = 2\pi f_p)\), and \( \omega_h \) is the electron angular gyrofrequency. \( N_e \) is the electron density, \( e \) is the electronic charge, \( m_e \) is the electron mass, \( \varepsilon_0 \) is the permittivity of free space, \( \nu \) is the collision frequency between electrons and neutrals which is dependent on the atmospheric density, and \( B \) is the imposed magnetic field. The subscripts \( L \) and \( T \) refer to longitudinal and transverse components relative to the direction of phase propagation. The \( \pm \) sign in \([2.3]\) indicates that two polarizations of the radiowave can exist, termed the ordinary and extraordinary waves.

The plasma frequency is the natural frequency of oscillation of electrons in a plasma; from \([2.3b]\) electron density \((m^3)\) and plasma frequency \((Hz)\) are related by

\[ N_e = \frac{\varepsilon_0 m_e \omega_p^2}{e^2} = 0.0124 f_p^2. \]  \[ 2.4 \]

This relates the electron density profile of the ionosphere to a plasma frequency profile; for instance the maximum electron density of the F region, \( N_eF2 \), has a corresponding plasma frequency, \( f_pF2 \).

Two simplifying cases can be considered to illustrate the basic principles of propagation in the ionosphere: negligible magnetic field and negligible collision frequency.

**Negligible magnetic field, negligible collision frequency**  In the simple case of an isotropic ionosphere with negligible magnetic field \((Y << 1)\) and no collisions \((Z = 0)\), the complex refractive index \([2.3]\) reduces to the real expression

\[ \mu^2 = 1 - X = 1 - \frac{f_e^2}{f_p^2}. \]  \[ 2.5 \]

For vertically propagating waves with frequency \( f_e \), reflection occurs when \( \mu = 0 \), i.e. \( X = 1 \), the plasma frequency and the wave frequency are equal, \( f_e = f_p \). For obliquely propagating waves, reflection occurs when

\[ \mu = \sin \varphi \]  \[ 2.6a \]

for a flat-earth approximation, or for a curved earth

\[ \mu = \left( \frac{R_y}{h + R_y} \right) \sin \varphi. \]  \[ 2.6b \]

where \( \varphi \) is the angle of incidence of the ray on the ionosphere, \( h \) is the height of reflection, and \( R_y \) is the radius of the earth. There is, then, a relationship between the frequency of an oblique wave, \( f_e \), and the frequency of a vertical wave, \( f_v \), that reflect at the same plasma frequency, \( f_p \) (from \([2.5]\) and \([2.6]\)):
where sec $\varphi$ is known as the secant factor. $k$ is a factor included to account for the curvature of the Earth and is dependent on the ground range of the propagation. For a flat earth approximation or short ground ranges, $k = 1$, and for extreme ground ranges, $k = 1.2$ (Sharp, 1959).

The maximum electron density in the ionosphere $N_m F_2$ has corresponding plasma frequency $f_m F_2$. Consequently, vertically propagating waves with $f > f_m F_2$, or obliquely propagating waves with $f > f_m F_2$ $k$ sec $\varphi$ will not be reflected and will propagate into space. The maximum frequency that will propagate along a path between two ground-based terminals is known as the maximum usable frequency or MUF.

**Significant collision frequency, negligible magnetic field** In the case that collisions are present, $Z > 0$, the refractive index becomes complex and attenuation of the radiowave, through dissipation of the wave energy by collisions between electrons and neutrals, becomes important. The refractive index is (from [2.3]):

$$n^2 = (\mu - i\chi)^2 = 1 - \frac{X}{1-iZ}. \quad [2.8]$$

When $\mu^2 >> \chi^2$, the absorption coefficient, the attenuation per unit distance, $\kappa = \omega \chi / c$, is

$$\kappa = \frac{\nu}{2\mu c \omega^2 + \nu^2}. \quad [2.9]$$

This is discussed more fully in §2.4.5.

**2.4.2 Radiowave paths in the ionosphere** There are several paths that a radiowave can take between transmitter and receiver, with different reflection heights, and possibly with intervening reflections from the ground. These paths are denoted *modes* and are labelled by the number of hops and the region of the ionosphere from which reflection occurs, as illustrated in FIGURE 2.9 (Davies, 1967).

The true path followed by the radiowave through the ionosphere can be represented as the equivalent path which consists of one or more triangular hops, each a straight ray to the path midpoint with elevation angle $\Delta$, a plane reflection with angles of incidence and reflection $\varphi$, and a straight ray to the ground (FIGURE 2.10). Reflections at the ground, between hops, are treated as plane reflections. Due to the curvature of the Earth, $\Delta = \varphi$.

When considering hops in this way the height of the reflection is not the true height of reflection, $h$, the height at which $\mu = \sin \varphi$, but the equivalent or virtual height, $h'$. $h'$ is the reflection altitude which is calculated from the return time, $\Delta t$, of a vertically propagating radiowave assuming free-space velocity $v = c$, hence $h' = c\Delta t/2$. However, in the ionosphere the refractive index $\mu < 1$, so $v < c$ and consequently $h' > h$. $h'$ can be found by integrating in height through the refractive index profile (determined from the electron density profile) of the ionosphere.
Figure 2.9 The nomenclature of various possible forward propagation modes (DAVIES, 1967).
Figure 2.10 The geometry of a radiowave propagating in the ionosphere, illustrating the true ray path and the equivalent ray path. $D$ is the 1-hop ground range, $h_r$ is the true height of reflection of the radiowave, and $h'$ is the virtual height of reflection. $\Delta$ is the angle of elevation and $\varphi$ and $\varphi_D$ are the angles of incidence at the point of reflection and in the D region. $h_D$ is the height of the D region and $R_e$ is the radius of the Earth.
Chapter 2  Theory and review

\[ h' = \frac{1}{\frac{1}{2} \mu} \]  \hspace{1cm} [2.10]

FIGURE 2.11 illustrates the virtual height of reflection as a function of frequency for a model electron density profile, composed of parabolic E and F regions and a linear transition region (BRADLEY and DUDENEY, 1973). Note that the electron density profile yields a plasma frequency profile which is a function of height. The virtual height, calculated from the plasma frequency profile, is a function of radiowave frequency.

The geometry of the ray also plays an important role in determining the height of reflection for a given frequency. \( h' \) is related to the ground range from transmitter to receiver, \( D \), the number of hops, \( n \), and the elevation angle, \( \Delta \), by

\[ h' = R_e \left\{ \tan(\Delta + \alpha) \sin \alpha + \cos \alpha - 1 \right\}. \]  \hspace{1cm} [2.11]

where \( \alpha \) is the angle made at the Earth’s centre by the launch point of the ray and the reflection point, which is given by

\[ \alpha = \frac{D}{2nR_e}, \]  \hspace{1cm} [2.12]

(see FIGURE 2.10). The angle of incidence of the ray to the normal to the ionosphere at the reflection point, \( \varphi \), is given by

\[ \tan \varphi = \frac{1}{\tan(\Delta + \alpha)}. \]  \hspace{1cm} [2.13]

These equations assume that each hop is identical in ground range and reflection height, an approximation which is valid provided the ionospheric profiles at each reflection point are similar.

Equations [2.11], [2.12], [2.13], and [2.7] were used to produce FIGURE 2.12a, illustrating the height of reflection and secant factor, \( f_s/f_e \), for 1-, 2-, and 3-hop rays as a function of elevation angle for paths with ground ranges of 2000 km and 3500 km. FIGURE 2.12b illustrates the elevation angle and secant factor as functions of reflection height. Certain combinations of ground range and reflection height are precluded by the curvature of the earth; 1-hop propagation with \( h' < R_e \left\{ \tan \alpha \sin \alpha + \cos \alpha - 1 \right\} \) (from [2.11] with \( \Delta = 0 \)) yields a negative elevation angle. In such cases 2-hop propagation is necessary, though this requires that the electron density conditions be met at two reflection points.

For a mode to be available for propagation it has to satisfy both the ionospheric criteria of equations [2.4], [2.7] and [2.10] and the geometric criteria of equations [2.11] to [2.13]. The method generally used to determine the solutions of these equations is graphical and involves superimposing transmission curves \( f_s \) as a function of virtual height determined for various values of \( f_e \) from [2.7], and [2.11] to [2.13]) on the graph of virtual height of reflection as a function of frequency (calculated from [2.10] for analytical electron density profiles or from experimentally determined ionogram traces). Such solutions are illustrated in FIGURE 2.13 for the same electron density profile as illustrated in FIGURE 2.11 for two propagation paths with ground ranges 2000 km and 3500 km. The points at which these curves intersect give the equivalent heights of reflection for a given frequency \( f_e \). There can
FIGURE 2.11 The relationship between plasma frequency profile (dotted curve), a function of altitude, and virtual height (solid curve), a function of frequency. The plasma frequency profile is constructed from two parabolic layers with a linear transition region.
FIGURE 2.12 (a - top panels) Virtual height of reflection (solid curves) and secant factor (dotted curves) for 1-, 2-, and 3-hop propagation as a function of elevation angle on a 2000 km and 3500 km path. (b - bottom panels) Elevation angle (solid curves) and secant factor (dotted curves) for 1-, 2-, and 3-hop propagation as a function of virtual height of reflection on a 2000 km and 3500 km path.
FIGURE 2.13 Transmission curves for 2000 km and 3500 km propagation paths superimposed on the virtual height as a function of frequency curve of FIGURE 2.11.
be several solutions for the same frequency in different ionospheric layers, hence E region and F region modes (e.g. $f_c = 11$ MHz on the 2000 km path), and in the same ionospheric layer, termed high angle and low angle rays to denote the solution with the higher and lower reflection height respectively (e.g. $f_c = 14$ MHz on the 2000 km path).

Some of the modes indicated by the transmission curve solution of the reflection height may not be available for propagation as they are screened or blanketed by the ionosphere below, especially the E region. This occurs when the radiowave passes through a region of the ionosphere where reflection takes place, but the reflection height is such that the ray returns to the ground away from the receiver, i.e. this reflection height is a solution of the ionospheric reflection criteria but is not a solution of the path geometry criteria.

The more obliquely a radiowave is incident on the ionosphere the higher the frequency that will be reflected from a given electron density (equation [2.7]). Consequently an E region mode will support a much higher frequency than an F region mode for similar maximum layer electron densities. An enhanced E region, for instance auroral E, can support propagation on frequencies much higher than those predicted for the F region, especially during periods when the F region electron density is depressed such as night-time conditions or during ionospheric storms. In this case an increased range of usable frequencies can be exploited to overcome transmission problems on lower frequencies (STEVENS, 1968).

Using these methods and a knowledge of the E and F region electron densities, determined from models such as the International Reference Ionosphere (IRI) (RAWER et al., 1978a, b; LINCOLN and CONKRIGHT, 1981; BILITZA, 1986, 1990) or ionosonde observations, the possible modes of propagation can be predicted.

2.4.3 Propagation in the presence of horizontal plasma gradients and sporadic E The equivalent path approximation considers propagation in the absence of any significant horizontal electron density gradients (ionospheric tilts) in the ionosphere. This leads to symmetric hops and great circle path propagation. Such assumptions can generally be made in the mid latitude ionosphere where horizontal gradients in electron density are small and can usually be ignored as they cause significant deviations from the plane reflection approximation for short ground ranges (less than 300 km) only. At high latitudes, however, very large gradients are encountered, such as the ionization walls of the mid latitude trough; BAKER and GREENWALD (1988) cite electron density gradients as essential for the explanation of the anomalous behaviour of propagation on a ~1000 km path in the high latitude ionosphere. Electron density gradients perpendicular to the path, or high electron density features to either side of the path, will give rise to non-great circle path (NGC) propagation. This can be of importance in the polar cap due to reflections from convecting plasma patches (JONES and WARRINGTON, 1992). Electron density gradients along the path will perturb the elevation angles at the receiver and transmitter.

In the presence of ionospheric tilts or horizontal structure in the ionosphere, mixed
modes are also possible, for instance the E-F mode which consists of a reflection from the E region, a reflection from the ground, and a final reflection from the F region (see FIGURE 2.9). Mixed modes are difficult to predict and are regarded as unimportant by FRICKER and DAVEY (1978), except perhaps in the case of localised sporadic E.

Propagation by sporadic E region modes alone requires E_g at all path reflection points; in the case of 2E, propagation over a 3000 km path the reflection points are 1500 km apart. FOLKESTAD (1968) reported that widespread sporadic E formations probably play an important role in sustaining propagation on high latitude paths; observations at Andøya, Norway, of 1- and 2-hop E_g ground backscatter echoes indicated that such formations could be up to 1500 km in extent. During February and March 1964 2E, ground backscatter echoes occurred most frequently at 02 UT, near local geomagnetic midnight. However, STEVENS (1968) found that the occurrence of E_g or F-E_g modes enhancing the MUF on a propagation path from Ottawa to Resolute was significant in summer months only. In January 1961 E_g was significant only 7% of the time, the diurnal maximum occurrence being less than 20% at night.

2.4.4 Scatter propagation Propagation of radiowaves is possible by forward scatter of the signals from magnetic field-aligned irregularities within the ionosphere. This scattering can give rise to off-great circle path propagation from features laterally displaced by great distances from the path midpoint. Irregularities are produced in the presence of steep electron density gradients perpendicular to the magnetic field, such as at the edges of electron density enhancements (polar cap patches and the auroral boundary blob), via the gradient drift instability, and by precipitation within the auroral oval (TSUNODA, 1988). Scatter propagation occurs as a consequence of coherent Bragg scatter from these irregularities. Scatter can only occur at points where raypaths are perpendicular to the geomagnetic field. At high latitudes, where the magnetic field is nearly vertically inclined, this requires that the ray is near reflection. The higher the frequency that is propagating the further the ray has to travel before it is refracted to perpendicularity with the magnetic field and the greater the group delay, producing a characteristic slant F trace on oblique forward-and back-scatter HF ionograms (e.g. BATES, 1960). In the case of forward scatter the signal is only observed at the receiver when the distances between the transmitter and the scatter volume and the receiver and the scatter volume are equal as the ray has to be refracted by an equal amount on the downward leg of the ray as on the upward leg to be incident on the receiver, i.e. scatter can only occur on a line perpendicular to the great circle path passing through the path midpoint. The scale length of the irregularities (the spatial Fourier component of the electron distribution), λ, responsible for the scatter propagation decreases as the radiowave frequency increases due to the Bragg condition:

\[ \lambda = \frac{\lambda}{2 \sin \theta} \]  

[2.14]

where λ is the radiowave wavelength (in the medium at the point of scatter), and θ is the
angle between the directions of the incident and scattered wave vectors measured at the point of scatter (Bauer, 1988). Also, higher frequencies correspond to scatter from larger off-great circle path bearings due to the requirement that the ray reaches perpendicularly. Consequently, $\theta$ increases and $\Lambda$ decreases with increasing signal frequency. The amplitude of the scattered signal is expected to decrease with increasing frequency as the occurrence distribution of irregularities as a function of irregularity wavelength (the irregularity spectrum) can generally be described by a power law, with decreasing occurrence at shorter wavelengths (Tsunoda, 1988).

2.4.5 Attenuation of Radiowaves

Collisions between free electrons and neutral gas molecules cause dissipation of the energy of the radiowave and consequent attenuation or absorption of the propagating signal. The attenuation of the wave, $L$ (Nepers; 1 Np = 8.68 dB), is found by integration of the absorption coefficient, $\kappa$ (Np m$^{-1}$), along the ray path

$$L = \int \kappa ds,$$

where (from [2.9])

$$\kappa = \frac{e^2}{2\varepsilon_0 c m_e} \frac{1}{\mu} \frac{N\nu}{\omega^2 + \nu^2}.$$  \hspace{1cm} [2.15a]

In general, the portions of the ray path in the lower ionosphere and near the reflection height contribute most to $L$: non-deviative and deviative absorption respectively.

Non-deviative absorption

In the D and lower E regions, $N\nu$ is a maximum and consequently $\kappa$ is large and absorption is high. However, in the lower ionosphere $N_\perp$ is sufficiently small that the refractive index is close to unity, $\mu = 1$, and attenuation in this region is known as non-deviative absorption. For $a^2 >> \nu^2$ the absorption coefficient for non-deviative absorption reduces to (from [2.15b])

$$\kappa = \frac{e^2}{8\pi^2 \varepsilon_0 c m_e} \frac{N_\perp \nu}{f^2}.$$  \hspace{1cm} [2.16]

The absorption is proportional to the electron density in the D region and inversely proportional to the square of the radiowave frequency: non-deviative absorption increases as the radiowave frequency decreases.

Deviative absorption

In the limit that $\mu \to 0$, [2.15b] indicates that absorption is high. As the radiowave is near reflection when $\mu$ is small, and there is marked curvature of the ray path, this is known as deviative absorption. In the case of oblique propagation reflection occurs at $\mu > 0$, the deviative absorption is low, and the major contribution to the attenuation of radiowaves is non-deviative absorption.

Intense absorption, especially non-deviative absorption due to the $f^2$ dependence, decreases the signal to noise ratio at the lowest frequencies to the extent that signals cannot be detected at the receiver, the limiting frequency being the lowest usable frequency or LUF. Consequently, the lowest frequency that can be detected is dependent on the level of
absorption and the power of the transmitter.

The non-deviative absorption of an oblique ray can be related to that of a vertically propagating ray passing through the same point in the D region by the factor sec $\phi_0$, where $\phi_0$ is the angle made by the oblique ray to the vertical in the D region.

$$\tan \phi_0 = \frac{1}{\tan(\Delta + \alpha_0)}.$$  \[
[2.17]\]

where $\alpha_0$ is the angle made at the centre of the earth between the ray at the ground and in the D region (see FIGURE 2.10). $\alpha_0$ can be found from the approximate altitude of the D region, $h_0$ (=80 km), and other parameters of the radiowave path, $h$, $\Delta$, and $\alpha$.

$$\cos(\Delta + \alpha_0) = \frac{R_0 + h}{R_0 + h_0}.$$  \[
[2.18]\]

Enhancement of the D region electron density, by solar illumination or electron precipitation, increases non-deviative absorption. In the case of auroral absorption the D region enhancement is produced by the precipitation of $>30$ keV electrons in the splash and drizzle precipitation regions. The location and level of auroral absorption has been determined from riometer observations (HARTZ et al., 1963) and empirically modelled (FOPPIANO and BRADLEY, 1983). The flux of auroral precipitation, and consequently the level of auroral absorption, is dependent on geomagnetic activity, and models require suitable parameterization of this behaviour. Sunspot number has been employed as a model input (FOPPIANO and BRADLEY, 1985), but this does not adequately reflect the much more rapidly varying level of magnetic disturbance at the Earth, and other indices have been employed: $A_p$ (HARGREAVES et al., 1987) and $K_p$ (FOPPIANO and BRADLEY, 1983). However, auroral absorption varies with the occurrence of substorms (JELLY and BRICE, 1967) on a timescale of ~10 minutes (REID, 1967) and consequently the 3-hourly indices $A_p$ and $K_p$ are poor indicators of absorption level.

2.4.6 Signal fading When several modes propagate simultaneously the radiowaves combine at the receiving antenna to give a resultant signal amplitude. In general the constituent radiowaves do not maintain a constant phase relationship due to Doppler shifts imposed by small ionospheric motions. This results in interference fading, a temporal variation in the amplitude of the combined signal. In-mode fading occurs as the ionosphere is not a perfectly smooth reflecting surface and consequently each mode comprises several signals reflected from nearby points in the ionosphere, each interfering with the others at the antenna. Interference fading can also occur between the ordinary and extraordinary modes, and between the high and low angle rays. In general, fading refers to any mechanism that causes variation in the received signal amplitude, which can include variations in absorption, focusing and defocusing of the radiowaves, and diurnal variations in the electron density. The period of the signal level fluctuation is dependent on the fading mechanism; fading produced by mode interference in a disturbed ionosphere can have frequencies of 10 to 100 Hz, whilst fading caused by focusing and defocusing of the radiowaves has periods
Fading, especially that due to Doppler shifting of the received components of the signal, gives rise to spreading of the signal in the frequency domain. The width of the spread spectrum indicates the magnitude of random motions of the ionosphere, the ionospheric disturbance. Doppler spreading of signals can be much larger on high latitude propagation circuits than at mid latitudes (Vincent et al., 1968; Hunsucker and Bates, 1969), especially within the auroral oval where precipitation perturbs the ionosphere to produce the F layer irregularity zone (Pike, 1971).

2.5 Summary

The range of frequencies within the HF band that will propagate on a given path are controlled by the electron density profile of the ionosphere. The MUF is determined by the electron density in the E and F regions; increasing the electron density in the reflecting region increases the MUF. The attenuation produced by a given D region electron density increases as the frequency of the radiowave decreases. The lowest frequency that can be received above the attenuation is the LUF. Increasing the D region electron density increases the level of absorption, and consequently the LUF. The width of the frequency spectra of received radiowaves gives an indication of the magnitude of random motions occurring in turbulent regions of the ionosphere.

The high latitude ionosphere is a highly disturbed propagation medium. Particle precipitation in the auroral zones increases electron densities at all altitudes in the ionosphere, producing irregularities in the F region, auroral E in the E region, and enhancing auroral absorption of radiowaves in the D region. The intensity of particle precipitation in the auroral zones is dependent on the level of geomagnetic activity and the occurrence of substorms. The auroral oval moves equatorwards and broadens during disturbed geomagnetic conditions. Within the polar cap, convection of dayside plasma into the nightside can increase electron densities in the dark F region. However, this convection can also lead to regions of F region electron density depletion, such as the mid latitude trough and the polar hole.

The morphological regions of the high latitude ionosphere are all influenced by changes in geomagnetic activity. Prolonged periods of disturbance, geomagnetic storms, can extend this influence to mid latitudes, producing global increases and decreases in the F region electron density.
Chapter 3
Data collection: the NONCENTRIC experiment

3.1 Experimental arrangement

The NONCENTRIC experiment was designed for the detailed investigation of the effects of the auroral oval and other high latitude phenomena on HF radio propagation. A number of high latitude propagation paths are monitored, which include both trans-auroral paths and paths contained entirely within the polar cap. The NONCENTRIC experiment covers most of the HF radio band (3 to 30 MHz), 14 frequencies between 3 to 23 MHz being employed. Two experimental campaigns were undertaken: the first during summer, 18 July to 11 August 1988, and the second in winter, 16 January to 10 February 1989.

3.1.1 Propagation paths

The experiment comprises one transmitter and five receivers, providing five propagation paths. The transmitter is located at Clyde River, Canadian NWT, situated within the polar cap at all times of day and for all geomagnetic conditions. One receiver is located at Alert, Canadian NWT, giving a path contained entirely within the polar cap. A second receiver near Fairbanks, Alaska, is located in the sub-auroral region during the day, giving a path which is trans-auroral. At night however, the site is located in the auroral zone. The remaining three receivers are located at Aberdeen and Leicester, UK, and Boston, USA, and monitor paths which are trans-auroral at all times of day. The positions of the receivers and the lengths of the paths are summarized in TABLE 3.1, and are illustrated in the top panel of FIGURE 3.1. The locations of four ionosonde stations, data from which are employed in chapter 6, are included also in FIGURE 3.1. In the two lower panels the position of the auroral oval (Kp = 3) is illustrated for 00 UT and 12 UT. The paths to Alert and Boston are approximately meridional, i.e. all points on the paths are at the same local time. The paths to Fairbanks, Aberdeen and Leicester, however, are extended in longitude and local time varies between transmitter and receiver. The Clyde River transmitter and the Aberdeen and Leicester receivers are separated by 4.5 hours of local time; the transmitter and the Fairbanks receiver are separated by 5.5 hours of local time.

<table>
<thead>
<tr>
<th>Site</th>
<th>Geographic location</th>
<th>Path length (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clyde River, NWT</td>
<td>70°N 291°E</td>
<td>-</td>
</tr>
<tr>
<td>Alert, NWT</td>
<td>82°N 298°E</td>
<td>1340</td>
</tr>
<tr>
<td>Fairbanks, Alaska</td>
<td>71°N 204°E</td>
<td>2960</td>
</tr>
<tr>
<td>Boston, USA</td>
<td>42°N 289°E</td>
<td>3120</td>
</tr>
<tr>
<td>Aberdeen, UK</td>
<td>57°N 358°E</td>
<td>3400</td>
</tr>
<tr>
<td>Leicester, UK</td>
<td>53°N 359°E</td>
<td>3800</td>
</tr>
</tbody>
</table>
FIGURE 3.1 (Top panel) A map of the locations of the NONCENTRIC transmitter, Clyde River, and receivers, Alert, Fairbanks, Boston, Aberdeen, and Leicester. The locations of four ionosonde stations, Resolute Bay, Churchill, Ottawa, and South Uist, are also indicated. (Bottom panel) The location of the auroral oval, $K_p = 3$, relative to the NONCENTRIC propagation paths at 00 UT and 12 UT.
The locations of the path midpoints are key parameters in the analysis of the NONCENTRIC data, for instance in determining solar zenith angles. The geographic and geomagnetic locations of the path midpoints are summarized in Table 3.2.

<table>
<thead>
<tr>
<th>Path</th>
<th>Geographic location</th>
<th>Geomagnetic location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alert</td>
<td>76°N 292°E</td>
<td>84°N 31°E</td>
</tr>
<tr>
<td>Fairbanks</td>
<td>76°N 249°E</td>
<td>82°N 281°E</td>
</tr>
<tr>
<td>Boston</td>
<td>56°N 290°E</td>
<td>66°N 9°E</td>
</tr>
<tr>
<td>Aberdeen</td>
<td>67°N 333°E</td>
<td>69°N 66°E</td>
</tr>
<tr>
<td>Leicester</td>
<td>65°N 336°E</td>
<td>67°N 67°E</td>
</tr>
</tbody>
</table>

3.1.2 Transmitter duty cycle Fourteen frequencies between 3 and 23 MHz were monitored by the experiment. The frequencies differed between the summer and winter campaigns but covered the same portion of the HF band. A coded signal was transmitted on each frequency once every hour for the duration of each campaign. To avoid overloading of the equipment the hourly duty cycle of the transmitter was divided into two portions each containing seven transmissions, each portion being separated by a rest interval of approximately 15 minutes. The first seven frequencies transmitted in each hour and the last seven are known as the primary and secondary frequencies respectively. During the summer campaign two separate transmitters were employed for the primary and secondary frequencies; during the winter campaign, one transmitter alone was used. The frequencies and times of transmission are noted in Table 3.3. During each campaign 8400 signals were transmitted (25 days, 14 frequencies each hour).

3.1.3 Transmission sequence Each transmitted signal was of two minutes duration. The signal was divided into six parts, each fulfilling a separate function (Table 3.4). A call sign was employed to aid recognition of the Clyde River signal at the receivers and to allow rejection of interference from other HF operators. The call sign (CZB) was transmitted in slow speed Morse as a 100% amplitude modulated sequence. The technique employed for recognition is described in §3.1.7. A PSK Barker code sequence was employed for analysis of mode structure; the analysis of these results is the subject of another study and is not a part of the present work. A CW carrier signal was transmitted to allow signal strength, signal to noise ratio, and Doppler spreading measurements to be made. Finally, the sign off, with no transmission, indicated that the transmission interval was over.

3.1.4 Transmitter and receiver antennas Different types of antenna were deployed at the transmitter and each receiver. Unfortunately, no detailed measurements of the antenna gain patterns were possible and only the modelled polar diagrams are available. This makes comparison between the behaviour of each propagation path difficult. The signal level at
Chapter 3 Data Collection

each site is influenced both by propagation and ionospheric effects and by changes in antenna gain due to variations in elevation angle of the ray paths. Also, the frequency dependence of the gain of each antenna is unknown. Each receiver antenna was aligned along the great circle path to the transmitter. Thus changes in azimuthal angle of arrival of signals due to off-great circle path propagation could affect the antenna gain by an unknown amount.

At the transmitter two similar antennas were employed in the summer, the primary and secondary frequencies being transmitted on each respectively. For the winter campaign, both were replaced by a single antenna which was resonant at ~7 and ~18 MHz. The receiver antenna at Leicester was also changed between the summer and winter campaigns.

<table>
<thead>
<tr>
<th>Frequency (MHz)</th>
<th>Transmission time after Summer campaign</th>
<th>Transmission time after Winter campaign</th>
<th>hour (mins)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary frequencies</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.185</td>
<td>3.185</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>4.900</td>
<td>4.455</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>6.800</td>
<td>6.800</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>9.941</td>
<td>9.941</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>13.886</td>
<td>13.886</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>18.204</td>
<td>18.204</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>20.900</td>
<td>20.900</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>Secondary frequencies</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.230</td>
<td>4.455</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>5.200</td>
<td>6.905</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>6.906</td>
<td>10.195</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>10.195</td>
<td>14.373</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>14.373</td>
<td>17.515</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>17.515</td>
<td>20.300</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>20.300</td>
<td>23.169</td>
<td>49</td>
<td></td>
</tr>
</tbody>
</table>

3.1.5 Receivers The receivers were programmed to automatically tune to the frequencies of the Clyde River transmissions at the times designated by the transmitter duty cycle. The internal clock of each receiver was accurate to within one second over the duration of a campaign. The receivers determined the real and imaginary components of the signal baseband frequency, so phase and amplitude of the received signals could be measured. The signal amplitude was recorded in decibels relative to 1 mV output from the antennas. The receiver gains were carefully calibrated and were found to be stable to within 1 or 2 dB throughout the campaigns. However, no calibration of the antennas, feeders or RF distribution systems was possible, so absolute values of field strength could not be determined. Consequently, absolute signal strengths measured at the various sites cannot be compared and the data are best suited to studies of the variations as a function of time of
Chapter 3

Data Collection

26

signal strength measured at each site.

<table>
<thead>
<tr>
<th>Time into transmission (s)</th>
<th>Duration (s)</th>
<th>Description</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10</td>
<td>Call sign</td>
<td>Signal recognition</td>
</tr>
<tr>
<td>10</td>
<td>30</td>
<td>Barker code</td>
<td>Mode structure determination</td>
</tr>
<tr>
<td>40</td>
<td>20</td>
<td>Call sign</td>
<td>Signal recognition</td>
</tr>
<tr>
<td>60</td>
<td>30</td>
<td>Carrier</td>
<td>Determination of Doppler spreading, signal strength and signal to noise ratio</td>
</tr>
<tr>
<td>90</td>
<td>20</td>
<td>Call sign</td>
<td>Signal recognition</td>
</tr>
<tr>
<td>110</td>
<td>10</td>
<td>Sign off</td>
<td>End of transmission indicator</td>
</tr>
</tbody>
</table>

A residual d.c. component present in the real and imaginary outputs of the receivers due to unwanted mixing of local oscillator signals within the receivers was discovered during laboratory tests. This component was found to occur with a magnitude comparable to that of weak signals and therefore, to remove any ambiguities regarding the identification of the transmitted signal, a +10 Hz shift was introduced on the signals at the receiver.

During the summer campaign the internal frequency references of the receivers were found to drift with time, causing problems in the analysis of the received signals after 1 August 1988 (the last 10 days of the campaign). For the winter campaign the receivers were attached to an external frequency standard to prevent recurrence of this problem.

The real and imaginary outputs of the receivers were sampled every 20 ms and recorded onto tape. 1000 sample pairs were recorded for the carrier portion of the transmission and 530 pairs were recorded for the final call sign portion of the signal.

3.1.6 Scaling of signal characteristics  To aid the analysis of the large quantity of data collected, three parameters were automatically scaled from the received amplitude data of each signal. The parameters determined were the peak signal amplitude, the spread index and the mean noise level.

Peak signal amplitude  The real and imaginary components of the received signal were combined in quadrature to give the signal amplitude as a function of time for the duration of the transmission and the maximum value determined. Below a signal level of -60 dB the signal was dominated by noise from the receiver and could not be detected.

Spread index  The spread index is a numerical representation of the spectral width of the received signals. The first step in the computation of the spread index is to determine the
normalized Fourier Transform, in the frequency range -25 to +25 Hz, of the received amplitude data. The transform generally has a peak at approximately +10 Hz due to the frequency shift described in §3.1.5. The area under the curve in the frequency range -25 to -12.5 Hz is calculated as this is assumed to be a measure of the noise content, with no signal contribution. If the noise level is assumed to be constant over the whole frequency range -25 to +25 Hz then the noise contribution to the area under the entire curve is four times this value. The area under the whole curve is found and the noise subtracted. The remaining area is a measure of the width of the spectrum, the area increasing with increasing width. A scaling factor of 20 is applied so that a normalized triangular spectrum of width 10 Hz at its base would have a spread index of 100 Hz (the area of such a triangle being $\frac{1}{2} \times 1 \times 10 = 5$). Figure 3.2 summarizes the procedure for calculating spread index.

Spread index is easy to compute and does not suffer from some of the disadvantages of other methods of spectral width determination; for example, fitting an assumed functional form to the signal spectrum fails when the spectrum is of an unexpected or disturbed shape.

Mean noise level The mean signal level measured in the frequency range -25 to -12.5 Hz of the Fourier spectrum is assumed to be the noise level over the whole spectrum and the signal to noise ratio is determined by means of this value.

3.1.7 Signal recognition A signal recognition procedure is necessary so that signals that are a) dominated by interference from sources other than the Clyde River transmitter or b) are measured in the absence of the Clyde River signal due to ionospheric absorption or penetration of the ionosphere, can be identified and rejected. Each signal is categorised by the outcome of two tests, the call sign and spread index tests.

Call-sign test The call sign portion of the transmitted signal has a characteristic spectral content determined by the frequency components of the Morse code keying sequence CZB. The top panel of Figure 3.3 illustrates the normalised variation of the received signal level on 17.515 MHz on the path to Leicester for 30 seconds of the transmission sequence. The first 20 seconds correspond to the carrier portion of the signal and the last 10 seconds correspond to the call sign portion of the signal; the call sign is transmitted twice, with an intervening period during which no transmission occurs. The Fourier spectrum of the call sign differs from that of the unmodulated carrier signal (both shown to the right of Figure 3.3) by the presence of modulation sidebands 6 dB less powerful than, and situated approximately 5 Hz either side of, the central peak. Software was written which automatically searched for this pattern in the spectrum of each signal. The peaks had to exceed the noise level by a factor of 2.5 for unambiguous recognition of the pattern.

Spread index test The spread index and signal amplitude of the carrier portion of the signal are employed in conjunction to provide a second recognition test. A negative spread index (due to the noise level being greater than the strength of the signal), or a low positive...
Chapter 3

Data Collection

1.0 Determine noise level between -25.0 and -12.5 Hz

-25 0 25
Frequency (Hz)

-25 0 25
Frequency (Hz)

1.0 Subtract estimate of total noise from spectrum

-25 0 25
Frequency (Hz)

1.0 Estimate spectral width from area under spectrum

-25 0 25
Frequency (Hz)

FIGURE 3.2 A schematic diagram of the calculation of spread index. The noise level is determined from the frequency range -25.0 to -12.5 Hz of the normalized Fourier transform of the received signal (top panel), and four times this value (middle panel) subtracted from the total area under the spectrum to give the spread index (bottom panel).
FIGURE 3.3 Three examples of received signal amplitude data (left panels) on the propagation path to Leicester, illustrating the carrier and call sign portions of the signal, and the Fourier transform of each portion (right panels).
spread index combined with a low signal amplitude, is indicative of a very low signal to noise ratio or a corrupt signal and the signal fails the test.

Signal categorisation One of five categories of signal recognition are awarded to each signal depending on the outcome of the two recognition tests. The categories are:

- **Signal** (red): Signals that passed both the call sign and spread index tests.
- **Failed spread index** (yellow): Signals that passed the call sign test but failed the spread index test.
- **Failed call sign** (orange): Signals that passed the spread index test, failed the call sign test and had a low signal amplitude.
- **Interference** (green): Signals that passed the spread index test, failed the call sign test and had a high signal to noise ratio. With a high SNR it would be expected that the call sign test would be passed unless an interfering signal was dominating the channel.
- **Noise** (blue): Signals that failed both call sign and spread index tests. This is most often a consequence of a low signal to noise ratio or an absence of the signal altogether due to penetration of the ionosphere by the signal. However, interfering signals sometimes produce this outcome if they are not positively identified as interfering signals.

The colours in parentheses represent these signal recognition categories in diagrams in the rest of this thesis. The additional category **no data** (white) indicates that the transmitter or receiver was inoperative at the time of the signal transmission. The analysis of the data concentrates primarily on signals that passed both signal recognition tests as these are the only signals which can be reliably attributed to a Clyde River transmission.

### 3.2 Limitations of the data

The NONCENTRIC experiment was not designed for the accurate measurement of signal strength or Doppler measurements and in many ways it is not ideally suited for this purpose. The following sections discuss the limitations of the experiment and ways in which they are overcome.

#### 3.2.1 Equipment limitations

As already stated, exact measurements of field strength are not possible, however, variations in the field strength can be measured accurately. The experiment is therefore best suited to the study of phenomena that cause such variations.

Unfortunately, the NONCENTRIC experiment has no angle of arrival measurement capability; it has to be assumed that rays propagate along the great circle paths from the transmitter to the receivers. Off-great circle path propagation might be poorly detected as
the antennas are directional in the great circle path direction. It is well known that side-scatter, from the ground and ionisation features in the high latitude ionosphere, is an important propagation effect.

The limited frequency range of the experiment, approximately 3-23 MHz in winter and 3-21 MHz in summer, means that the behaviour of the highest propagating frequencies cannot be investigated: in winter the expected MUF on the paths to Boston and Leicester is in excess of 30 MHz.

Equipment problems were encountered at the receiving site at Aberdeen, resulting in a low signal to noise ratio throughout both experimental campaigns. Consequently the occurrence of correctly recognized signals at this site was very low, especially during the winter campaign when only 10% of the signals received could be identified as Clyde River transmissions.

3.2.2 Time resolution Each of the fourteen frequencies is sampled once every hour. This time resolution is insufficient for the investigation of many transient ionospheric phenomena, including travelling ionospheric disturbances (TIDs) and rapid temporal variations in auroral absorption. Such disturbances produce uncertainties in the data: each data point will have a superimposed component of the disturbance which is unknown and unpredictable. Due to the uncertainty associated with each data point geophysical changes on individual days are difficult to investigate. The only way to overcome this limitation is by averaging data over each campaign or sub-periods of the campaigns in order to determine a mean diurnal variation or other mean behaviour. Gross variations in the data, however, can be identified without averaging and the overall response of the data to large geophysical phenomena, such as geomagnetic storms, are readily apparent.

Pairing frequencies The poor time resolution can in part be overcome by pairing frequencies. For each primary frequency transmitted in the first half of the hour there is a corresponding secondary frequency transmitted in the second half of the hour which is different by only ~100 kHz (the frequency difference is less than 6% in the worst case and ~1-2% in most). If each pair of frequencies is taken together, the fourteen frequencies combined to form seven frequency bands, then the sampling rate is doubled (a sampling resolution of half an hour) or the number of samples in each frequency range is doubled (for periods when there is not much data or binning of the data reduces sample sizes). Care must be taken, however, as the primary and secondary frequencies are transmitted on separate, though identical, antennas; correspondence between the data recorded on the primary and secondary frequencies in each frequency pair is generally good, however. Consequently, for periods when data is sparse or for analyses which involve binning the data such that sample sizes are small, data from both frequencies are employed.

3.2.3 Sampling interval Each measurement of the received signal is made over a period
of 30 seconds and from this the peak signal amplitude is deduced. A sampling interval of 30 seconds is, however, generally accepted to be very short for a representative measurement of signal amplitude because the signal level fluctuates randomly (so called fading) with periods from seconds to minutes, especially at high latitudes (DAVIES, 1990). The period of the fading depends on the fading mechanism: flutter fading experienced in the auroral regions during disturbed conditions can have frequencies of 10 to 100 Hz, whilst fading due to focusing or defocusing of the propagating rays has periods of 15 to 30 minutes. In polar regions fading can have a peak-to-trough variation in signal amplitude of 15 dB. The upper two panels of FIGURE 3.3 illustrate signal amplitude data received at Leicester on 17.515 MHz at 11 UT and 12 UT on 24 January 1989. At 11 UT the carrier and call sign portions of the signal are modulated by ~1 Hz fading. The corresponding frequency spectra are narrow. At 12 UT the fading is much more rapid and the corresponding frequency spectra are spread.

3.3 Interference

External sources transmitting at frequencies close to the NONCENTRIC channels can dominate the wanted signal. The bottom panel of FIGURE 3.3 illustrates the variation in amplitude of an interfering signal received at Leicester on 6.905 MHz. The frequency spectra of this signal are illustrated to the right of the diagram and are of a very different form than that produced by NONCENTRIC signals (top two panels of FIGURE 3.3). The signal amplitude and noise level produced by the interference are ~30 dB greater than those measured for NONCENTRIC signals during periods when the interference is absent. The parameters calculated from such signals are not representative of the state of the ionosphere on the Clyde River to Leicester path and the data have to be excluded from the analysis.

3.3.1 Local interference Interference is commonly observed at 4.455 MHz and 6.905 MHz during the winter campaign on the paths to Leicester and Aberdeen; since these are the only sites to observe this interference the source must be fairly localized. The diurnal variation of the peak signal strengths measured at 4.455 MHz and 6.905 MHz are reproduced in FIGURE 3.4. At 6.905 MHz the interference is received between 07/08 UT and 18 UT. During the periods when interference is absent the NONCENTRIC signal is frequently observed. At 4.455 MHz the interference begins slightly earlier at 06 UT and continues after the end of interference on 6.905 MHz, until 00 UT when the signal level begins to drop. The interference spans a greater period at 4.455 MHz than at 6.905 MHz as the ionosphere can support the interfering signal with a lower critical frequency. During the day the strength of the interfering signal exhibits a \( \cos \chi \) dependence, indicating that the interference must be propagating by a sky wave, attenuated by the diurnal variation of the ionospheric D region electron density, and is not a local ground wave. If the interfering signal has a fairly local source, its daytime signal strength variation should be governed by
FIGURE 3.4 Mean diurnal variation of signal level at 4.455 and 6.905 MHz on the path to Leicester, winter campaign (solid curves). Also illustrated are the cosine of the solar zenith angle, $\cos \chi > 0$, (scaled by eye) for the Clyde River to Leicester path midpoint (dashed curves) and local to Leicester (dotted curves). The portion of the day during which an interfering signal dominates the channel is indicated in each panel.
cos \chi measured locally, not at the path midpoint from Clyde River. In FIGURE 3.4 the daytime values of cos \chi (\chi < 90°) have been plotted, calculated for 54°N, 1°W, the midpoint between Aberdeen and Leicester, and for the midpoint of the path from Clyde River. The curves have been scaled to achieve the best fit by eye. The local cos \chi variation fits the data with a high degree of accuracy, the Clyde River path midpoint cos \chi variation being displaced by several hours. Also, the amplitudes of the local cos \chi curves are found to be in the ratio of the frequencies squared, \((4.455/6.905)^2\), as would be expected for D region non-deviative absorption.

At 6.800 MHz interference is experienced between 12 UT and 22 UT. At 23 UT interference ceases abruptly and reception of the Noncentric signal recommences. This interference is less intense than at 4.455 MHz and 6.905 MHz; the peak signal level is only \(-10\) dB higher than the Noncentric signal. Consequently the NONCENTRIC signal is sometimes correctly recognized during this period as variations in the ionosphere cause the signal or the interference to dominate. The same increase in signal strength and poor reception of the NONCENTRIC signal is found on 6.800 MHz at Aberdeen. At 13.886 MHz interference is received at 2 hours only, 09 UT and 14 UT.

3.3.2 Global interference Stronger sources of interference can be detected at all five receiving sites. For instance, all sites experience an enhanced signal and noise level on 9.941 MHz between 23 UT and 04 UT. On most paths the return of the noise level to the background level at 05 UT is very abrupt. As this noise enhancement is observed simultaneously on all paths it is unlikely that this is an ionospheric or propagation effect and is most likely interference from a source other than the Clyde River transmitter. Signals received at these times are poorly recognized, generally failing just the call sign test (Fairbanks) or both recognition tests (Alert, Aberdeen, Boston and Leicester).

3.4 Summary

The NONCENTRIC experiment to investigate the propagation of HF signals through the auroral ionosphere has been outlined. Several different types of path are available, which enable different regions (e.g. auroral oval, polar cap) to be investigated. Some of the limitations of the experiment in regard to absolute signal strength measurements and time resolution have been identified. The techniques adopted for signal recognition and measurement of the spread in the Doppler frequency spectrum have been discussed in detail.
Chapter 4

4.1 Introduction

From a preliminary investigation of the NONCENTRIC data the median behaviour of HF radiowave propagation at high latitudes, and the mechanisms which are important in producing deviations from this median behaviour are identified. Detailed investigation of these mechanisms are presented in later chapters. In the present chapter the diurnal and seasonal variation of correct signal recognition times and the change in signal behaviour during periods of disturbed geomagnetic activity are described. A similar analysis of spread index and signal level is then presented.

4.2 Recognition of signals

The diurnal pattern of the periods of correct signal recognition gives information concerning the ionospheric control of the propagation on each of the experimental paths. The occurrence of recognized signals, times at which communication is most reliable, and the reasons for failure of recognition are investigated.

4.2.1 Diurnal variation of signal recognition

The top panel of FIGURE 4.1 illustrates the results of the signal recognition tests as a function of time and frequency for the path to Boston for 26-31 January 1989, six typical days of the winter campaign (note that in this and similar figures, the frequency scale is not linear). The colours represent the signal recognition categories described in §3.1.7. The bottom panel of FIGURE 4.1 illustrates the variation of \( \cos \chi \) at the path midpoint for the same period. Local noon (unless specified otherwise, local times refer to the path midpoint) occurs at 17 UT when \( \cos \chi \) is a maximum, indicated by the triangles at the top of both panels of FIGURE 4.1. Sunrise and sunset at the ground occur at 12 UT and 22 UT respectively, when \( \chi = 90^\circ, \cos \chi = 0 \), indicated by vertical dashed lines.

Signals at frequencies above 10 MHz are not detected at night. At night the F region electron density \( N_e F_2 \approx f_r F_2^2 \) is too low to support reflection of these frequencies and the signals penetrate the ionosphere and propagate into space. Consequently no signal is detected at the receiver and only noise is measured (blue). Correct recognition (red) on these frequencies begins at sunrise (12 UT) when photo-ionization of the atmosphere increases the F region electron density to a level sufficient to support the propagation. Recognition does not stop at sunset (22 UT) but continues for several hours as a consequence of slow recombination rates in the F region maintaining \( f_r F_2 \) above the level necessary for propagation. The time at which propagation finally ceases after sunset varies
FIGURE 4.1 Signal recognition results for six days, 28–31 January 1989, on the path to Boston, winter campaign (top panel), and cos χ for the same period, determined for the path midpoint. Vertical dashed lines indicate sunrise and sunset (cos χ = 0). Arrows at tops of both panels indicate local noon.
Signals at frequencies below 10 MHz are predominantly recognized at night. The F region electron density remains at levels sufficient to support propagation on these frequencies throughout the day and night. Night-time propagation on frequencies below 10 MHz is less reliable than daytime propagation on frequencies above 10 MHz; a high proportion of low frequency signals fail one or both of the signal recognition tests, in general as a consequence of interference from external sources. Low frequency signals fail both recognition tests near local noon, most clearly observed on 27 and 28 January. On the lowest frequencies (~4.5 MHz) recognition fails between sunrise and sunset, but the duration of recognition loss decreases as frequency increases. At 10-14 MHz signals are lost near noon only. Signal recognition fails on low frequencies during the day due to attenuation of the signals by absorption in the solar-enhanced D region (see §4.4 and chapter 5).

Figure 4.2 illustrates the diurnal variation of the mean signal level of all signals received in six frequency bands on the path to Boston during the winter campaign. At frequencies of 14 MHz and above the signal level increases abruptly at sunrise (12 UT) as propagation of the signal commences; the weak signal levels measured during the night at these frequencies are an indication of the noise level. The slow decrease in mean signal level near sunset (22 UT) is a consequence of the variable time of signal loss in the evening; on individual days the decrease in signal level is abrupt when propagation stops (chapter 6). Between sunrise and sunset the signal level remains relatively constant. At frequencies of 10 MHz and below the mean signal level remains approximately constant during the night. After sunrise the signal level decreases during the morning, reaches a minimum at local noon, and increases again in the afternoon, in response to the photo-ionization of the D region during the day. The level of daytime absorption decreases with increasing frequency due to the $f^2$ frequency dependence of non-deviative absorption. At frequencies of 17 MHz and above no daytime variation of signal level occurs as the non-deviative absorption is negligible. The diurnal variation of signal level at intermediate frequencies (~14 MHz) is a combination of increasing signal level at sunrise and decreasing level at sunset, as for higher frequencies, and a decrease in signal level near local noon as a consequence of absorption.

Recognition of signals on frequencies of 10 MHz and below ceases once the daytime D region absorption becomes great enough that the call sign modulation pattern can no longer be distinguished from the noise. As the level of absorption decreases with increasing frequency the recognition of lower frequencies fails for a greater proportion of the day than higher frequencies which fail recognition near noon only, when absorption is a maximum.

Thus the diurnal variation of signal recognition is predominantly controlled by solar photo-ionization of the ionosphere, the recognition of higher frequencies depending on the F region electron density variation and the recognition of the low frequencies depending on the variation of the electron density of the D region.
FIGURE 4.2  Mean diurnal variation of signal level in six frequency bands for the path to Boston, winter campaign. Vertical dashed lines indicate sunrise and sunset and the arrow indicates local noon. The each curve is displaced by 50 dB for clarity.
4.2.2 Seasonal variation of signal recognition  FIGURE 4.3 illustrates the percentage occurrence of the five signal recognition categories on the 14 experimental frequencies for each path during the summer (left panels) and winter (right panels) campaigns; the overall occurrence is indicated at the right of each panel. In summer the distributions are generally singly peaked, with the maximum occurrence of correctly recognized signals occurring at a frequency near 10 MHz or above. In winter the distributions for the longer ground ranges (Boston, Aberdeen, and Leicester) are bi-modal with one maximum near 7 MHz and the second at a frequency of 14 MHz or above. The winter distribution at Alert is difficult to interpret due to interference problems at 6.800 and 9.941 MHz. The difference between the summer and winter distributions is a consequence of the seasonal differences in the diurnal variation of $\phi F_2$. In summer, at high and mid latitudes, the diurnal variation in $\phi F_2$ is small, $\phi F_2$ remaining near 5-7 MHz throughout most of the day due to the seasonal anomaly (see chapter 6). The top panel of FIGURE 4.4 illustrates the daytime and night-time frequency distribution of correctly recognized signals during the summer campaign on the path to Boston. The two distributions have maxima at 14.373 MHz (day) and 9.941 MHz (night), and their sum gives a singly-peaked distribution. In winter, the $\phi F_2$ variation is much larger, between 4 MHz at night and 12 MHz during the day. This is reflected in the shifting of the daytime and night-time signal occurrence distributions to higher and lower frequencies, 18.204 and 6.800 MHz respectively (bottom panel of FIGURE 4.4). The sum of these two distributions gives a bi-modal distribution.

The bi-modal nature of the winter distributions is reinforced by the frequency dependent gain of the Clyde River transmit antenna, resonant at ~7 and ~18 MHz. FIGURE 4.5 illustrates the mean signal level of correctly recognized signals on each path as a function of frequency for the winter campaign. A maximum in signal level is observed at 18 MHz on all paths, irrespective of ground range. A maximum in signal level is observed at 7 MHz on the paths to Alert, Fairbanks and Boston and at 5 MHz on the paths to Leicester and Aberdeen.

The period of solar illumination changes between summer and winter and this is reflected in the relative occurrences of correctly recognized signals in summer and winter. More signals are recognized on the lower frequencies in winter than in summer as a consequence of the shorter sunlit period when D region absorption is significant; this is most obvious at 6.800 MHz on the paths to Fairbanks, Boston, and Leicester where the occurrence of correct recognition is much greater in winter than in summer (FIGURE 4.3). The noon maximum of $\phi F_2$ is greater in winter than in summer and consequently there is a greater daytime occurrence of recognized signals at frequencies above 18 MHz during the winter campaign (FIGURE 4.4). However, over the whole day, the occurrence of correctly recognized signals at frequencies above 18 MHz is not much different in winter and summer (FIGURE 4.3) as the winter daylight period is shorter.

The maximum occurrence of recognized signals on each path is tabulated in TABLE 4.1, along with the frequency at which this maximum occurs. These results are uncertain in
FIGURE 4.3 Signal recognition occurrence distribution as a function of propagation frequency for all NONCENTRIC paths during the summer (left panels) and winter (right panels) campaigns. At the right of each panel is indicated the total occurrence of signal recognition categories.
**Figure 4.4** Occurrence distribution of correctly recognized signals on the path to Boston, during the summer (upper panel) and winter (lower panel) campaigns. Distributions during the day (solid curves), during the night (dashed curves) and sum of day and night (shaded histogram) are indicated.
Winter campaign

![Graph showing mean signal level as a function of frequency for five NONCENTRIC propagation paths.]

FIGURE 4.5 Mean signal level as a function of frequency for the five NONCENTRIC propagation paths.
some cases as interference affects the distributions, for instance at 6.800 and 9.941 MHz on the path to Alert, winter campaign, where interference is identified as noise by the signal recognition tests (FIGURE 4.3). In general, the maximum of the distributions moves to higher frequencies in the summer, and the maximum occurrence is greater in summer than in winter. This is a consequence of the greater overlap of the daytime and night-time occurrence distributions in the summer than in the winter (FIGURE 4.4).

<table>
<thead>
<tr>
<th>Path</th>
<th>Frequency (MHz)</th>
<th>Occurrence (%)</th>
<th>Path</th>
<th>Frequency (MHz)</th>
<th>Occurrence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alert</td>
<td>6.905</td>
<td>97</td>
<td>Alert</td>
<td>6.800-9.941</td>
<td>82</td>
</tr>
<tr>
<td>Fairbanks</td>
<td>9.941</td>
<td>96</td>
<td>Fairbanks</td>
<td>6.905</td>
<td>68</td>
</tr>
<tr>
<td>Boston</td>
<td>9.941</td>
<td>88</td>
<td>Boston</td>
<td>6.800</td>
<td>61</td>
</tr>
<tr>
<td>Aberdeen</td>
<td>13.886</td>
<td>51</td>
<td>Aberdeen</td>
<td>6.905</td>
<td>26</td>
</tr>
<tr>
<td>Leicester</td>
<td>13.886</td>
<td>70</td>
<td>Leicester</td>
<td>14.373</td>
<td>47</td>
</tr>
</tbody>
</table>

1 Uncertain due to interference on 6.800 and 9.941 MHz.

4.2.3 Dependence of signal recognition on path length  TABLE 4.1 indicates that the frequency at which the greatest number of correct identifications are observed increases with increasing ground range (the paths are tabulated in order of increasing ground range). This is especially true during the summer campaign where the distributions are singly peaked. The secant factor of a path becomes larger with increasing ground range, and a given F region critical frequency will therefore reflect a higher signal frequency. Consequently the maximum in occurrence of signal recognition is shifted to higher frequencies as ground range increases. The occurrence of correct recognition at the distribution maximum decreases with increasing ground range; this trend is not followed by Aberdeen, due to equipment problems (see §3.2.1).

4.2.4 Dependence of signal recognition on geomagnetic activity There is a decrease in the occurrence of correctly recognized signals during periods of high geomagnetic activity. FIGURE 4.6 illustrates the occurrence of different signal recognition categories for each day of the winter campaign for 9.941, 13.886, 17.515, and 18.204 MHz on the path to Fairbanks. \( \Sigma K_p \), the sum of the eight daily \( K_p \) values, is indicated for each day also (note that the scale of \( \Sigma K_p \) is inverted to aid comparison). Periods of high \( \Sigma K_p \) are clearly associated a lower occurrence of correctly recognized signals and a higher occurrence of the noise recognition class, especially 20-21 January and 3-4 February. The number of days during which propagation is adversely affected during disturbed periods increases with increasing propagation frequency. At 9.941 MHz only 1-2 days exhibit decreased recognition during disturbed periods (21 January and 3-4 February). At 18.204 MHz, however, 5-6 days at a time have decreased signal recognition (20-24 January and 1-6
FIGURE 4.6 Daily occurrence of signal recognition categories at four frequencies on the path to Fairbanks during the winter campaign. Daily $\Sigma K_p$ is also indicated (solid curve), on an inverted scale.
April. Recognition decreases on lower frequencies due to an increase in auroral absorption (see chapter 5). Signal recognition fails on high frequencies due to a decrease in the F region electron density during negative ionospheric storms (see chapter 7).

The general behaviour on all paths is similar but each can exhibit individual features. On the Alert path the decrease in recognition on the higher frequencies is observed, but not the decrease on the lower frequencies as the auroral oval is far equatorward of the path and so no auroral absorption affects the propagation. The correspondence between high $\Sigma K_p$ and the decrease in recognition of signals on lower frequencies is greater for the path to Boston than for the path to Fairbanks, reflecting the relative influence of the oval on the two paths. The decrease in high frequency propagation with increasing $\Sigma K_p$ is not as great on the Boston path as on the Fairbanks path, as the decrease of recognized signals during the day is compensated by an increase in night-time anomalous propagation during disturbed periods (see chapter 6). Also, the daytime MUF is greater than the highest frequency of the experiment (23 MHz) on the longer paths to Boston and Leicester, and so decreases in the MUF have to be very large before influencing the experimental frequencies.

4.3 Spread index

Spread index is an indicator of the frequency spread of a radiowave produced by random motions in the ionosphere, especially near the path reflection points (§2.4.6). A high spread index indicates considerable distortion of the signal in the frequency domain as a consequence of multi-path or Doppler spreading of the signal. The diurnal variation of the spread index on different frequencies and the variation of spread index with $K_p$ gives an indication of the factors producing disturbance in the ionosphere. Factors influencing the spread index on each path are now described.

4.3.1 Path dependence of spread index behaviour  The variation of spread index on each propagation path is described in turn, though the path to Aberdeen is omitted due to the equipment problems experienced and its similarity to the path to Leicester.

Figure 4.7 illustrates the mean diurnal variation of spread index for each of the propagation paths during the summer and winter campaigns, the vertical bars indicating one standard deviation above and below the mean. Both mean and median behaviour have been investigated, and both produce similar results. Only signals which passed both signal recognition tests are included in the analysis. Two frequency bands are illustrated for each path, where possible one each from the lower and higher portions of the frequency range of the experiment. Figure 4.8 illustrates the diurnal variation of the position of the auroral oval ($K_p = 3$) relative to each of the paths. The contours indicate the diurnal variation of the solar zenith angle $\chi$ along the paths during the winter campaign.

Alert  In summer the spread index is low at all times and there is little diurnal variation on
FIGURE 4.7  Mean diurnal variation (curve) and standard deviation (vertical bars) of spread index on the paths to Alert, Fairbanks, Boston, and Leicester during the summer (upper panels) and winter (lower panels) campaigns. Two frequency bands are indicated for each path and campaign. Arrows indicate local noon at the path midpoint.
FIGURE 4.8  Diurnal variation of the location of the FELDSTEIN oval (Kp = 3) along the propagation paths (transmitter at top and receiver at bottom of each panel) to Alert, Fairbanks, Boston and Leicester. The variation of solar zenith angle $\chi$ along the paths during the winter campaign is also indicated.
any frequency. The diurnal variation in winter is larger and the mean spread index is larger than during the summer campaign. The diurnal maximum occurs during the morning sector. The minimum occurs near 17 UT, local noon, when the path is solar-illuminated. There is no change in the diurnal variation as $K_p$ increases.

**Fairbanks** During the summer campaign the diurnal variation of spread index on frequencies above 10 MHz is a maximum at 22 UT, near local noon (20 UT). The auroral oval is coincident with the equatorward 2-hop reflection point between 20-04 UT. During this period spread index increases with increasing $K_p$. Throughout the rest of the day the oval is equatorward of the reflection point and the path is predominately enclosed within the polar cap the spread index is low. At these times, spread index shows no trend with $K_p$ or even a decrease as $K_p$ increases, possibly reflecting the diminishing influence of the oval as it moves to lower latitudes away from the path. On frequencies below 10 MHz there is little diurnal variation in spread index. During the winter campaign there is little diurnal variation in spread index on any frequency, the level remaining high at all times.

**Boston** During the summer, on propagation frequencies which are predominantly received at night (below 7 MHz), spread index is a maximum near local midnight (05 UT) when the midpoint of the path is within the auroral oval. Spread index increases with $K_p$ at this time. Higher frequencies (9-14 MHz) also have a maximum in spread index in the midnight sector. As $K_p$ increases a second maximum in spread index appears near local noon (16-19 UT) on frequencies between 9-18 MHz. This is possibly caused by a disturbance of the ionosphere by precipitation in the cusp region of the dayside auroral oval.

In winter spread index tends to increase with increasing $K_p$. The cusp dependence of spread index is smaller, and this possibly reflects a seasonal modulation of cusp precipitation as the northern hemisphere is oriented towards and away from the sun during the summer and winter respectively. At night (23-06 UT) the variation of spread index with $K_p$ is more complicated than a simple increase during disturbed conditions. Figure 4.9 illustrates the mean variation of spread index as a function of $K_p$ for 6.800 and 6.905 MHz between 23-10 UT. Also indicated is the diurnal variation of the position of the auroral oval relative to the path for four levels of $K_p$; the equatorward portion of the oval is emphasised in dark grey. At 23 UT spread index increases with all values of $K_p$ as the oval approaches, and is then coincident with, the path midpoint. At 00-01 UT spread index increases with increasing geomagnetic activity, for $K_p < 4$, as the oval moves from a point poleward of the path midpoint to reach the midpoint when $K_p = 4$. At $K_p > 4$ spread index decreases as the path midpoint moves nearer the centre of the auroral oval. This possibly indicates that the maximum in spread index is associated with the low latitude boundary of the auroral oval. Between 02-04 UT the path midpoint is within the auroral oval for $K_p > 0$ and the low latitude boundary moves equatorward, away from the path midpoint, with increasing geomagnetic activity; at this time spread index decreases as $K_p$ increases. The behaviour
**Chapter 4**

**Preliminary Analysis**

**Winter Boston**

Mean spread index of recognised signals

6.800, 6.905 MHz

**Figure 4.9** Variation of mean spread index (solid curves) and standard deviation (vertical bars) as a function of $K_p$ on the path to Boston during the winter campaign, for hours between 23 UT and 10 UT (top panels). The diurnal variation of the position of the FELDSTEIN oval along the Boston path for four levels of $K_p$ is illustrated (bottom panels). The low latitude portion of the oval is indicated in dark grey.
between 05-08 UT is similar to the period 00-01 UT as the low latitude boundary moves towards and then away from the path midpoint with increasing $K_p$. After 09 UT the spread index variation returns to a monotonic increase with increasing $K_p$ as the oval has retreated polewards to a position similar to that at 23 UT.

During the winter campaign there is less diurnal variation of spread index than during the summer, spread index remaining high at all times. On lower frequencies spread index is a maximum at night. The night-time maximum is wider in winter than in summer indicating the variation is associated with the length of the day as well as the auroral oval.

**Leicester** In winter the diurnal variation of spread index changes markedly with frequency, indicating that different frequency regimes are controlled by different mechanisms. At ~7 MHz spread index is a maximum during the night at 06 UT and a minimum during the day between 10-20 UT, when solar illumination of the path is at its greatest. The maximum and minimum correspond to times when the path is predominantly in darkness and sunlight respectively. At higher frequencies (>13 MHz) the spread index is a minimum at 06-08 UT and a maximum at 23-00 UT when the path midpoint is coincident with the auroral oval. The longitudinal extent of the path separates the solar control and auroral control of spread index in time. Spread index generally increases with $K_p$ at all times and on all frequencies. Between 18-00 UT, however, the spread index decreases on the higher frequencies for $K_p > 4$ when the low latitude boundary of the oval has moved equatorward of the path midpoint, in a similar manner to the path to Boston.

Mean values of spread index are lower in summer than in winter. The diurnal variation of spread index in summer is similar to the winter patterns but the variation is smaller. The diurnal variation of spread index on the higher frequencies is controlled by the position of the auroral oval, and on the lower frequencies is controlled by the level of illumination, as during the winter. On the higher frequencies between 10-14 UT there is evidence for an increase and then decrease in spread index as $K_p$ increases, corresponding to the movement of the cusp relative to the poleward 2-hop reflection point.

### 4.3.2 Auroral oval dependence of spread index

Figure 4.10 illustrates the mean variation of the spread index of correctly recognized signals as a function of $K_p$ (two frequency bands are illustrated for each path for both the summer and winter campaigns). Polar cap paths, the paths to Alert and Fairbanks, show little variation of spread index with increasing geomagnetic activity. The trans-auroral paths to Boston and Leicester, however, show a positive correlation between spread index and $K_p$. The auroral oval produces disturbance in the ionosphere, the level of disturbance being related to the level of geomagnetic activity. Radiowaves passing through this region of the ionosphere therefore experience Doppler spreading and an increase in spread index.

Trans-auroral paths display an increase in spread index with $K_p$ at all times, though the increase is especially large when the auroral oval is coincident with a path reflection point.
Chapter 4

Preliminary Analysis

Summer campaign

<table>
<thead>
<tr>
<th>Alert</th>
<th>Fairbanks</th>
<th>Boston</th>
<th>Leicester</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.941 MHz</td>
<td>10.195 MHz</td>
<td>6.800 MHz</td>
<td>6.905 MHz</td>
</tr>
<tr>
<td>13.300 MHz</td>
<td>14.373 MHz</td>
<td>9.941 MHz</td>
<td>10.195 MHz</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Alert</th>
<th>Fairbanks</th>
<th>Boston</th>
<th>Leicester</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.800 MHz</td>
<td>6.905 MHz</td>
<td>6.800 MHz</td>
<td>6.905 MHz</td>
</tr>
<tr>
<td>18.204 MHz</td>
<td>19.300 MHz</td>
<td>20.300 MHz</td>
<td>20.900 MHz</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Alert</th>
<th>Fairbanks</th>
<th>Boston</th>
<th>Leicester</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.800 MHz</td>
<td>6.905 MHz</td>
<td>6.800 MHz</td>
<td>6.905 MHz</td>
</tr>
<tr>
<td>18.204 MHz</td>
<td>19.300 MHz</td>
<td>20.300 MHz</td>
<td>20.900 MHz</td>
</tr>
</tbody>
</table>

Winter campaign

<table>
<thead>
<tr>
<th>Alert</th>
<th>Fairbanks</th>
<th>Boston</th>
<th>Leicester</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.800 MHz</td>
<td>6.905 MHz</td>
<td>6.800 MHz</td>
<td>6.905 MHz</td>
</tr>
<tr>
<td>18.204 MHz</td>
<td>19.300 MHz</td>
<td>20.300 MHz</td>
<td>20.900 MHz</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Alert</th>
<th>Fairbanks</th>
<th>Boston</th>
<th>Leicester</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.800 MHz</td>
<td>6.905 MHz</td>
<td>6.800 MHz</td>
<td>6.905 MHz</td>
</tr>
<tr>
<td>18.204 MHz</td>
<td>19.300 MHz</td>
<td>20.300 MHz</td>
<td>20.900 MHz</td>
</tr>
</tbody>
</table>

Figure 4.10 Mean variation (curve) and standard deviation (vertical bars) of spread index as a function of $K_p$ on the paths to Alert, Fairbanks, Boston, and Leicester during the summer (upper panels) and winter (lower panels) campaigns. Two frequency bands are indicated for each path and campaign.
Chapter 4  Preliminary Analysis

Two regions in the oval are of significance for interpreting the experimental data: the cusp region and the midnight sector oval. The increase in spread index with $K_p$ is better correlated with the position of the low latitude boundary of the midnight sector oval relative to the path reflection points than with the whole width of the oval; this region is associated with the auroral boundary blob, a region of enhanced F region plasma density (Rino et al., 1983; Weber et al., 1985; Senior et al., 1987), and the poleward edge of the trough, where diffuse precipitation and spread F are commonly found (e.g. Jones, 1980; Rodger and Pincock, 1982). These regions are described in more detail in chapter 2. The cusp region produces greater variations in spread index during the summer than in the winter, due to the seasonal variation of its orientation relative to the sun. The oval controlled variation in spread index is greatest on the higher frequencies (>10 MHz).

4.3.3 Solar-controlled diurnal and seasonal variation of spread index  Spread index is higher during the night than during the day. At night, when recombination processes dominate, the ionospheric reflecting surfaces (iso-ionic contours) are less uniform than during daytime. This results in several independent propagation sub-modes between the transmitter and receiver, producing mutual interference (fading). This situation also produces Doppler spreading of the received signal as the rate of change of phase path length of each sub-mode will differ. A marked increase in spread index should thus occur between day and night. The solar controlled variation of spread index is greatest on lower frequencies, below ~10 MHz. Spread index is generally higher in winter than in summer.

The difference between the solar controlled variation on the lower frequencies and the oval controlled variation on the higher frequencies is most obvious on the Leicester path as the longitudinal extent of the path separates the two behaviours in time. On other paths the two effects are merged as the oval is most prominent during the night.

4.4 Signal level

As discussed in §4.2.1 two mechanisms are responsible for causing recognition of signals to fail. On higher frequencies rays can penetrate the ionosphere and consequently are not detected at the receiver. In this case the signal level measured is an indication of the background noise level and there is no contribution from the NONCENTRIC signal. On lower frequencies the rays are reflected to the receiver but are attenuated by the D region to such an extent that the signal to noise ratio is low and recognition fails. Despite having failed the signal recognition, the signal level gives information about the propagation conditions. Ignoring low frequency signals which failed the recognition tests would bias the data towards signals which experienced less absorption. Consequently, only correctly recognized signals are employed in determining mean signal levels on frequencies above 10.195 MHz; at 10.195 MHz and below, mean signal levels are determined from all signals, irrespective of signal recognition outcome.
Both mean and median signal level variations were determined. Both methods give nearly identical results. The mean of the decibel values of signal level are adopted, as opposed to the mean of the power. The latter method biases the results towards the higher signal levels and does not give a satisfactory indication of variations in the data.

4.4.1 Path dependence of signal level behaviour The diurnal variation of signal level and the variation of signal level with $K_p$ for each of the paths are described below. Figure 4.11 illustrates the mean diurnal variation of the signal level in two frequency bands for each path during both the summer and winter campaigns. The path to Aberdeen is omitted due to equipment problems and its similarity to the path to Leicester.

Alert In summer, on the lower frequencies ($<7$ MHz) the signal level follows a diurnal variation with a maximum at night, between 02-06 UT, and a minimum at noon, 17 UT. This is a consequence of an increase in non-deviative absorption due to photo-ionization of the D region during the day. At 9.941 and 10.195 MHz the decrease in signal level at noon becomes negligible due to the $f^2$ dependence of non-deviative absorption. At higher frequencies ($>10$ MHz) signals are received during the day only and have constant signal level. Signal level is independent of $K_p$, except at the highest frequencies (17.515 and 18.204 MHz) where signal level decreases slightly with $K_p$.

In winter there is very little diurnal variation in signal level as the path is in darkness at all times and there is no change in signal level with $K_p$.

Fairbanks In summer, on propagation frequencies below 10 MHz the signal level is a maximum at midnight, 08 UT, and a minimum at noon, 20 UT. Above 10 MHz there is little diurnal variation as non-deviative absorption is low. In winter there is little diurnal variation in signal level as the path is in darkness at all times except near noon, and there is little variation with $K_p$.

Boston In summer the diurnal variation of low frequency ($<10$ MHz) signal level follows the pattern expected due to solar illumination of the D region, with a minimum at noon, 17 UT, and a maximum during the night. Between 00-13 UT, from ~5 hours before midnight to ~8 hours after midnight, signal level decreases as $K_p$ increases. At other times there is little correlation between signal level and $K_p$. Higher frequencies ($>10$ MHz), received predominantly during the day only (10-00 UT) have constant signal level. A few signals are received at night (02-08 UT), associated with anomalous propagation events (see chapter 6). The mean signal level for night-time propagation is lower than the signal level of signals propagating during the day. Signal level decreases with increasing $K_p$ during the day.

During the winter campaign the diurnal variation of signal level on frequencies above and below 10 MHz follows the patterns observed in summer. The decrease in signal level associated with the daytime solar-illumination of the D region is narrower during the winter.
FIGURE 4.11  Mean diurnal variation (curve) and standard deviation (vertical bars) of signal level on the paths to Alert, Fairbanks, Boston, and Leicester during the summer (upper panels) and winter (lower panels) campaigns. Two frequency bands are indicated for each path and campaign. Arrows indicate local noon at the path midpoint.
campaign than during the summer as the period of solar illumination is shorter. The signal level of low frequency signals received between 00-11 UT (night) and high frequency signals received between 12-00 UT (day) show a negative correlation with $K_p$.

**Leicester** At frequencies below 10 MHz the diurnal variation of signal level is more complicated than the behaviour on the path to Boston due to the longitudinal extent of the Leicester path. The signal level is constant during the night and begins to increase at sunrise as more modes begin to propagate, due to the increase in F region electron density. The signal level then displays a daytime minimum which is a consequence of solar-controlled non-deviative absorption. After sunset the signal level decreases to the nighttime level as propagation modes are lost. This behaviour is best observed during the winter campaign. The maximum in signal level is greater at sunset than at sunrise due to the time lag of the F region electron density, which can remain high for several hours after sunset. Signal levels have a negative correlation with $K_p$ during the night. At higher frequencies the signal level is constant during the day, and has a negative correlation with $K_p$. The same behaviour is observed in winter.

**4.4.2 Solar and auroral control of absorption** On lower frequencies, generally below 10 MHz, the variation of signal level is predominately controlled by solar illumination of the D region and the consequent increase of non-deviative absorption during the day. In summer the absorption starts earlier and ends later than in the winter, reflecting the changes in the length of the day and the times of sunrise and sunset between seasons. If the absorption was solely dependent on solar illumination, signal levels should be independent of geomagnetic activity, but variations in signal level correlated with $K_p$ are observed, especially in the night and morning sectors where the drizzle and splash precipitation regions (chapter 2) are most predominant; this is due to enhancement of the D region by precipitation of energetic particles, auroral absorption. Consequently, a correlation is only observed on paths which are trans-auroral, *i.e.* Boston and Leicester (figure 4.12). Auroral absorption will be investigated in greater detail in chapter 5.

**4.4.3 Dependence of signal level on fading** Both high and low frequencies display a decrease in signal level with increasing geomagnetic activity. On lower frequencies (<10 MHz) on trans-auroral paths auroral absorption is the main contributing factor to this decrease. On higher frequencies, however, auroral absorption is negligible due to the frequency dependence ($f^2$) of non-deviative absorption. Possible mechanisms for the decrease of signal level in the absence of non-deviative absorption are fading and Doppler spreading of the signal (§2.4.6). Fading and Doppler spreading are indicated by a high spread index, and so a negative correlation between signal level and spread index is an indication of fading-related signal level decrease.

Signal level is negatively correlated with spread index on all paths at most times.
**Chapter 4**

Preliminary Analysis

**Summer campaign**

![Graphs showing mean variation and standard deviation of signal level as a function of Kp on the paths to Alert, Fairbanks, Boston, and Leicester during the summer campaigns for two frequency bands.]

**Winter campaign**

![Graphs showing mean variation and standard deviation of signal level as a function of Kp on the paths to Alert, Fairbanks, Boston, and Leicester during the winter campaigns for two frequency bands.]

**Figure 4.12** Mean variation (curve) and standard deviation (vertical bars) of signal level as a function of $K_p$ on the paths to Alert, Fairbanks, Boston, and Leicester during the summer (upper panels) and winter (lower panels) campaigns. Two frequency bands are indicated for each path and campaign.
However, as discussed above, on many paths signal level is negatively correlated with $K_p$ and spread index is positively correlated with $K_p$, usually due to the influence of the auroral oval. Consequently, in such cases, a negative correlation between signal level and spread index is expected and does not indicate any direct link between the two parameters. On the path to Alert during the winter campaign, however, where little diurnal variation of signal level is observed and spread index and signal level are independent of $K_p$ (top and middle panels of FIGURE 4.13), negative correlation between signal level and spread index is found (bottom panel of FIGURE 4.13). This is possibly a consequence of a) a decrease in signal level due to the nature of the fading present, or b) a decrease in the peak signal level as the signal power is distributed across a frequency range (broadening of the frequency spectrum) by Doppler spreading. However, assuming a triangular amplitude-frequency spectrum ($\S 3.1.7$) and constant signal power indicates that a peak signal level decrease of approximately 10 dB is expected for a spread index of 100 Hz. Consequently, mechanism b) cannot totally account for the relationship observed between signal level and spread index.

4.5 Summary

The measured characteristics of the propagation paths are as anticipated from the behaviour of the ionosphere: the propagation is predominantly solar-controlled, as would be expected for mid latitude paths. This is manifested as an increase in the MUF during the day, due to an increase in $f_pF2$ by solar illumination, and an increase in the LUF, due to an increase in the D region electron density, also produced by solar illumination. Fading causes an increase in the frequency spread of signals and this can cause a lowering of signal level. However, the level of geomagnetic activity plays an important role in modifying this behaviour due to precipitation in the auroral oval, and it is the effect on the low frequency and high frequency ends of the HF spectrum that the investigation will address in chapters 5 and 6 respectively.
FIGURE 4.13  Signal level (top panel) and spread index (middle panel) as a function of $K_p$ at 6.800 MHz, 19 UT, on the path to Boston during the winter campaign. Signal level as a function of spread index (bottom panel) for the same data set. Lines of best fit and coefficients of correlation are indicated in each panel.
Chapter 5
Radiowave absorption in the high latitude ionosphere

5.1 Introduction
Throughout the two experimental campaigns, the ionospheric electron density remains sufficiently high to allow radiowaves with frequencies below ~10 MHz to be propagated along the NONCENTRIC paths at all times of day. The received signal level does not remain constant, however, decreasing around local noon and during periods of geomagnetic activity. The following chapter describes two mechanisms controlling the attenuation of signal levels: the daytime photo-ionization of the D region and D region enhancement by auroral precipitation. The D region enhancement is modelled and the mechanism generating the precipitation, the magnetospheric substorm, is discussed. Its role in producing short-lived but intense absorption events is investigated in detail.

5.2 Daytime variation of HF signal level
The signal levels of radiowaves with frequencies below ~10 MHz decrease when the propagation path is illuminated during the day and remain generally constant during the night (FIGURE 4.2). The attenuation of the radiowaves is produced by non-deviative absorption, dependent on the D region electron density and collision frequency. The electron density is controlled primarily by solar illumination. However, the ion chemistry of the lower ionosphere and neutral atmosphere is complex and several factors influence the electron density height distribution and its diurnal and seasonal variations.

5.2.1 Solar control of D region electron density and HF signal level
The solar control of the D region electron density can be estimated from a brief overview of the main chemical processes occurring in the D region (RATCLIFFE, 1972). At sunrise, ion-electron pairs are predominantly produced by photo-detachment of O$_3^-$ ions. During the day photo-ionization, of NO by Lyman α radiation (1216 Å) and O$_2$ and N$_2$ by X-rays (<10 Å), becomes the primary production mechanism (DAVIES, 1990). The rate of production is proportional to the amount of solar radiation falling per unit area of the atmosphere, proportional to the cosine of the solar zenith angle, $\chi$. The electron production rate is then

$$ q = q_0 \cos \chi $$

where $q_0$ is the rate of production when the sun is directly overhead. In the lower ionosphere, where ions are predominantly molecular, the loss processes are dominated by dissociative recombination, the following reactions being the most important (DAVIES, 1990):
Chapter 5  Radiowave absorption

O\(_2^+ + e^- \rightarrow O + O\),  N\(_2^+ + e^- \rightarrow N + N\),  NO\(^+ + e^- \rightarrow N + O\).

The ion and electron loss rate is equal to the product of the mean dissociative recombination coefficient and the number of electrons and positive ions present, -\(\alpha N_e N_i^+\). Under conditions of charge neutrality \(N_i^+ = N_e^* = (1 + \lambda)N_e\), where \(\lambda\) is the ratio of negative ions to electrons, \(N_i^*/N_e\). Electrons are also lost by attachment with neutral molecules to form negative ions (e.g. \(O_2 + e^- + M \rightarrow O_2^- + M\), a three-body reaction (HARGREAVES, 1979)) at a rate \(-L\). The rate of change of electron density in the D region is then the sum of the production and loss rates:

\[
\frac{dN_e}{dt} = q - \alpha_e (1 + \lambda)N_e^2 - L. \quad [5.2]
\]

The negative ions themselves obey a continuity equation, relating their production, at rate \(L\), and their loss by ion-ion neutralization with positive ions (e.g. \(O_2^- + N_2^+ \rightarrow O_2 + N_2\)) at a rate \(-\alpha_i N_e^* N_i^-\)

\[
\frac{dN_i^-}{dt} = \lambda \frac{dN_e}{dt} = L - \alpha_i \lambda (1 + \lambda)N_e^2. \quad [5.3]
\]

Combining [5.2] and [5.3] to give the total continuity equation yields

\[
\frac{dN_e}{dt} = \frac{q}{(1 + \lambda)} - (\alpha_e + \lambda \alpha_i)N_e^2. \quad [5.4]
\]

The rates of electron production and loss are rapid in comparison to the rate of change of \(\cos \chi\) and consequently quasi-equilibrium is achieved, \(dN_e/dt \approx 0\), and the electron density is given by

\[
N_e = \left(\frac{q}{(\alpha_e + \lambda \alpha_i)(1 + \lambda)}\right)^{1/2}. \quad [5.5]
\]

Generally, \(\lambda \alpha_i < < \alpha_e\) (HARGREAVES, 1979) and [5.5] simplifies to

\[
N_e = \left(\frac{q}{(1 + \lambda)\alpha_e}\right)^{1/2}. \quad [5.6]
\]

If it is assumed that \(\lambda\) varies significantly at sunrise and sunset only and \(\alpha_e\) and \(\alpha_i\) remain generally constant during the day, then from [5.1] and [5.5] the daytime D region electron density is predicted to be proportional to \(\cos^{1/2} \chi\) (CHAPMAN, 1931). In general, however, it is found that the electron density varies as \(\cos^n \chi\) where 0.2 < \(n < 0.9\) (DAVIES, 1990) due to diurnal variations in \(\alpha_e\), \(\alpha_i\), and \(\lambda\). At night no photo-ionization takes place and the D region electron density is maintained at a low, fairly constant level by cosmic radiation diverted from the dayside by the terrestrial magnetic field (RATCLIFFE, 1972), or a rapidly varying level by auroral precipitation.

An HF radio wave passing through the D region is attenuated by an amount proportional to the product of the D region electron density and the collision frequency (§2.4.5). The collision frequency profile does not, in general, vary appreciably, except during artificial modification of the ionosphere by high-power radiowave heating (e.g. DAVIES, 1990) and possibly auroral precipitation when the electron temperature is perturbed.

Signal strength on the path to Boston displays a dependence on the cosine of the solar
zenith angle when the ionosphere is illuminated (cos $\chi > 0$), consistent with a $\cos^n\chi$ dependence of D region electron density. Figures 5.1a and b include scatter plots of signal level as a function of $\cos \chi$ determined at the path midpoint, for the summer and winter campaigns respectively. The left panels illustrate the morning hours, 00-11 LT, the middle panels the afternoon hours, 12-23 LT, and the right panels indicate the mean variation for the morning (dashed curve) and afternoon (solid curve). Vertical dashed lines indicate $\cos \chi = 0$, corresponding to sunrise and sunset. When the D region is in darkness ($\cos \chi < 0$) the signal level remains generally constant; this is especially noticeable during the winter campaign when the path is in darkness for most of the day.

The value of the exponent $n$ can be determined by plotting signal level as a function of $\cos^n\chi$ ($\chi < 90^\circ$) for several values of $n$ ($0.15 < n < 1.5$) and determining the coefficient of correlation, $r$, with a linear fit (see Figure 5.2). This method yields a best fit value for $n$, though in general $r$ is high over a large range of $n$ (in Figure 5.2 $r = -0.98$ for $0.30 < n \leq 0.75$), and few conclusions can be drawn.

### 5.2.2 Comparison of low and high latitude signal level variation

Workers at middle and low latitudes (e.g. SCHWENTEK, 1969; VICE, 1969; LASTOVICKA, 1977) have noted a discrepancy between morning and afternoon signal levels: absorption lags solar zenith angle by up to 30 minutes,

$$L(t) = A \cos^n\chi(t - \tau) \tag{5.7}$$

where $\tau$ is the lag time, $0 < \tau < 30$ mins, though SCHWENTEK (1969) reports that $\tau = 0$ on 65% of days. This lag between the electron production maximum and electron density maximum is predicted by APPLETON (1953) as a consequence of the finite rates of electron production and recombination. The predicted relaxation time lag is

$$\tau = \frac{1}{2\alpha N_e}, \tag{5.8}$$

where $\alpha$ is the recombination rate and $N_e$ is the electron density; typical values of $N_e$ and $\alpha$ give $\tau \sim 20$ minutes (APPLETON and PIGGOTT, 1954). Accordingly the electron density, and hence absorption, is greater in the afternoon than in the morning for the same value of $\chi$.

This is not the case in the present study. The mean signal level measured in the morning sector is lower than that in the evening sector at corresponding solar zenith angles, especially for trans-auroral paths (right hand panels of Figures 5.1a and b). Also, comparison of the left and middle panels of Figures 5.1a and b indicate that a population of low signal level measurements exists in the post-midnight sector that is not present in the evening sector (see especially 6.800 and 9.941 MHz in Figure 5.1b). The reduced signal level in the morning sector is a consequence of absorption of the signals in the auroral zone, and this feature will be investigated in the rest of this chapter.
**Chapter 5**  
Radiowave absorption

**Summer campaign**  
**Boston**

**Figure 5.1a**  
Signal level as a function of \( \cos \chi \) (determined at the path midpoint) for four frequencies during the period of the summer campaign on the path to Boston. Left and middle panels indicate individual signal level measurements for the morning (00-11 LT) and afternoon (12-23 LT) periods respectively. The right panel indicates the mean signal level variation for the morning (dashed line) and afternoon (solid line). The vertical dashed lines at \( \cos \chi = 0 \) indicate sunrise and sunset.
Chapter 5  Radiowave absorption

Winter campaign  Boston

FIGURE 5.1b  Signal level as a function of $\cos \chi$ (determined at the path midpoint) for four frequencies during the period of the winter campaign on the path to Boston. See FIGURE 5.1a for further description.
Summer campaign  Boston

**Figure 5.2** Mean signal level at 6.800 MHz during the summer campaign on the path to Boston as a function of $\cos^2 \chi$ for $0.15 \leq n \leq 1.35$, $\chi < 90^\circ$ (i.e. between sunrise and sunset). Morning and afternoon signal levels are indicated by dashed and solid lines respectively. The coefficient of correlation is indicated in the top right hand corner of each panel.
Chapter 5

Radiowave absorption

5.3 Auroral absorption

Enhanced absorption is observed on HF signals propagating in the high latitude region. Auroral absorption occurs predominantly in the midnight and morning sectors at the equatorward edge of the auroral oval (e.g. HARTZ et al., 1963 (see FIGURE 2.5)) and is attributed to D region enhancement by high energy (>30 keV) electron precipitation, as inferred from satellite-borne particle detectors (e.g. HARTZ and BRICE, 1967; CRAVEN, 1970). HARTZ and BRICE (1967) divided the precipitation into two broad categories: drizzle precipitation occurring predominantly in the morning sector and splash precipitation occurring in the midnight sector (see chapter 2).

The contribution of auroral absorption to the diurnal variation of signal level is investigated by means of the data collected on the trans-auroral propagation path to Boston. The meridional Boston path is selected in preference to the longitudinally extended Leicester path to avoid ambiguity in determining local time. The frequency 6.800 MHz is chosen for the investigation because it propagates for long periods of the day and moreover is relatively free of interference from external sources. All other frequencies below 10 MHz propagating on the path to Boston are dominated by interference for some or all of the day.

5.3.1 Geomagnetic dependence of auroral absorption

FIGURE 5.3 illustrates the mean diurnal variation of signal level at 6.800 MHz for different levels of geomagnetic activity, the upper panel for the summer campaign and the lower panel for the winter campaign. The abscissa spans the day from midday to midday; local midnight is indicated by the vertical line. The curves are generally symmetrical about local midnight, decreasing towards local noon due to the solar controlled absorption. During the night (01-08 UT in summer, 22-11 UT in winter) the mean signal level decreases with increasing geomagnetic activity. The decrease in signal level with increasing $K_p$ is more pronounced in the post-midnight sector than in the pre-midnight sector and at geomagnetic activities of $K_p \geq 3$. During the winter campaign, the absorption maximum appears to move from an earlier (06 UT, $K_p = 3$) to a later (07-08 UT, $K_p = 4$) time with increasing geomagnetic activity. During the winter campaign, the absorption for $K_p = 5$ decreases markedly between 07-09 UT contradicting the trend for $K_p < 5$; this feature is examined in more detail later and is shown to be spurious.

5.3.2 The spatial and temporal pattern of auroral absorption

The spatial pattern of auroral absorption has been empirically modelled by FOPPIANO and BRADLEY (1983), the modelled parameter being $Q_0$, the percentage occurrence of >1 dB absorption measured on a 30 MHz riometer. In this scheme two overlapping regions of absorption are attributed to splash and drizzle precipitation, the magnitude and spatial extent of the absorption regions being related to geomagnetic activity, quantified by $K_p$. FIGURE 5.4 illustrates the location of the splash precipitation (panel a) and drizzle precipitation (panel b) regions for $K_p = 0$, and the sum of the contributions from the two regions for four levels of $K_p$ (panels c to f).
FIGURE 5.3 Mean signal level at 6.800 MHz on the path to Boston as a function of UT for different levels of $K_p$, during the summer campaign (top panel) and winter campaign (bottom panel). The abscissa spans the day from local noon to local noon at Boston; the vertical line indicates local midnight.
Chapter 5  Radiowave absorption

Figure 5.4 The FOPPIANO and BRADLEY (1983) model of $Q_1$, the percentage occurrence of >1 dB auroral absorption on a 30 MHz riometer. Panels $a$ and $b$ illustrate the contributions from the splash and drizzle precipitation regions respectively, for $K_p = 0$. Panels $c$ to $f$ illustrate the sum of the splash and drizzle region contributions for four levels of $K_p$. The position of the FELDSTEIN oval is indicated for $K_p \leq 4$. The location of the Boston propagation path at 12 UT and the 0.25 and 0.75 ground range points are indicated for reference. The coordinates are geomagnetic latitude and MLT.
determined from the FOPPIANO and BRADLEY model. The locations of the Feldstein oval and the Boston propagation path for 12 UT are included for comparison.

The predicted diurnal variation of auroral absorption on the Boston path was determined from this model for different levels of $K_p$, assuming 1-hop propagation and consequently two passes of the D region by the radiowave, at approximately 0.25 and 0.75 of the ground range. FIGURE 5.5 illustrates the sum of the absorption probability from the two D region passes with contributions from both splash and drizzle precipitation (solid curves). The dotted curves with maximum at 06-07 and 11 UT indicate the contributions from the splash and drizzle regions respectively for $K_p = 0$. FOPPIANO and BRADLEY argue that median absorption $A_n$ (dB) and $Q_i$ (%) are related by $A_n \sim 0.02 Q_i$, though many different relationships have been formulated, each with large differences in predicted $A_n$ (FOPPIANO and BRADLEY, 1983). An additional factor $(30/6.8)^2 \sec \phi_0 = 60$ (see equations [2.13] to [2.15]) transforms from a 30 MHz vertically propagating radiowave, on which the FOPPIANO and BRADLEY model is based, to a 6.800 MHz oblique radiowave propagating from Clyde River to Boston. In general the auroral absorption increases with increasing geomagnetic activity and is greatest in the post-midnight sector, consistent with the Non-centric observations illustrated in FIGURE 5.3. The maximum in absorption moves to later local times as $K_p$ increases, due to the equatorwards motion of the drizzle precipitation region relative to the path D region incidence points as geomagnetic activity increases. This motion also causes the predicted absorption to be greater for $K_p = 0$ than for $K_p = 1$ or 2 in the post-midnight sector, despite a monotonic increase in the level of absorption in the splash and drizzle precipitation regions with increasing $K_p$. Both the movement of the absorption maximum to later times with increasing $K_p$ and the similarity of the $K_p = 1$ and $K_p = 2$ absorption levels are consistent with the Non-centric observations, especially during the winter campaign.

The close correspondence between observations and model predictions emphasises the value of the splash and drizzle precipitation region paradigm and this will be adopted again later in this study.

5.3.3 Comparison of $K_p$ and $D_n$ as indicators of absorption level Geomagnetic activity is quantified by a variety of indices, e.g. $K_p$, $D_n$, and $AE$, each based on a different aspect of ionospheric-magneto-spheric coupling to indicate the level of disturbance on a local or a global scale. $AE$ measures the magnitude of the auroral electrojet on timescales of minutes to 1 hour, $K_p$ indicates the range of variation of magnetic fluctuations at mid latitudes in a three hour period, and $D_n$ measures the magnitude of the ring current averaged over a one hour period. A comparison was conducted between $K_p$ and $D_n$ (at the time of writing $AE$ is unavailable for the period of the experiment) to determine the index most suitable for prediction of the level of auroral absorption under different geomagnetic conditions.

The winter campaign signal level data (6.800 MHz, Boston path) were collected into 24 bins (00-01 UT, 01-02 UT, etc.) and in each bin a least squares fit was calculated to the two
FIGURE 5.5 The diurnal variation of $Q_1$ determined for the Boston path from the FOPPIANO and BRADLEY model for 6 levels of $K_p$ (solid lines). The contributions from the splash and drizzle precipitation regions for $K_p = 0$ are also indicated (dashed lines). The vertical dashed line indicates local midnight.
geomagnetic indices. The coefficient of correlation, $r$, and the gradient of the fit, the rate of change of signal level with each index, $m$, are plotted as a function of UT in Figure 5.6, the upper panel for $r$ and the lower panel for $m$. As $K_p$ increases and $D_s$ decreases with increasing geomagnetic activity $r$ and $m$ are generally of opposite sign for the two indices. Consequently, $m_{Dst}$, $r_{Dst}$ and $-m_{Kp}$, $-r_{Kp}$ are plotted to aid comparison. The coefficient of correlation is high ($-0.7$) throughout the night ($01-11$ UT). During the day the correlation is much lower, due to domination of D region production by solar photo-ionization which is independent of geomagnetic activity. The gradient of the least squares fit is greatest at night also, though it is greater in the post-midnight sector ($06-11$ UT) than the pre-midnight sector, indicating greater levels of absorption in the drizzle precipitation region than in the splash precipitation region.

The correlation between the temporal variations of $r$ calculated for each index and $m$ for each index is striking. This suggests that both indices, though measuring different aspects of the magnetospheric-ionospheric coupling system, react in similar ways to increasing geomagnetic activity. Indeed, the coefficient of correlation between $K_p$ and $D_s$ is -0.64 for the winter campaign (Figure 5.7); however, the correlation is only -0.48 during the less disturbed summer campaign.

The correlation between $D_s$ and absorption might be expected to be better than between absorption and $K_p$ due to its higher temporal resolution. However, the ring current, measured by $D_s$, predominantly comprises protons since electron and proton bounce rates in the ring current are in the ratio $(m_1/m_2)^{1/2} \approx 40$ (Parks, 1991) and consequently electrons are lost rapidly to the ionosphere after injection from the magnetotail. $D_s$ is, then, an indicator of a slowly varying trend in the number of protons in the ring current as opposed to instantaneous changes in the electron precipitation rate. Electron injections are indicated by small decreases in $D_s$ but these are superimposed on the much larger background proton contribution.

$K_p$ has a temporal resolution of three hours, which is similar to the mean recurrence period of substorms of $\sim 2.4$ hours (Borovsky et al., 1993) or $2-4$ hours (Lui, 1991). Consequently each time bin is approximately associated with an individual substorm and perhaps indicates the level of disturbance produced by this substorm. However, it is a poor indicator of the absorption variation within a three hour period, as the duration of the absorption associated with a substorm is expected to be $\sim 30-60$ mins (Akasofu, 1968; see also §5.5.3 and §5.6).

Consequently, neither $K_p$ nor $D_s$ are ideally suited to predict auroral absorption on short timescales as neither is a direct measure of the level of precipitation producing the D region enhancement. The maximum correlation, $-0.7$, determined between both indices and absorption level probably places an upper limit on the accuracy of auroral absorption prediction employing these indices.
Chapter 5  Radiowave absorption

Figure 5.6 The diurnal variation of the coefficient of correlation (panel a) and the gradient of the line of best fit (panel b) between signal level and $K_p$ (solid line) and $D_n$ (dashed line), for the winter campaign.
Chapter 5  
Radiowave absorption

**Figure 5.7** $D_a$ as a function of $K_p$ for the summer campaign (top panel) and winter campaign (bottom panel). The line of best fit is indicated by the dashed line and the coefficient of correlation is indicated in the bottom left hand corner of each panel.
5.4 Model of the energy flux of >30 keV electron precipitation in the auroral zone

The level of auroral absorption due to precipitating electrons increases with geomagnetic activity, and consequently it is expected that the flux of precipitating electrons must increase during disturbed periods also. An empirical model of the energy and number fluxes of precipitating electrons (HARDY et al. 1985; 1987) was extended in order to model the pattern of high energy electron flux over the high latitude region and to determine the variation of this flux with geomagnetic activity.

5.4.1 The HARDY model of auroral precipitation Three satellites were employed in the construction of the HARDY et al. (1985) model: F2 and F4 from the Defence Meteorological Satellite Program (DMSP), and P78-1 from the Space Test Program. F2 and F4 had circular sun-synchronous orbits at an altitude of 840 km, F4 at an inclination of 97.4° in the 10-22 LT meridian and F2 in the dawn-dusk meridian, though this precessed toward the 08-20 LT meridian in its 2.5 year lifetime. The P78-1 satellite was in a 600 km circular sun-synchronous orbit in the noon-midnight meridian. 15 months of data from F2 and F4 and 12 months of data from P78-1 were employed to give an even coverage of season and geomagnetic activity; the data were collected between 1977 and 1980. The satellite electron detectors comprised 16 energy channels which produced energy spectra in the range 50 eV to 20 keV. Average total number and energy fluxes, \( n_T \) and \( E_T \), were binned in MLT, CGMLAT and geomagnetic activity (\( K_p \)). The latitudinal variation of \( n_T \) and \( E_T \) across the auroral oval were represented by an Epstein transition function, the coefficients of which varied slowly in MLT and so were in turn represented by Fourier series to reduce the number of coefficients needed. \( n_T \) and \( E_T \) (el cm\(^{-2} \) s\(^{-1} \) sr\(^{-1} \) and eV cm\(^{-2} \) s\(^{-1} \) sr\(^{-1} \) respectively) were assigned minimum values of \( 10^7 \) and \( 10^6 \) poleward and equatorward of the oval respectively. FIGURE 5.8 illustrates the variation of \( n_T \) (panels a, c, e) and \( E_T \) (panels b, d, f) as functions of MLT and CGMLAT for three levels of \( K_p \). The positions of the Boston propagation path at 12 UT and the FELDSTEIN oval are also indicated for reference.

5.4.2 Extension to the HARDY model The minimum energy of precipitating electrons necessary to produce ionization at D region altitudes, \( E_{\text{min}} \), is \(~30 \text{ keV} \) (REES, 1963; BERGER and SELTZER, 1970). The greatest correlation between riometer measurements of auroral absorption and observed electron fluxes is achieved for energies of 40-80 keV which ionize the D region in the altitude range 85-95 km (COLLIS and HARGREAVES, 1980). The number of precipitating electrons with \( E > E_{\text{min}} \) can be estimated from a first approximation energy spectrum and \( E_T \) and \( n_T \) given by the HARDY model and provides an indication of the location of D region enhancement produced by the precipitation for different levels of \( K_p \).

Assuming a Maxwellian energy spectrum for the precipitation, the number of electrons precipitating with energies in the range \( E \) to \( E+\text{d}E \) is
FIGURE 5.8  The total electron number flux, $n_T$, (panels a, c, and e) and total electron energy flux, $E_T$, (panels b, d, and f) determined from the HARDY et al. (1985; 1987) model, for three levels of $K_p$. The contours indicate the $\log_{10}n_T$ and $\log_{10}E_T$ flux levels. $n_T$ and $E_T$ have units of $\text{el cm}^{-2} \text{ s}^{-1}$ and $\text{eV cm}^{-2} \text{ s}^{-1}$ respectively. The locations of the FELDSTEIN oval and the Boston path at 12 UT are also indicated. The coordinates are geomagnetic latitude and MLT.
Chapter 5  Radiowave absorption  50

\[ n(E) dE = n_0 E e^{-E/E_0} dE \]  [5.9]

where \( n_0 \) and \( E_0 \) are constants defining the spectrum (BAROUCH, 1981; SHARBER, 1981).

Integrating over the total energy of the spectrum gives

\[ n_\tau = \int_0^\infty n(E) dE = n_0 E_\tau^3, \quad E_\tau = \int_0^\infty n(E) E dE = 2n_0 E_0^3 \]  [5.10, 5.11]

and after rearrangement to find \( n_0 \) and \( E_0 \) in terms of \( n_\tau \) and \( E_\tau \):

\[ n_0 = \frac{4n_\tau^3}{E_\tau^4}, \quad E_0 = \frac{E_\tau}{2n_\tau}. \]  [5.12, 5.13]

Values of \( n_\tau \) and \( E_\tau \) obtained from the HARDY model allow the spectrum to be characterised in terms of \( n_0 \) and \( E_0 \). Only electrons with \( E > E_{\text{min}} \) enhance the D region, so integrating over the high energy tail of the Maxwellian spectrum [5.9] gives the energy flux into the D region, \( E_{\text{D}} \):

\[ E_{\text{D}} = \int n(E) E dE = n_0 E_0 \left[ E_{\text{max}}^3 + 2E_0 E_{\text{min}} + 2E_0^3 \right] e^{-E_{\text{D}}/E_0}. \]  [5.14]

This is only a first approximation as the electron detectors aboard the satellite operate in the range 50 eV to 20 keV only, and thus \( n_\tau \) and \( E_\tau \) are not the total number and energy fluxes but are the contributions from electrons with energies in this range. Consequently [5.10] to [5.13] are only approximate. A more accurate estimation of the true total fluxes can be found by integrating a Maxwellian spectrum between \( E_1 \) and \( E_2 \), the minimum and maximum detected energies respectively.

\[ E_{\text{D}} = \int_{E_1}^{E_2} n(E) E dE = -n_0 E_0 \left[ \left( E_1^2 + 2E_0 E_1 + 2E_0^2 \right) e^{-E_1/E_0} - \left( E_2^3 + 2E_0 E_2 + 2E_0^3 \right) e^{-E_2/E_0} \right]. \]  [5.15]

Equations [5.15] and [5.16] are not easy to solve for \( n_0 \) and \( E_0 \) but rearrangement gives

\[ E_{\text{D}} = \frac{E_{\text{D}}}{2n_\tau} + \frac{n_\tau E_0}{2n_\tau} \left( E_1^2 e^{-E_1/E_0} - E_2^3 e^{-E_2/E_0} \right), \]  [5.17]

\[ n_\tau = -\frac{E_{\text{D}}}{E_0} \left[ \left( E_1 + E_0 \right) e^{-E_1/E_0} - \left( E_2 + E_0 \right) e^{-E_2/E_0} \right]. \]  [5.18]

Equations [5.12] and [5.13] provide initial estimates of \( n_0 \) and \( E_0 \) and then [5.17] and [5.18] are iterated to yield more accurate values. Generally 5 iterations are sufficient for \( n_0 \) and \( E_0 \) to converge. Equations [5.12] and [5.13] underestimate the flux level obtained by iteration of [5.17] and [5.18] in the cusp region by a factor of ~6; otherwise the approximation holds quite accurately. This indicates that the Maxwellian energy spectrum approximation is not valid in the cusp region (see §5.4.3).

It must be noted that the results of the model are an extrapolation of the DMSP observations as the maximum energy detected by the satellite and the minimum energy required for D region enhancement are 20 keV and 30 keV respectively.

5.4.3 Model results  FIGURE 5.9 illustrates the energy precipitation into the D region for
FIGURE 5.9  The total energy of electrons precipitating into the D region for 5 levels of $K_p$, determined from the extension to the HARDY model. The contours indicate the $10^1$, $10^3$, $10^5$, $10^7$, and $10^9$ eV cm$^{-2}$ s$^{-1}$ flux levels. The coordinates are geomagnetic latitude and MLT.
different levels of $K_p$ estimated by the extended HARDY model. The location of the FELDSTEIN oval and the Boston propagation path at 12 UT are also illustrated. A band of precipitation is coincident with the equatorward edge of the morning sector auroral oval, with little precipitation occurring in the evening sector, consistent with the splash and drizzle precipitation zone paradigm of HARTZ and BRICE (1967) (c.f. HARTZ et al. (1963) riometer observations in FIGURE 2.5 and FOPPIANO and BRADLEY model in FIGURE 5.4) and the average location of $>$30 keV precipitation determined by CRAVEN (1970). The precipitation pattern moves equatorward by several degrees with increasing $K_p$, consistent with the motion of the FELDSTEIN oval. At $K_p \geq 3$ a precipitation feature appears in the pre-midnight sector, consistent with the expected location of the westward travelling surge or WTS (e.g. AKASOFU, 1968), indicating that as $K_p$ increases the precipitation in the WTS becomes more statistically significant.

There are generally two maxima in the energy flux, one at noon corresponding to an influx from the magnetospheric cusp, and a second due to drizzle precipitation near 06 MLT. The cusp maximum is not observed by HARTZ and BRICE (1967), and indeed the cusp is a region of low energy (<2 keV) electrons which produce F region, not D region, enhancement (DEEHR et al., 1980). The cusp maximum is an artefact of the model, produced by an inaccuracy of the Maxwellian energy distribution approximation in the cusp region. HARDY et al. (1985) note that the electron energy spectrum softens and becomes non-Maxwellian at noon.

FIGURE 5.10 illustrates the magnitude of the drizzle precipitation energy flux maximum as a function of $K_p$. The energy flux level increases with increasing $K_p$ and consequently produces a greater enhancement of the D region electron density. The increase in mean absorption with increasing $K_p$ (identified in §5.3.1) is, then, a consequence of increasing mean precipitation energy flux into the D region during disturbed periods.

5.5 Relationship between substorm occurrence and HF absorption; substorm correlated absorption bursts (SCABs)

The patterns of auroral absorption and particle precipitation determined by workers such as HARTZ et al. (1963), HARTZ and BRICE (1967), CRAVEN (1970), FOPPIANO and BRADLEY (1983) and HARDY et al. (1985) are temporal averages of a highly dynamic feature. Auroral absorption can vary greatly over distances of ~100 km and within timescales of ~10 minutes (REID, 1967). This variability is emphasised by FIGURE 5.11 which illustrates the developing spatial pattern of electron precipitation following a substorm expansion phase onset assuming two representative electron energies: 5 keV, responsible for luminous features, and 50 keV, responsible for auroral absorption (AKASOFU, 1968). After the onset the midnight sector auroral oval thickens to form the auroral bulge and the westward travelling surge appears. These features have an expected lifetime of ~15 mins. The region of 50 keV electron precipitation expands eastwards to
FIGURE 5.10  The maximum total electron energy flux into the D region in the morning sector as a function of $K_p$, determined from the extension to the HARDY model.
FIGURE 5.11 Four schematics of the typical development of the 5 keV and 50 keV electron precipitation patterns in the hour following a substorm expansion phase onset ($t = 0$ min). The panels represent the periods $t = 0-5$ min, $t = 5-10$ min, $t = 10-30$ min, and $t = 30-60$ min. (After AKASOFU (1968)).
extend over the morning sector auroral oval, this feature having an expected lifetime of 30 mins to 1 hour. Consequently, the occurrence of auroral absorption will be linked to the occurrence of individual substorms, and the lifetime of the absorption will be short in comparison to the timescales of geomagnetic indices such as $K_p$.

To determine the relationship between substorm occurrence and auroral absorption, the occurrence times of substorms were determined for the winter campaign and compared with NONCENTRIC signal level variations.

5.5.1 Determination of substorm occurrence

Two methods, both involving geosynchronous satellite data, were employed to identify the onset of substorm expansion phases during the winter campaign. The first is based on changes in the terrestrial magnetic field and the second on the trapped particle fluxes.

Field dipolarization: GOES-7 satellite data

The GOES-7 satellite is geosynchronous over the west coast of the USA (~265° E). Onboard magnetometers provide the components $H_p$ (parallel to the satellite spin axis, approximately parallel to the earth’s rotation axis), $H_e$ (eastward) and $H_s$ (earthward) of the terrestrial field. From the values of $H_p$ and $H_e$, the angle of the field to the earth-satellite line in the N-S meridian, $\theta = \arctan(H_s/H_e)$, was determined. This parameter provides information regarding changes in the configuration of the magnetosphere under the influence of the solar wind and IMF. The magnetosphere deviates from the dipole approximation, being compressed on the dayside by the solar wind pressure and extending to form the magnetotail on the nightside. Consequently measured on the dayside is typically higher than the value calculated for a dipolar field (~68°, determined from the dip angle given by $\tan I = 2 \tan \Phi$, where $\Phi$ is the geomagnetic latitude of the spacecraft) and is lower on the nightside. During a substorm growth phase the magnetotail becomes more extended as the plasma sheet is compressed by the accumulation of open flux in the magnetotail lobes and consequently $\theta$ measured on the nightside decreases (e.g. BAKER et al., 1981; LUI, 1991). At the onset of the substorm expansion phase tail reconnection occurs, the magnetotail relaxes to a more dipolar configuration and $\theta$ rapidly increases (NAGAI, 1982; LUI, 1991). This effect is known as a dipolarization. Substorms can only be detected in this way when the satellite is within the region of the terrestrial field that is influenced by changes in the magnetotail, i.e. within several hours of local midnight. However, the GOES-7 satellite is sufficiently close to the longitude of the Boston path that it is a good indicator of the occurrence of substorms when the path is within the substorm injection region.

Trapped particle fluxes: LANL satellite data

Data from the Charged Particle Analyzers (CPAs) of three LANL geosynchronous satellites (1982-019, 1987-097, 1984-129) provided an estimate of the fluxes of trapped electrons (30-300 keV) and protons (80-600 keV) at 6.6 $R_p$. Sudden increases of an order of magnitude in the fluxes of electrons and protons are
taken to indicate a substorm particle injection. The particle flux increase is often preceded by a short-lived decrease corresponding to \textit{plasma sheet thinning} and the movement of the low density plasma mantle into the vicinity of the satellite (e.g. \textsc{Sa}u\textsc{va}d and \textsc{W}inckler, 1980). The flux increase corresponds approximately to the onset of the substorm expansion phase, but uncertainties arise due to the longitudinal drift time of the particles from the midnight sector injection region to the satellites. The satellites are arranged at approximately equidistant points around the equator (the satellites are at approximate longitudes of 70°E, 195°E, 325°E), however, to ensure as good coverage as possible. Expansion phase onset times have been determined to the nearest 30 mins.

\textbf{Figure 5.12} illustrates the variation of $\theta$ measured at the GOES-7 satellite and the fluxes determined from the LANL data for 24 and 23 January 1989. 00-12 UT 24 January is a quiet period; the field angle is 10-20° lower than the dipole approximation within the magnetotail and the fluxes observed at the three LANL satellites remain approximately constant. Three substorms occur between 12-24 UT. During this period the GOES-7 satellite is on the dayside and is not affected by changes in the magnetotail and the field angle measured approaches the dipole approximation. In the LANL data, however, three large flux increases are observed, almost simultaneously at each spacecraft. The increase in flux is largest at the first spacecraft to the east of local midnight. The delay between observation at the other spacecraft is a consequence of the eastwards longitudinal drift time of the injected electrons. The approximate times of these flux injections are indicated in the field angle panel by arrows. The 23 January is a more disturbed period, and five substorms are identified in the LANL data. Two of these substorms occur when GOES-7 is within the magnetotail, at ~0330 UT and ~0800 UT. A gradual decrease is observed in $\theta$ during the substorm growth phase. This is followed by an abrupt return to large $\theta$, the dipolarization, at the expansion phase onset. The initial dipolarization at 0330 UT is rapidly followed by a second smaller decrease and then increase in $\theta$, possibly due to an intensification of the substorm or the onset of a second expansion phase.

\textbf{Figure 5.13} indicates the occurrence of substorms for the winter campaign as a function of UT and the methods by which each was detected: a field dipolarization measured by GOES-7, a particle flux increase measured by the LANL spacecraft, or simultaneous observations of both. The detection of substorms by the GOES-7 satellite is concentrated between 01 and 08 UT as it is during this time that the spacecraft is within the magnetotail. The detection rate of GOES-7 decreases towards the dawn and dusk sectors where the magnetic field is less affected by the extension and dipolarization of the tail. The LANL spacecraft achieve 360° longitudinal coverage. 45 substorms were identified with the GOES-7 data and 131 with the LANL data; in 38 cases the simultaneous observation of LANL particle injections and a GOES-7 dipolarization were attributed to an individual substorm. A total of 138 substorms were identified during the 24 day campaign, giving a mean time between substorm onsets of ~4.2 hours. In the region where the coverage of the GOES-7 and
FIGURE 5.12  The field angle $\theta$ determined at the GOES-7 satellite (top two panels) and the electron fluxes measured at the three LANL spacecraft (bottom six panels) as a function of UT for the 24 and 23 January 1989. The horizontal dashed lines in the two top panels indicate the dipole approximation to $\theta$; arrows indicate the times of particle injections determined from the LANL data. The local time of each LANL satellite is indicated at the top of each LANL panel. The top left panel includes an inset schematic illustrating the angle $\theta$ and a dipolar (solid) and stretched (dashed) field line. The top right panel includes definitions of the dipolarization parameters $\theta_{\text{min}}$ and $\Delta \theta$. 
FIGURE 5.13 The diurnal variation of substorm occurrence for the winter campaign and the method of identification: a GOES-7 dipolarization, a LANL particle injection, or both. The vertical dashed line indicates local noon at the Boston path.
5.5.2 Correlation between substorm occurrence and auroral absorption

Figure 5.14 illustrates the signal level measured at 6.800 MHz on the path to Boston on 24 and 23 January 1989 (panels a and c) and also the variation in the magnetic field angle measured by the GOES-7 satellite and the times of particle injections determined from the LANL data (panels b and d). No substorms occur on 24 January during the period that the Boston path is in darkness (23-12 UT) and the signal level measured on this path follows the typical quiet-time diurnal variation. On 23 January three substorms occur when the Boston path is within the midnight and morning sectors, at 0330 UT, 0800 UT, and 1100 UT. Occurring simultaneously with these substorms are decreases in NONCENTRIC signal level of ~20 dB. During the day (13-22 UT) ~40 dB absorption produced by solar illumination of the D region dominates over the ~20 dB substorm correlated absorption and consequently it is impossible to determine the relationship between substorm occurrence and absorption in this time sector.

Twenty-three such substorm correlated absorption bursts (SCABs) were identified during the winter campaign, summarised in Table 5.1. Several potential events were rejected as they could not be positively identified as absorption related to substorm occurrence; for instance, at 19 UT on 23 January a decrease in signal level appears correlated with the occurrence of a substorm. However, this is not identified as a SCAB as comparison with the daytime signal level variation on 24 January reveals that the signal level at 18 UT 23 January is abnormally high, and at 19 UT no absorption has occurred.

5.5.3 Local time distribution of SCAB occurrence

Figure 5.15 illustrates the diurnal variation of SCAB occurrence throughout the winter campaign, the number of substorms with associated absorption events compared to the number of substorms during which absorption is not observed, plotted as a function of UT. Also indicated is the diurnal variation of the proportion of substorms with correlated absorption events. The absorption events occur between 01 UT (~20 LT, approximately the maximum westward extent of the westward travelling surge (AKASOFU, 1968) and the splash precipitation region (see Figure 5.4)) and 12 UT (sunrise). Between sunrise and sunset (12-21 UT, see Figure 5.15) SCABs cannot be distinguished from the solar-controlled diurnal variation in signal level. No absorption events are found to be associated with substorms which occur when the Boston path is between sunset and the region associated with the westward travelling surge and splash precipitation region (between 21 and 01 UT).
FIGURE 5.14  The diurnal signal level variation at 6.800 MHz on the path to Boston (panels a and c) and the GOES-7 magnetic field angle \( \theta \) (panels b and d) for 24 (panels a and b) and 23 (panels c and d) January 1989. Arrows in panels b and d indicate the times of particle injections identified in the LANL data. Vertical dashed lines indicate the correspondence between substorm occurrence and decreased signal level.
Chapter 5

Radiowave absorption

Figure 5.15: The diurnal variation in the occurrence of substorms with and without associated SCABs observed on the Boston path during the winter campaign. The dark line indicates the diurnal variation in the percentage occurrence of substorms with observed SCABs. Local midnight occurs at 05 UT. Sunrise (SR) and sunset (SS) are indicated by vertical dashed lines.
Between 02 UT and 10 UT, in the period that SCABs are observed (01 UT and 11-12 UT are not included as these are on the extremes of the observable period), SCABs are observed to be associated with 41% of substorms. Absorption is not observed with every substorm, even when the propagation path is within the midnight and morning sectors, due to spatial uncertainties regarding the co-location of the precipitation region and the absorption points of the propagation path (the points where the path passes through the D region), and temporal uncertainties regarding the duration of the precipitation. The precipitation regions move rapidly, both in longitude (westward travelling surge and the eastwards expansion of the drizzle precipitation region (FIGURE 5.11)) and in latitude (the auroral oval moves equatorward during the substorm growth phase and then expands poleward during the expansion phase) and are not necessarily coincident with the absorption points of the path at the time that the once-hourly measurements of signal level are made. Also, the duration of absorption is limited by the duration of the precipitation, as any enhancement of the D region electron density decays rapidly due to the high recombination rate in the lower ionosphere. The duration of the precipitation is generally less than one hour (AKASOFU, 1968) and so SCABs can occur within the interval between the once-hourly signal level measurements, in which case they are not observed.

There are absorption events during which absorption is observed on two consecutive signal level measurements, for instance 23 January at 03-04 UT and 08-09 UT (FIGURE 5.14), suggesting that the absorption had a duration greater than 1 hour or that two absorption events occurred in rapid succession. In such cases there is generally evidence for multiple expansion phase onsets during the absorption periods. On 23 January, two dipolarizations occur between 03-04 UT, one at 0315 UT and a second at 0400 UT, and a double-stepped flux injection occurs between 08-09 UT. Consequently, the absorption associated with one substorm has a duration shorter than 1 hour, consistent with AKASOFU (1968).

The duration of the absorption depends on the local time of the observations and hence the probability of observation of an absorption event will change accordingly. The splash and drizzle precipitation regions have lifetimes of ~15 mins and 30-60 mins respectively (AKASOFU, 1968; in AKASOFU these precipitation regions give rise to N type (explosive, short duration) and M type (gradual, long duration) absorption respectively). These durations are related to the local time distribution of SCAB observation: in the pre-midnight sector (02-05 UT) and post-midnight sector (06-10 UT) the proportion of substorms with observed SCABs is 33% and 56% respectively (FIGURE 5.15). This leads to the conclusion that most substorms produce appreciable levels of absorption, but the relatively low proportion of substorms during which SCABs are observed is predominantly a consequence of the low (1 hour) time resolution of the NONCENTRIC experiment.

This has implications for the interpretation of temporally-averaged empirical models of auroral absorption. The variation in mean absorption with local time is weighted by the lifetimes of the two precipitation regions, potentially underestimating the level of absorption in the midnight sector relative to the absorption in the morning sector. The westward
travelling surge, related to the splash precipitation region, has a harder and more intense energy spectrum than the drizzle precipitation region and consequently produces a higher level of absorption, typical levels of absorption on a 30 MHz riometer being 6.0 dB and 3.25 dB in the midnight and morning sectors respectively (Miller and Vondrak, 1985). Care must be taken, then, in interpretation of the Foppiano and Bradley (1983) model of riometer absorption which has a much higher value of $Q_i$ in the morning sector than in the midnight sector (Figure 5.4).

5.5.4 Substorm intensity and associated absorption level Two methods for quantifying the magnitude of the substorms occurring during the winter campaign have been developed, one using the field angle determined from the GOES-7 satellite, and the other using the magnitude of the flux injections determined at the LANL spacecraft. This allows investigation of the relationship between substorm magnitude and the NONCENTRIC signal level variation associated with SCABS.

The stress in the magnetotail prior to the onset of the expansion phase of a substorm can be quantified by the minimum field angle measured prior to the dipolarization, $\theta_{min}$. Lopez et al. (1988) have shown that the magnetotail stress and the level of relaxation of the magnetotail during the dipolarization, $\Delta \theta$, are related. $\theta_{min}$ and $\Delta \theta$ are defined in Figure 5.12. At the onset of the expansion phase the magnetospheric cross tail current at $-10 R_e$ is diverted to the auroral ionosphere in the substorm current wedge, the ionospheric consequence of which is a westward electrojet (e.g. McPherron, 1979) whose intensity can be measured by the auroral electrojet index, $AE$. The maximum magnitude of $AE$ during a substorm is related to $\theta_{min}$ (Lopez and Rosenvinge, 1993) indicating that the strength of the electrojet is related to the stress in the magnetotail and the energy released in the magnetotail relaxation. At the time of writing $AE$ was unavailable for the period of the present study. However, $\theta_{min}$ and $\Delta \theta$ were employed as indicators of substorm magnitude. Figure 5.16 indicates SCAB signal level plotted as a function of both $\theta_{min}$ and $\Delta \theta$, for those events where a dipolarization was observed. In both cases there is a trend to lower signal level as the magnetotail stress and relaxation increase, especially for $\Delta \theta$ where $r = 0.83$, indicating that the level of precipitation in the ionosphere during a substorm is related to the magnitude of the dipolarization of the magnetotail.

The magnitude of the particle injection measured by the LANL spacecraft is also a possible indicator of substorm intensity. However, poor correlation is found between SCAB signal level and the LANL flux level measured immediately following the injection or the change in flux level during the injection. This is possibly due to uncertainty in the location of the particle trapping region (the major proportion of the particles comprising the ring current are trapped in the altitude range 4-6 $R_e$ (e.g. Frank, 1967)) relative to the spacecraft (in geosynchronous orbit at 6.6 $R_e$). Also, the trapped electron flux decays during the eastwards drift from the midnight sector injection region to the spacecraft by an amount that cannot be determined. Consequently, it is uncertain how well the flux levels measured at
Chapter 5  Radiowave absorption

Figure 5.16  SCAB signal level plotted as functions of the dipolarization parameters $\theta_{\text{min}}$ (top panel) and $\Delta \theta$ (bottom panel) for substorms with observed dipolarizations. The lines of best fit and the coefficients of correlation are indicated in each panel.
5.6 Substorm occurrence, SCAB magnitude, and the relationship with $K_p$

Within the splash and drizzle precipitation regions, mean absorption levels increase and mean signal levels decrease with increasing $K_p$ (§5.3). This absorption is produced by the precipitation of >30 keV electrons, the flux of which increases with increasing $K_p$ (§5.4). However, this precipitation and auroral absorption is a consequence of the occurrence of substorms (§5.5). The decrease in mean signal level observed by the NONCENTRIC experiment, then, must be a consequence of an increasing occurrence of substorms with increasing $K_p$, or an increasing substorm magnitude and SCAB absorption with increasing $K_p$, or a combination of both.

During the winter campaign, substorms occur with greater frequency during disturbed periods, as indicated by FIGURE 5.17 which illustrates the number of hours for which $K_p = 1, 2, ..., 6$, the number of substorms identified per $K_p$ bin, and the corresponding number of substorms occurring per hour. The absorption experienced during individual SCABs increases with increasing $K_p$ also, as indicated by FIGURE 5.18 which illustrates the signal level measured during each SCAB plotted as a function of $K_p$. Such a correlation is expected as $K_p$ is correlated with $A E$ (ROSTOKER, 1991) which is itself correlated to substorm magnitude (LOPEZ and ROSENVINGE, 1993). However, the correlation is not as good as that determined between SCAB signal level and the parameters of magnetotail dipolarizations $\theta_m$ and $\Delta \theta$ (§5.5.4) and considerable spread is observed in each $K_p$ bin. The magnitude of the absorption during an individual SCAB is, then, better correlated with rapid changes in the magnetotail than with 3-hourly $K_p$. The overall decrease in mean signal level with increasing $K_p$, though, is a consequence of increasing substorm occurrence and substorm magnitude during disturbed periods.

In §5.3.1 it was noted that the mean signal level variation between 07-09 UT at $K_p = 5$ during the winter campaign contradicted the expected trend (FIGURE 5.3). This is a consequence of the low occurrence of $K_p = 5$ conditions between these hours. The three-hour $K_p$ intervals 06-09 UT and 09-12 UT have activity of $K_p = 5$ only on 22 and 17 January respectively. The signal level variation and geomagnetic conditions for 17 and 22 January are illustrated in FIGURE 5.19, in which the two $K_p$ bins in question are highlighted in grey. The numbers at the bottom of panels b and d indicate the $K_p$ level in each three-hour bin. Major substorm activity occurred in the two highlighted intervals, at 10-11 UT on 17 January and at 06 UT on 22 January. At these times auroral absorption occurred, as indicated in panels a and c. However, no substorms occurred at 09 UT on 17 January or 07-08 UT on 22 January and consequently the signal level measured at these times was not affected by auroral absorption, despite the $K_p = 5$ conditions. These three signal level measurements produced the anomaly in the $K_p = 5$ curve of FIGURE 5.3. This emphasises the problematic nature of employing 3-hourly indices to predict or categorize the highly
FIGURE 5.17 The number of hours and the number of substorms identified in each $K_p$ bin during the winter campaign. The line indicates the number of substorms occurring per hour as a function of $K_p$. 
FIGURE 5.18  SCAB signal level plotted as a function of $K_p$ for the Boston path during the winter campaign. The line of best fit and coefficient of correlation are indicated.
Figure 5.19  The diurnal signal level variation at 6.800 MHz on the path to Boston (panels a and c) and the magnetic field orientation $\theta$ determined at GOES-7 (panels b and d) for the 17 (panels a and b) and 22 (panels c and d) January 1989. Arrows in panels b and d indicate the times of particle injections determined in the LANL data. The numbers at the bottom of panels b and d indicate the $K_p$ level in each three hour bin. The grey highlighted periods are explained in the text.
dynamic, short timescale absorption events associated with the auroral zone.

5.7 Summary

The major diurnal variation in signal level is controlled by solar photo-ionization of the D region ionosphere which leads to absorption of >40 dB at 7 MHz at noon. Trans-auroral paths suffer auroral absorption at night, between 20 LT and sunrise, the absorption being greatest in the post-midnight sector. The level of absorption increases with increasing geomagnetic activity, so at $K_p = 4$ the mean level of absorption during the night is ~20 dB at 7 MHz. This absorption is coincident with the pattern of high energy electron precipitation into the D region in the auroral zone. Such electron precipitation is a consequence of enhancement of the ring current by particle injections occurring during substorms. Substorms detected by geosynchronous orbit particle and magnetic field observations give rise to SCABS, absorption events in the midnight and morning sector auroral zones. The duration of these absorption events appears to be dependent on the local time of observation, ~15 mins prior to midnight and ~30-60 mins after midnight. The intensity of the absorption events can be correlated with the magnitude of the extension of the magnetotail prior to the substorms and the size of the dipolarization during the substorms. The absorption during these events can be ~30 dB at 7 MHz. The decrease in mean signal level with increasing geomagnetic activity is a consequence of increasing substorm occurrence and magnitude during these periods.
### Table 5.1 SCABs observed during the winter campaign

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Signal level (dB)</th>
<th>Flux increase factor</th>
<th>Flux level</th>
<th>$\theta_{min}$ (°)</th>
<th>$\Delta \theta$ (°)</th>
<th>$K_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>16 Jan</td>
<td>04</td>
<td>-48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17 Jan</td>
<td>07</td>
<td>-44</td>
<td>2</td>
<td>$1 \times 10^3$</td>
<td>10</td>
<td>27</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>-44</td>
<td>13</td>
<td>$5 \times 10^3$</td>
<td>-</td>
<td>-</td>
<td>5</td>
</tr>
<tr>
<td>18 Jan</td>
<td>01</td>
<td>-27</td>
<td>-</td>
<td>-</td>
<td>25</td>
<td>19</td>
<td>4</td>
</tr>
<tr>
<td>19 Jan</td>
<td>08</td>
<td>-23</td>
<td>2</td>
<td>$7 \times 10^6$</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>21 Jan</td>
<td>02</td>
<td>-38</td>
<td>20</td>
<td>$4 \times 10^8$</td>
<td>-</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>22 Jan</td>
<td>02</td>
<td>-22</td>
<td>-</td>
<td>-</td>
<td>30</td>
<td>15</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>04</td>
<td>-29</td>
<td>5</td>
<td>$2 \times 10^7$</td>
<td>-</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>05</td>
<td>-46</td>
<td>10</td>
<td>$4 \times 10^7$</td>
<td>22</td>
<td>19</td>
<td>4</td>
</tr>
<tr>
<td>23 Jan</td>
<td>03</td>
<td>-34</td>
<td>13</td>
<td>$4 \times 10^7$</td>
<td>33</td>
<td>25</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>08</td>
<td>-36</td>
<td>7</td>
<td>$4 \times 10^7$</td>
<td>42</td>
<td>19</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>-30</td>
<td>10</td>
<td>$4 \times 10^7$</td>
<td>-</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>25 Jan</td>
<td>06</td>
<td>-23</td>
<td>-</td>
<td>-</td>
<td>43</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>26 Jan</td>
<td>07</td>
<td>-26</td>
<td>6</td>
<td>$2 \times 10^7$</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>28 Jan</td>
<td>03</td>
<td>-22</td>
<td>-</td>
<td>-</td>
<td>38</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>30 Jan</td>
<td>02</td>
<td>-33</td>
<td>5</td>
<td>$2 \times 10^7$</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>31 Jan</td>
<td>02</td>
<td>-25</td>
<td>5</td>
<td>$3 \times 10^7$</td>
<td>30</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>05</td>
<td>-40</td>
<td>8</td>
<td>$3 \times 10^7$</td>
<td>28</td>
<td>20</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>09</td>
<td>-41</td>
<td>-</td>
<td>-</td>
<td>27</td>
<td>19</td>
<td>3</td>
</tr>
<tr>
<td>1 Feb</td>
<td>06</td>
<td>-33</td>
<td>7</td>
<td>$2 \times 10^7$</td>
<td>-</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>3 Feb</td>
<td>01</td>
<td>-53</td>
<td>30</td>
<td>$3 \times 10^7$</td>
<td>-</td>
<td>-</td>
<td>6</td>
</tr>
<tr>
<td>5 Feb</td>
<td>02</td>
<td>-36</td>
<td>-</td>
<td>-</td>
<td>18</td>
<td>22</td>
<td>4</td>
</tr>
<tr>
<td>6 Feb</td>
<td>02</td>
<td>-40</td>
<td>15</td>
<td>$3 \times 10^7$</td>
<td>18</td>
<td>30</td>
<td>5</td>
</tr>
</tbody>
</table>

**Date** indicates the date and **Time** the hours of each day that SCABs were identified during the winter campaign. **Signal level** gives the signal level measured at 6.800 MHz during the SCAB. **Flux increase factor** and **Flux level** indicate the increase in flux during the injection and the flux level following the injection of 30-300 keV electrons for SCABs which had associated particle injections measured by the LANL spacecraft. **$\theta_{min}$** and **$\Delta \theta$** indicate the minimum polar angle of the terrestrial field prior to the dipolarization and the change in field angle during the dipolarization for SCABs which had associated dipolarizations measured by the GOES-7 satellite. **$K_p$** indicates the geomagnetic activity at the time of the SCAB.
Chapter 6

Variations in the Maximum Usable Frequency

6.1 Introduction

The maximum usable frequency (MUF), the highest frequency that will propagate along a path, varies as the electron density profile of the ionosphere along the propagation path changes. Higher electron densities in the ionosphere, especially in the E and F regions, will allow the propagation of higher frequencies.

The main contributing factor to the variation of electron density in the ionosphere is the diurnal variation of solar illumination, electron density being a maximum during the day and a minimum at night. In §6.2 the mean diurnal variation of the E and F region electron density during the two NONCENTRIC campaigns is obtained from the observations of four vertical ionosonde stations. The relationship between the measured ionospheric parameters and the maximum usable frequencies on the oblique propagation paths is investigated.

The electron density of the solar-produced high latitude ionosphere is perturbed by the magnetosphere-ionosphere coupling processes manifested in the auroral oval, the mid-latitude trough, and the polar cap, with the level of perturbation generally being dependent on the level of geomagnetic activity. Variations in electron density produced by these regions are localised and consequently their influence on each propagation path can be different. Precipitation in the auroral oval together with plasma convection in the polar cap, act to increase the MUF a) during the night on the Boston, Aberdeen and Leicester paths and b) during the day on the Alert and Fairbanks paths. These mechanisms and their influence on the oblique MUFs are investigated in §6.3 and §6.4. The decrease in the MUF measured on the Boston, Aberdeen and Leicester paths, resulting from the depletion in electron density due to the mid latitude trough is investigated in §6.5.

6.2 Ionosonde observations of ionospheric parameters and comparison with NONCENTRIC observations

The following ionospheric parameters were measured during the NONCENTRIC campaigns by four vertical ionosondes located in the area of interest: \( f_{F2} \), the F region critical frequency, \( h'F \), the virtual height of the bottom of the F region, \( h'F2 \), the virtual height of the F region maximum, \( f_E \), the E region critical frequency, and \( f_{E} \), the sporadic E critical frequency. Details of these observations are now presented.

6.2.1 The ionosonde stations

The four ionosonde stations are located over a range of mid and high latitudes, between 45°N and 75°N geographic latitude and 56°N and 84°N geomagnetic latitude (TABLE 6.1). The Ottawa and South Uist stations are sub-auroral,
Chapter 6 Maximum usable frequency

Churchill is at auroral latitudes, and Resolute Bay is within the polar cap. The locations of the ionosonde stations relative to the five NONCENTRIC propagation paths and the auroral oval are indicated in FIGUREs 3.1 and 6.1. The Ottawa ionosonde is located near the Boston receiver, the South Uist ionosonde is located near the Aberdeen and Leicester receivers. The Churchill ionosonde is located at approximately the latitude of the Boston path midpoint, though 1.5 hours behind in local time. The Resolute Bay ionosonde is located one third of the distance between the Clyde River transmitter and the Fairbanks receiver; this is the latitude of the Alert path midpoint, though 2 hours behind in local time. The Churchill ionosonde is situated underneath the auroral oval during the night, the exact time depending on geomagnetic activity: between 03-10 UT for $K_p = 0$ and between 23-14 UT for $K_p = 4$.

<table>
<thead>
<tr>
<th>Station</th>
<th>Geographic coordinates</th>
<th>Geomagnetic coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolute Bay</td>
<td>74.7°N 265.1°E</td>
<td>83.7°N 314.7°E</td>
</tr>
<tr>
<td>Churchill</td>
<td>58.8°N 265.8°E</td>
<td>69.3°N 329.8°E</td>
</tr>
<tr>
<td>Ottawa</td>
<td>45.4°N 284.1°E</td>
<td>56.6°N 359.4°E</td>
</tr>
<tr>
<td>South Uist</td>
<td>57.4°N 352.7°E</td>
<td>55.5°N 75.2°E</td>
</tr>
</tbody>
</table>

6.2.2 Mean ionospheric parameters FIGURES 6.2a and b illustrate the variation of the scaled ionospheric parameters from the four ionosonde stations for six typical days of the summer (20-25 July 1988) and winter (16-21 January 1989) campaigns respectively. The mean diurnal variation of $f_E, f_{F2}, h'F$ and $h'F_2$ at each ionosonde station during both campaigns is indicated in FIGURE 6.3, vertical bars indicating the standard deviation. The mean levels of $f_E, f_{F2},$ and $h'F_2$ at each ionosonde station at local noon and midnight during the summer and winter campaigns are tabulated in TABLES 6.2a and b. The method of scaling $f_E$ and $f_{Ei}$ differs slightly from station to station. Where both exist, $f_E$ refers to the solar-produced E region and $f_{Ei}$ to the sporadic E layer when they can be distinguished. At South Uist, only $f_{Ei}$ is scaled; this records the maximum critical frequency of the E region whether it is solar-produced or sporadic in nature. Also, a minimum value of $f_E$ of 1.5 MHz is recorded at South Uist, even when no E region exists; this is most noticeable during the winter campaign.

During the summer campaign $f_{F2}$ varies little between day and night as a consequence of the seasonal anomaly (DAVIES, 1990). The mean daytime maximum of $f_{F2}$ during the summer campaign is 6 MHz and the mean night-time minimum is 4.5 MHz. During the winter campaign, at South Uist, Ottawa, and Churchill, $f_{F2}$ varies between 12 MHz at local noon and 4 MHz at local midnight. At Resolute Bay, within the polar cap, the mean variation is much lower, between 6.5 MHz during the day and 3.5 MHz at night; Resolute Bay is in constant darkness during the winter campaign. However, at Resolute Bay $f_{F2}$ varies randomly by up to 5 MHz from hour to hour, a much larger variation than at other
FIGURE 6.1 The location of the auroral oval ($K_p = 3$) relative to the five NONCENTRIC propagation paths and the four ionosonde stations for six UTs. Coordinates are geographic latitude and LT, the + indicating the geographic pole and the x indicating the geomagnetic pole.
Summer campaign  Ionogram parameters $h'F_2$ $h'F$ $f_{F2}$ $f_E$ $f_{Es}$

FIGURE 6.2a  Ionospheric parameters $f_{F2}$ (red curves), $f_E$ (light green curves), $f_{Es}$ (dark green curves), $h'F$ (light blue curves) and $h'F_2$ (dark blue curves) measured at each ionosonde station, for 20–25 July, summer campaign.
Winter campaign, ionogram parameters $h'F2$, $h'F$, $f_{F2}$, $f_E$, $f_{Es}$

**FIGURE 6.2b** Ionospheric parameters $f_{F2}$ (red curves), $f_E$ (light green curves), $f_{Es}$ (dark green curves), $h'F$ (light blue curve) and $h'F2$ (dark blue curve) measured at each ionosonde station, for 16-21 January, winter campaign.
FIGURE 6.3 Mean diurnal variation of ionospheric parameters $f_{F2}$ (red curves), $f_E$ (green curves), $h^{'F}$ (light blue curves) and $h^{'F2}$ (dark blue curves) measured at each ionosonde station. Standard deviations are indicated by vertical bars. At Ottawa and Resolute Bay during the winter campaign, data bins are 2 hours wide (as opposed to 1 hour wide) due to a sparsity of data.
Chapter 6  Maximum usable frequency

stations, where the typical difference from hour to hour is approximately 1 MHz. These differences are reflected in the high standard deviation in $f_{F2}$ at Resolute Bay during the winter campaign (FIGURE 6.3). This is possibly an indication of convection of high electron density plasma patches across the polar cap (see §6.4.3).

The variation of $f_{E}$, in the absence of sporadic $E$, at the stations in both summer and winter is consistent with a solar zenith angle dependence of electron density (e.g. LEFTIN, 1976). The E region generally disappears at night, though a small E region ($f_{E} = 2$ MHz) remains at Resolute Bay during the summer as the station is illuminated at all times of day. Random variations in $f_{E}$ are very small, reflected in the low standard deviation indicated in FIGURE 6.3 (at South Uist the mean variation in $f_{E}$ has been determined from $f_{E}$ and consequently the standard deviation is much higher).

In summer $h'F2$ varies from ~400 km during the day to ~300 km at night at all stations. In winter $h'F2$ data is available for the Ottawa ionosonde only, but this indicates that the F region layer height is greatest at night (~300 km) and decreases during the day (~270 km).

<table>
<thead>
<tr>
<th>Station</th>
<th>$AE$ (MHz)</th>
<th>$f_{E}$ (MHz)</th>
<th>$f_{F2}$ (MHz)</th>
<th>$h'F2$ (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Day</td>
<td>Night</td>
<td>Day</td>
<td>Night</td>
</tr>
<tr>
<td>Resolute Bay</td>
<td>3.0</td>
<td>2.0</td>
<td>5.5</td>
<td>5.0</td>
</tr>
<tr>
<td>Churchill</td>
<td>4.0</td>
<td>-</td>
<td>6.5</td>
<td>4.5</td>
</tr>
<tr>
<td>Ottawa</td>
<td>4.0</td>
<td>-</td>
<td>7.0</td>
<td>3.5</td>
</tr>
<tr>
<td>South Uist</td>
<td>4.0</td>
<td>-</td>
<td>6.5</td>
<td>4.5</td>
</tr>
</tbody>
</table>

**TABLE 6.2a** Mean day and night values of $f_{F2}$, $f_{E}$, and $h'F2$, summer campaign

<table>
<thead>
<tr>
<th>Station</th>
<th>$AE$ (MHz)</th>
<th>$f_{E}$ (MHz)</th>
<th>$f_{F2}$ (MHz)</th>
<th>$h'F2$ (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Day</td>
<td>Night</td>
<td>Day</td>
<td>Night</td>
</tr>
<tr>
<td>Resolute Bay</td>
<td>2.0</td>
<td>-</td>
<td>7.0</td>
<td>4.0</td>
</tr>
<tr>
<td>Churchill</td>
<td>3.0</td>
<td>-</td>
<td>12.5</td>
<td>4.0</td>
</tr>
<tr>
<td>Ottawa</td>
<td>3.5</td>
<td>-</td>
<td>12.5</td>
<td>4.0</td>
</tr>
<tr>
<td>South Uist</td>
<td>3.0</td>
<td>-</td>
<td>11.5</td>
<td>3.0</td>
</tr>
</tbody>
</table>

**TABLE 6.2b** Mean day and night values of $f_{F2}$, $f_{E}$, and $h'F2$, winter campaign

*Very variable

6.2.3 Comparison of mean ionospheric parameters with NONCENTRIC observations

The NONCENTRIC experiment cannot determine the exact value of the maximum usable frequency, the MUF, as only 14 discrete frequencies are transmitted; however, bounds of ±1.5 MHz can generally be placed on its value.

Secant factors for various modes on the paths to Alert, Fairbanks, Boston and Leicester are tabulated in TABLE 6.3, determined by the method described in §2.4. The paths to Boston and Leicester have secant factors of ~3 for the 1F mode. In winter the ionosphere near the midpoints of these paths has an F region critical frequency of approximately 12
Maximum usable frequency

MHz, giving a predicted MUF of ~36 MHz, much higher than the highest frequency of the experiment. In summer, at local noon $f_{F2} \approx 6$ MHz, and the predicted MUF is ~20 MHz, near the maximum experimental frequency. Consequently, the daytime variation of the MUF on these paths is difficult to determine, especially during the winter campaign. The night-time values of $f_{F2}$ are similar during the summer and winter campaigns, and give a predicted MUF in the range 10-13 MHz.

FIGURE 6.4 illustrates the signal recognition results for the paths to Boston and Leicester on 23 and 24 July 1988, during the summer campaign. On the path to Boston, typically the highest frequency propagating during the night is 10.195 MHz (03-09 UT 24 July) and during the day is ~20 MHz (14-00 UT 23 July), both consistent with the predicted MUF. During the day propagation on frequencies below the MUF and above the LUF is very reliable. Deviations from the predicted behaviour occur: between 03-08 UT during the night of 23 July propagation occurs on frequencies (~10 MHz) higher than the predicted night-time MUF. Such anomalous propagation events (APEs) are discussed in more detail in §6.3.

The behaviour is more complicated on the Leicester and Aberdeen paths, which span 4.5 hours of local time. The solar zenith angle varies considerably along these paths, leading to significant horizontal gradients in the electron density along the direction of the propagation. This leads to a much less clearly defined diurnal variation in the MUF and a less predictable and lower probability of signal propagation. On the path to Leicester, the day (~08-21 UT) is marked by an increased occurrence of correctly recognized signals on frequencies above 10 MHz (bottom panel, FIGURE 6.4). However, daytime propagation is much less reliable than on the path to Boston, and the variation in the MUF is difficult to discern. During the night, sporadic signal recognition occurs on frequencies above the predicted MUF (~10 MHz), though not as coherent features such as the APE observed on the

<table>
<thead>
<tr>
<th>Mode</th>
<th>Reflection height (km)</th>
<th>Alert</th>
<th>Fairbanks</th>
<th>Boston</th>
<th>Leicester</th>
</tr>
</thead>
<tbody>
<tr>
<td>1E</td>
<td>110</td>
<td>4.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2E</td>
<td>110</td>
<td>3.0</td>
<td>4.9</td>
<td>5.1</td>
<td>5.3</td>
</tr>
<tr>
<td>3E</td>
<td>110</td>
<td>2.2</td>
<td>3.9</td>
<td>4.1</td>
<td>4.6</td>
</tr>
<tr>
<td>1F</td>
<td>280</td>
<td>2.3</td>
<td>3.4</td>
<td>3.4</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>350</td>
<td>2.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.1</td>
</tr>
<tr>
<td>2F</td>
<td>280</td>
<td>1.5</td>
<td>2.5</td>
<td>2.6</td>
<td>2.9</td>
</tr>
<tr>
<td></td>
<td>350</td>
<td>1.4</td>
<td>2.1</td>
<td>2.2</td>
<td>2.5</td>
</tr>
<tr>
<td>3F</td>
<td>280</td>
<td>1.3</td>
<td>1.9</td>
<td>2.0</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td>350</td>
<td>1.2</td>
<td>1.7</td>
<td>1.7</td>
<td>1.9</td>
</tr>
</tbody>
</table>
FIGURE 6.4 Signal recognition results from 23–24 July, summer campaign, on the paths to Boston and Leicester.
Chapter 6

Maximum usable frequency
64

path to Boston.

Times of depressed \( f_{F2} \) (during negative ionospheric storms) measured at the
ionosonde stations are correlated with depressed MUFs on the oblique paths. FIGURE 6.5
illustrates the signal recognition on the path to Fairbanks and ionogram parameters
determined from the Churchill ionosonde, for 19-24 January, winter campaign. During the
daytime of 20 January \( f_{F2} \) is much reduced and the daytime maximum in the MUF on the
path to Fairbanks is absent. This is investigated in more detail in chapter 7.

6.3 Enhancements of the MUF on trans-auroral paths: anomalous
propagation events (APEs)

The predicted night-time MUF on the paths to Boston, Aberdeen and Leicester is
approximately 10 MHz, and this typically corresponds well with the highest frequencies
observed propagating on these paths during the night. However, during some of the nights,
periods occur during which signals with frequencies up to 10 MHz higher than the predicted
MUF can propagate. These are termed anomalous propagation events (APEs) and are now
discussed in detail.

6.3.1 Anomalous propagation events An example of an APE observed on the Boston
path during the summer campaign is illustrated in the top panel of FIGURE 6.4, and five
examples from the winter campaign are reproduced in panels b to f of FIGURE 6.6; panel a
represents an undisturbed night which is included for reference. These events are typically
of 2-4 hours duration and can occur on frequencies up to the maximum frequency of the
experiment, i.e. 10 MHz higher than the predicted night-time MUF. During the six APEs
illustrated in FIGURES 6.4 and 6.6, signal recognition fails at 6.905 and 9.941 MHz. This is
a feature common to most nights, however, and is a consequence of co-channel interference
experienced at these times. In panel a of FIGURE 6.6 the signal recognition fails on 9.941
MHz between 01-04 UT, even in the absence of an APE. It is interesting to note that APEs
can occur during periods when the F region electron density is depressed as a consequence
of ionospheric storm effects, when otherwise no signals propagate (e.g. 02-06 UT, or
possibly 02-11 UT, 21 January, panel c, FIGURE 6.6).

The mean signal level during anomalous propagation events is \(-7\) dB lower than typical
daytime values and the Doppler spreading is lower than typical daytime spread. Note
however that anomalous propagation events occur during periods of high geomagnetic
activity when daytime Doppler spreading is enhanced. Very low signal level APEs are often
observed on 17.515 and 18.204 MHz only (panel e, FIGURE 6.6). This is a consequence of
the low signal level combining with the frequency dependent gain of the antennas, which
are most sensitive at \(-18\) MHz (see chapter 3), to produce a signal to noise ratio sufficient
for correct signal recognition at these frequencies only.

APEs also occur on the Leicester and Aberdeen paths, though they are more apparent
FIGURE 6.5  Signal recognition results from 19–24 January, winter campaign, on the path to Fairbanks, and ionogram parameters from Churchill for the same period.
Winter Boston
Signal Recognition

FIGURE 6.6 Five examples of anomalous propagation events (panels b–f) and a quiet night for reference (panel a) on the path to Boston, winter campaign.
during the winter campaign as the summer behaviour is less well defined (bottom panel of FIGURE 6.4). Leicester APEs generally appear between 23 UT and 08 UT. This is, in some cases, well correlated with APEs on the Boston path, e.g. 21 January, as indicated in FIGURE 6.7, an APE appearing at Leicester at 00-01 UT and at 03-05 UT at Boston. The 3-4 hour delay between observations of the APEs could be due to the local time separation of the two paths.

6.3.2 Geomagnetic dependence and local time distribution of APE occurrence  The three most pronounced examples of anomalous propagation events observed during the winter campaign occur on the nights following sudden storm commencements (SSCs) or during periods of large negative $D_s$ (17, 21 January, and 1 February, 1989) which suggests that they are a storm-time effect, or occur predominantly under very disturbed geomagnetic conditions. FIGURE 6.8 illustrates the proportion of nights on which anomalous propagation events are observed as a function of $K_p$ during the winter campaign. At $K_p < 4$ significant anomalous propagation events are not seen, but the occurrence of APEs increases as $K_p$ increases for $K_p \geq 4$.

FIGURE 6.9 (top panel) illustrates the diurnal occurrence of anomalous propagation (i.e. when propagation occurs on frequencies above the quiet night-time MUF) for the Boston path during the winter campaign. The diurnal occurrence maximum occurs at 03 UT and corresponds to a percentage occurrence over the whole campaign of ~50%. The lower panel of FIGURE 6.9 illustrates the diurnal variation in the position of the auroral oval relative to the Boston path for $K_p = 3$. The occurrence of anomalous propagation coincides with the times that the path midpoint is within the oval and the occurrence rate increases as the path midpoint approaches the middle of the oval. The midpoint of the path is within the oval between 00-08 UT for $K_p = 1$ and 21-10 UT for $K_p = 4$, which corresponds closely with the times when anomalous propagation events are observed (22-08 UT). The energy flux in the continuous aurora is a maximum in the middle of the oval (WHALEN, 1983) and this corresponds to where enhancement of the E and F regions by particle precipitation is greatest. This increase in electron density leads to an increase in the path MUF.

6.3.3 Possible mechanisms for APE production  The geomagnetic activity dependence of APE occurrence and the correspondence between the local time distribution of APE occurrence and the position of the auroral oval relative to the propagation path suggest that APEs are produced by enhancement of the ionospheric electron density by auroral precipitation at E or F region altitudes. Night-time enhancements of the E and F regions between 00-05 UT on 1 February 1989 were observed by the European Incoherent Scatter radar (EISCAT), 6 hours ahead of the Boston path in local time. An APE was observed on the Boston path between 03-06 UT on 1 February (panel f of FIGURE 6.6). The F region electron density measured by EISCAT exceeded $6\times10^{11} m^{-2}$ at times ($f_{F2} = 7 MHz$). If an electron density of this magnitude occurred at the Boston path midpoint the ionosphere would
FIGURE 6.7 Signal recognition results from 20–21 January, winter campaign, on the paths to Boston and Leicester, illustrating an example of near-simultaneous observation of APEs on the two trans-auroral paths (03–06 UT on Boston path, 00–01 UT on Leicester path.)
FIGURE 6.8 The number of nights with $K_p = 2, 3, 4, 5,$ and 6 (white bars), and the occurrence of nights with APEs (grey bars), on the path to Boston during the winter campaign. The proportion of nights with APEs as a function of $K_p$ is also indicated.
FIGURE 6.9 The diurnal occurrence of propagation on frequencies above the quiet time MUF (APEs) on the path to Boston during the winter campaign (top panel), and the diurnal variation of the position of the auroral oval ($K_p = 3$) along the propagation path (bottom panel).
support frequencies of up to \(-24\) MHz by the IF propagation mode. The Churchill ionosonde indicates night-time enhancements in \(f,F2\) possibly associated with precipitation, e.g. 02 and 07 UT 20 January, FIGURE 6.2b, and there are also indications of enhanced \(f,F2\) at South Uist possibly produced by precipitation during very disturbed conditions, e.g. 04-06 UT 18 January and 00-02 UT 21 January, FIGURE 6.2b. The second of these examples is correlated with the APE observed at Leicester illustrated in the bottom panel of FIGURE 6.7. However, the occurrence of F region electron density enhancements is not sufficient to account for all APEs observed.

APEs occur predominantly in the pre-midnight sector, \(-80\%\) of them occurring before local midnight (FIGURE 6.9), thus following the auroral E temporal distribution given by PITTINGER and GASSMANN (1971). E region propagation on the path to Boston requires 2-hop modes due to the length of the path and the curvature of the earth. The secant factor for 2E propagation on the path to Boston is \(5.1\), and consequently auroral E with critical frequency \(f,E\) = 4 MHz is sufficient to provide 20 MHz propagation and \(f,E\) = 8 MHz will support 40 MHz propagation. The auroral E is, however, required at both reflection points for propagation to be possible; combined F-E modes are unlikely to occur since with this path geometry the F region reflection point has a lower secant factor than the IF mode, which itself does not propagate. Good correlation can be found between the occurrence of auroral E at the Churchill ionosonde and the occurrence of APEs on the Boston path. FIGURE 6.10 illustrates the signal recognition at Boston and the ionosonde parameters determined at Churchill and Ottawa on 26-27 July, during the summer campaign. Simultaneous with both APEs observed at Boston on these two nights, large increases in \(f,E\) are observed at Churchill and on 26 July these occur as far south as Ottawa. Also, as mentioned in §6.3.1, APEs can occur during periods of depressed \(f,F2\), emphasising the role of the E region in the formation of APEs. It is concluded that APEs can be produced by E or F region enhancements in electron density, but are predominantly a consequence of auroral E.

It must be noted that \(f,F\) indicates the maximum critical frequency within the sporadic E layer, which can be quite patchy. The background critical frequency within the layer is given by \(f,F\), the blanketing E critical frequency \((f,E < f,F)\), which is the frequency below which no F region traces are observed on ionograms as total reflection occurs within the E layer (PIGGOTT and RAWER, 1961). Thus \(f,E\) generally gives an overestimate of the highest frequencies which will propagate via sporadic E layers.

6.3.4 Auroral and sporadic E occurrence and relationship to APE occurrence

The occurrence of sporadic E during the summer and winter campaigns has been established from the four ionosonde stations referred to in §6.2.1. The occurrence of night-time sporadic E is more significant for oblique propagation than daytime sporadic E since at night the maximum frequency of propagation supported by the F region is low. TABLE 6.4 indicates the percentage of nights on which significant \((f,E > 3\) MHz) sporadic E was
FIGURE 6.10 Signal recognition results from 26–27 July, summer campaign, on the path to Boston (top panel), and ionogram parameters determined at Churchill and Ottawa for the same period (bottom panels). The occurrence of APEs and auroral or sporadic E (dark green curves) are correlated.
observed during the summer and winter campaigns at Ottawa, South Uist, and Churchill (Resolute Bay is not included as the occurrence of sporadic E is low at all times and there is no distinction between day and night during the two campaigns). Sporadic E was observed at all stations, predominantly during the summer campaign. At Ottawa and South Uist the percentage of nights on which sporadic E is observed is ~40% and ~10% during the summer and winter campaigns respectively. Sporadic E is observed at Churchill on over 95% of nights during both the summer and winter campaigns and this is attributed to enhancement of the E region by precipitation i.e. auroral E. The diurnal variation of the percentage occurrence of sporadic E for each of the stations in the summer and winter campaigns is illustrated in Figure 6.11, for $f_E > 4$ MHz, $f_E > 6$ MHz and $f_E > 8$ MHz. During the summer campaign, at Ottawa and South Uist the occurrence of sporadic E with $4 < f_E < 6$ MHz is greatest during the day. At this time, however, the solar-produced E region critical frequency, $f_E$, is approximately 4 MHz so sporadic E with critical frequencies in this range is not very significant. At Churchill the occurrence of auroral E was greatest at night when the station is under the auroral oval; the night-time distribution of auroral E occurrence has a greater local time extent in winter than in summer.

<table>
<thead>
<tr>
<th>Station</th>
<th>Percentage occurrence sporadic E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ottawa</td>
<td>Summer 46% Winter 4%</td>
</tr>
<tr>
<td>South Uist</td>
<td>Summer 36% Winter 17%</td>
</tr>
<tr>
<td>Churchill</td>
<td>Summer 100% Winter 96%</td>
</tr>
</tbody>
</table>

Figures 6.12a and b illustrate the diurnal variation of the percentage occurrence of sporadic E with $f_E > 4$ MHz, $f_E > 6$ MHz and $f_E > 8$ MHz for $K_p < 3$ and $K_p \geq 3$ (in time intervals two hours wide to maintain statistical significance) during the summer and winter campaigns. At all stations the occurrence of daytime sporadic E decreases with increasing geomagnetic activity. At Ottawa and South Uist the (relatively low) occurrence of night-time sporadic E remains unchanged with increasing geomagnetic activity. At Churchill, the occurrence of night-time auroral E increases with increasing geomagnetic activity, consistent with the increase in APE occurrence on the path to Boston at these times (§6.3.2).

The local time extent during which auroral E is observed increases with increasing geomagnetic activity also, reflecting the longer period of time that the station remains under the auroral oval during disturbed periods: between 03-10 UT for $K_p = 0$ and between 23-14 UT for $K_p = 4$ (local midnight is 06 UT). At Resolute Bay the occurrence of sporadic E decreases with increasing geomagnetic disturbance at all times, reflecting the behaviour of polar cap sporadic E reported by BESPROZVANNAYA et al. (1980).

APEs are more common on the Boston path than on the Leicester path. The Boston path is shorter than the Leicester path, and so the auroral oval covers a larger proportion of the
Chapter 6

Maximum usable frequency

Figure 6.11 The occurrence of sporadic and auroral E with $f_E > 4, 6, \text{and } 8 \text{ MHz}$ at the four ionosonde stations during the summer (top panels) and winter (bottom panels) campaigns.
Figure 6.12a  The occurrence of sporadic and auroral E with $f_0E_s > 4$, 6, and 8 MHz at the four ionosonde stations during the summer campaign for $K_p < 3$ (top panels) and $K_p \geq 3$ (bottom panels).
FIGURE 6.12b  The occurrence of sporadic and auroral E with $f_{D}E_{s} > 4$, 6, and 8 MHz at the four ionosonde stations during the winter campaign for $K_{p} < 3$ (top panels) and $K_{p} ≥ 3$ (bottom panels).
length of the Boston path than the Leicester path (FIGURE 4.8). Consequently, auroral E can more easily extend to cover both 2-hop reflection points of the Boston path than of the Leicester path. Even though auroral E is observed at Churchill on over 90% of nights, APEs are observed on the Boston path less often (~50% of nights). This is an indication that although auroral E is present at most times, it is usually of insufficient spatial extent to cover both path reflection points. The increase in occurrence of APEs on the Boston path with increasing geomagnetic activity is probably a combination of an increased occurrence of high $f_{\text{r}}E$, auroral E and also a more equatorward auroral oval, increasing the probability that the auroral E will cover both 2-hop reflection points.

As an example, 26 July, FIGURE 6.10, where sporadic E is observed at both Churchill and Ottawa simultaneously, indicating a large region of sporadic E covering most of the Boston path.

6.4 Enhancements of the MUF on polar cap paths

Propagation conditions on the paths to Alert and Fairbanks vary greatly between summer and winter as a consequence of large seasonal changes in the polar cap ionosphere. In summer the polar cap is under constant solar illumination and consequently the F region electron density remains nearly constant. In winter the polar cap is in constant darkness, and the background F region electron density becomes low. However, the convection of dayside plasma into the nightside becomes important in winter, maintaining higher F region electron densities in the form of a tongue of ionization (e.g. Knudsen, 1974).

In the following sections the MUF variations on the paths to Alert and Fairbanks are considered in turn for each campaign, summer and winter. No conclusions can be drawn for the Alert path during the winter campaign since poor reception on the higher frequencies, as a consequence of interference, makes determination of the MUF impossible.

6.4.1 Alert, summer campaign

In the polar cap in summer the F region critical frequency remains nearly constant throughout the day, varying predominantly between 5 to 6 MHz, mainly as a consequence of constant solar illumination, though with a diurnally varying solar zenith angle. The virtual height of the F region maximum varies, however, from 280 km when Resolute Bay is in the pre-midnight sector at ~05 UT to 450 km at local noon, 18 UT. The increase in virtual height near local noon is partly a consequence of the increasing E region critical frequency at this time, which varies between 2 MHz at 06 UT to 3.5 MHz at 18 UT.

FIGURE 6.13 illustrates the signal recognition results for the path to Alert and the ionogram parameters from the Resolute Bay ionosonde for 18-23 July 1988. TABLE 6.5 summarises the critical layer frequencies measured at Resolute Bay and the corresponding oblique frequencies which can be supported by great-circle path propagation for the path to Alert during the day and night. At night (02-10 UT) the highest frequency observed propagating is ~11 MHz, corresponding to the expected frequency supported by the 1F
FIGURE 6.13 Signal recognition results from 18-23 July, summer campaign, on the path to Alert, and ionogram parameters from Resolute Bay for the same period.
Chapter 6  Maximum usable frequency

mode. During the day the highest expected frequency is \(-16\) MHz, propagating via the 1E mode. However, during undisturbed periods, propagation is possible on frequencies in excess of \(21\) MHz (e.g. 18-20 and 23 July 1988, FIGURE 6.13). Great circle path propagation at \(21\) MHz would require \(f_{c}F2 = 12.5\) MHz \((h'F2 = 450\) km\), \(f_{c}F2 = 9\) MHz \((h'F2 = 280\) km\), or \(f_{c}E = 4.5\) MHz, all greater than the ionospheric parameters observed at Resolute Bay. Consequently, off-great circle path propagation is necessary to explain the observed MUF variation on the Alert path.

On disturbed days, when \(f_{c}F2\) is depressed during F region ionospheric storms, e.g. 21 and 22 July 1988, the daytime MUF is depressed from the undisturbed level of \(>21\) MHz to 14-18 MHz, the level expected for daytime 1E propagation; the polar cap E region electron density is generally unaffected during storm-times. Consequently, during the day frequencies up to \(-16\) MHz propagate via the 1E mode and frequencies between \(-16\) MHz and the enhanced MUF propagate via the F region. Also, the maximum propagating frequency experienced occasional drop-outs to \(-14\) MHz for short periods of \(-1\) hour duration, e.g. 21 UT 19 July and 12 UT 20 July 1988, indicating that the source of the MUF enhancement is not present at all times. The time of occurrence of the enhanced MUF varies from day to day, commencing at 14-16 UT on 18 July and at 10 UT on 20 July, and ending at 01 UT on 20 July (following on from 19 July) and at 23 UT on 20 July (FIGURE 6.13).

The off-great circle path propagation necessary to explain the enhanced MUF is possibly a consequence of scatter from field-aligned irregularities produced in the auroral zone by precipitation and in the polar cap at the edges of convecting patches (see §2.4.4). Assuming a vertically inclined magnetic field in the polar cap, the scattering volume must be \(1500-2000\) km from the transmitter and receiver for rays of frequency \(20\) MHz to reach perpendicularity with the field by refraction within the background F region (critical frequency \(-6\) MHz). With a great circle path length of \(1300\) km this corresponds to an off-great circle path bearing of approximately \(\pm 65^\circ\). As propagation frequency increases the scatter volume moves to greater distances from transmitter and receiver, and the off-great circle path bearing increases. Such scatter propagation is not possible when the background F region critical frequency is depressed (e.g. 21-22 July, FIGURE 6.13) as the ray refraction is then insufficient for the radiowaves to reach perpendicularity with the magnetic field. Similarly, scatter propagation is not present at night because the E region critical frequency is low and cannot satisfy the perpendicularity requirement. The variability of the time of occurrence each day of the scatter propagation is either dependent on the presence of irregularities in the polar cap, the location of the irregularities relative to the propagation path, or the times at which perpendicularity with the magnetic field can be achieved. On days when the start of propagation was delayed (e.g. 18 and 23 July, FIGURE 6.13) \(h'F2\) was greater than its mean daytime value, and propagation only commenced once \(h'F2\) returned to lower values; at times of increased \(h'F2\) less refraction of the radiowaves occurs and perpendicularity with the magnetic field is not possible.

The mean signal level of the scatter propagation on frequencies above \(-17\) MHz is \(-10\)
Chapter 6

Maximum usable frequency

70

dB lower than the mean signal level of the E region propagation at ~14 MHz. Above this transition the mean signal level increases again, in contradiction to an expected decrease in signal level of irregularity scatter propagation with increasing frequency (see §2.4.4). However, this increase in signal level could be due to the approach of the scatter volume to the auroral oval as the off-great circle path bearing increases with increasing frequency. The occurrence of irregularities is expected to increase near the auroral oval as it is associated with the F layer irregularity zone (PIKE, 1971) and intense irregularity backscatter (WAGNER and PIKE, 1972; MÖLLER, 1974).

6.4.2 Fairbanks, summer campaign The variation of the MUF on the path to Fairbanks during the summer campaign displays two maxima, one during the day at 18 UT, which occurs on 80% of days, and a less uniform second maximum generally pre-midnight near 04 UT, which occurs on 60% of days. For example, during the period 24-26 July (FIGURE 6.14) a maximum in MUF at 18 UT is present on 24 and 25 July, but is absent on 26 July, and a maximum at 04 UT is present on 26 July, but is absent on 24 and 25 July. The two maxima are a consequence of the changing importance of the E and F regions for HF propagation as the layer heights and electron densities vary throughout the day. TABLE 6.6 indicates the frequencies predicted to be supported by the E and F regions during the day and night. At night the E region critical frequency is low and the highest frequency that can propagate by the 2E mode is ~10 MHz. Although the F region critical frequency remains nearly constant throughout the day, the decreased F region layer height at night increases the secant factor for the 1F mode and the highest frequency propagated by this mode becomes a maximum at ~17 MHz. Near local noon, however, the F region layer height has increased, decreasing the secant factor for the 1F mode so the highest supported frequency is ~15 MHz. The E region critical frequency has increased at this time so that the 2E mode will support frequencies similar to those supported by the 1F mode, ~15 MHz. The daytime maximum of the MUF (18-20 MHz) is higher than this prediction and this is possibly a consequence of scatter propagation. Both maxima are dependent on the F region critical frequency, as can be seen during periods of depressed $f_{c}F2$ when both the night-time and daytime MUF are depressed (e.g. 26-27 July, FIGURE 6.14). However, the daytime MUF does not decrease below 14 MHz and the night-time MUF does not decrease below 10 MHz, the frequencies

<table>
<thead>
<tr>
<th>Mode</th>
<th>Critical frequency (MHz)</th>
<th>Secant factor</th>
<th>Frequency supported (MHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day</td>
<td>1F</td>
<td>4.0</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td>1E</td>
<td>3.5</td>
<td>110</td>
</tr>
<tr>
<td>Night</td>
<td>1F</td>
<td>2.0</td>
<td>280</td>
</tr>
<tr>
<td></td>
<td>1E</td>
<td>2.0</td>
<td>110</td>
</tr>
</tbody>
</table>
FIGURE 6.14 Signal recognition results from 24–29 July, summer campaign, on the path to Fairbanks, and ionogram parameters from Resolute Bay for the same period.
Superimposed on the E region critical frequency variation measured at Resolute Bay is
the sporadic E layer. There is little correlation between the occurrence of sporadic E with
changes in the MUF indicating that this sporadic E is patchy (more so than the auroral E
observed at Churchill) and is unable to support propagation at both 2E reflection points.

<table>
<thead>
<tr>
<th>Mode</th>
<th>Reflecting layer characteristics</th>
<th>Secant factor</th>
<th>Frequency supported (MHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Critical frequency (MHz)</td>
<td>Virtual height (km)</td>
<td></td>
</tr>
<tr>
<td>Day</td>
<td>IF</td>
<td>6.0</td>
<td>450</td>
</tr>
<tr>
<td></td>
<td>2E</td>
<td>3.0</td>
<td>110</td>
</tr>
<tr>
<td>Night</td>
<td>IF</td>
<td>5.0</td>
<td>280</td>
</tr>
<tr>
<td></td>
<td>2E</td>
<td>2.0</td>
<td>110</td>
</tr>
</tbody>
</table>

6.4.3 Fairbanks, winter campaign  The diurnal variation of the MUF on the Fairbanks
path during the winter campaign is similar to the mean diurnal variation of $f_fF2$ at Resolute
Bay, i.e. MUF and $f_fF2$ are greatest between 15-08 UT (FIGURE 6.15). This corresponds to
when the propagation path and the Resolute Bay ionosonde are a) nearest local noon (18
UT), when the solar-produced plasma density is at its greatest (15-21 UT), and b) are within
the dusk cell of the polar cap convection pattern (22-08 UT). As shown by SOJKA and
SCHUNK (1987), under IMF southward conditions with the effect of corotation added, the
dusk convection cell draws in flux tubes from lower latitudes on the dayside than the dawn
convection cell and consequently maintains a higher F region plasma density in the evening
sector than in the morning sector.

The MUF on the Fairbanks path and $f_fF2$ at Resolute Bay in winter are characterised by
large random increases and decreases of the order 1-5 MHz on timescales of ~1 hour. Such
variations in $f_fF2$ were interpreted by BUCHAU et al. (1983) as evidence for patches of
enhanced electron density convecting overhead. High electron density plasma from the
dayside can be carried across the polar cap into the low electron density nightside
ionosphere within the polar cap convection flow (e.g. BUCHAU et al., 1983; 1985; WEBER
et al., 1984). This usually takes the form of patches of enhanced F region plasma with
diameters of ~1000 km and electron densities as large as 10 times the background electron
density, drifting in the convection electric field at ~1 kms⁻¹. As the strength and sense of the
convection flow is dependent on the $B_z$ component of the IMF, polar cap patches occur most
predominantly during periods of $B_z < 0$ or disturbed geomagnetic activity (WEBER et al.,
1984). No enhancement of the E region electron density is observed as at E region altitudes
rapid recombination leads to erosion of the plasma on timescales shorter than the drift time
across the polar cap. At F region altitudes the recombination rate is low enough that plasma
density enhancements can persist for several hours. Polar cap patches are of most
FIGURE 6.15 Signal recognition results from 3–8 February, winter campaign, on the path to Fairbanks, and ionogram parameters from Resolute Bay for the same period.
importance in winter when they are observed against the very low density of the ambient polar cap ionosphere. 

Weber et al. (1984) observed such ionization patches by their photo-emission and reported their predominance during moderately disturbed geomagnetic activity ($K_p \geq 4$). In contrast, at the Resolute Bay ionosonde and on the Fairbanks path, they are observed at all levels of geomagnetic activity. A comparison of IMF $B_z$ determined by the IMP-8 satellite, and NONCENTRIC signal recognition (Figure 6.16) indicates a possible correlation between enhancement of the MUF and $B_z < 0$ conditions. On 24 January $B_z$ is predominantly small and slightly positive, corresponding to a low MUF (10-14 MHz) between 08-15 UT. On 25 January, $B_z$ is generally negative and the MUF is enhanced (14-18 MHz) between 08-15 UT. On 26 January $B_z$ is small and slightly negative and the MUF is has a medium value (14 MHz). On 27 January the MUF is enhanced, corresponding to negative $B_z$, apart from a short period of strongly positive $B_z$ between 06-11 UT when the MUF decreases to 10 MHz.

The correlation between $f$O2 measured at Resolute Bay and the MUF on the Fairbanks path is not one-to-one, indicating that if convecting patches are supporting great circle path propagation then they have low spatial correlation over distances of ~400 km, the distance from Resolute Bay to the path midpoint. However, if the mechanism for propagation is scatter from irregularities formed by the steep electron gradient at the edges of the patches (Tsunoda, 1988), then off-great circle path propagation from patches not observed overhead by the ionosonde is possible and exact correlation is not expected.

During storm periods (e.g., 20-21 January) patch activity is observed but with decreased $f$O2 critical frequencies and MUF. This is expected if the source of the enhanced plasma, the dayside mid latitude F region, is depleted due to negative F region storm effects.

6.5 The influence of the mid latitude trough on the time of evening signal loss

The time at which propagation on the higher frequencies ceases at the end of the day on the path to Boston is dependent on the geomagnetic activity during the evening period. Figure 6.17 illustrates the variation of signal level on three days, 19 January, 5 February and 2 February, on the path to Boston at 20.300 MHz, during the winter campaign. Vertical dashed lines indicate sunrise and sunset. Signal levels of ~60 dB correspond to system noise and indicate that signals are not propagating from the Clyde River transmitter to the Boston receiver. Signals begin to propagate at 12 UT, sunrise, irrespective of the geomagnetic activity at this time. However, propagation ceases earlier in the day as geomagnetic activity increases. Figure 6.18 illustrates mass plots of the last hour at which propagation is observed each day of the winter campaign as a function of $K_p$ for the path to Boston, for four frequencies 23.169, 20.900, 20.300, and 18.204 MHz. Also marked is the average sunset time at the ground (dashed line). As geomagnetic activity increases the propagation ceases at earlier local times. At lower values of $K_p$, propagation can continue
FIGURE 6.16 Signal recognition results from 23-26 January, winter campaign, on the path to Fairbanks (top panel), and ionogram parameters from Resolute Bay (middle panel) and IMF Bz determined by the IMP-8 satellite for the same period (bottom panel).
Figure 6.17  Diurnal variation of signal level on 20.300 MHz for 19 January and 2 and 5 February, winter campaign, on the path to Boston. Signal level increases from the noise level at sunrise (SR) as propagation commences. The time at which propagation ceases and the signal level returns to the noise floor depends on $K_p$, and can be prior to sunset (SS) for high $K_p$ and after sunset for low $K_p$. 
FIGURE 6.16 Mass plots of the time of last propagation as a function of $K_p$ for the path to Boston on 18.204, 20.300, 20.900, and 23.169 MHz. Also indicated are the time of sunset (horizontal dashed line) and the time at which the mid latitude trough opens at the path midpoint (solid curve), determined from the HALCROW and NISBET (1977) trough model.
for several hours after sunset, but at higher levels of $K_p$ propagation ceases before sunset.

Propagation stops when the electron density of the F region becomes insufficient to produce reflection of the radiowave. As the sun sets the major source of electron production in the ionosphere is removed and the F region begins to deplete as loss processes dominate. However, the rate of recombination is low enough that the electron density can be maintained at levels sufficient for propagation of frequencies as high as 20 MHz to continue for several hours after sunset. Consequently, the time at which signals cease to propagate is not directly linked to the solar zenith angle or the time of sunset, but is dependent on the loss processes that govern the rate of electron removal from the F region. The mid latitude trough is a region of depleted F region electron density in the night-time sub-auroral ionosphere, at latitudes similar to the Boston path midpoint, and can adversely affect high frequency radio propagation by allowing rays to penetrate the ionosphere or by modifying ray trajectories so that they are no longer incident on the receiver (THRANE, 1981). The local time at which the dusk wall or leading edge of the trough appears in the evening is dependent on $K_p$ (HALCROW and NISBET, 1977; WHALEN, 1987, see §2.3.7) and so the time at which propagation is disrupted by the trough should be $K_p$ dependent also.

Superimposed on each panel of FIGURE 6.18 is the time at which the trough opens at the path midpoint as a function of $K_p$ (solid line) determined from the empirical trough model of HALCROW and NISBET (1977). The trough position is not shown for $K_p = 6$ as under such disturbed conditions the trough is located equatorward of the path midpoint and would be expected to have less effect on the propagation. The predicted trough opening time correlates well with the time of signal loss. Also, as predicted, the time of last propagation for $K_p = 6$ is later than would be expected from extrapolation of the trend for $K_p \leq 5$ (20.900 and 23.169 MHz), as a consequence of the trough moving equatorward of the path midpoint. The spread in local times at which the trough disrupts propagation at a specific $K_p$ is expected since $K_p$ is not an ideal indicator of trough location (RODGER et al., 1986). The distributions of time of signal loss at each of the four frequencies illustrated in FIGURE 6.18 are shifted to slightly later local times as propagation frequency decreases. The F region electron density necessary to support 18 MHz propagation is only 60% of that necessary to support 23 MHz propagation and consequently propagation on lower frequencies can continue after higher frequencies are lost. At 18.204 MHz at higher levels of $K_p$ there are occasions when propagation continues until much later than the predicted time of trough appearance. This is a consequence of propagation being maintained on E region modes during anomalous propagation events (APEs).

The path to Leicester displays the same trend. Propagation ceases between 15 UT and 21 UT, dependent on $K_p$, corresponding to the time at which the trough crosses the path midpoint and the equatorward 2-hop reflection point. FIGURE 6.19 illustrates the signal recognition results from Leicester during 6 days of the winter campaign. Arrows indicate the time each day at which propagation ceases suddenly on all frequencies above ~10 MHz for 1-2 hours, corresponding to the appearance of the trough. As the night progresses the
FIGURE 6.19 Signal recognition results from the path to Leicester, 26 January – 2 February, winter campaign. Arrows at the top of the panel indicate the approximate time at which propagation ceases on high frequencies as a consequence of the trough opening. The trough subsequently moves equatorwards of the path midpoint and propagation can recommence 1–2 hours later.
trough moves equatorwards, away from the path midpoint, and IF propagation can recommence for 1-3 hours, to follow the undisturbed diurnal variation.

6.6 Summary

Enhancements of the MUF above the usual diurnal variation can occur on propagation paths traversing the high latitude region. On trans-auroral paths, during geomagnetically disturbed conditions, auroral E can give rise to periods of 2-4 hours duration of greatly enhanced (~10 MHz) night-time MUF. In the polar cap several mechanisms can enhance the MUF, the particular cause and effect depending on path geometry and season. Convection of high electron density patches across the polar cap in winter allows sporadic MUF enhancement, and scatter propagation from irregularities in the polar cap and auroral oval allow off-great circle path propagation on higher frequencies than are supported by normal great circle path modes.

The mid latitude trough causes disruption of propagation on frequencies above 15 MHz in evening sector, the time of disruption depending on the time of trough appearance, which is itself dependent on geomagnetic activity. At high geomagnetic activities the propagation can cease before local sunset.
Chapter 7

The response of HF propagation to geomagnetic storms

7.1 Introduction

The global ionosphere does not respond instantaneously to periods of enhanced geomagnetic activity and the response can persist for hours or days after quiet magnetic conditions return. Consequently, individual values of $K_p$ will not necessarily accurately quantify the state of the ionosphere and such measurements should be viewed within the context of longer term variations in the solar terrestrial environment, i.e. magnetospheric and ionospheric storms.

Magnetospheric storms are triggered by changes in the solar wind and interplanetary magnetic field (IMF). Shock waves in the solar wind give rise to distortion of the geomagnetic field. In addition, changes in orientation of the IMF (generally the north-south component of the IMF, $B_z$, is most important) influence the efficiency of the energy coupling between the IMF and the magnetosphere. Increased coupling leads to accumulation of energy in the magnetosphere which subsequently increases substorm activity, resulting in enhanced energy deposition in the high latitude ionosphere by precipitation and joule heating. The extent and strength of the high latitude convection pattern also increase as a consequence of the increase in IMF-magnetosphere coupling. Such changes in the high latitude ionosphere have repercussions for the mid and low latitude ionosphere, where storm-time changes in global wind patterns and atmospheric composition give rise to increases and decreases in $F$ region electron density, i.e. the positive and negative ionospheric storm effects respectively (see §2.3.8).

The influence of magnetospheric and ionospheric storms on HF propagation at high latitudes is now investigated.

7.2 Geomagnetic storms during the experimental campaigns

The severity and evolution of geomagnetic storms are often monitored by the index $D_n$, a measure of the horizontal component of the earth's magnetic field at the equator. A geomagnetic storm can be identified in $D_n$ by its characteristic three phase evolution, the initial, main, and recovery phases (MATSUSHITA, 1967). The $D_n$ variation for the storm period 19-25 January 1989, winter campaign, is illustrated in FIGURE 7.1. The $D_n > 0$ nT initial phase is a consequence of an increase in the horizontal magnetic field strength at the equator as the magnetosphere is compressed by an increase in pressure in the solar wind. Increased coupling between the solar wind and magnetosphere leads to enhancement of the ring current during geomagnetic storms (FRANK, 1967) as magnetotail plasma is injected into near-earth space by substorm activity. The magnetic effect of an enhancement of the
Figure 7.1 The $D_s$ variation from 19-25 January, during the winter campaign, indicating the initial, main and recovery phases of a magnetospheric storm.
ring current is to produce a decrease in the horizontal component of the equatorial magnetic field at the earth’s surface and consequently $D_n$ becomes increasingly negative, the main phase. The maximum negative excursion of $D_n$ during the main phase is an indication of the magnitude of the enhancement of the ring current and consequently of the storm itself, and during very large geomagnetic storms this can be several hundred nT. $D_n$ slowly returns to less negative values as ring current electrons and protons are lost to the auroral ionosphere by precipitation during the recovery phase. The $D_n$ variation produced by a geomagnetic storm can last from several hours to a few days.

Figures 7.2a and b illustrate the variation in $D_n$, $K_p$, $P_{sw}$ (the solar wind pressure, $P_{sw} = \rho_{sw}v_{sw}^2$, where $\rho_{sw}$ and $v_{sw}$ are the solar wind density and velocity), and $B_z$, for the summer and winter campaigns respectively. Solar wind and IMF parameters were determined by the IMP-8 satellite. Included in Figure 7.2b is the daily substorm occurrence for the winter campaign, determined in chapter 5.

The winter campaign starts on 16 January 1989 during the recovery phase of a major storm, with $D_n \sim -125$ nT; $D_n$ returns to quiet conditions (-25 nT) over a period of 4 days. As quiet conditions are attained a sudden storm commencement (SSC) occurs at 1129 UT on 20 January (Solar Terrestrial Prompt Reports) and after the initial phase, $D_n$ returns to -125 nT during the 5 hour duration main phase. Recovery to less negative values of $D_n$ takes approximately 4 days. After 30 January $D_n$ becomes increasingly negative again for a period of 8 days: $D_n < -50$ nT, indicating mild storm activity, and reaches -75 nT on several occasions. This period is not preceded by a storm commencement and does not exhibit the short duration initial and main storm phases. This type of storm is generally associated with periods of predominantly southward IMF, during which enhanced substorm activity (see bottom panel of Figure 7.2b) leads to growth of the ring current. The quietest period during the winter campaign is 25-30 January, when $D_n$ rarely exceeds -25 nT and the mean daily substorm occurrence is 4.1. The occurrence of substorms is greater during storm periods (6.3 during 16-18 January and 6.0 during 21-23 January) and is greatest (9.5) during the period 3-6 February. A pulse in the solar wind pressure at 15 UT 16 January is accompanied by a sudden increase (18 UT) and then decrease in $D_n$; this is possibly evidence for a new storm commencement during the recovery phase of the previous storm, the increase and decrease in $D_n$ corresponding to the initial and main phases of the storm superimposed on the already present storm-time $D_n$ variation.

$D_n$ is generally less negative during the summer campaign than during the winter campaign (minimum $D_n$ values of -65 and -125 nT respectively). This is reflected in the campaign mean $K_p$ values of 2 and 3.5 for the summer and winter respectively. The periods 18-20 July, 21-23 July, 26-27 July, and 9-10 August exhibit the characteristic storm-time $D_n$ variation (initial, main and recovery phases), though the low values of $D_n$ (-25 to -50 nT) indicate very mild storm conditions. The 21-24 July storm period is initiated by a double pulse in the solar wind pressure (Figure 7.2a) and a turning of the IMF from northward to southward ($B_z \sim -2$ nT), producing the initial and (small, $D_n \sim -25$ nT) main phases in the $D_n$
FIGURE 7.2a  The $D_{st}$ variation for the summer campaign (upper panel). Also illustrated are $K_p$ (upper middle panel), solar wind pressure (lower middle panel), and the $B_z$ component of the IMF (bottom panel) for the same period. Storm periods I and II are also indicated.
FIGURE 7.2b The $D_{st}$ variation for the winter campaign (upper panel). Also illustrated are $K_p$ (upper middle panel), solar wind pressure (middle panel), the $B_z$ component of the IMF (lower middle panel), and the daily substorm occurrence (bottom panel) for the same period. Storm periods III and IV are also indicated.
variation and a peak in $K_p$. An SSC occurs at 0311 UT 21 July (Solar Terrestrial Prompt Reports). There follows a significant decrease in $D_m$ (to $-50$ nT) at 02 UT and a second, longer duration peak in $K_p$ on 22 July corresponding to a further increase in the southward component of the IMF ($B_z \sim -7$ nT). The storm commencement and intensification are separated by approximately one day.

During both campaigns increases in $K_p$ are well correlated with the storm-time variation in $D_m$, though the maximum value of $K_p$ is not necessarily directly proportional to the minimum value in $D_m$ (e.g. compare storm periods 21-24 July and 9-10 August 1988 in FIGURE 7.2a). In §5.3.3 the correlation between $D_m$ and $K_p$ was investigated, and a coefficient of correlation of approximately -0.6 was determined for the winter campaign. Although for the most part $D_m$ and $K_p$ are inversely proportional, the maximum in $K_p$ often occurs during the storm initial phase when $D_m$ is itself positive, e.g. 12-14 UT 20 January and 03-07 UT 21 July. This accounts for the populations of $D_m > 0$, $K_p \sim 3$ (smaller storms) and $D_m > 0$, $K_p \sim 6$ (larger storms) data points in FIGURE 5.7, and the relatively low coefficient of correlation despite the good correspondence between storm-time $D_m$ and $K_p$ variations.

<table>
<thead>
<tr>
<th>Storm</th>
<th>Storm interval</th>
<th>Minimum $D_m$ (nT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>21 - 24 July</td>
<td>-60</td>
</tr>
<tr>
<td>II</td>
<td>26 - 27 July</td>
<td>-40</td>
</tr>
<tr>
<td>III</td>
<td>20 - 24 January</td>
<td>-120</td>
</tr>
<tr>
<td>IV</td>
<td>31 January - 6 February</td>
<td>-75</td>
</tr>
</tbody>
</table>

### 7.3 Ionosonde and oblique path observations during geomagnetic storms

Four storm periods have been selected for further investigation, two from each of the summer and winter campaigns. These intervals are indicated in FIGURES 7.2a and b and are summarized in TABLE 7.1.

The deviation of $f_sF2$ (measured at the four ionosonde stations employed in chapter 6) from its mean diurnal variation and the changes in the propagation conditions on the NONCENTRIC paths during storm periods are discussed for each of the four storm intervals. The variation in $f_sF2$ during the four storm intervals is illustrated in FIGURES 7.3a and b, and the variation in signal recognition on the NONCENTRIC paths is illustrated in FIGURES 7.4a-d. Where possible changes in the MUF on the oblique paths are determined, though in many cases these are difficult to measure due to the sporadic nature of the propagation, the limited frequency range of the experiment, and the discrete, non-linear frequency scale. As an indicator of the storm-time change in propagation conditions, the variation in the number of correctly recognized signals received per day (summed over all 14 experimental frequencies), $N_r$, during the storm intervals is illustrated in FIGURES 7.5a and b. This
FIGURE 7.3a  The $D_n$ variation for storm periods I and II, summer campaign (top panels), and $f_0F2$ (thick curves) and the campaign mean diurnal variation of $f_0F2$ (thin curves) determined at the ionosonde stations at Resolute Bay, Churchill, Ottawa, and South Uist for the same periods (bottom panels).
FIGURE 7.3b The $D_s$ variation for storm periods III and IV, winter campaign (top panels), and $f_pF_2$ (thick curves) and the campaign mean diurnal variation of $f_pF_2$ (thin curves) determined at the ionosonde stations at Resolute Bay, Churchill, Ottawa, and South Uist for the same periods (bottom panels).
FIGURE 7.4a The $D_s$ variation for storm period I, summer campaign (top panel), and the signal recognition results for the paths to Alert, Fairbanks, Boston and Leicester (bottom panels) for the same period.
FIGURE 7.4b The $D_s$ variation for storm period II, summer campaign (top panel), and the signal recognition results for the paths to Alert, Fairbanks, Boston and Leicester (bottom panels) for the same period.
FIGURE 7.4c The $D_\text{st}$ variation for storm period III, winter campaign (top panel), and the signal recognition results for the paths to Alert, Fairbanks, Boston, and Leicester (bottom panels) for the same period.
FIGURE 7.4d The $D_s$ variation for storm period IV, winter campaign (top panel), and the signal recognition results for the paths to Alert, Fairbanks, Boston and Leicester (bottom panels) for the same period.
FigURE 7.5a  The $D_{st}$ variation for storm periods I and II, summer campaign (top panels), and the recognition ratio (middle panels) and spread index ratio (bottom panels) for the Alert, Fairbanks, Boston and Leicester propagation paths for the same periods.
**FIGURE 7.5b** The $D_n$ variation for storm periods III and IV, winter campaign (top panels), and the recognition ratio (middle panels) and spread index ratio (bottom panels) for the Alert, Fairbanks, Boston and Leicester propagation paths for the same periods.
variation is shown as the recognition ratio $N_r/N_n$, where $N_r$ is the number of correctly recognized signals received on a typical undisturbed day. The quietest periods of the summer and winter campaigns are identified from the $D_n$ variations (see FIGURES 7.2a and b) as 29-31 July 1988 and 25-29 January 1989; the percentage of all signals received that were correctly recognized during these two quiet periods are indicated in TABLE 7.2. The percentage occurrence of correctly recognized signals on the path to Alert during the winter campaign is much lower than on the other propagation paths partly as a consequence of interference on 6.800, 9.941 and 13.886 MHz. The recognition ratio parameter takes no account of the frequency dependence of storm-time effects discussed in §4.2.4, though this is necessary for the reasons outlined above.

The observations from the four storm periods are described in §7.3.1-4 and these are compared and discussed in §7.3.5.

<table>
<thead>
<tr>
<th>Site</th>
<th>Summer campaign</th>
<th>Winter campaign</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alert</td>
<td>43%</td>
<td>19%</td>
</tr>
<tr>
<td>Fairbanks</td>
<td>37%</td>
<td>33%</td>
</tr>
<tr>
<td>Boston</td>
<td>46%</td>
<td>40%</td>
</tr>
<tr>
<td>Leicester</td>
<td>41%</td>
<td>31%</td>
</tr>
</tbody>
</table>

7.3.1 Storm 1 Following the start of the initial phase of the storm at -01 UT 21 July, the daytime $f_F2$ at Ottawa is reduced by 1.5 MHz, 0.5 MHz and 1.0 MHz on 21, 22, and 23 July respectively; night-time $f_F2$ is reduced by $\sim1$ MHz (FIGURE 7.3a). At South Uist $f_F2$ is unaffected on 21 July but reduced by 1.5-2.0 MHz and 0.5-1.0 MHz on 22 and 23 July respectively. There is little data at Churchill during this period, but a reduction in $f_F2$ of 0.5-1.0 MHz is suggested; there is a 3 MHz reduction during the evening of 22 July. At Resolute Bay $f_F2$ is reduced by up to 2 MHz especially during the night-time and during 22 July. By 24 July the diurnal $f_F2$ variation has returned to quiet time conditions at all ionosondes. At Ottawa the storm-associated reduction in $f_F2$ is observed on 21 July, but at South Uist the reduction in $f_F2$ is not observed until 22 July, one day after the storm commences. Ottawa is in the midnight sector where the composition disturbance is first carried equatorward by thermospheric winds (see chapter 2). At the same time South Uist is located in the morning sector where no composition disturbance is present until rotation of the earth transports it eastwards; the composition disturbance is encountered at South Uist when it next approaches local midnight on 22 July.

Propagation to Boston and Leicester is disturbed on 22 July, and fewer than usual correctly recognized signals are received (FIGURE 7.4a). Anomalous propagation events (APEs - see chapter 6) occur at Boston and Leicester at 00-04 UT on 22 July corresponding
Chapter 7  Geomagnetic storms  79
to auroral E observed at Churchill; there are also APEs at Boston on 21 and 23 July. There is almost no daytime propagation on the path to Aberdeen (not shown). The daytime MUF on the path to Alert is reduced on 21 and 22 July.

FIGURE 7.5a illustrates the variation of the number of correctly recognized signals per day for each of the propagation paths during the storm. All paths start near quiet time propagation conditions (recognition ratio of 1.0) on 20 July. After the storm commencement on 21 July the recognition ratio decreases on all paths. The recognition ratio is a minimum on the 22 July for all paths (except Alert where the minimum occurs on 21 July), as opposed to 21 July when the storm commences. This delayed minimum in the propagation reliability is possibly a consequence of the local time dependence of ionospheric storm effects mentioned above, or is a consequence of the storm intensification caused by the increased southward component of $B_y$ on 22 July (see §7.2). That most paths have minimum recognition ratios on 22 July, despite a large local time difference between them, suggests the latter explanation. After 22 July the propagation conditions begin recovery to quiet time conditions. The ionosonde variations of $f_o F_2$ have returned to normal by 24 July, but propagation conditions remain disturbed at Boston and Leicester until 26 July when the next storm commences.

7.3.2 Storm II At Ottawa a positive storm effect is observed during the afternoon of 26 July, $f_o F_2$ increasing by 1.5 MHz over the quiet time variation (FIGURE 7.3a). $f_o F_2$ is depressed by 1.5 MHz during the following night. A second positive storm effect occurs during the afternoon of 29 July following a (small) onset signature in $D_n$. At South Uist $f_o F_2$ is unaffected during the day of 26 July but is depressed by 2.5 MHz during the following night. At Churchill $f_o F_2$ is depressed during the evening and night of the 26 July and the night of 27 July. There are also small (0.5-1 MHz) positive effects observed during the afternoons of 26, 27 and 28 July. At Resolute Bay $f_o F_2$ is depressed by 1-2 MHz during most of 26 and 27 July.

A positive storm effect is observed at Ottawa during the afternoon of 26 July, but no corresponding increase in the MUF is observed on the path to Boston (FIGURE 7.4b). The MUF is depressed during the day of the 26 July. At this time at Churchill, near the Boston path midpoint, $f_o F_2$ is depressed. The positive storm effect observed at Ottawa is confined to a region equatorward of the path reflection point and consequently higher frequency propagation is not supported. APEs appear on the Boston path during the nights of 26 and 27 July as a consequence of auroral E observed at Ottawa. There is no data at Boston on 28 July. Propagation to Leicester and Aberdeen is disturbed during 26 and 27 July and at Fairbanks a large decrease in the MUF is observed during this period, especially between 00-10 UT, 27 July. The daytime MUF at Alert is depressed on 26-29 July.

The decrease in the reception of correctly recognized signals is illustrated in FIGURE 7.5a. At Fairbanks and Leicester the degradation in propagation maximises on 27 July, one day after the storm commencement. Little decrease in recognition ratio is observed at
Chapter 7  Geomagnetic storms

Boston. The decrease in the recognition ratio at Alert maximises on the 28 July, and recovery to quiet time conditions is very slow.

7.3.3 Storm III The daytime $f_p F_2$ peaks at Resolute Bay and Churchill are absent on 20 January (FIGURE 7.3b). The Resolute Bay $f_p F_2$ at this time is reduced to 3-4 MHz. At Churchill the day-time value of $f_p F_2$ varies randomly near its night-time mean value of 5-6 MHz, a reduction of nearly 7 MHz from the mean daytime value! The Churchill daytime peak has returned by 21 January but $f_p F_2$ decreases early in the afternoon, possibly as a consequence of an extended mid latitude trough. At Ottawa and South Uist the changes to $f_p F_2$ are much less pronounced, the South Uist peak $f_p F_2$ is unchanged. Both decrease 1-2 hours early in the afternoon as a consequence of the mid latitude trough. The night-time $f_p F_2$ at Ottawa is reduced by ~2 MHz on 21 January. A peak in the night-time $f_p F_2$ at South Uist, 2 MHz above the night-time mean $f_p F_2$, at 00-01 UT on 21 January is possible evidence for a very extended auroral oval enhancing the F region electron density at relatively low latitudes. The two higher latitude ionosondes, Resolute Bay and Churchill, are most greatly affected during this storm period, indicating the greatest decrease in $f_p F_2$ near the auroral oval and within the polar cap. At the mid latitude stations Ottawa and South Uist, little negative storm effect is expected as these stations are in the morning sector when the storm commences (JONES, 1971).

The greatest decrease in the MUF is observed at Fairbanks, with the daytime peak on the 21 January totally absent, corresponding to the decrease in $f_p F_2$ observed at Resolute Bay (FIGURE 7.4c). During quiet times the highest frequency propagating to Fairbanks at noon is typically 20 MHz; at noon on 21 January this is reduced to between 7 and 10 MHz. Ignoring layer height variations this corresponds to a $>50\%$ decrease in $f_p F_2$ and a decrease of $N_p F_2$ to less than 70\% of quiet time values. Alert also suffers a decrease in the highest propagating frequency, though the level is difficult to determine due to interference problems. Daytime propagation to Boston and Leicester is disturbed on 20 January, corresponding to the decreased Churchill $f_p F_2$ observed at this time. The night-time MUF on the 21 January, excluding the occurrence of APEs, is very low (~7 MHz at Leicester) or propagation is almost totally absent (Boston). APEs are observed at Boston during the nights of the 21-23 July, though these become smaller each night. APEs are observed at Leicester and Aberdeen at 00-01 UT 21 January, corresponding to the night-time increase in $f_p F_2$ observed at South Uist at this time.

FIGURE 7.5b illustrates the decrease in correct signal recognition during the storm period; this is calculated from 12 UT each day to 11 UT on the following day as the initial storm disturbance occurs at approximately 12 UT on 20 January. The recognition ratio is a minimum on the 20 January, in the 24 hour period immediately following the storm commencement, followed by slow recovery to quiet time conditions which are attained on 23-24 July.
7.3.4 Storm IV  There is a 1-4 MHz reduction in the daytime $f_F2$ observed at South Uist on 1, 2, 3 and 6 February, the greatest reduction occurring on 3 February corresponding to the period of minimum $D_n$ (Figure 7.3b). There is also evidence for periods of enhanced night-time $f_F2$, 2-3 MHz above the night-time mean, on 2, 4, and 5 February, possibly due to precipitation in an extended auroral oval. At Resolute Bay $f_F2$ is more variable during this period than during quiet periods. The background level of $f_F2$ is reduced below the mean diurnal variation, though with short, intermittent periods during which $f_F2$ is enhanced over the mean value by 4-5 MHz, especially near the dayside. This is possibly a consequence of convection of enhanced F region electron density plasma from lower latitudes into the dark polar cap. However, this enhanced ionization is not in the form of a well defined tongue of ionization covering large regions of the polar cap as predicted by Knudsen (1974), Sojka and Schunk (1983), and other workers, but as ionization patches (Buchau et al., 1983; 1985; Weber et al., 1984). During the interval that $D_n$ is a minimum (2 and 3 February) the ionization patches have a lower $f_F2$ than during the rest of the storm interval (1, 4 and 5 February). This is expected if the source of the enhanced electron density plasma is the mid latitude dayside, as here $f_F2$ is depressed due to the mid latitude negative storm effect. At Ottawa there is no obvious divergence of $f_F2$ from the mean diurnal variation, though the variation is poorly defined during the daytimes of 2 and 3 February. No data is available for Churchill at this time.

The variation in the MUF is difficult to determine at Alert due to interference problems, but at Fairbanks the MUF is depressed throughout the storm period and below the MUF the occurrence of correctly recognized signals is much reduced, especially on the 3 February (Figure 7.4d). This indicates a very low background electron density with sporadic occurrence of patches of enhanced density within the polar cap. APEs occur at Boston on 1, 3 and 6 February, and at Boston and Leicester the daytime propagation is disturbed.

The recognition ratio from each path (Figure 7.5b) is a minimum on 3 February, corresponding to the minimum in $D_n$. Unlike storm periods I, II, and III, there is no sudden commencement of the storm but a gradual decrease in $D_n$ due to increased substorm activity. This is reflected in the variation in $f_F2$ observed at the ionosonde stations and the recognition ratio on the propagation paths.

7.3.5 Comparison of observed storm effects  During storm period II the Ottawa ionosonde shows the classical mid-latitude ionospheric response to a magnetospheric storm: a positive storm effect during the afternoon, followed by a negative storm effect during the night (Figure 7.3a). The positive storm effect is expected to be more pronounced at Ottawa than at South Uist as it is at a lower geographic latitude and consequently the solar ionization rate is higher, allowing a rapid increase in electron density (see chapter 2). The occurrence of positive storm effects at Ottawa does not give rise to enhancements of the MUF on the Boston path, however, as the increase in electron density is confined to latitudes
Chapter 7

Geomagnetic storms

equatorward of the path midpoint; the propagation behaviour is controlled by the ionospheric storm response near the Churchill ionosonde (i.e. at auroral latitudes). Increases in the MUF are observed on the Boston and Leicester paths, but these occur during the night as a consequence of ionization by enhanced precipitation in the auroral oval, i.e. APEs. Occasionally the enhancement occurs in the F region (e.g. 00-01 UT 21 January, simultaneous observation of fF2 increase at South Uist and APEs on the paths to Leicester and Aberdeen), but more usually the enhancement occurs in the E region, observed as auroral E at Churchill. This is the storm-time auroral zone electron density enhancement predicted by SOJKA and SCHUNK (1983), though occurring more intensely in the E region than at altitudes considered by the SOJKA and SCHUNK model (160 km and above).

All four storms periods gave rise to disturbed propagation conditions, lowering the proportion of signals received that were correctly recognized. This is due to a) the disturbed nature of the ionosphere which increases Doppler spreading and fading of the signals (see below) and b) the narrowing of the frequency range over which propagation can occur, mainly as a consequence of a decrease in the F region electron density. The occurrence of recognized signals during storm periods can be as low as 0.25 to 0.5 of quiet time levels, even during the mild storm conditions experienced during the two experimental campaigns (minimum $D_I$ between -40 and -120 nT). The severity of the propagation degradation does not appear to be strongly dependent on the minimum $D_I$ occurring during a storm (see TABLE 7.3 and FIGURE 7.6): even the mildest storm observed ($D_I = -40$ nT) produces a propagation degradation to 0.4 of quiet time levels on the Leicester and Fairbanks paths. Recovery to quiet time levels generally occurs 2-4 days after the storm commencement.

<table>
<thead>
<tr>
<th>Storm</th>
<th>$D_I$ (nT)</th>
<th>Alert</th>
<th>Fairbanks</th>
<th>Boston</th>
<th>Leicester</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>-50</td>
<td>0.7</td>
<td>0.4</td>
<td>0.6</td>
<td>0.3</td>
</tr>
<tr>
<td>II</td>
<td>-40</td>
<td>0.5</td>
<td>0.4</td>
<td>-</td>
<td>0.4</td>
</tr>
<tr>
<td>III</td>
<td>-120</td>
<td>0.9</td>
<td>0.3</td>
<td>0.8</td>
<td>0.5</td>
</tr>
<tr>
<td>IV</td>
<td>-75</td>
<td>0.8</td>
<td>0.4</td>
<td>0.5</td>
<td>0.6</td>
</tr>
</tbody>
</table>

The degradation of the propagation on the Boston path appears low in comparison to Fairbanks and Leicester (FIGURES 7.5a and b). This is a consequence of the offset of the daytime degradation of propagation by the increased occurrence of night-time APEs during disturbed conditions. FIGURE 7.7 illustrates the number of correctly recognized signals propagating on frequencies above 14 MHz during APEs (the APE "magnitude") on the path to Boston during the winter campaign. The APE magnitude increases during storm times (16-17 January, 21-22 January, 1-3 and 6 February). The largest APE observed occurred on 17 January following the storm intensification at 18 UT 16 January (see §7.2). A significant APE occurred on 27 January during a period when no storm-time variation is observed in
FIGURE 7.6  A comparison of minimum recognition ratio as a function of minimum $D_{st}$ during the four storm periods for the paths to Alert, Fairbanks, Boston and Leicester.
FIGURE 7.7 The $D_x$ variation for the winter campaign (top panel), and APE "magnitude", the number of correctly recognized signals on frequencies above 14 MHz propagating during APEs, on the path to Boston for the same period (bottom panel).
Dn, though $K_p$ reaches 4 at this time. These results indicate that APEs are not solely storm-time phenomena. Recalculating the Boston recognition ratio for the daytime only (12-23 UT) so as to remove the contribution from night-time APEs, indicates that the decrease in signal recognition during the day approaches the value for the Leicester path. Figure 7.8 illustrates the re-determined Boston recognition ratio for storm period III, together with the original Boston recognition ratio and the Leicester recognition ratio. It must also be noted that on the Boston and Leicester paths, especially during the winter campaign, the predicted MUF is much greater than the maximum frequency of the NONCENTRIC experiment. Consequently, decreases in the MUF are not observed on these two paths and the decrease in the recognition ratio will be an underestimate of the true decrease in the available HF band.

The recognition ratio results from Alert during the summer campaign are difficult to interpret as they are dominated by the presence or absence of scatter propagation on the higher frequencies (see chapter 6), which seems dependent of factors other than simple storm/quiet conditions. During storm period I the scatter propagation ceases, and returns once quiet conditions are re-attained (23 July). However, although the occurrence of scatter propagation diminishes during storm period II, it does not reappear until 30 July, long after the propagation on the other paths has returned to normal.

The path to Fairbanks is severely affected during storm times, the MUF being greatly reduced in agreement with the decrease in $f_pF_2$ observed at Resolute Bay. Sporadic increases in the MUF above the storm-reduced level can occur during these times, indicating the presence of polar cap patches, but these are generally limited to frequencies below the quiet-time MUF. The large increase in the polar cap electron density (the tongue of ionization) predicted by SOJKA and SCHUNK (1983) as a consequence of enhanced convection during storm times is not observed.

The storm periods I and IV indicate that the level of propagation degradation is dependent on the level of IMF-magnetosphere coupling and the consequent level of magnetosphere-ionosphere coupling. It does not so much depend on the initial disturbance of the geomagnetic field produced by the solar wind shock wave. During storm period I the propagation deterioration is greatest on the second day of the storm following the storm intensification produced by the increased southward component of $B_z$. During storm period IV the propagation degradation maximises on 3 February when $D_n$ is a minimum. The decrease in $D_n$ during this period and the storm time disturbance are not a consequence of a solar wind shock, but are due to enhanced magnetosphere-ionosphere coupling manifested by a high occurrence of substorms.

Signals received during storm times are in general characterized by high spread indices, reflecting the correlation between spread index and $K_p$ determined in chapter 4. Figures 7.5a and b illustrate the spread index ratio, the ratio of the mean spread index of correctly recognized signals on a disturbed day and a quiet day, for the four storm periods. In all cases, except storm period IV, the spread indices of each path increase on the first or second day following the storm commencement. Such signal distortion contributes to the poor
FIGURE 7.8 The recalculated recognition ratio to remove contribution from APEs (see text for details) for Boston, and the original recognition ratios for Boston and Leicester.
signal recognition experienced during storm periods. The variation in spread index is less clear than the variation in recognition ratio during storm periods.

7.4 Storm-time absorption in the polar cap

On the polar cap path to Alert, during the summer campaign, most signals received on frequencies of ~10 MHz and below are correctly recognized. However, during both storm periods I and II recognition fails on frequencies below ~6 MHz. A similar, though less severe, effect is observed on the polar cap path to Fairbanks, but not on the trans-auroral paths to Boston and Leicester. These effects are not observed during the winter campaign storm periods III and IV, even though the $D_s$ minimum during these two periods is lower than during periods I and II. The observations from storm periods I and II are described in §7.4.1 and §7.4.2 and these are then discussed in §7.4.3.

7.4.1 Storm I

FIGURE 7.9a illustrates the variation in signal recognition (top panel) and the variation in signal level (blue curves), noise level (red curves) and the mean diurnal variation of the signal level (green curves) observed on four frequencies at the lower end of the HF spectrum (central panels) for the path to Alert during storm period I. The mean diurnal signal level variation was determined from the quiet periods immediately preceding and following the storm. Also indicated is the $D_s$ variation for the same period (bottom panel). During 20-21 July recognition fails on frequencies between 3 and 7 MHz, indicating an increase in the path LUF. The LUF begins to increase 20 hours before the associated SSC (0311 UT, 21 July) at 07 UT 20 July, and returns to its pre-storm level 42 hours later. The failure to recognize the lower frequencies is a consequence of increased absorption within the polar cap, the level of absorption over the mean signal level reaching ~40 dB on 5.200 MHz at 20 UT on 20 July (i.e. before the SSC). On the lowest frequencies the signal levels saturate at the noise floor (e.g. between 12 UT 20 July and 18 UT, 21 July on 3.185 MHz) and consequently the true level of absorption at these times cannot be determined. The level of absorption decreases dramatically at ~00 UT 21 July (especially at ~5 MHz) due to a diurnal variation of the absorption which is more obvious during storm period II (see §7.4.2). The absorption increases again at 04 UT 21 July, probably as a consequence of SSC-associated absorption (HOLT, 1968). The path to Fairbanks, also a predominantly polar cap path though at a lower latitude, displays an increase of its nighttime LUF as well, though the level of increase is difficult to determine (FIGURE 7.9b, the maximum in signal level during the night is absent on 21 July at frequencies 6.800 and 6.905 MHz, and is decreased on 10.195 MHz). Here again, the increase in LUF occurs several hours before the SSC.

7.4.2 Storm II

FIGURE 7.9c illustrates the variation in signal recognition and signal level during storm period II. From the start of the storm, 26 July, the LUF increases, the daytime
FIGURE 7.9a The signal recognition results for the path to Alert during storm period I, summer campaign (top panel), and signal level (blue curves), mean diurnal signal level variation (green curves), and noise level (red curves) for 3.185, 3.230, 4.900 and 5.200 MHz (middle four panels) and the $D_n$ variation (bottom panel) for the same period.
FIGURE 7.9b The signal recognition results for the path to Fairbanks during storm period I, summer campaign (top panel), and signal level (blue curves), mean diurnal signal level variation (green curves), and noise level (red curves) for 5.200, 6.800, 6.905 and 10.195 MHz (middle four panels) and the $D_n$ variation (bottom panel) for the same period.
FIGURE 7.9c The signal recognition results for the path to Alert during storm period II, summer campaign (top panel), and signal level (blue curves), mean diurnal signal level variation (green curves), and noise level (red curves) for 3.185, 3.230, 4.900 and 5.200 MHz (middle four panels) and the $D_s$ variation (bottom panel) for the same period.
maximum of the LUF maximising on 27 and 28 July, and gradually decreases throughout the rest of the storm period. Also, signal levels are depressed from their undisturbed mean variation at all times of day on all four frequencies shown. The absorption increase starts ~12 hours after the commencement of the storm (~00 UT, 26 July) at approximately 12 UT, and reaches a maximum during the following day (~14 UT, 27 July) before decreasing gradually to pre-storm levels over a period of 4 days (~00 UT, 1 August). The depth of the absorption is impossible to determine on the lowest frequencies due to saturation, but at 5.200 MHz on the path to Alert the absorption is ~30 dB at its maximum. The level of absorption displays a diurnal variation being a minimum near 05 UT, local midnight, and a maximum near 17 UT, local noon. This variation is due to the attachment of D region electrons to molecules such as O\textsubscript{2} during the night and their rapid photo-detachment under conditions of high insolation during the day (DAVIES, 1990; HOLT, 1968). This event is again observed at Fairbanks, but not at Boston or Leicester.

7.4.3 Comparison and discussion of absorption observations The storm-time absorption events described affect polar cap paths more strongly than trans-auroral paths, which suggests a link with polar cap absorption events (PCAs). PCAs are produced by ionization of the polar cap D region by very energetic solar wind protons, produced by the sun at the same time as the solar wind shock waves which can produce geomagnetic storms (HOLT, 1968). These energetic protons can stream ahead of the solar wind shock wave and be incident on the earth hours before the storm commencement. PCAs are usually identified as absorption events producing complete black-out of the HF spectrum, but this is not the case in the two events observed. Consequently, this indicates the existence of low intensity PCAs which do not produce complete HF black-out, but which can produce a significant narrowing of the available HF band, especially when combined with a storm-time decrease in the MUF. Major PCAs occur approximately once a month at solar maximum and less often at other times (COLLINS et al., 1961), but the two minor PCAs observed commence within one week of each other, and occur during very mild storm conditions.

The two events observed differ considerably in their duration and time of commencement relative to their corresponding storm onsets. The first PCA precedes the SSC by several hours, the second does not begin until after the onset of the storm (as determined from the $D_n$ variation). The first PCA lasts for a short period compared to the duration of the ($D_n$ determined) recovery phase of the storm. Although there is a considerable ring current intensification at 02 UT 22 July shortly before the end of the PCA (as a consequence of an increased southward component of the IMF), the absorption is unaffected. The duration of the second PCA is many days longer than the storm recovery phase, even though storm II is weaker than storm I. $D_n$, then, is a poor indicator of the severity or duration of PCA events. The duration is determined by the “storage” of solar particles in interplanetary space. Irregularities in the interplanetary magnetic field cause scatter of the solar protons (mean free path ~0.05 AU) which then propagate to the earth by diffusion (HOLT, 1968). The
Chapter 7

Geomagnetic storms

86

travel time of the particles to the earth can be hours or days giving rise to a long duration PCA. The difference in the times of the maxima of the two PCAs relative to the storm commencements is possibly a consequence of the positions on the sun of the two solar flares or coronal mass ejections (CMEs) initiating the storms. Due to the hosepipe spiral structure of the solar magnetic field, particles ejected from the western edge of the solar disc can stream ahead of the solar wind shock front along the magnetic spiral and be incident on the earth; in such cases PCA maximum occurs before the storm commencement. Solar particles ejected from the eastern edge of the solar disc, however, are carried away from the earth by the magnetic spiral. In this case, only particles trapped behind the shock front are carried to the earth giving rise to PCAs which maximise after the storm commencement (HAURWITZ et al., 1965; HOLT, 1968).

No PCAs are observed during the winter campaign. No PCA is expected during storm period IV as the disturbance is attributed to southward orientation of the IMF and not to the occurrence of a solar wind shock wave with associated energetic solar protons. Storm period III, however, might be expected to display a PCA. However, during winter the polar cap is in constant darkness and the negative ions present in the D region do not photodissociate to increase the daytime D region electron density as inferred in §7.4.2. Consequently, the D region electron density remains relatively low and the absorption is low.

7.5 Summary

$D_s$ provides a good indication of storm periods when degradation of HF propagation can occur. However, it gives a poor estimation of the intensity or duration of the storm time degradation of the propagation. The magnitude and nature of MUF reductions or enhancements and the duration of PCA events are poorly correlated with $D_s$, minimum values or the duration of the estimated $D_s$ recovery phase. The response of the ionospheric-magnetospheric system to solar wind and IMF inputs is highly complex and it appears that no two storms perturb high latitude HF propagation in exactly the same manner. Consequently, only general conclusions can be drawn from this study.

Propagation is disturbed during storm times, the proportion of signals which are correctly recognized decreasing, especially during the day. The occurrence of APEs increases during storm times, especially on the path to Boston, as a consequence of increased precipitation within the auroral oval. APEs increase the time available for propagation on higher frequencies (>10 MHz) during the night. Positive storm effects, observed at the lower latitude ionosonde stations of Ottawa and South Uist, occur equatorward of the midpoints of the oblique propagation paths and consequently do not produce enhancements in the MUF. During storm times the mid latitude trough becomes extended, decreasing the time available for trans-auroral higher frequency (>10 MHz) propagation during the day. Within the polar cap the F region electron density is depressed
during storm conditions, decreasing the MUF. Patches of enhanced plasma convecting from
the lower latitude dayside can enhance the maximum frequency of propagation above this
depressed level, but predominantly during the winter and only for short periods of time (~1
hour). Significant absorption of lower frequencies propagating within the polar cap
increases the LUF and further reduces the frequencies available for propagation.

The level of storm-time disturbance appears most dependent on the strength of the
coupling between the IMF and the magnetosphere, controlled by \( B_z \), as opposed to the initial
disturbance produced in the magnetosphere by the solar wind shock. This is observed
during storm period I when the propagation disturbance increases after the storm
intensification produced by an increase in the southward component of the IMF. Also, storm
disturbances can be present in the absence of a solar wind shock, due to an enhanced level
of IMF-magnetosphere coupling, such as during storm period IV.

Storm time disturbance of high latitude HF propagation appears to be the rule as opposed
to the exception, over 50% of both summer and winter campaigns experiencing disturbed
propagation conditions. Even during mild storm conditions the proportion of the HF
spectrum available for communication per day can decrease to 0.3-0.5 of that available
during quiet, non-disturbed periods.
Chapter 8
Summary and suggestions for further work

8.1 Introduction
The objective of the NONCENTRIC experiment was to investigate the behaviour of HF radio propagation in the high latitude ionosphere. The experiment comprised five propagation paths with ground ranges of between 1200 km and 3800 km. Two of the propagation paths, from Clyde River to Alert and Fairbanks, were contained predominantly within the polar cap, and three, from Clyde River to Boston, Aberdeen and Leicester, were trans-auroral at all times. Measurements of signal level, noise level, and Doppler spreading were made on fourteen frequencies in the HF band between 3 and 23 MHz. A simple signal recognition test, based on the Morse code call sign CZB, was applied in order to establish whether the received data consisted of a NONCENTRIC transmission, noise, or interference. Two 24-day experimental campaigns were conducted in the summer of 1988 and the winter of 1989.

The data were correlated with several geophysical data sets such as the geomagnetic indices $D_{st}$ and $K_p$, vertical ionosonde observations from four high and mid latitude stations, geosynchronous orbit particle and magnetic field observations from the LANL and GOES-7 spacecraft, and solar wind and interplanetary magnetic field observations from the IMP-8 spacecraft. These data sets allowed the state of the magnetospheric-ionospheric system to be determined.

8.2 Summary and conclusions
The propagation characteristics on all paths monitored follow a diurnal pattern governed by solar illumination of the ionosphere. However, deviations from the solar-controlled behaviour were observed and these differ from path to path, the greatest differences being between trans-auroral and polar cap propagation paths.

8.2.1 General diurnal behaviour
The diurnal behaviour of the propagation on the five experimental paths is controlled predominantly by the solar illumination of the ionosphere. The increase in the daytime electron density of the ionospheric F region allows propagation on the longer paths of frequencies up to 20 MHz in the summer and much in excess of 20 MHz in winter. A similar daytime increase in the D region electron density increases the level of non-deviative absorption and consequently the lowest usable frequency becomes a maximum (typically ~10 MHz on the longer paths) at local noon.

On polar cap paths the MUF can be greater than expected due to the presence of
ionospheric irregularities which can support off-great circle path forward scatter propagation. On trans-auroral paths, where the path midpoints are at lower latitudes, the variation in the MUF is consistent with observations of the E and F region critical frequencies.

The behaviour, both solar-controlled and otherwise, is more complicated and less predictable on longitudinally extended paths. On meridional paths, propagation on frequencies between the LUF and MUF is very reliable, especially during the day. On longitudinally extended paths, however, the daytime propagation is much more sporadic, a lower proportion of signals in the available HF band being correctly recognized.

8.2.2 Trans-auroral paths Precipitation within the auroral oval produces enhancement of the ionospheric electron density at D, E or F region altitudes, depending on the energy of the precipitating particles. E and F region electron density enhancements can be beneficial to trans-auroral propagation, increasing the MUF and hence the frequency band over which communication can take place. In contrast, D region enhancements can be detrimental, increasing absorption on lower frequencies, and raising the LUF.

Precipitation within the auroral oval produces disturbance of the auroral ionosphere and consequent Doppler spreading of signals. The spread index, a measure of the frequency range over which the signal's energy is spread, increases with increasing geomagnetic activity on trans-auroral paths.

The mean signal strength of the lower frequency (<10 MHz) radiowaves decreases with increasing $K_p$. This is a consequence of the enhancement of the D region electron density by >30 keV electrons, producing an increase in non-deviative absorption of trans-auroral radiowaves. The spatial pattern and geomagnetic behaviour of >30 keV electron precipitation modelled by an extension of the HARDY et al. (1985) electron precipitation model, are consistent with riometer observations of auroral absorption and precipitation observations. The model predicts an increase in the level of precipitation with geomagnetic activity, similar to that observed experimentally.

The intensity of this high energy precipitation, whose source is the ring current trapping region, is closely correlated with the occurrence of substorms and consequently produces substorm correlated absorption bursts (SCABs), periods of high absorption (~30 dB at 7 MHz) lasting 15-60 minutes. SCABs occur predominantly in the midnight and morning sectors of the auroral zone, consistent with previous observations of the local time distribution of auroral absorption (e.g. HARTZ et al., 1963) and the drizzle and splash precipitation regions (HARTZ and BRICE, 1967; CRAVEN, 1970). The level of absorption experienced during a SCAB is correlated with the magnitude of changes in the magnetotail field angle at geosynchronous orbit during the substorm expansion phase and also with $K_p$. The frequency of occurrence of substorms also increases with $K_p$. Consequently, it is concluded that periods of high geomagnetic activity are a consequence of both an increased occurrence of substorms and an increased substorm magnitude.
Chapter 8

Summary

90

Models of auroral absorption (e.g. FOPPIANO and BRADLEY, 1983), generally parameterized by 3-hourly $K_p$ or other low time-resolution geomagnetic indices, are time-averaged and consequently underestimate the peak absorption which is possible during SCABS and overestimate the absorption experienced at other times.

Precipitation of 1-10 keV particles from the magnetospheric plasma sheet produces an enhancement of the E region, so called auroral E, and, less often, the F region. This allows propagation on frequencies above the mean diurnal MUF, especially during the night when precipitation in the auroral oval is most intense. Such propagation can occur on frequencies up to 10 MHz above the quiet MUF and lasts typically for periods of 2-4 hours. These are referred to as anomalous propagation events (APEs). On average APEs occur during ~50% of nights, though the occurrence increases with geomagnetic activity and during geomagnetic storm periods. Propagation by means of APEs on 3000 km paths requires patches of auroral or sporadic E of over 1500 km in diameter. Auroral E is observed under the auroral oval on almost 100% of nights; the lower occurrence of APEs indicates that auroral E formations are less likely to be sufficiently widespread to support propagation. As APEs are supported by the E region, they are unaffected by depletions in the F region electron density which occur during ionospheric storms. The signal levels of APEs are lower than typical daytime signal levels. The spread indices of APEs are low also, even though APEs occur during geomagnetically disturbed periods when daytime spread indices are typically high. This is further evidence of E region propagation as opposed to the normal F region reflections.

The mid latitude trough plays a major role in controlling the availability of higher frequencies for propagation, especially in the evening sector. On paths where the midpoint is close to the latitude of the mid latitude trough, propagation ceases when the trough opens. If the trough moves equatorwards of the path midpoint as the evening progresses, propagation can recommence while the F region electron density is sufficient to support it. The time at which the trough affects the propagation is dependent on geomagnetic activity, propagation being disrupted earlier in the evening or afternoon as the level of disturbance increases.

8.2.3 Polar cap paths In winter, the polar cap is in constant darkness and the background electron density is low, resulting in a low MUF. Patches of higher F region electron density plasma are carried into the polar cap from lower latitudes by the polar cap convection pattern, allowing sporadic propagation on frequencies higher than the predicted MUF. Such propagation occurs predominantly during periods of geomagnetic disturbance and southward IMF. During geomagnetic storms when the mid latitude F region electron density is depressed the electron density of the patches decreases.

During the summer, the MUF on polar cap paths can be enhanced above the level predicted from ionosonde observations of the F region electron density, by scatter propagation from irregularities in the polar cap ionosphere. During geomagnetic storms,
when the polar cap F region electron density is depressed, scatter propagation becomes less common. On shorter polar cap paths the maximum frequencies propagating via the E and F regions (discounting scatter propagation) are similar, and consequently during storms, when the F region electron density is depressed but the E region remains unaffected, E region modes can maintain the propagation on frequencies above the MUF predicted for the depleted F region. The very common occurrence of scatter propagation, especially on the path to Alert, indicates an abundance of irregularities within the polar cap.

During even mild storm periods, absorption on polar cap propagation paths increases on the lowest frequencies. This appears to be due to weak PCA events.

8.2.4 Storm-time propagation Prolonged periods of enhanced $K_p$ are correlated with characteristic geomagnetic storm variations in $D_s$. Such periods are correlated with global disturbances in the ionosphere known as ionospheric storms. These are manifested as decreased F region electron densities in the polar cap and at mid latitudes, and hence by decreased daytime MUFs. In the polar cap the storm disturbance can be accompanied by an increase in absorption on the lower frequencies, leading to an increase in the LUF. Even very mild storm conditions are accompanied by a significant narrowing of the HF band available for propagation, decreasing the number of correctly recognized signals per day by over 50% in many cases. Within this narrow HF band, Doppler spreading of signals increases and the propagation becomes less reliable. Storm-times are also accompanied by an increased occurrence of auroral E, increasing trans-auroral night-time MUFs. This can offset the degradation in daytime propagation.

In the case of storms initiated by a shock in the solar wind, the propagation degradation maximum occurs within 1-2 days of the storm commencement and is followed by a more gradual recovery to quiet conditions after 2-4 days. In the case of storms produced by enhanced coupling between the IMF and the magnetosphere, the level of propagation degradation is dependent on the level of coupling. The recovery of oblique propagation quality to quiet time conditions lags behind the recovery of the F region electron density observed with vertical ionosondes.

8.2.5 Summary The propagation characteristics of the propagating signals, e.g. the level of auroral absorption experienced and the occurrence of APEs, are very dependent on the location of the path relative to the auroral oval. Hence, it is possible that better selection of transmitter and receiver locations can make propagation paths less susceptible to auroral disturbances.

The high latitude ionosphere is almost constantly in a disturbed state, even mild geomagnetic disturbances having a profound effect on HF propagation within the polar regions. However, beneficial effects do exist, such as APEs, which can be exploited to overcome propagation degradation during disturbed conditions. As knowledge of the behaviour of the high latitude ionospheric morphology and its effects on HF propagation
improve, remedial measures can be employed - re-routing, re-scheduling, or frequency management of transmissions - to improve high latitude HF communications.

This study has succeeded in relating the observed HF propagation conditions to the underlying changes in the ionosphere and to the basic physical mechanisms which produce these variations. Although the general properties of the high latitude ionosphere are well known, this is the first detailed study in which the characteristics of a number of quite different HF propagation paths have been accounted for in terms of a consistent model of the polar ionosphere.

8.3 Further work

The results from the NONCENTRIC experiment indicate that the effect of the high latitude ionosphere on HF radiowave propagation is very dependent on the location of the propagation path. Consequently, to determine more specifically the extent of these effects a greater number and more extensive configuration of propagation paths is needed, for example paths tangential to the auroral oval. One hour resolution is insufficient to resolve the variations produced in the data by the rapidly varying ionosphere and a much higher time resolution (~1 min) is necessary. Also, the high latitude ionosphere abounds in horizontal electron density gradients (for instance the auroral oval, the mid latitude trough, and polar cap patches) and plasma irregularities, which can produce off-great circle path propagation. Direction finding (DF), determining the direction of arrival of the received signals, would be a very useful tool to aid determination of the location of morphological regions in the high latitude ionosphere. Absolute measurements of signal level need to be made so that signal levels on different paths and frequencies can be compared.

Such an experiment, the SPRINGBOARD experiment, is now being conducted. Direction finding, true signal level determination, and a higher time resolution are features of the new experiment. The higher time resolution is achieved by rationalising the content of the transmitted signals, thereby reducing the duration of transmissions to 30 seconds, and by sacrificing the frequency coverage of the experiment, by employing only three frequencies, from the upper, middle, and lower portions of the HF frequency band (approximately 5, 10, and 15-20 MHz). This allows a time resolution of 90 seconds on each of the three frequencies. A variety of trans-auroral and polar cap paths, and paths tangential to the auroral oval, nine in total, are provided by an arrangement of three transmitters and nine receivers (see FIGURE 8.1).

These improvements should allow much better investigation of many of the features discovered with the NONCENTRIC experiment. The higher time resolution will allow the temporal evolution and duration of auroral absorption to be better determined, and will provide better relative timings with the onset of the substorm activity which is the cause of the absorption. With more propagation paths it will also be possible to determine the evolution of the spatial extent of the absorption. The improved time resolution will also
FIGURE 8.1 The arrangement of propagation paths in the SPRINGBOARD experiment. The bottom two panels illustrate the location of the auroral oval, for $K_p = 3$ at 00 UT and 12 UT, relative to the propagation paths.
establish whether auroral E provides prolonged, undisturbed propagation, or sporadic, rapidly varying propagation conditions.

The direction of arrival of a radiowave forward scattered from a polar cap ionization patch is known to change as the ionization feature is carried along in the convection flow pattern (JONES and WARRINGTON, 1992). Direction finding in the new experiment will allow the investigation of propagation from such ionization patches. DF will also enable the direction of arrival of propagation from auroral E and auroral F to be measured, and so establish if these provide off-great circle path propagation on trans-auroral paths or paths tangential to the auroral oval. It will also help identify the regions of the polar cap from which scatter propagation is possible.

The initial study which the NONCENTRIC experiment has identified the major deviations of high latitude HF propagation from solar-controlled behaviour. The limitations of the NONCENTRIC experiment have been identified and these will be rectified, and the present work consolidated and expanded upon, in the newly deployed SPRINGBOARD experiment.
References


BAKER, D. N., E. W. HONES, P. R. HIGBIE, R. D. BELJAN, and P. STAUNING, Global properties of the magnetosphere during a substorm growth phase: a case
References


References


References


References


References


References

PARKS, G. K., Physics of Space Plasmas, Addison-Wesley, California, 1991.


References


References


THRANE, E. V., Media effects on systems in the high latitude region, in Propagation effects on military systems in the high latitude region, AGARD Conf. Proc., 1981.


WHALEN, J. A., A quantitative description of the spatial distribution and dynamics of the energy flux in the continuous aurora, J. Geophys. Res., 88, pp. 7155-69,
References

1983.


