

The effect of the high latitude ionosphere on superresolution direction finding

Thesis submitted for the degree of
Doctor of Philosophy

Carmine Rizzo
Department of Engineering
University of Leicester

June 2002

UMI Number: U149912

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U149912

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

In memory of my father, Luciano

The effect of the high latitude ionosphere on superresolution direction finding

Carmine Rizzo

Abstract

The direction-of-arrival of HF signals (9.292 MHz) propagated on a 2100 km South-to-North path in the high latitude ionosphere were investigated on three geomagnetically quiet days in January 1996. A simple Beam forming and some superresolution direction finding (SRDF) algorithms, Loaded Capon, Iterative Null Steering (INS) and MUSIC, were utilised to measure the azimuth, elevation and amplitude of this signal. The validity of the SRDF Spread Maximum Likelihood (SML) algorithm was tested. The array aperture and shape was changed by using 12 and 8 antenna arrays. The performance of the five algorithms was compared under a wide variety of propagation conditions.

The effect of the integration time and the time over which the estimates are averaged on the measurements has been evaluated. For a fixed integration time value, an increase in the time/average resulted in a decrease in the standard deviation of the bearing estimates for all algorithms and array geometries.

Simulations of enhanced ionisation structures (patches) drifting across the great circle path agreed well with the observed azimuth and elevation swings. The simulated trajectories were often consistent with the direction of the convection flow.

The behaviour of experimental observations was reproduced by two simulations. In the first, the effect of one signal spread in azimuth was studied. In the second, two spread signals were modelled with a variety of azimuth spreads, angular separation and relative amplitude. While Loaded Capon and INS were able to resolve the two spread signals at angular separations close to and above the signal spread, MUSIC succeeded for unspread signals only.

For Loaded Capon and INS secondary energy arose which closely tracked the stronger signal - a feature commonly observed in the data. The theory of the effect of non-zero bandwidth signals on the two largest eigenvalues of the covariance matrix (Zatman, 1998) has been applied to the case of an angular spread source and numerical results from this compared to simulations.

Acknowledgments

The first person I would like to mention is my supervisor Dr. Alan Stocker. I benefited greatly from his expertise and his patience in correcting my English. Thanks.

I would also like to thank Prof. Mark Lester and Dr. Declan Bates from Leicester University, and Robin Frewer, my manager at Nortel Networks, for their disinterested help.

Many of my old friends have played a crucial role in my academic career, and in particular Andrea Francini, Daniele Romano, Franco Seminaroti, Marco Siccardi, Salvatore Nuzzo and Tonino Romano. Thanks.

At Leicester University I met many new friends who helped me a lot to make my time in Leicester enjoyable. A collective thanks to my University, football and pub friends Adrian Beck, Clive Jackson, Cristina Bates, Darren and Helena Wright, David Brown, David Neudegg, Dell Baines, Desmond Sim, Emma Woodfield, Eric and Deb Bradshaw, Hina Khan, Igor Nasyrov, Jackie Davies, James Dobson, Jim Wild, Jo Swingler, John Bagley, John Sutherland, John Williams, my "meate" Jonny Rae, Kathryn McWilliams, Ken & Wendy, Lee Solomon, Lisa Baddeley, Lorraine & Michelle, Mark Rowlinson, Martin Roderick, the beautiful Mrs. Arlene, Neil Arnold, Nigel Wade, Nikolay Zaalov, Our Mick, Pat & Sherryl who made me travel a lot, Paul Eglitis the Egg, Peter Chapman, Phil Kirk, Ranvir Dhillon, Rob Dickens the Arctic, Rob Kelly, Selim Chunara, Simon Draper, Shuk Kwan Liu, Stan P. Parker, Steve Hopkins, Steve and Gabby Milan, Tim and Emma Yeoman, and others who I am probably forgetting at the moment.

This thesis is dedicated to my parents Tina and Luciano, who always supported me in my path.

Last but on top of my thoughts, Roberta. Without her constant love and encouragement this work would not have been possible. I love you.

CONTENTS

Chapter 1. Introduction

1.1	Introduction	1.1
1.2	The ionosphere	1.1
1.3	Propagation of HF signals through the ionosphere	1.2
1.4	The high latitude ionosphere	1.7
1.5	The effect of drifting large-scale density structures on the direction of arrival of HF signals	1.11
1.6	Ionospheric sounding	1.13
1.7	Summary	1.16

Chapter 2. Review of direction finding techniques and superresolution direction finding algorithms

2.1	Introduction	2.1
2.2	History of direction finding techniques	2.1
2.3	Resolution techniques	2.5
2.4	Beam forming process	2.7
2.5	Superresolution direction finding (SRDF) techniques	2.9
2.6	Superresolution direction finding (SRDF) algorithms	2.10
2.6.1	The Maximum Likelihood Methods	2.11
2.6.2	The Maximum Entropy Methods	2.16
2.6.3	The Eigenanalysis Methods	2.17
2.7	MUSIC, Iterative Null Steering (INS), Loaded Capon and Spread Maximum Likelihood (SML) algorithms	2.20
2.8	Comparison between the SRDF algorithms utilised for this research	2.27
2.9	Visibility of multiple signals	2.31
2.10	SRDF for non-zero bandwidth and finite angular spread signals	2.32
2.11	Summary	2.40

Chapter 3. Data collection method and overview of the measurements

3.1	Introduction	3.1
3.2	Data acquisition system	3.2
3.3	Observations on 22 nd January 1996	3.5
3.4	Observations on 23 rd January 1996	3.7
3.5	Observations on 24 th January 1996	3.10
3.6	Simulation of a drifting over-dense plasma structure	3.12
3.7	Summary	3.16

Chapter 4. Behaviour of the INS, Loaded Capon, MUSIC and SML algorithms

4.1	Introduction	4.1
4.2	Algorithm and array geometry dependence of the measured azimuth of arrival	4.1
4.3	The effect of integration time on the observed bearing spread	4.6
4.4	Summary	4.9

Chapter 5. Modelling of one spread signal

5.1	Introduction	5.1
5.2	Simulation of a spread signal	5.1
5.3	The effect of changing the spread energy distribution model	5.5
5.4	Behaviour of simulated azimuth as a function of angular spread	5.6
5.5	Comparison between modelling and observations	5.8
5.6	A comparison of measured azimuth spread with that observed from ionogram and simulation results	5.10
5.7	Theory of two-signal representation from one signal spread in azimuth	5.12
5.8	Comparison between theoretical and simulation results of two-signal representation from one signal spread in azimuth	5.17
5.9	Summary	5.19

Chapter 6. Modelling of two spread signals

6.1	Introduction	6.1
6.2	Modelling of two spread signals	6.1
6.3	Visibility of two signals with varying amplitude ratio	6.2
6.4	Summary	6.6

Chapter 7. Conclusions and future work

7.1	Introduction	7.1
7.2	Conclusions	7.1
7.3	Future work	7.7

References

1. INTRODUCTION

1.1 Introduction.

This chapter aims to introduce the ionosphere as a propagation medium for high frequency (HF) radio signals. The ionospheric regions are described, with emphasis on their importance for radiowave propagation. In particular, the problems related to the propagation of HF signals at high latitudes are discussed. A discussion is presented about the convection of plasma at high latitudes and the effects of large-scale electron density structures on HF direction finding. These structures can reflect HF signals causing large variations in bearing. The accuracy of direction finding (DF) systems estimates is typically of 0.1° , whereas the deviations of the directions of arrival of HF signals at the receiver can be as much as $\pm 100^\circ$ or more at high latitudes. It is then clear that the performance of DF systems at high latitudes are mainly limited by the propagation environment.

The ionospheric sounding method for investigating the propagation conditions is introduced.

A detailed description of ionospheric radiowave propagation can be found in many textbooks, in particular Davies (1990). For a review of auroral and polar cap ionospheric effects on radio propagation, the reader is referred to a paper by Hunsucker (1992).

1.2 The ionosphere.

The Earth's atmosphere consists of several distinct regions or layers with varying attributes, and is described by properties such as level of ionisation, pressure, density, temperature and composition (see, for example, Hargreaves, 1992). The region of the atmosphere known as ionosphere is of interest for the work detailed in this thesis.

The ionosphere is regarded as the ionised part of the atmosphere, to be distinguished from the background neutral gas (Banks and Kockarts, 1973). For the purposes of radio communication, the ionosphere is defined as that part of the upper atmosphere where ionisation sufficient to affect the propagation of radiowaves exists (Davies, 1990). Solar radiation causes gases to dissociate above about 50 km. As a result, free electrons are produced in the atmosphere.

The ionosphere is divided into regions with an alphabetic nomenclature, introduced by Sir Edward Appleton (Silberstein, 1959). Some of the ionospheric regions are identified in Figure 1.2.1. The D region is the region of lowest altitude which is important in radio propagation because it absorbs energy

from waves at MF, HF and VHF, and it reflects LF and VLF signals. The D region extends from 50 to 90 km and is produced by Lyman, α and hard X-rays. The E region, covering the height interval 90 km to about 140 km, is produced by solar soft X-rays and by EUV (Lyman β radiation at 1025.7 Å and Lyman continuum at 910-980 Å). Under certain conditions a thin layer of this region becomes exceptionally ionised and is then known as sporadic E (E_s). Sporadic E is important because it can reflect VHF radiowaves up to frequencies of about 100 MHz, whilst the normal E layer can only reflect HF radiowaves. The F region (140 km and above) is very sensitive to solar radiation, and it is where the peak in electron density usually occurs. It is subdivided into the F1 and F2 layers, which are produced by solar extreme ultraviolet light (EUV). The F region also reflects HF radiowaves. Frequencies up to the critical frequency are reflected at vertical incidence and higher frequencies at oblique incidence (see Section 1.3).

Above the F region peak of electron density is the topside ionosphere, and above that is the protonosphere or plasmasphere. The altitude of the base of the protonosphere (2000-3000 km) varies with latitude and between day and night. The ionosphere can, therefore, be considered as lying between about 50 and 2000-3000 km.

The electron density depends then on the solar radiation and as such is subject to diurnal, seasonal and an 11-year solar cycle variation. With some exceptions the electron density is statistically higher in daytime than nighttime (Figure 1.2.2). For seasonal variations, the nighttime and daytime situations have to be distinguished. The nighttime F layer tends to be higher in summer than in winter, this tendency being accentuated at lower latitudes. The maximum electron density and the total electron content of the nighttime F region are greater in summer than in winter. The main feature of the daytime situation is that at middle latitudes, the peak density is considerably larger in winter than in summer; this being called the seasonal anomaly. In summer months, at middle latitudes, the F layer splits into the F1 and F2 layers. Finally, the electron density follows the trend of the number of sunspots in the 11-year sunspot solar cycle, thus affecting the range of frequencies reflected from the ionosphere (Figure 1.2.3).

1.3 Propagation of HF signals through the ionosphere.

The ionosphere is widely used as a propagation medium for HF signals, i.e. those of frequencies 3-30 MHz. The propagation of radiowaves is affected by the presence of free electrons in the ionosphere. A transmitter sends a signal to the receiver at a given frequency, through a reflection by the ionosphere. If the

transmitter and the receiver are not at the same site, the radiowave will propagate obliquely. The transmission and reflection of radio signals is better interpreted with some prior understanding of the radio refractive index of the ionosphere n , which is expressed by the Appleton formula (Ratcliffe, 1959). Ignoring the presence of the Earth's magnetic field and the presence of electron collisions, the refractive index is given by

$$n^2 = 1 - X = 1 - \left(\frac{f_N}{f}\right)^2 \quad (1.3.1)$$

where

$$X = \frac{Ne^2}{\epsilon_0 m \omega^2}$$

and

$$f_N = \frac{e}{2\pi} \sqrt{\frac{N}{\epsilon_0 m}} \quad \text{is the plasma frequency.}$$

N is the electron density, e is the charge on an electron, ϵ_0 is the permittivity of free space, m is the mass of an electron and $\omega = 2\pi f$ is the angular wave frequency.

From Equation 1.3.1 we can see that the square of the refractive index is less than unity, and therefore the same applies for n , as its value must be real. Furthermore, the refractive index decreases with increase of electron density and, for a fixed value of electron density, increases with increase of frequency.

Let us consider a radiowave incident at an angle i_0 on a plane refracting layer (see Figure 1.3.1). The electron density increases as the wave penetrates into the layer, and the wave normal follows the Snell's law:

$$n_0 \sin i_0 = n_1 \sin i_1 \quad (1.3.2)$$

where i_0 is the angle between the wave normal and the perpendicular to the surfaces of constant refractive index n_0 , and i_1 and n_1 are the corresponding values at the base of the layer.

The ionospheric medium can be regarded as a number of slabs, each of which has a uniform refractive index (n_1, n_2, n_3, \dots). Applying Snell's law to each boundary, we have

$$\begin{aligned}
n_0 \sin i_0 &= n_1 \sin i_1 \\
n_1 \sin i_1 &= n_2 \sin i_2 \\
n_2 \sin i_2 &= n_3 \sin i_3 \\
&\vdots
\end{aligned}
\tag{1.3.3}$$

Since the plasma frequency increases with height, n becomes smaller and the ray gradually bends until it becomes horizontal. Therefore, when the wave is reflected, $i = 90^\circ$, so we have

$$n_r = \sin i_0 \tag{1.3.4}$$

therefore, for vertical propagation (i.e. $i_0 = 0$)

$$n_r = 0 \tag{1.3.5}$$

From Equations 1.3.1 and 1.3.5 we can infer that reflection occurs when the plasma frequency f_N equals the wave frequency f . Hence, when the wave frequency exceeds the maximum plasma frequency (also known as critical or penetration frequency, see Section 1.6) of the region, the wave will penetrate the layer. Nevertheless the wave could be reflected by a further region, in which a higher electron density is present.

Equation 1.3.1 is not valid when the magnetic field cannot be neglected, in which case Equation 1.3.6 applies

$$n^2 = 1 - \frac{2X(1-X)}{2(1-X) - Y_T^2 \pm [Y_T^4 + 4(1-X)^2 Y_L^2]^{1/2}} \tag{1.3.6}$$

where

$$Y_T = \frac{eB_T}{m\omega} \quad \text{and} \quad Y_L = \frac{eB_L}{m\omega}$$

The subscripts T and L refer to the transverse and longitudinal components of the imposed magnetic field B , with reference to the direction of the wave normal. When the magnetic field effects are negligible (i.e. $Y \ll 1$), we obtain Equation 1.3.1.

The effect of the magnetic field is to split the incident wave into two waves, which are taken into account in Equation 1.3.6 by the \pm sign. The wave with the positive sign in 1.3.6 is called the 'ordinary' wave, since it is reflected at the same height as it would be in the absence of the magnetic field. In the case of vertical propagation ($n = 0$) the positive sign gives

$$X_r = 1 \tag{1.3.7}$$

and the negative sign gives

$$X_r = 1 - Y \quad (1.3.8)$$

for $f < f_{B,e}$ (the electron cyclotron frequency) or

$$X_r = 1 + Y \quad (1.3.9)$$

The waves with the negative sign in 1.3.6 are called extraordinary waves. The ordinary and extraordinary waves travel different paths through the ionosphere.

Several theorems relate certain propagation characteristics of two waves, one reflected with oblique incidence and the other reflected with normal incidence, from the same true height: the 'secant law', the 'Breit and Tuve's theorem' and the 'Martyn's equivalent path theorem'.

Figure 1.3.2 shows a ray incident on a plane ionosphere at an angle ϕ_o . The secant law, given by the Equation 1.3.10, relates the frequency f_{ob} of a wave reflected obliquely at a given real height to the frequency f_v of a wave incident vertically and reflected at the same real height.

$$f_{ob} = f_v \sec \phi_o \quad (1.3.10)$$

The frequency f_v is called the 'equivalent vertical frequency' corresponding to f_{ob} . It is apparent from Equation 1.3.10 that the ionosphere can reflect signals at a higher frequency obliquely than it can at vertical incidence.

The secant law needs to be modified for distances greater than 500 km as effects due to the Earth's curvature become important. A correction factor k is introduced in Equation 1.3.11 to compensate for this. The correction factor has values between 1.0 and 1.2, depending on the electron density profile and path length (Smith, 1939).

$$f_{ob} = kf_v \sec \phi_o \quad (1.3.11)$$

For a flat earth and a horizontally-stratified ionosphere with no magnetic field, Breit and Tuve's theorem states that the equivalent (or virtual) path for propagation between a transmitter T and a receiver R is given by the length of the equivalent triangle TAR (Figure 1.3.2). The time taken to traverse the actual (curved) path TBR is the same as that which would be taken to traverse the equivalent (triangular) path TAR in free space. Note that the true height of reflection (at B) is always less than the equivalent height at A and that both T and R must lie outside the ionosphere.

Martyn's equivalent path theorem expresses the important relationship that the virtual height of reflection of an obliquely incident wave is the same as that of the equivalent vertical wave.

However, the ionosphere has a structure in constant evolution and is affected by various disturbances, which are often associated, either directly or indirectly, with events on the sun. These occurrences can affect the performance of HF systems, for example by changing the maximum usable frequency (MUF) of a transmission path (see Sections 1.4 and 1.6). The ionosphere cannot be regarded as a smooth reflecting surface, and thus it does not behave as a simple mirror for HF signals. Furthermore, its structure can cause the signal to be reflected more than once from a given layer and often from different layers before reaching the receiver. The various modes in Figure 1.3.3 are named by means of a symbolism referring to the number of intermediate ground reflections and to the relevant reflecting layers (Davies, 1990). The characteristics of the received signal differ for each mode, with the elevation being particularly affected. Furthermore, more than one mode can occur simultaneously. An example of two-moded propagation is illustrated in Figure 1.3.4. When the received signal results from more than one propagation path, it will have different delays and its characteristics will be affected (such as elevation, elevation spread, azimuth, azimuth spread, Doppler and Doppler spread), and it will also be affected by fading, which consists of a temporal variation in its amplitude. We can distinguish between interference and in-mode fading. Interference fading occurs when several propagating modes with different path lengths combine at the receiving antenna to give a resultant signal amplitude. In-mode fading occurs when a single mode comprises several signals reflected from nearby points in the ionosphere, which happens because the ionosphere is not a perfectly smooth reflecting surface. Fading can have a duration from a fraction of a second to several hours, depending on the causes that originated it.

Some of the causes of change of ionospheric conditions are ionospheric storms and sudden ionospheric disturbances. Ionospheric storms are associated with geomagnetic, auroral and magnetospheric storms. Ionospheric storms are the most disruptive disturbances on radio frequencies, in particular those occurring in the F2 region. This is because they last for several days and they extend into the highly populated middle latitudes, where most ionospheric radio communications are conducted. Ionospheric forecasts are therefore of high value to both civilian and military operators, who may face situations where the need of a communication medium is crucial. Sudden ionospheric disturbances also heavily affect radio communications in the HF band, but are less disruptive than ionospheric storms because they are relatively short lived, from a few

minutes to several hours. For further information concerning the effect of solar disturbances on radio propagation, the reader is referred to the books by Gassman (1963), Agy (1970), Mitra (1974) and to the SESC Glossary of Solar-Terrestrial Terms (SESC 1988).

In order to take account of the effects of the ionospheric disturbances described above, a single ionospheric mode is often modelled as a single ray specularly reflected from a smooth ionosphere (specular component) surrounded by a cone of rays produced by the roughness of the ionosphere (scattered or diffracted component). This effect is known as coning and can be clearly seen in Figure 1.3.5. Ionospheric drifts, vertical movements and Travelling Ionospheric Disturbances (TIDs) (Davies, 1990) cause Doppler shifts on signals propagating through the ionosphere. For example, consider Figure 1.3.6, in which irregularities drift in an east to west direction. Because of this, positive Doppler shifts will be imposed on those scattered components to the east of the Great Circle Path (GCP) direction (i.e. the shortest path between the transmitter and the receiver), where the motion of the scatterers is in a direction tending to shorten the path. Negative Doppler shifts will be imposed on those scattered components to the west of the Great Circle Path direction, where the motion of the scatterers is in a direction that lengthens the path. Likewise, downward vertical plasma movements will cause positive Doppler shifts, upward movements will cause negative Doppler shifts.

1.4 The high latitude ionosphere.

At high latitudes the ionosphere may be in continuous sunlight or darkness for long periods, depending upon the season. In these circumstances the ionosphere will be typical of daytime or nighttime conditions, respectively. An example of ionosphere structure at high latitudes is shown in Figure 1.4.1. The shaded region indicates the position of the auroral oval (see also later in this section) for low magnetic activity. Inside the oval is the polar cap. The location of these regions varies considerably, and therefore the plot is derived on a statistical basis. The movements of the regions take place in a clockwise direction during the day. The average extension varies according to seasonal changes, whilst sudden alterations may occur because of sudden ionospheric disturbances like magnetic storms. Feldstein and Starkov (1967) demonstrated that the extension of the auroral oval is proportional to the intensity of magnetic storms.

Figure 1.4.2 shows an example of the location of the auroral oval at different values of the geomagnetic index K_p , which quantifies the level of geomagnetic activity. The K_p index is derived from observatory K indices for a given 13

magnetic observatory network. Each K index is a 3 h range measure of the irregular ionospheric variations associated with magnetic field disturbances. Each observatory assigns an integer from 0 (very quiet) to 9 (very disturbed) to each of eight 3 h UT intervals (00:00-03:00, 03:00-06:00,..., 21:00-24:00). The K index from each observatory is then mapped to a common scale that is intended to eliminate diurnal, seasonal and observatory differences. This scale is usually represented as 0, 0+, 1-, 1..., 8+, 9-, 9, and the values are known as K_s (sometimes a 0-27 scale of integer numbers is used, in order to avoid the + and - tags). Finally these K_s are averaged over the network of observatories to give K_p . The subscript 'p' means planetary, and therefore designates a global magnetic activity index. The index is unitless and is quasi-logarithmic. An equivalent index on linear scale, called A_p is often used in the place of K_p (US Department of Commerce, National Geophysical Data Center Web Site). For further details on this index the reader is referred to Parkinson (1983). The crucial importance of the level of geomagnetic activity in determining propagation conditions was stressed by Roesler and Bliss (1988). They concluded that ionospheric phenomena can either degrade or enhance the performance of radio propagation at high latitudes. The latter situation can even occur when the signal strength decreases, in the case that the noise decreases by a higher factor, this causing the signal to noise ratio to increase.

The high latitude ionosphere is a region subject to many disturbances, and the auroral oval region in particular. Therefore, if the reflection point occurs inside the auroral oval, the signal transmission may be highly affected. As seen in Section 1.3, many ionospheric disturbances affect HF propagation at any latitudes. Furthermore, at high latitudes other disturbances occur, such as polar cap absorption (PCA) and auroral absorption (AA) events, and strong fading. As a consequence, the overall ionospheric effects on radio propagation are more pronounced than at middle and low latitudes. This causes the ionospheric layers not to be well defined at high latitudes, and also very unstable. Therefore the high latitude ionosphere is often regarded as a rough reflecting surface for obliquely propagating HF radiowaves, its roughness being enhanced by the fluctuations discussed above. As a result of this roughness, signals can arrive at the receiver over a wide range of angles in both azimuth and elevation. Furthermore, ionospheric movements are more frequent and faster than at middle and low latitudes, thus the Doppler shift and Doppler spread imposed onto the signal are often larger at high latitudes.

Another important factor that affects the ionosphere structure at high latitudes is the magnetosphere, which derives from the interaction between the solar and geomagnetic fields, as in Figure 1.4.3. The Earth's magnetic field lines near the

poles extend thousands of kilometres into space, where they connect with the interplanetary magnetic field (IMF), which is generated by the sun. In the upper atmosphere, protons ejected from the sun spiral down the magnetic field lines, where they interact with the atmosphere causing several effects, including northern and southern lights (i.e. visible aurorae). The region of open magnetic field lines is referred to as polar cap, which is effectively open to protons from the solar wind. In particular, a solar flare causes the migration of protons, which reach the Earth in 30 minutes to several hours. The collision between protons and atmospheric gases produces ionisation, which occurs over most of the polar cap and causes intense absorption of HF radio signals (polar cap absorption, PCA). The occurrence of PCA events is proportional to the sunspot number (Landmark, 1968).

The boundary region of open field lines surrounding the cap which connect with the IMF and the closed field lines is called auroral oval (mentioned earlier in this section). In the latitudes typical of the auroral oval, a localised phenomenon named auroral absorption (AA) takes place, which occurs only over a few hundred kilometres. Its intensity varies very rapidly and it is caused by precipitated electrons with different energies depending on the latitudes (Sharp and Johnson, 1968). These electrons cause regions of increased ionisation when they penetrate down the D layer. AA usually accompanies visible aurora, but the two are not always correlated in position.

At high latitudes the rate of fading tends to be higher than at middle latitudes (Vincent *et al*, 1968; Hunsucker and Bates, 1969). Goodman J.M. (1992) showed that fades are often associated with spread F conditions. In particular flutter fading, that is fades with periods of less than 1 second, often occurs on trans-auroral paths, where particle precipitation perturbs the ionosphere to produce F layer irregularities (Pike, 1971).

The average circulation of ionospheric plasma at high latitudes is referred to as 'convection' (Lockwood, 1993). The convection controls the plasma distribution and its temporal variations, and therefore it affects the operation of HF systems, which make use of ionospheric reflections in order to obtain over-the-horizon propagation. In the E region, the plasma lifetime is of the order of a few seconds, during which free charges move distances typically less than one kilometre. As a result, plasma decays about at the same location where it was produced. However, in the F region the plasma lifetime is considerably higher, of the order of several hours, during which the plasma can travel over several thousand kilometres. Even though the plasma will move across regions of enhanced plasma production (the dayside and the auroral zones) and loss (the nightside), its density will change quite slowly, because of the long lifetime. The

plasma spatial and temporal distribution in the F region will therefore be much more complex than in the E region, and will be controlled by the convection pattern. High density plasma in the polar cap (of particular interest for this thesis) is produced by photoionisation on the dayside and subsequently convected poleward, its motion strongly depending on the season (Lockwood, 1991a). Density variations affect the operation of HF systems in a number of ways. For example, the MUF (see Section 1.6) of a point-to-point HF link can change dramatically, as it depends on the plasma density at the F layer peak (Blagoveshchenskij *et al*, 1992). Furthermore, horizontal gradients of the isotonic contours cause deviations of HF ray paths (Davies and Rush, 1985), thus causing errors in signal locations obtained by means of direction finding systems (Tedd *et al*, 1985).

The shape of the convection pattern is strongly dependent on the orientation of the IMF embedded with the solar wind flow (Richmond *et al*, 1988). Figure 1.4.4 shows sketches of the polar cap convection flow for six orientations of the IMF: the northward direction along the Earth's magnetic dipole axis is given by the positive values of the B_z component of the field, the B_x component is sunward and B_y is duskward. The centre of the six frames represents the magnetic pole ($\Lambda = 90^\circ$), and the flow patterns are as would be seen by an observer who is co-rotating with the Earth. The convection patterns in Figure 1.4.4 represent long-term averages (Lockwood, 1991b). However, Hapgood *et al* (1991) have shown that the IMF is stable for more than two hours only 15% of the time, therefore steady convection will be relatively rare, and the convection pattern at any instant will be considerably different from those illustrated in Figure 1.4.4.

Associated with the convection flow are large-scale electron density structures (Lockwood and Carlson, 1992), which are common features in the polar cap F region ionosphere. These feature are known as patches, which occur during periods of southward directed IMF ($B_z < 0$), and arcs, occurring with low geomagnetic activity and when the IMF is directed northward ($B_z > 0$). Patches and arcs have somewhat different characteristics. Patches drift antisunward at speeds typically of a few kilometres per second, with electron density enhancements of up to a factor of 10 above the background electron density in the F-region ionosphere (Weber *et al*, 1984; Buchau *et al*, 1983). Arcs are Sun-Earth aligned plasma striations having an electron density factor of 2-3 above the background electron density in the F-region ionosphere (Carlson *et al*, 1984). Arcs drift from dawn to dusk (perpendicular to their alignment), their speed being typically a few hundreds metres per second, thus much lower than the speed of patches.

1.5 The effect of drifting large-scale density structures on the direction of arrival of HF signals.

It is known that the performance of HF direction finding systems is related to the mode content of the incoming signal and the frequency of operation, both of which depend on the electron density distribution in the ionosphere, which can be highly variable at high latitudes (Davies, 1990). Therefore, the performance of direction finding systems can be very limited at high latitudes, where signals can arrive at the receiver with bearings displaced from the Great Circle Path by up to $\pm 100^\circ$ or more (Warrington *et al*, 1997a). Large deviations from the Great Circle direction for HF signals propagating within the polar cap can be caused by large-scale over-dense plasma structures, such as the commonly occurring patches and arcs, drifting across the polar cap (see Section 1.4).

The large deviations in bearing caused by these polar cap features have serious implications for the operation of position location systems operating within the HF band at high latitudes. It is important to emphasise that the accuracy of bearing measurements as an indication of the positioning of a target is usually limited by ionospheric propagation effects and not by the DF instrumental accuracy, which is typically of the order of 0.1° . The understanding of both the propagation environment and the associated behaviour of direction finding systems is of primary importance, to be able to identify and reject bearings generated or affected by the above mentioned structures. Measurements of bearing errors are shown in Table 1.5.1 (see next page), for ionospheric regions ranging from the polar cap to mid-latitudes. The magnitude of bearing errors within the polar cap in winter is of particular importance for this work, since the analysed data were collected on a polar cap path and in wintertime. From Table 1.5.1 it is evident that in the polar cap, at wintertime, the errors in bearing can be as much as $\pm 100^\circ$ and are caused by patches and arcs travelling across the polar cap. These generally occur during periods of quiet geomagnetic activity. However, bearings are very disturbed and scattered during geomagnetically active periods.

When the signal is reflected by patches and arcs, the directions of arrival are likely to cover the solid angles subtended by the moving features, and thus are not an accurate indicator of the transmitter direction. Some examples were shown by Warrington *et al* (1997b) for three high latitude paths (Figure 1.5.1). The receiver was at Alert, in the Canadian North West Territories, where a goniometric direction finding system was deployed. The signals were received from Halifax, Nova Scotia (4180 km), from Thule, Greenland (670 km) and from Iqaluit, North West Territories (2100 km). Bearing measurements taken on 18th November 1990 for the paths Halifax-Alert and Thule-Alert are illustrated in

Figures 1.5.2 and 1.5.3 respectively, and measurements taken from 21st to 24th February 1994 for the Iqaluit-Alert (I-A) path in Figure 1.5.4. The I-A path is particularly relevant, since the data presented in this thesis were also collected on this path (see Section 3.2). These examples show very large, rapid bearing swings. The bearing deviations are not random, but exhibit a periodic structure, and the number of swings is greater in the pre-noon period than in the post-noon period. These swings were then attributed to reflections from large, drifting, electron-density structures such as patches and arcs. It is important to consider that although it is not possible to correct for the type of bearing errors caused by drifting electron-density structures, it is possible to predict the periods during which they occur. These predictions could be used to assign ionospherically based weighting factors to the individual line of bearing measurements to improve the accuracy of the target position estimate.

Table 1.5.1 Magnitude of HF DF bearing errors measured in the various ionospheric regions and the geophysical conditions during which these errors occur (Warrington *et al*, 1997a).

Region	Magnitude of maximum bearing error	Cause	Times of occurrence
Polar Cap	±100° (Winter)	Travelling blobs and patches. Sun-aligned arcs	During magnetically quiet periods. For active times, bearings very disturbed and scattered.
	±10° (Summer)	Winter effects reduced due to enhanced MUFs	Irregularities comparable to ambient.
Auroral region	±100°	Tilts associated with auroral oval. Signals often lost due to high absorption	Magnetically active.
	<±5°	For dayside propagation	Magnetically quiet
Sub-auroral	±100°	Tilts associated with both polar and equatorward walls of the sub-auroral trough. Path geometry relative to the trough is important.	When $f \geq$ MUF at night when trough forms. Times of occurrence depend on A_p .
	<±5°	For dayside propagation	Mid-latitude conditions.
Mid-latitude	±2°	Tilts due to TIDs and Sunrise/set terminator	TIDs any time. Dawn/Dusk

Data relevant to the I-A path have also been analysed by Dumas (1997), who presented some results from 23rd January 1996. The same data have been analysed in the work detailed in this thesis. In the data interval investigated by Dumas, the trace of the detected signal often showed large variations in

bearings as much as 100° over 4.5 hours. This is shown in Figure 1.5.5, which illustrates the bearings of the received signal from about 05:15 to 09:45 UT on 23rd January 1996. The azimuth angle of arrival varies from a minimum of about 80° below the true bearing after 07:00 to a maximum of about 30° above the true bearing at about 09:30. Several swings in azimuth are clearly visible. The gaps correspond to the time intervals when no signal was present because the transmitter was off to allow sounder sequences along the propagation path. Dumas developed a simulation of the motion of patches, which indicated that these swings were a result of a series of successive patches of over-dense plasma drifting along approximately the same path. The speed of these structures ranged between 210 m/s and 1350 m/s in the data interval analysed by Dumas, and the direction of their motion was antisunward, as expected by patches. Dumas showed that the results of the simulation were in good agreement with those obtained from the experimental data, a result consistent with the signal being reflected by moving structures of over-dense plasma. Dumas presented evidence of multiple traces (Figure 1.5.6), most of them being closely separated, which he considered as caused by patches. When three traces were present at the same time (particularly apparent at around 06:10 and between 06:20 and 06:30) Dumas stated that three patches were contributing to the signal. However, although it is possible for drifting patches to cause large deviations from the true bearing of a received signal, it is not necessarily the case that multiple traces each represent a real signal or mode, particularly when they are closely separated. In fact, Dumas did not consider the effect of spread reflections from the ionosphere, which it is important to take into account, especially at high latitudes. Furthermore, Dumas processed the data with only one array geometry, i.e. the full 12-antenna array deployed at Alert (see Section 3.2). Multiple traces can arise from a single spread mode, the precise details depending on algorithm and array geometry (see Chapter 4 for experimental data and Chapters 5 and 6 for modelled data).

1.6 Ionospheric sounding.

The height at which a radio signal propagating through the ionosphere at a given frequency is reflected depends upon the electron density and the angle of incidence. Ionospheric disturbances affect the stability of the layers, in particular at high latitudes, where the layers are highly variable, and consequently the electron density of the ionospheric regions may often change by large amounts. Therefore, the first fundamental step in overcoming propagation problems is to gain a good knowledge of the ionospheric structure and propagation conditions along the transmission path.

For this purpose a technique known as ionospheric sounding is performed. One such sounding method consists of transmitting a signal from the ground to the ionosphere, varying its frequency in a systematic way, and recording the time-delay of the reflected signal. The trace thereby obtained is called an ionogram. When the receiver and the transmitter are either both at the same site or located near to each other, the trace will show a vertical ionogram. Vertical reflection occurs when the plasma frequency f_N equals the wave frequency f (see Section 1.3). Two examples of vertical ionograms for a middle-latitude location are illustrated in Figures 1.6.1a and b.

Figure 1.6.1a shows a summer daytime ionogram with ordinary wave traces only. Reflections are from the E layer, the F1 layer, and F2 layer at progressively higher heights and on progressively higher frequencies. Notice the group retardation near the critical (penetration) frequencies (see Section 1.3) f_oF1 and f_oF2 . Group retardation occurs when the frequency of the signal approaches the peak of electron density of each ionospheric layer, which causes its speed to decrease. This leads to an increase of the delay, hence of the virtual height (but not of the real reflection height, see Section 1.3). Figure 1.6.1b illustrates an ionogram taken on a summer evening showing one-hop and two-hop echoes from the F layer. The signals are split into two separate modes, ordinary waves (O) and extraordinary waves (X) (see Section 1.3). The ordinary and extraordinary waves have critical frequencies f_oF and f_xF respectively. Again we can see the increase of the virtual height caused by retardation in underlying ionisation at the lower frequency end of the extraordinary trace (1 hop, F layer).

Vertical and oblique ionograms enable the number of modes which are present on a particular frequency to be determined (Figure 1.6.1b), which may correspond to how many directions of arrival are detected by a direction finding system. Oblique ionograms differ from the vertical ones in that the transmitter and the receiver are deployed at two separate sites. Oblique sounding is used to infer the state of the ionosphere at the reflection point or points of an oblique propagation path, and is obtained by send a signal across a range of frequencies from the transmitter to the receiver. The trace of the delay as a function of frequency provides important information about the range of frequencies suitable for use over the signal path. If the two ends are time synchronised, the delay represents the absolute time of flight. Oblique ionograms provide information about the ionospheric layers and the presence of sporadic E ionisation, which can be utilised to determine possible modes of propagation. Furthermore, information about layer heights allows the calculation of the range of possible elevation angles of arrival. In general,

oblique ionospheric sounding is of most use for direction finding systems operators when the propagation environment is difficult for radiowaves in the high frequency band, and the properties of the reflection points are different from what can be regarded as normal propagation conditions. However, usually oblique ionograms cannot be taken over the path of interest, except in research experiments where the position of the transmitter is known, such as for this work.

Oblique ionograms are also used to identify two important characteristics of a transmission path, the Operational Maximum Usable Frequency (OPMUF) and the Basic Maximum Usable Frequency (MUF), which are defined as follows. The OPMUF is the highest frequency that would permit acceptable operation of a radio service between given terminals, at a given time and under specified working conditions (such as antennas, transmitter power, class of emission, information rate, and required signal-to-noise ratio). The MUF is the highest frequency radiowave, which can propagate between given terminals, on a specified occasion, by ionospheric refraction alone (junction frequency). An example of the diurnal variation of both frequencies is depicted in Figure 1.6.2, which refers to a sub-auroral path. Commonly, the optimum working frequency will be around 85% of the OPMUF (Davies, 1990).

An oblique ionogram taken in middle latitudes is shown in Figure 1.6.3. The traces generated by one-hop, two-hop and three-hop F2 reflections are clearly visible. The one-hop F trace shows a low-angle trace (lower portion of the trace) and a high-angle trace (higher portion of the trace), which join at the junction frequency (JF). For the one-hop F trace, the basic MUF (or F-layer junction frequency, FJF) is the same as the maximum observed frequency (FMOF), i.e. the highest frequency on which the signal is shown by the ionogram, independently of the path traversed (which occurs via an F-layer reflection, in this circumstance). The high-angle two-hop trace shows splitting into ordinary and extraordinary traces near the junction frequency (2FJF).

High latitude ionograms often show the presence of features, which are regarded as patches or arcs of enhanced ionisation (see, for example, Rogers *et al*, 2001) (see Section 1.5). Figure 1.6.4 illustrates an example of oblique ionogram with such a feature. The trace is very spread between about 4 MHz and 8.5 MHz, which means that irregularities were present when the ionogram was taken. Furthermore, there are two features at frequencies between about 8.5 MHz and 11 MHz, at relative delays of about 2.5 ms and 3.5 ms. These features are possibly caused by patches or arcs of enhanced ionisation, which can offer a favourable path to the signal, in particular in case of disturbed propagation environment. As a result, if this feature is located away from the GCP, the

signal will arrive at the receiver at an azimuth angle different from the true bearing. This over-dense plasma feature will be also likely to spread the signal in azimuth and elevation, since its volume may offer a wide range of reflection points. Furthermore, its movements are likely to shorten or lengthen the propagation path, thus causing signal Doppler shift.

Oblique ionograms repeated at regular intervals can show the evolution of the propagation conditions over a certain path. For example, movements of the above-mentioned features around the GCP can be inferred by the interpretation of a sequence of ionograms. This can cause the signal to 'follow' the feature, thus resulting in a continuous variation of the direction of arrival of the signal at the receiver. Observations of this effect will be discussed in more detail in Chapters 3 and 4.

1.7 Summary.

The ionosphere is an ionised part of the atmosphere that can be used as a propagation medium for radio signals in the high frequency band. These can be reflected by the ionosphere, which allows operators to achieve beyond-line-of-sight communication. The frequencies at which signals can propagate are largely determined by the electron density profile of the ionosphere. At high latitudes, the ionosphere is subject to very strong disturbances, which cause the ionospheric layers to be highly variable. As a result, the high latitude ionosphere can be considered to be a rough reflecting surface for HF signals, and consequently the propagation can be heavily affected.

At high latitudes the electron density distribution in the ionosphere can often be affected by drifting large-scale over-dense plasma structures, such as patches and arcs (frequently reported in the scientific literature). Patches and arcs movements can cause large deviations to HF signals propagating within the polar cap, thus heavily affecting their direction of arrival. This occurrence severely limits the performance of HF DF systems, since their accuracy is typically of 0.1° , whereas the deviations of the directions of arrival of HF signals at the receiver can be as much as $\pm 100^\circ$ or more at high latitudes. Ionospheric sounding is used to investigate the propagation conditions and to infer the evolution of the ionospheric environment.

In the next chapter direction finding techniques will be described. Superresolution direction finding algorithms will be discussed, with emphasis on those utilised for this research. Measurements and results obtained with these algorithms are presented in Chapters 3 to 6.

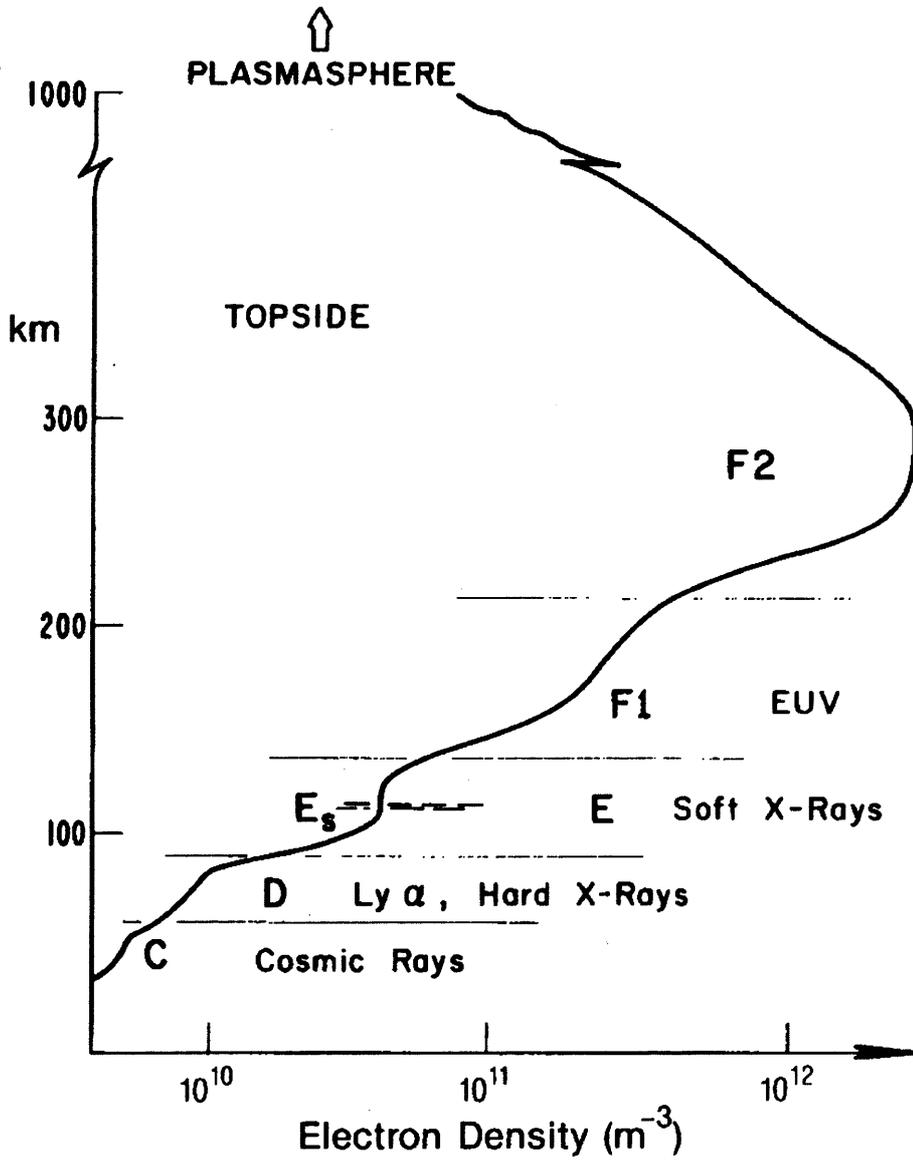


Fig. 1.2.1 Ionospheric structure on a summer day in a middle latitude, and the main bands of solar and cosmic ionising radiations (Davies, 1990).

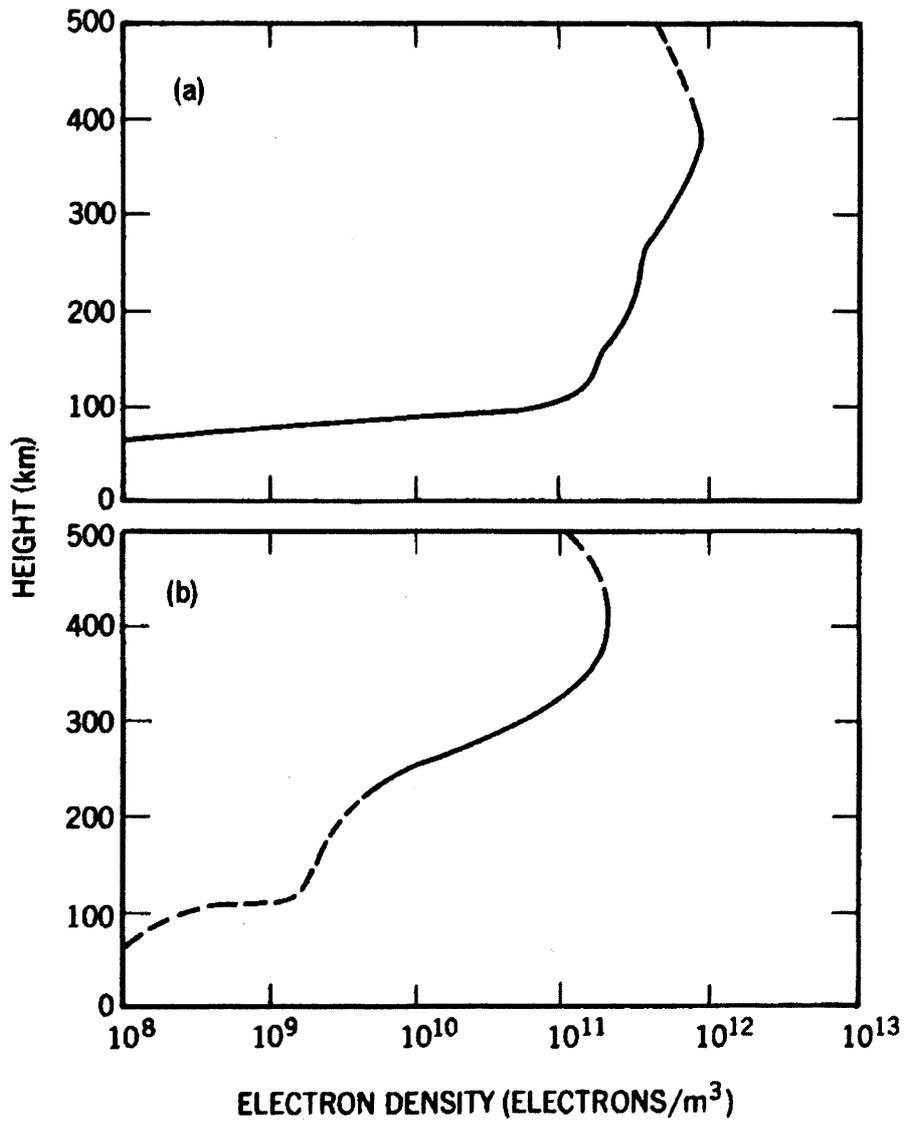


Fig. 1.2.2 Sample electron-density profiles: (a) Daytime, (b) Nighttime (Davies, 1990).

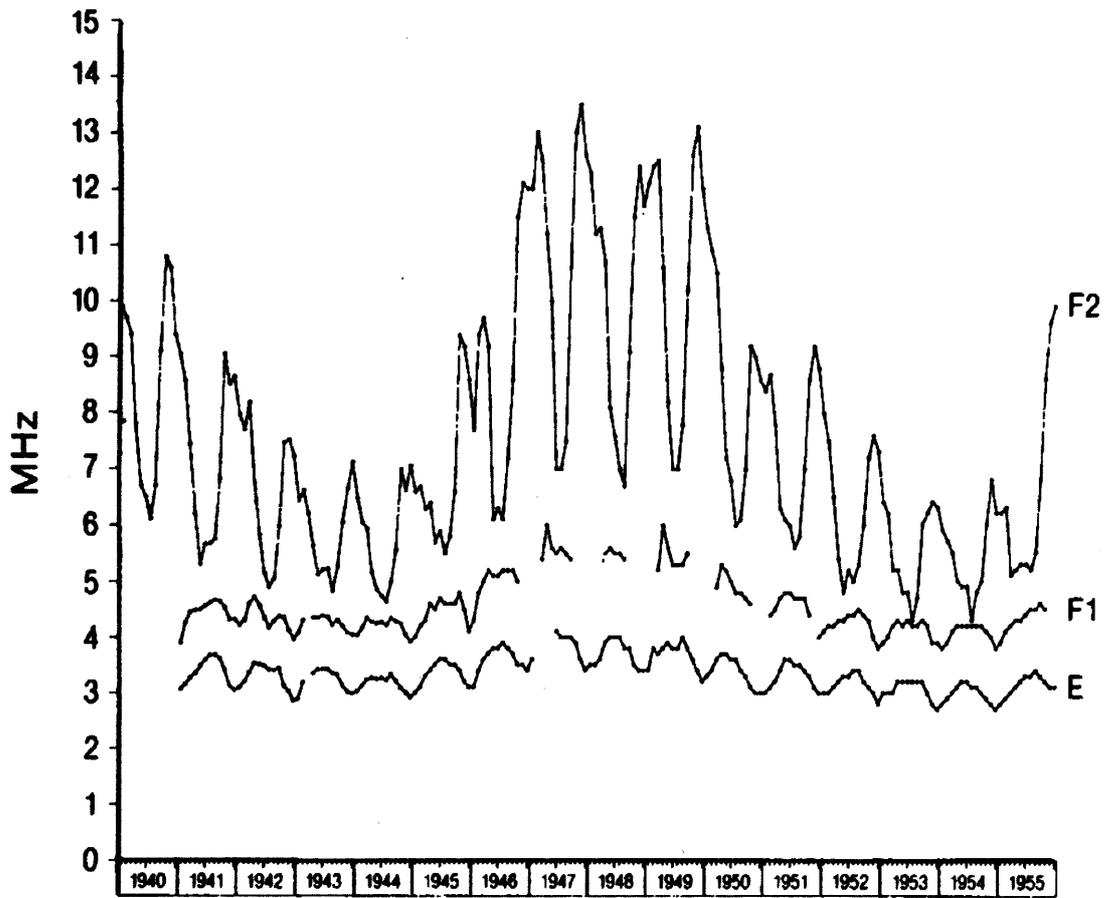


Fig. 1.2.3 Variation of median critical frequencies with solar cycle (Davies, 1990).

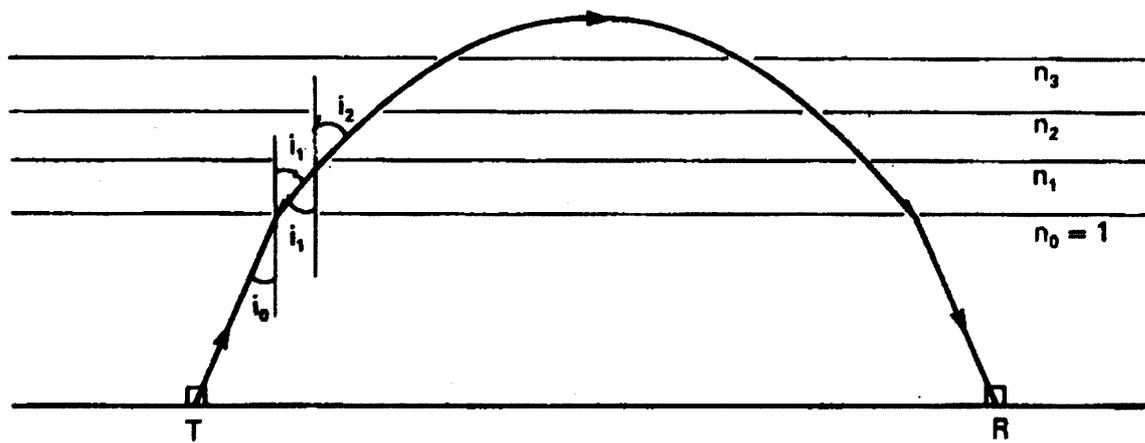


Fig. 1.3.1 Refraction of a radio wave in the ionosphere (Hargreaves, 1992).

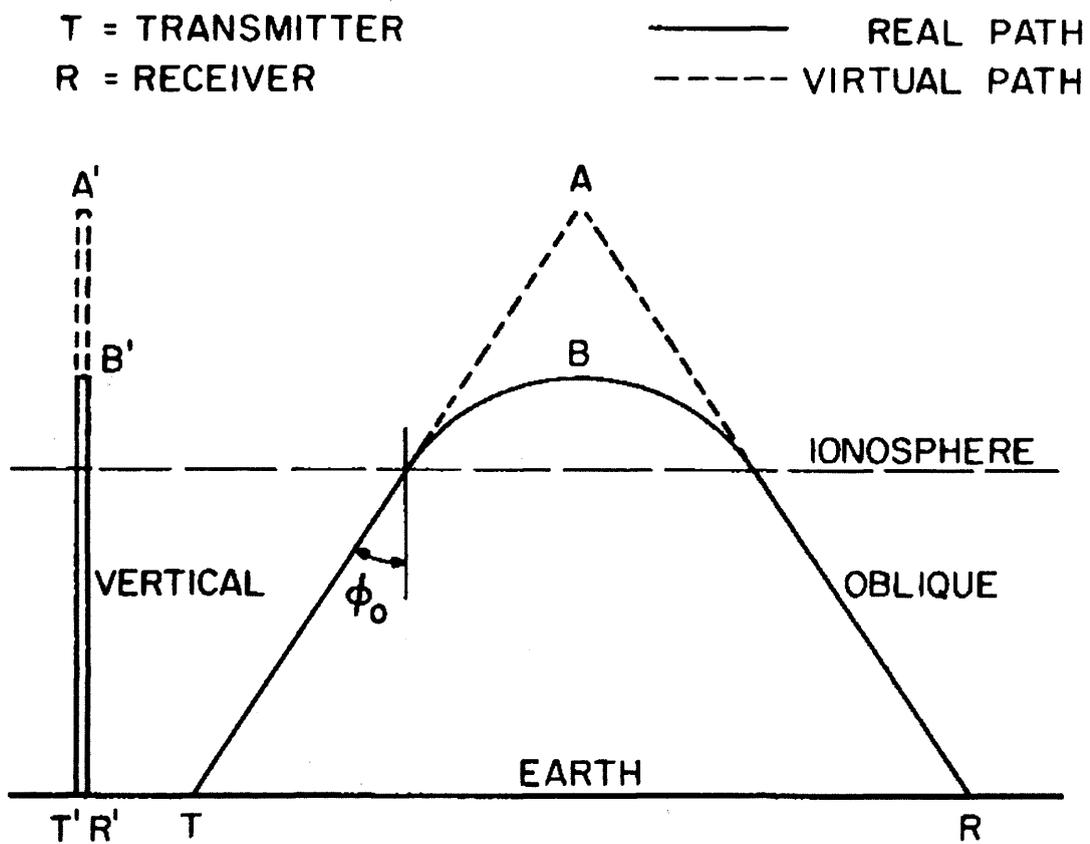


Fig. 1.3.2 Equivalent waves reflected obliquely and vertically at the same real height (B and B') and the same virtual heights (A and A'), plane ionosphere and plane earth (Davies, 1990).

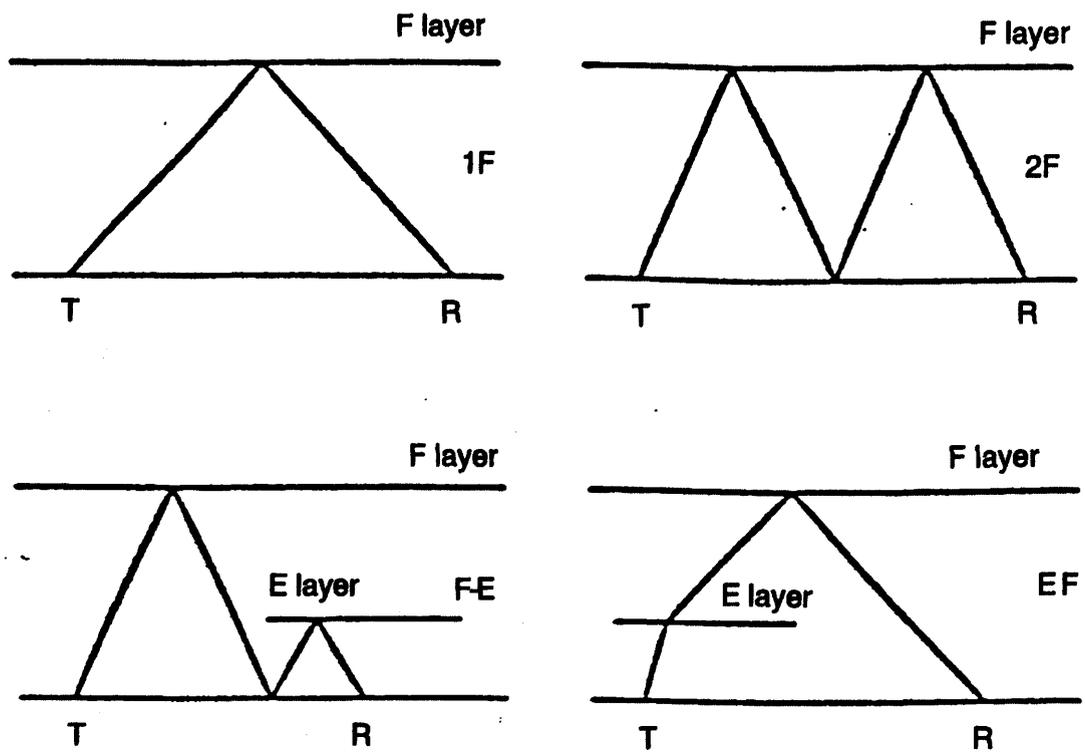


Fig. 1.3.3 Examples of different propagation paths producing different elevation angles at the receiver (Warrington, 1986).

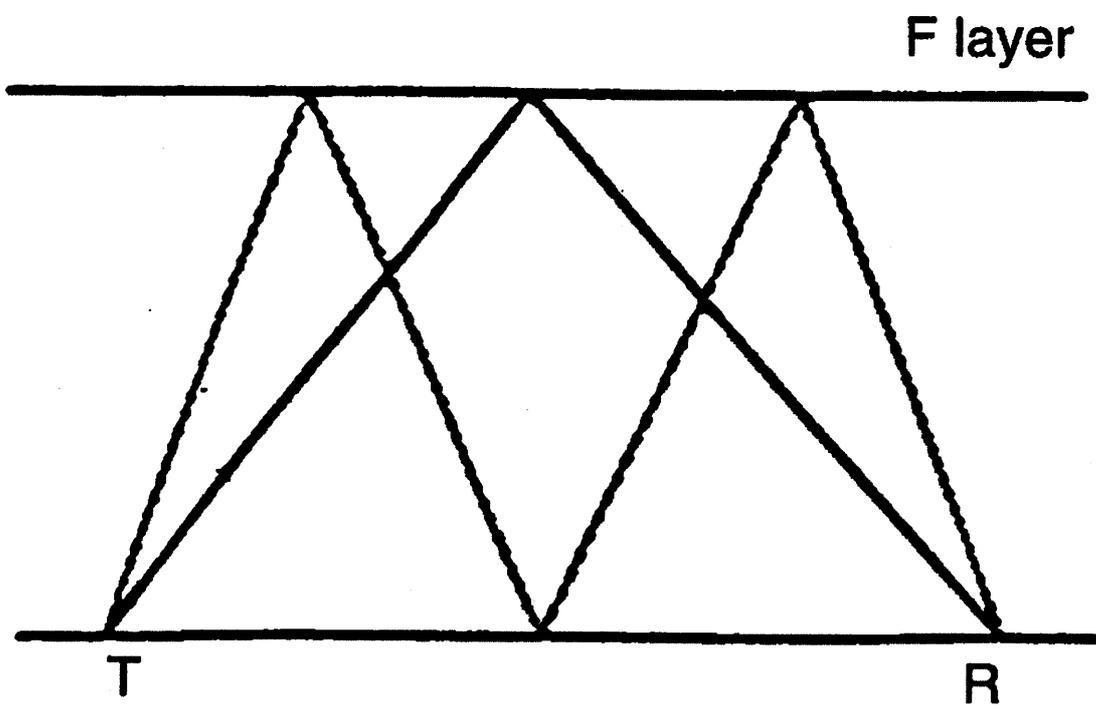


Fig. 1.3.4 Example of two-moded propagation (Warrington, 1986).

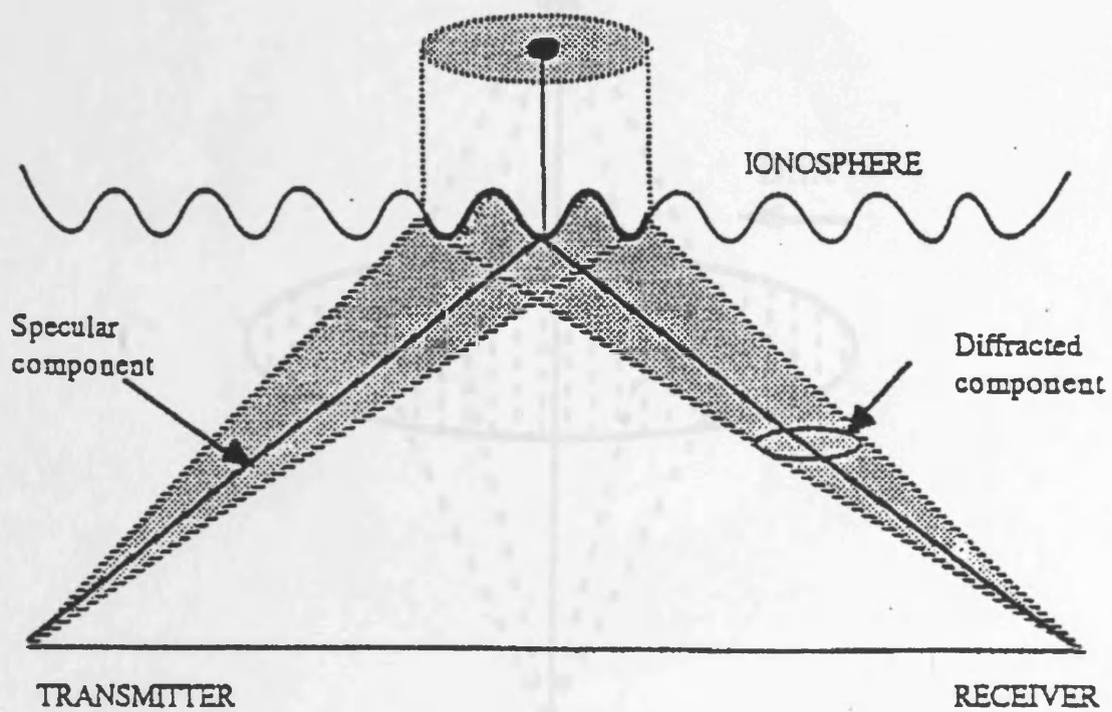


Fig. 1.3.6 Schematic diagram showing the presence of γ -rays along the great circle path (solid line) and via scattered components (dashed line) from a rough ionospheric reflector in which the irregularity drift has an east to

Fig. 1.3.5 Representation of signals being reflected from nearby points in a rough reflector (Warrington, 1986).

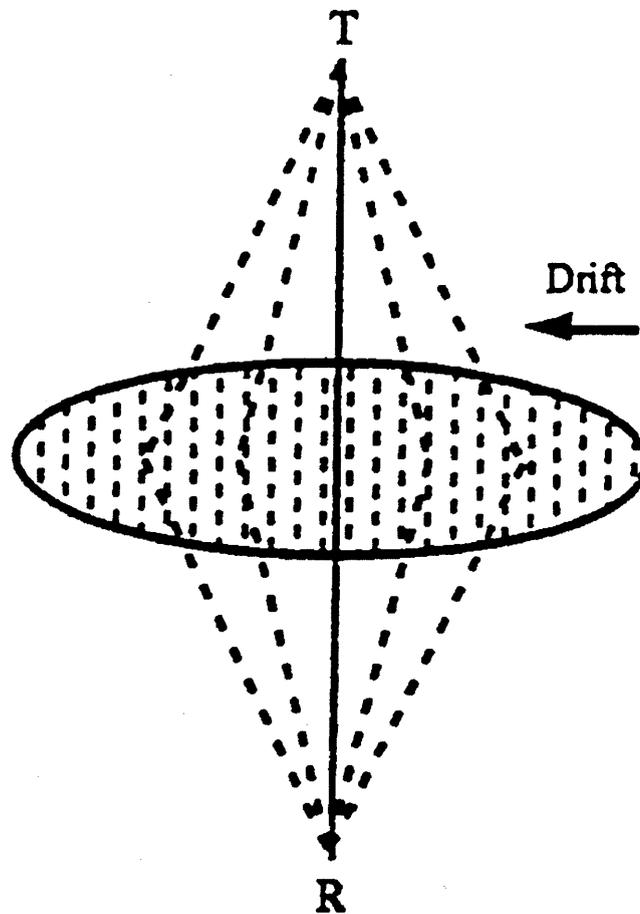


Fig. 1.3.6 Schematic diagram showing the presence of propagation along the great circle path (solid line) and via scattered components (dashed lines) from a rough ionospheric reflector in which the irregularity drift has an east to west component perpendicular to the direction of propagation (Warrington, 1998).

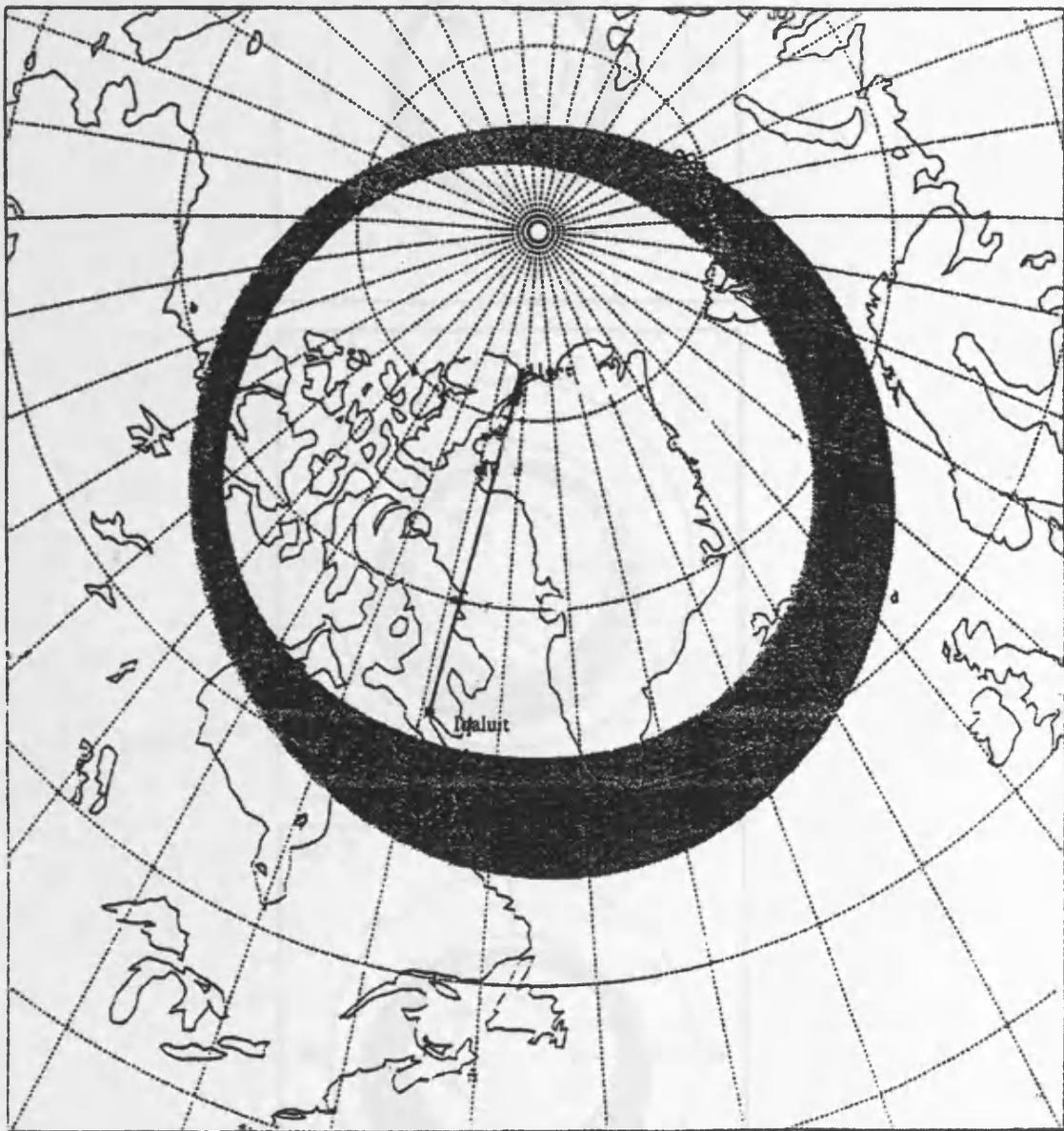


Fig. 1.4.1 Mean position of the visible aurora for low magnetic activity at 00UT. The arrow shows the Iqaluit to Alert path.

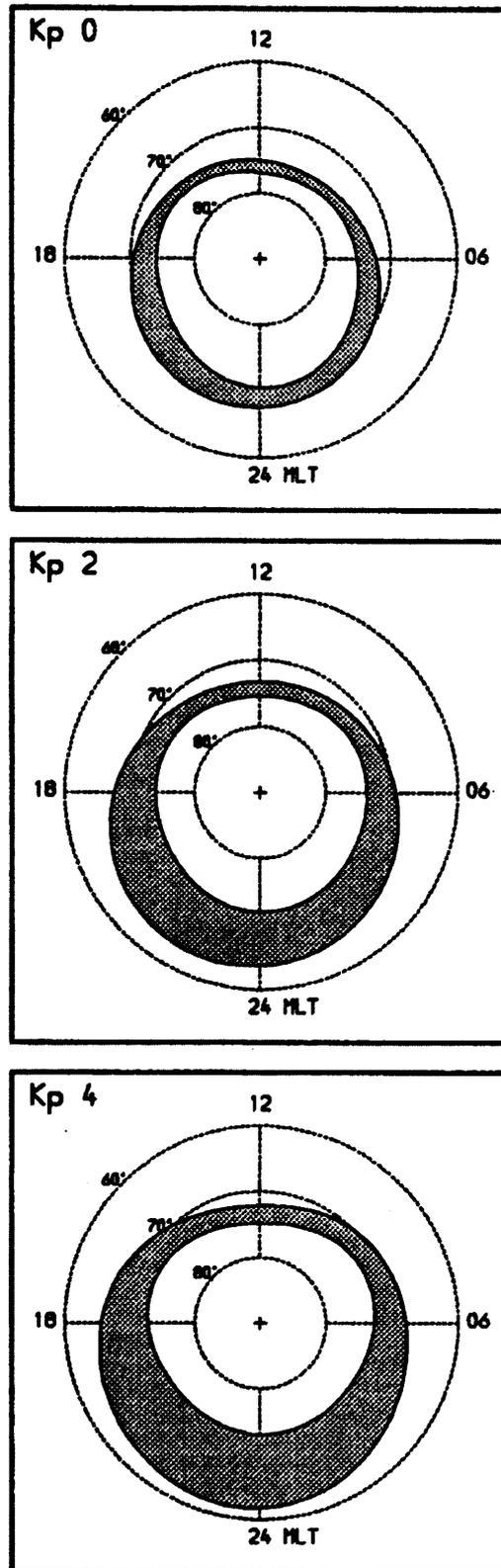


Fig. 1.4.2 The location of the auroral oval at different degrees of geomagnetic activity, quantified by the geomagnetic index K_p . The coordinates are geomagnetic latitude and magnetic local time (Milan, 1994).

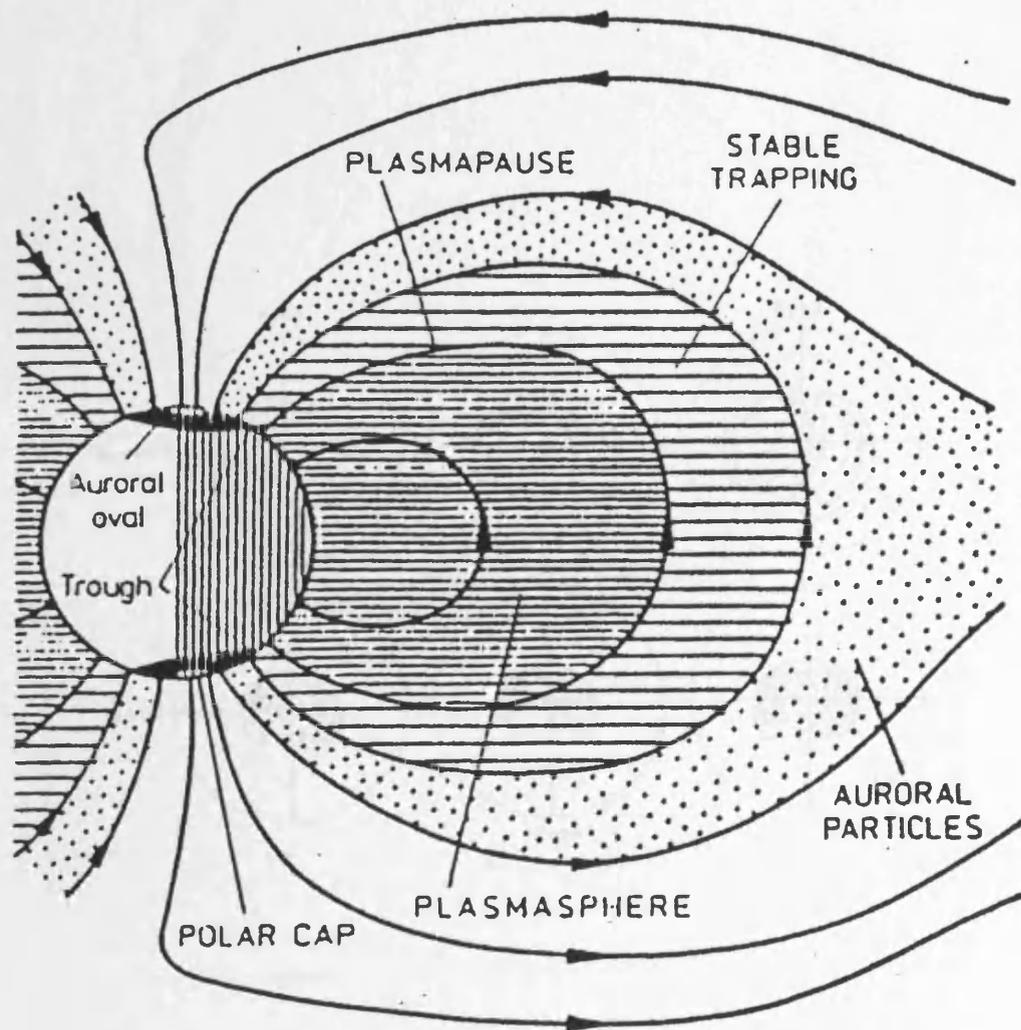


Fig. 1.4.3 Schematic diagram illustrating the high latitude ionosphere and the formation of the auroral oval, polar cap and high latitude trough (Cannon, 1989).

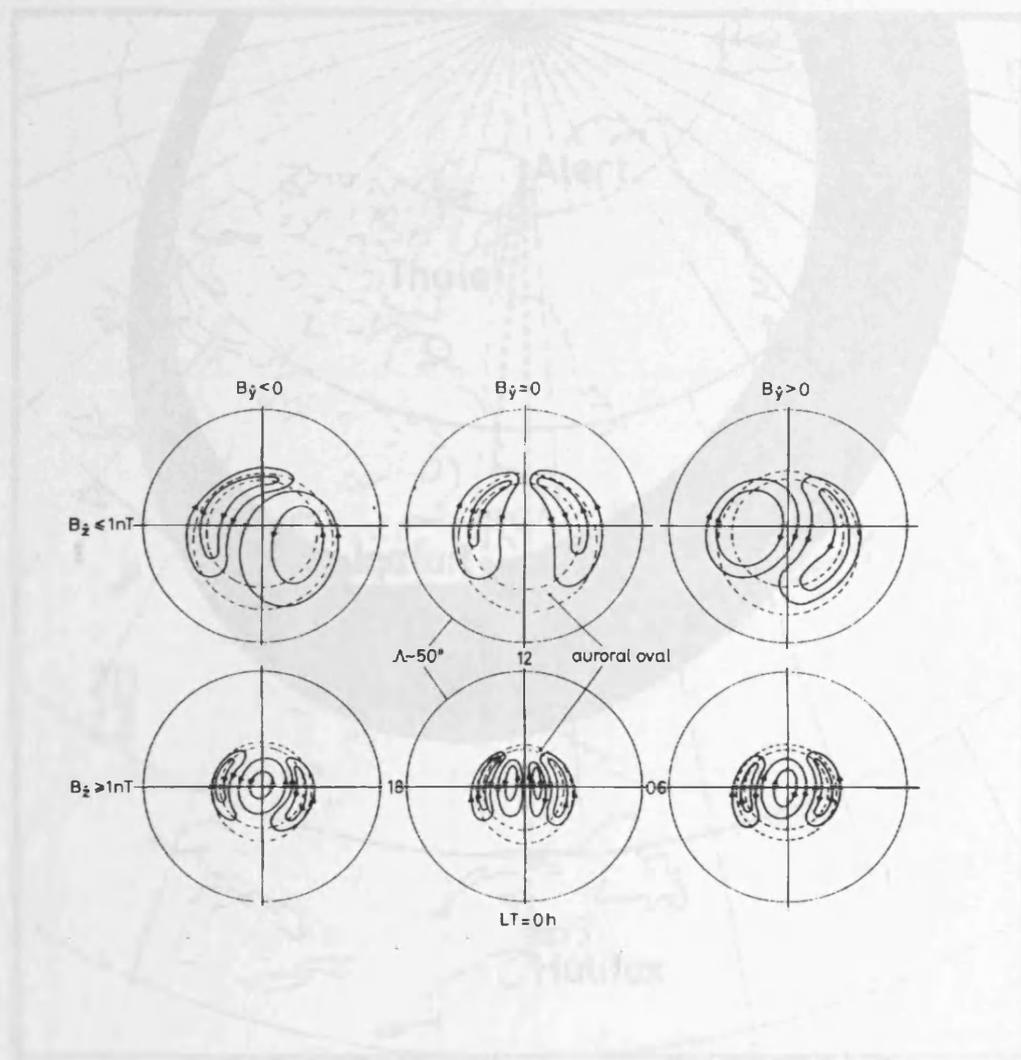


Fig. 1.4.4 Model of crosspolar-cap convection flow (Lockwood, 1993).

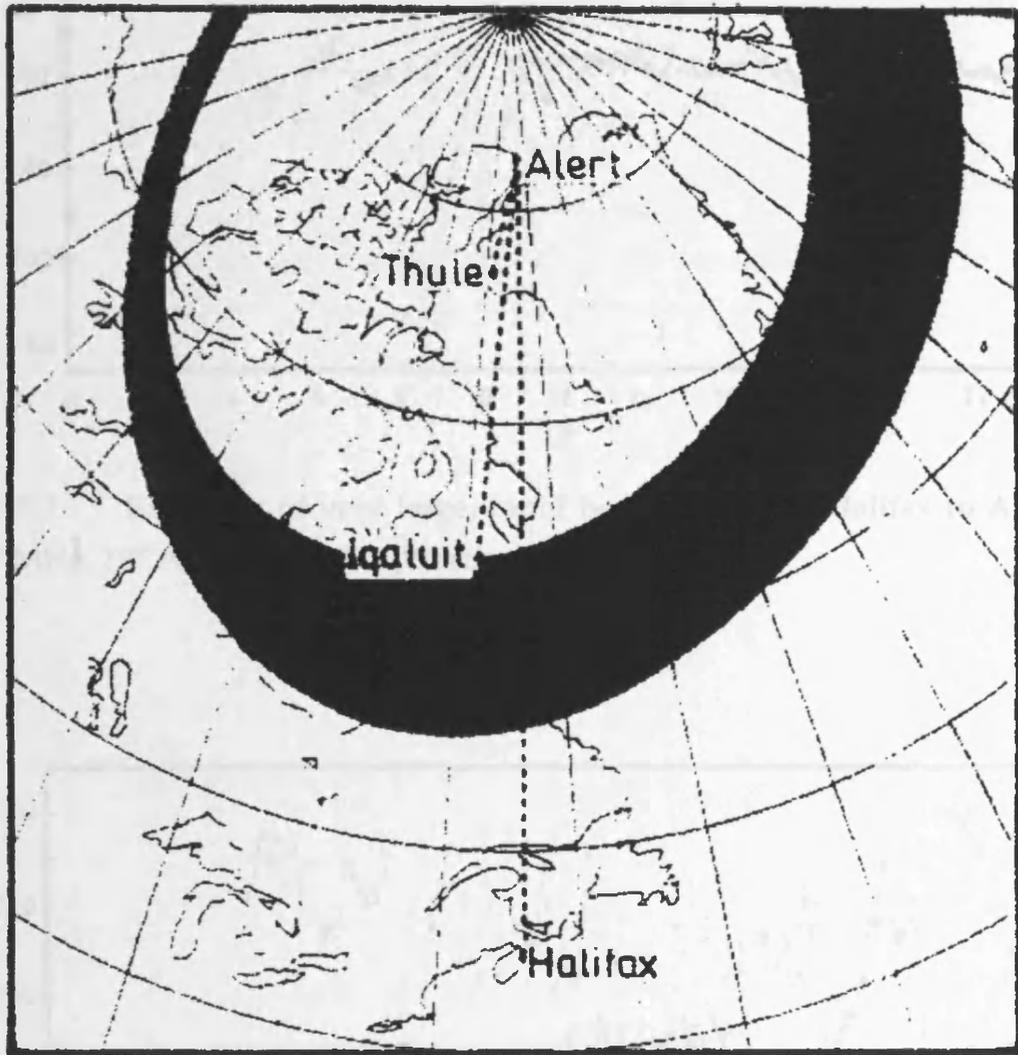


Fig. 1.5.1 Location of transmitting and receiver sites. A modelled position of the auroral oval at 00:00 UT and $K_p=3$ is also shown (Warrington *et al*, 1997b).

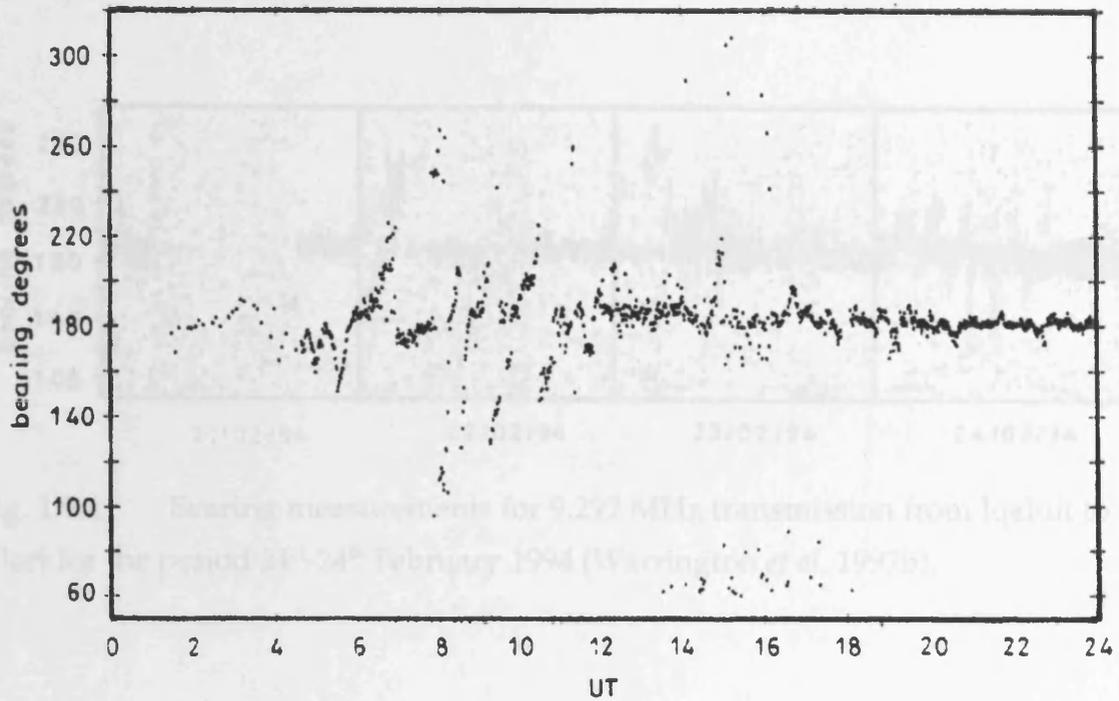


Fig. 1.5.2 Examples of very large, rapid bearing swings. Halifax to Alert, 8.697 MHz, 18th November 1990 (Warrington *et al*, 1997b).

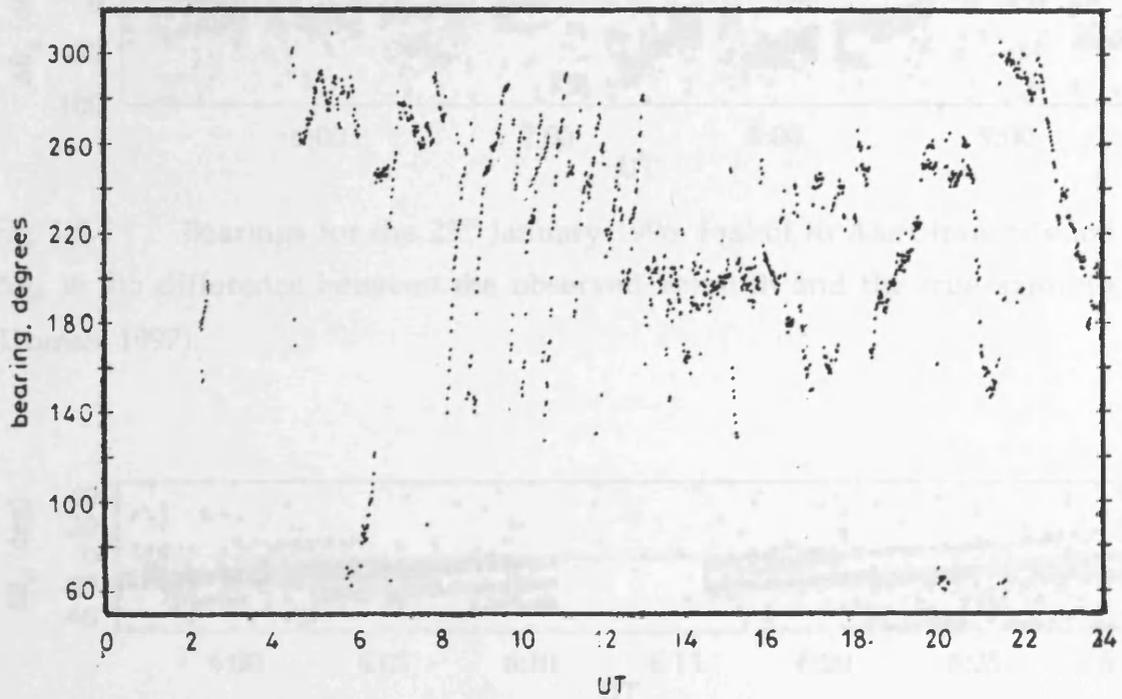


Fig. 1.5.3 Examples of very large, rapid bearing swings. Thule to Alert, 8.050 MHz, 18th November 1990 (Warrington *et al*, 1997b).

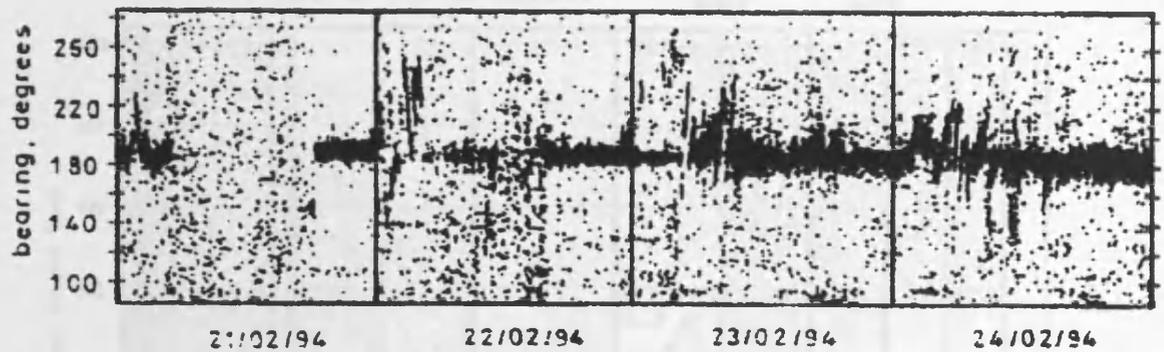


Fig. 1.5.4 Bearing measurements for 9.292 MHz transmission from Iqaluit to Alert for the period 21st-24th February 1994 (Warrington *et al*, 1997b).

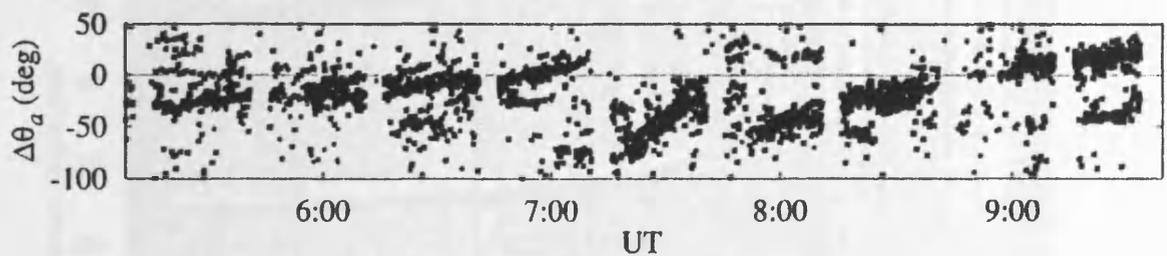


Fig. 1.5.5 Bearings for the 23rd January 1996, Iqaluit to Alert transmission. $\Delta\theta_a$ is the difference between the observed azimuth and the true azimuth, (Dumas, 1997).

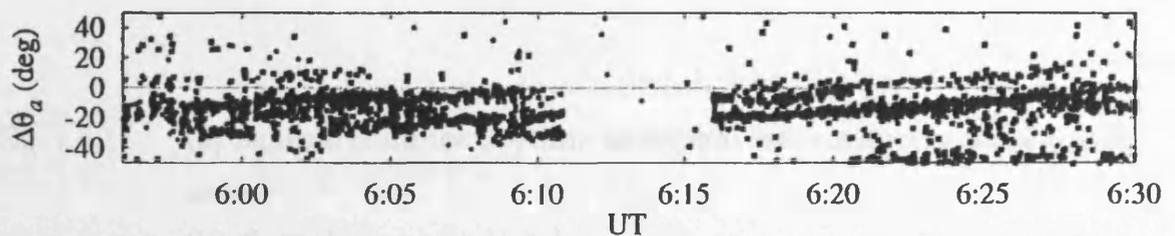


Fig. 1.5.6 An expanded view of the 23rd January 1996, Iqaluit to Alert transmission (Dumas, 1997).

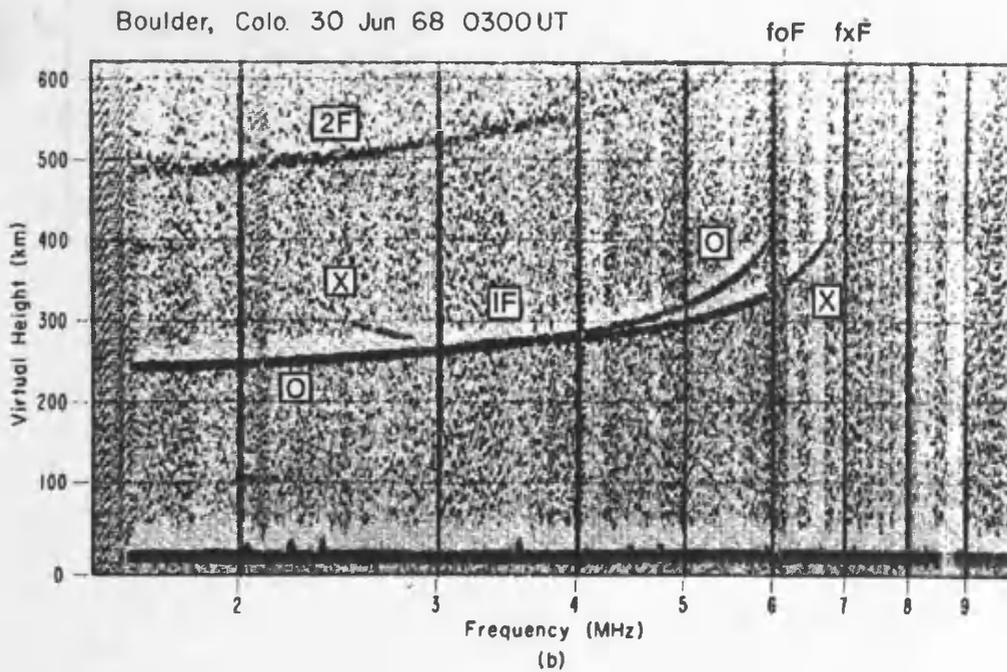
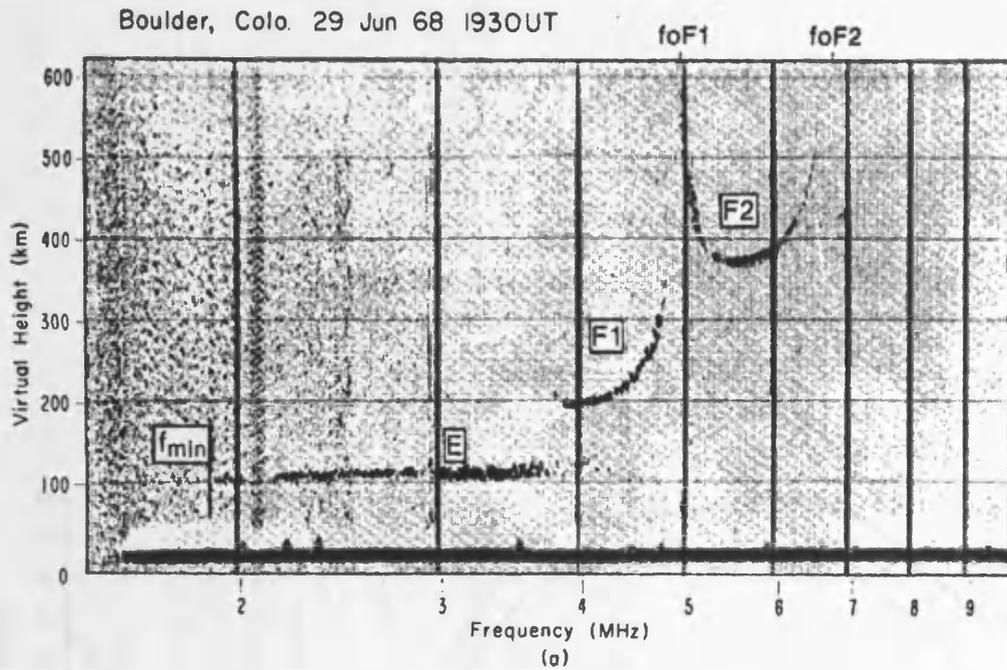


Fig. 1.6.1 (a) Sample summer daytime ionogram with ordinary wave traces only.
 (b) Sample ionogram taken on a summer evening in middle latitudes, showing one-hop and two-hop echoes from the layer (Davies, 1990).

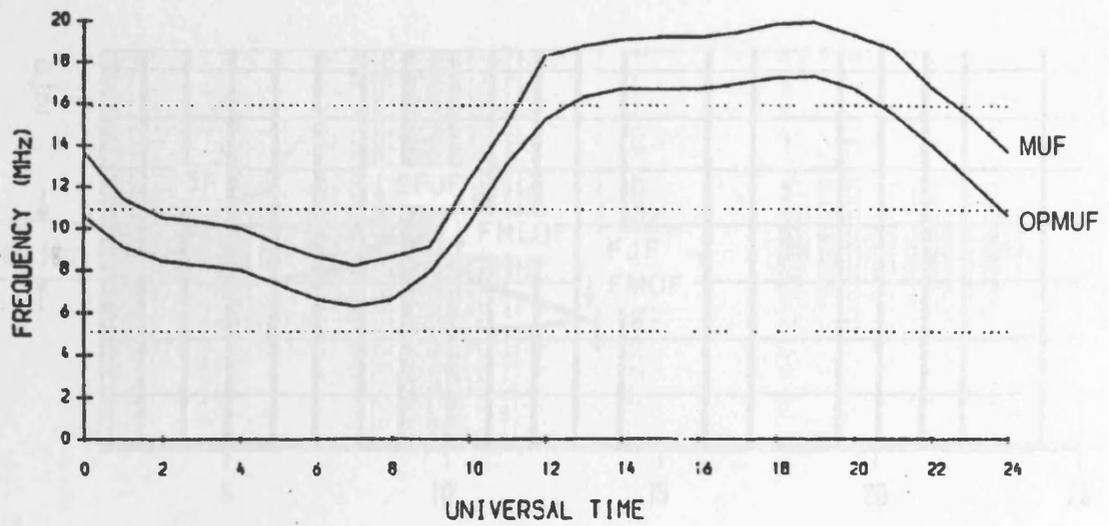


Fig. 1.6.2 ITU-R 533a computer model predictions of MUF and OPMUF on the Halifax to Cheltenham path for March 1994. The horizontal lines indicate the fixed frequencies of operation close to 5, 11 and 16 MHz of the Halifax transmitter (Rogers *et al*, 1996).

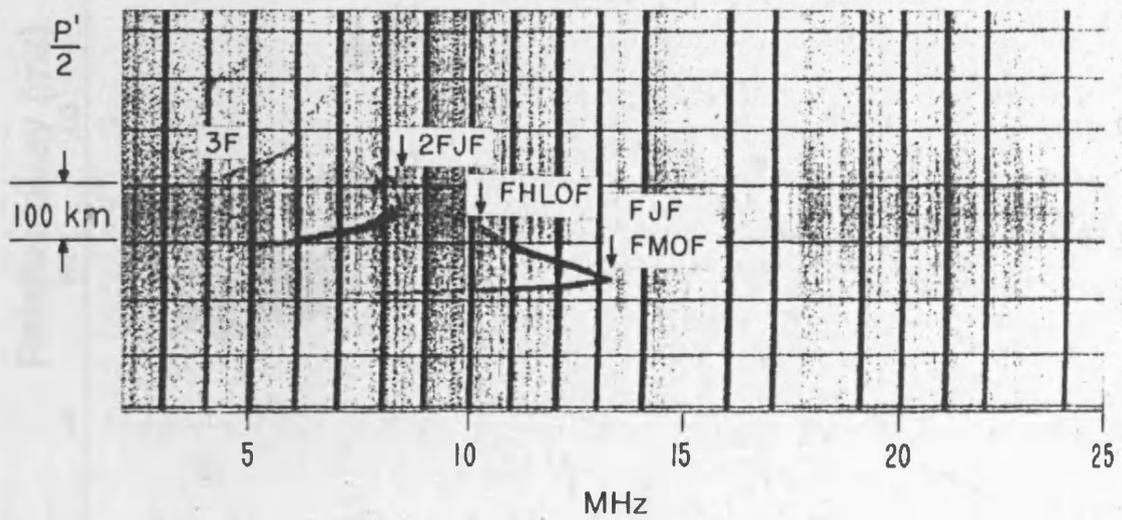


Fig. 1.6.3 Oblique ionogram for the 2370 km, geographic approximately east-west path from Boulder, Colo., to Sterling, Va., 1st September 21:12 UT (90° WTM) (Davies, 1990).

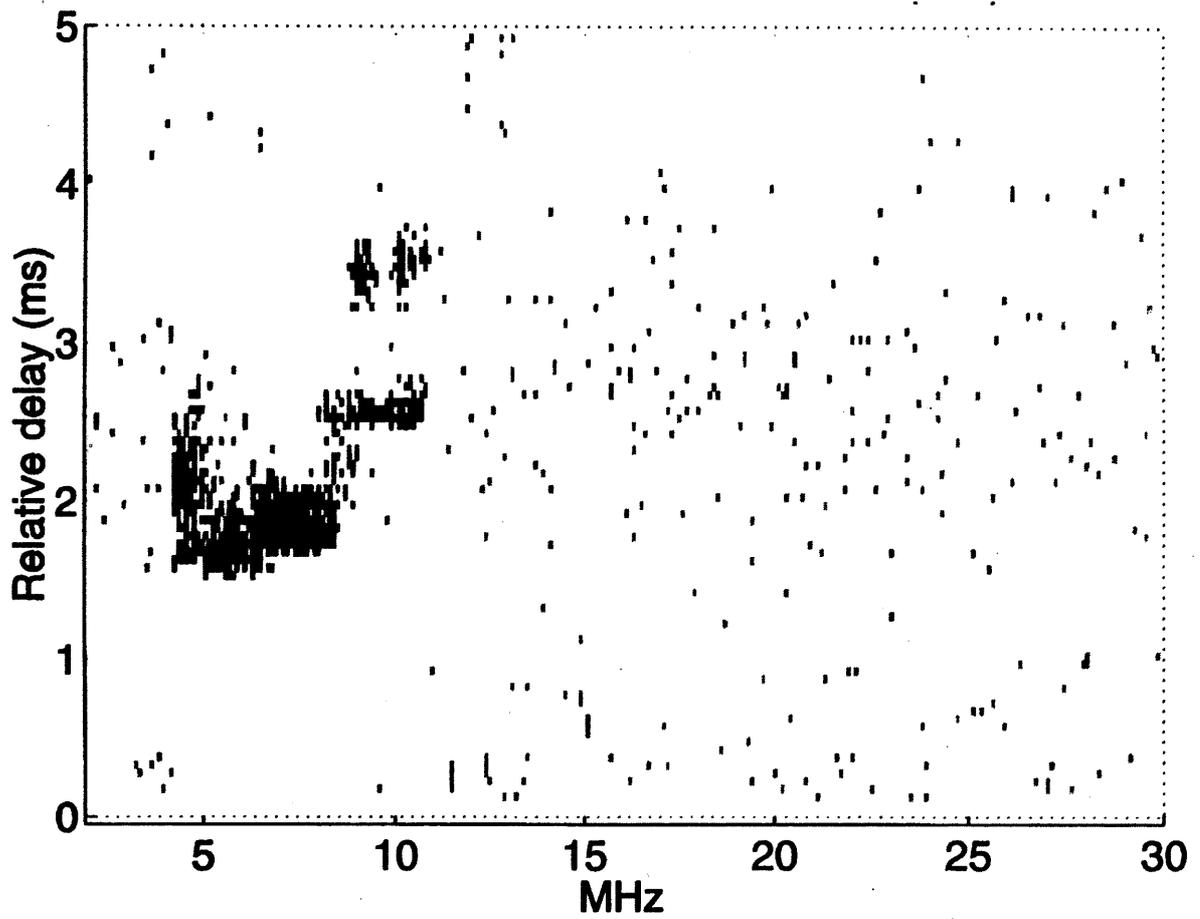


Fig. 1.6.4 Example of an oblique ionogram taken over the polar cap path from Iqaluit to Alert on the 23rd January 1996 between 09:41 and 09:46 UT.

2. REVIEW OF DIRECTION FINDING TECHNIQUES AND SUPERRESOLUTION DIRECTION FINDING ALGORITHMS

2.1 Introduction.

Fixed land-based stations are used to measure the direction of arrival of signals, usually transmitted from an unknown site (except for research and testing purposes). Researchers have focused their efforts on the testing and improvement of various types of direction finders, which are classified according to their size. Broadly, the bigger the size, the higher the resolving power of the direction finder. However, at the same time the cost of the site increases as its size. A compromise between low resolution and high cost is to adopt sites of average size of a few hundred metres, which can cover the HF band (3-30 MHz) (see next section). The advent of multichannel receivers and digital computers has opened new opportunities. The main challenge has become to enhance the performance of direction finders by means of modern signal processing techniques. Superresolution direction finding algorithms are currently widely used to achieve this aim, because they are remarkably better than conventional techniques in terms of resolving power (see Section 2.7). Those used for this work are presented.

In the last three decades, many researchers focused their efforts in the improvement of superresolution direction finding algorithms, and in the development of new ones. A review of the performance comparisons of various algorithms with both real and simulated data, in different propagation environments, is presented, with a special interest paid to the algorithms utilised in this research.

A comparison between the performances of four array configurations and two DF algorithms is discussed (Jenkins, 1997) in terms of visibility of a weaker point source signal in presence of a stronger spread source signal.

Theoretical studies by Zatman (1998) and Hayward (1997) are presented, that show that multiple traces can arise from one signal, under certain circumstances.

2.2 History of direction finding techniques.

The aim of this section is to review the history of land-based stations for direction finding and the relevant techniques. A direction finder would ideally be able to work over 360° of azimuth, 90° of elevation and a wide frequency range. When a fixed land-based transmitting station is used, the DOA and the

frequency of the received signal can be compared to the expected ones. Direction finders capable of measuring both the azimuth and the elevation of a signal are referred to as three-dimensional direction finders (Adcock, 1959). Direction finders are defined according to their size. A wide-aperture direction finder (WADF) is defined in the British Standard (1971) as one 'in which the arriving wavefront is sampled simultaneously (or nearly so) by an array of aerials extending over a distance comparable to or greater than one wavelength'. When the aperture is a few tenths of a wavelength, we have a narrow-aperture direction finder (NADF). A very wide-aperture direction finder (VWADF) is an antenna array with aperture of more than ten wavelengths. The system used to collect the data analysed in this research is a WADF.

The first direction finder used at the beginning of the 20th century was a simple vertical loop, rotated about a vertical axis. Adcock (1919) created a NADF that would be much less affected by polarisation errors than the loop. The Adcock was adopted and developed during the 1930s and, because of the potential applications of direction finders in a military environment, it was extensively used during the Second World War, which gave a strong stimulus to research. The sources of instrumental and site errors were identified and reduced (see, for example, Smith-Rose and Ross (1947)). Crampton (1947) realised a combination of two or more Adcocks to reduce site imperfections and wave-interference in larger systems.

A common type of WADF is the circularly disposed antenna array (CDAA). For this array geometry, two concentric rings are necessary to cover the whole 10:1 ratio of the HF band from 3 MHz to 30 MHz, because a single ring maintains a satisfactory radiation pattern up to 4:1 frequency range. The outer ring is used for the lower part of the band and the inner ring is used for the upper part of the band. An example is shown in Figure 2.2.1. There are three main types of CDAA. The Wullenweber (Rindfleisch, 1956) forms the desired directional pattern by combining the outputs from a sector of the whole ring. A goniometer consisting of 24 elements is described by Hockley (1973). A capacitive switch works as a spinning goniometer, selecting which elements are to be used at any moment, which enables the Wullenweber to scan and direct the antenna beam. 24 stator plates are connected to the 24 elements, whereas 8 rotor plates rotate to collect 8 signals from an arc of 8 elements at any moment (see Figure 2.2.2). In practise a smoother commutation is obtained by increasing the number of rotor plates. Either a sum or a difference pattern can be used to estimate the DOA of a signal. If, for example, A is the sum of four signals and B the sum of the four adjacent signals, the two outputs from the goniometer are a sum

pattern (A+B) and a difference pattern (A-B), also known as rabbit-ears pattern (Figure 2.2.2). For those algorithms using the difference pattern, the azimuth of the arriving ray is given by the pointing angle of the goniometer where is a minimum signal between the two maxima of the rabbit-ears pattern. Other algorithms select the maximum of the sum pattern as the direction of arrival of the signal, corresponding to the direction of the minimum of the difference pattern. The second type of CDAA is the Doppler DF. Whale (1954) used this system, in which a single element is moved rapidly round the perimeter of a circle, thus imposing a frequency modulation on the signal. More usually, a simulated Doppler effect is obtained by smooth commutation round a ring of fixed elements. In the third type of CDAA the azimuth is calculated by phase comparison between successive pairs of elements (Earp and Godfrey, 1947). This system is also referred to as commutated antenna direction finder (CADF). These systems were often adopted after the Second World War, the Wullenweber being too expensive. However, it was soon realised that the performance of CADFs was poor when dealing with weak signals in comparison of strong interfering signals. Therefore interferometer systems were developed (Ross *et al*, 1951), which also had the advantage of providing information about the elevation angle. However, the Wullenweber proved to be the system giving the most precise measurements in terms of azimuth of the arriving signal. In the early seventies two research groups built two linear VWADFs. An array of 256 vertical elements over an aperture of 2.5 km belonging to Stanford University is described by Sweeney (1970). An array for the Communication Research Centre, Canada, is described by Rice (1971); this array has a 1.2 km main arm and a shorter crossarm for elevation measurements. Attempts to measure the elevation angle of the received signal were also made with the Wullenweber (see Jones *et al*, 1966). For this purpose an alternative is to deploy a separate vertical array (Gething *et al*, 1969).

The problems related to the dimensions of an array aperture are a limiting factor for the resolving power of DF systems (see next Section). An aperture of about 100 km is needed to achieve a resolution of 0.1° . This is clearly virtually impossible. The usual measure of an array aperture is not bigger than 300 m. An acceptable suppression of site errors can be achieved with an aperture of 1 wavelength, whilst 4-5 wavelengths are necessary to effectively reduce the wave interference. This difference arises because site re-radiators produce unwanted signals well separated in azimuth, whereas wave interference involves closely separated modes, which could arise from the target signal being reflected, for example, by the E and F ionospheric layers.

This research concerns with the ability of modern signal processing techniques to enhance the performance of WADFs in multiray wavefields, that is when rays can arrive at the receiver from unwanted transmitters and from re-radiating objects in the proximity of the DF array. A multiray scenario also occurs when the original signal propagates via two widely separated paths. The direction finder is said to have resolved the wavefield when it provides measurements, which allow the correct calculation of the ray parameters. Whereas goniometric systems can be capable of resolving two signals which are separated in azimuth more than the natural beamwidth, modern signal processing techniques can allow two signals which are separated in azimuth less than the natural beamwidth to be resolved.

In the last decades new techniques have increased dramatically the performance of direction finders. Multichannel receivers and the application of direction finding algorithms by means of digital computers have created modern techniques for signal processing. Elements of an antenna array, each connected to a receiver, simultaneously take successive measurements of amplitude and phase of the signal, thus providing a sequence of space samples of the wave-field, commonly called snapshots. The evolution of the wave-field pattern then provides information about the DOA of the incoming signal. An early approach to this technique, known as wavefront analysis, was made by Bain (1956).

As a preliminary assessment, computer simulations are extremely helpful in providing useful information about performances of different array configurations and digital signal processing schemes. The performance of a direction finding system has to be tested in complex multiray environments, but not when the propagation environment is disturbed such that the wave-field is virtually impossible to solve. On the other hand, simple single ray situations are not challenging, because any direction finder should be able to give a correct answer. The ability to discriminate between two or more signals is a key quality factor for direction finding algorithms. Cooper (1973) suggested that, when the wanted signal is the weaker of two received signals, the signal-to-noise ratio can be improved by a form of automatic DF system. Lim (1977) achieved this by placing a null in the direction of arrival of the unwanted signal. A null could be generated and steered electronically to any required direction in a circular array. Gabriel (1976), L.J. Griffiths (1977), J.W.R. Griffiths (1983) and Hudson (1981) studied the performance of adaptive arrays, in which the nulls are positioned automatically in the selected DOAs. The advent of these new DF techniques provided a new approach to a number of aspects of

array design, especially those involving the control of nulls (Gething, 1973; Gething and Haseler, 1974; Cawsey, 1974).

2.3 Resolution techniques.

The resolving power of fixed land-based stations depends heavily on its dimension, the bigger the array size the higher the resolving power. The azimuthal resolution is determined by the natural beamwidth of the antenna array, which is given by

$$\text{beamwidth} = 2 \sin^{-1} \left(\frac{\lambda}{2a} \right) \quad (2.3.1)$$

where λ is the wavelength of the incoming signal and a is the aperture of the antenna array in the broadside direction (see, for example, Baden Fuller, 1993). The elevation angle resolution is higher with vertical arrays than with horizontal arrays, for physical reasons, and needs to be carefully evaluated for every single pattern (see Gething, 1991).

As an example, consider that we wish to resolve two modes separated by 1° in azimuth to an accuracy of 0.1° . A NADF with an aperture of a few tenths of a wavelength is incapable to resolve the two modes, whereas a VWADF with an aperture of a few tens of a wavelength can face this task, but is very expensive. WADFs lie between these two extremes. They have a higher resolving power than NADFs and are less expensive than VWADFs. In the last decades many researchers have focused their efforts in the attempt to enhance the resolving power of WADFs by means of modern signal processing techniques. In fact, the advent of multichannel receivers and digital computers has dramatically enhanced the performance of WADFs. The processor of the data collected by a multichannel receiver is typically a computer algorithm. The aim is to utilise a processor, which works perfectly in single-ray conditions and in the best possible way in a multiray environment.

There are four main types of DFs suitable for the measurement of the direction of arrival in the HF band.

- The circularly disposed antenna array, with a spinning goniometer and a single-channel receiver. It allows for azimuth measurements only.
- The crossed-arm interferometer, with a twin-channel receiver used as a phase meter, for measurements of both azimuth and elevation of the arriving signal. It is a system with multilobed radiation patterns, which can have several shapes, as described by Sherrill (1971). Elements are usually arranged in two or more straight arms, to form a cross, triangle or L-shape.

- A vertically stacked array of horizontally polarised elements, for elevation measurements only. As an example, the elements may form an array of several equally spaced loops on a vertical tower.
- Any horizontally disposed set of n elements used with an n -channel receiver, for measurements of both azimuth and elevation. It might be arranged in a circle, in a crossed-arm form or in a more irregular shape. Regardless of its shape, it is different from a circularly disposed antenna array because the outputs of the elements are not combined in a beam-forming network to produce a directional radiation pattern. It also differs from a crossed-arm interferometer because the interferometer usually selects its elements in pairs. Instead, it records digitally the amplitudes and phases of the element voltages for subsequent analysis. It is a system of this type on which the data in this thesis were collected. The vertical array generally uses the same procedure.

The DF processors can be based on either amplitude comparison or on phase difference measurements. The amplitude comparison is used in radar monopulse systems, which compare the amplitudes from two antenna beams by means of a twin-channel receiver and oscilloscope display, as shown in Figure 2.3.1. In the simplest case, the antennas have identical radiation patterns and coincident phase centres. In this figure, as an example, the oscilloscope shows a straight line of inclination 45° . This means that the two signals are equal, suggesting that a single ray is received on the axis of symmetry.

It is important to consider that between two patterns there might be several intersections, therefore any displayed angle can correspond to more than one direction of arrival. Furthermore, the amplitude ratio (for a single ray) is a function of azimuth and elevation angle, which are impossible to determine without any additional measurements. The amplitude ratio is not affected by changes in the elevation angle on the boresight direction, so this system does not allow for discrimination between one mode and several modes in or very near this plane. As a result of the comparative sensitivity to azimuth changes and insensitivity to elevation changes near boresight, this system is at a clear disadvantage when the objective is ray resolution, whereas it can be an advantage when only the azimuth measurement is required.

The phase difference measurements are used in interferometers. In this system, shown in Figure 2.3.2, the phase centres of the elements of the directive arrays are separated by a distance d . The beam patterns are identical and point in the same direction. Let a and b be the output phasors for a single ray; these phasors will have the same amplitude but different phase by ϕ , where

$$\phi = \frac{2\pi d}{\lambda} \cos\theta \quad (2.3.2)$$

and θ is the angle between the ray and the line joining the phase centres. In this system, the twin-channel receiver and the oscilloscope are used as a phase meter by displaying $(a+b)$ against $(a-b)$ changed in phase by $\pi/2$.

As shown in Figure 2.3.3, the angle ϕ is given by

$$\phi = 2\gamma \quad (2.3.3)$$

where γ is the inclination of the line display on the oscilloscope. Ross *et al* (1951) demonstrated that the line becomes an ellipse when a and b have not the same amplitude, but the above relationship will still be valid, with γ as the inclination of the major axis of the ellipse.

In this systems an error in the measurement of θ can occur because various combinations of azimuth α and elevation Δ can give the same value of θ as shown by the equation

$$\cos\theta = \cos\alpha \cos\Delta \quad (2.3.4)$$

Let us consider a linear array, i.e. one arm of an interferometer. An error in the determination of θ can also result from an error in the measurement of ϕ , as can be seen in the equation

$$\frac{d\phi}{d\theta} = -\frac{2\pi d}{\lambda} \sin\theta \quad (2.3.5)$$

The maximum accuracy can be obtained by making $d\phi/d\theta$ as large as possible, to make sure that big changes in the value of ϕ would result in small changes in the value of θ . Therefore, maximum instrumental accuracy can be obtained with d as large as possible, and $\theta = \pi/2$; whereas accuracy drops to zero when $\theta = 0$. It follows that the horizontal array has to be as large as possible and as close as possible to perpendicularity to the ray, to make accurate azimuth measurements. Whereas a vertical array is more desirable when elevation measurements are a priority.

As a comparison, we can say that in the amplitude-comparison system, using squinted beams with coincident phase centres, directive patterns of known shape are essential; whereas in the phase difference system, using parallel beams and separated phase centres, directivity is optional. More sophisticated systems combine the two methods.

2.4 Beam forming process.

The conventional beam steering method (simply named in this thesis as Beam algorithm) differs from the superresolution ones (see Section 2.6) in that the

signals from the antennas are simply combined to form a natural beam in a specific look direction. The look direction is then scanned in both azimuth and elevation, and the direction(s) of strongest reception taken as the estimate(s) of the direction(s) of arrival. The performance of the Beam algorithm has been compared to that of some superresolution algorithms.

Summing the signals from each antenna can form a beam in the directional sensitivity pattern of an antenna array. The signals arriving at each antenna have different phases, because of the physical distance between the single antennas. Phase changes need therefore to be applied such that signal components from a specified direction have equal phases. To achieve this, either an RF beam forming network is utilised or the outputs of the individual receivers connected to each antenna are sampled and processed in a computer. The latter approach (an n -channel technique) is that taken in the work described in this thesis.

The beam forming process may be described mathematically by

$$s(\alpha, \epsilon) = a(\alpha, \epsilon)^H x \quad \text{or} \quad s(\alpha, \epsilon) = x^H a(\alpha, \epsilon) \quad (2.4.1)$$

where

- (α, ϵ) is the look direction (azimuth and elevation)
- $a(\alpha, \epsilon)$ is the steering vector (a column vector containing the expected complex signal parameters for a constant amplitude signal incident onto the array from the specified look direction)
- x is a column vector containing the signal samples
- H indicates a Hermitian transpose.

Neglecting the individual directional sensitivity patterns of the elements and any constant factors, the power received in the direction (α, ϵ) is therefore given by

$$p(\alpha, \epsilon) = s(\alpha, \epsilon)^2 \quad (2.4.2)$$

which may be written as

$$p(\alpha, \epsilon) = a(\alpha, \epsilon)^H x x^H a(\alpha, \epsilon) \quad (2.4.3)$$

For one or more sample sets, $x x^H$ can be replaced by the covariance matrix, R , such that

$$p(\alpha, \epsilon) = a(\alpha, \epsilon)^H R a(\alpha, \epsilon) \quad (2.4.4)$$

where

$$R = \frac{1}{N} XX^H \quad (2.4.5)$$

X is a matrix containing one or more instantaneous sample sets, one in each column

N is the number of sample sets in X .

The directional sensitivity pattern of the array will generally contain one or more side lobes, which may have significant amplitudes compared to the main lobe. This occurs when the main lobe and the side lobe directions are close in signal space, that is when the directions for which the steering vector of the main lobe direction and the steering vectors associated with the side lobe direction are similar.

2.5 Superresolution direction finding (SRDF) techniques.

Superresolution direction finding is a term applied to techniques, which are capable of resolving two simultaneous signals, whose angular separation is less than the natural beamwidth of the antenna array. Flexibility in array geometry, superior accuracy, multiple co-channel signal operation, robustness, and the requirement of only a few samples to provide accurate bearings are the main features of the superior performance by the SRDF algorithms with respect to the conventional ones (Schmidt, 1986).

The conventional techniques, reviewed earlier in this chapter, can only satisfactorily cope with a single signal and, for many reasons such as the increasing congestion of the electromagnetic spectrum and various disturbances, it cannot now be assumed that only one signal is present in a frequency channel. Broadly, the resolving power of conventional techniques could be improved by increasing the aperture of the antenna array. On the other hand, a compromise between cost and performances is necessary. SRDF techniques aim to enhance the resolving power of conventional techniques by means of the intelligent analysis using modern signal processing algorithms.

SRDF employs signal samples obtained simultaneously from a spaced aperture antenna array. The signals feed multiple coherent digital receivers, with resulting data processed in fast DSP (Digital Signal Processing). The advances in DSP hardware and the developments of SRDF algorithms are constantly increasing the performances of SRDF systems. The signal samples are used to obtain the covariance matrix (see previous section). The DF function of each SRDF algorithm will contain the relevant covariance matrix to produce the bearing estimates. The output of the estimator can be shown in different ways.

The estimate of the direction of arrival of a signal, repeated for a given amount of time, will produce a trace in the case that a signal is detected at the receiver. The output of each estimate at a given instant is a graph of the angular spectrum, as shown in Figure 2.5.1, in which we consider a two-ray problem. In Figure 2.5.1a the two angles of arrival are given by the two maxima in the pattern. In Figure 2.5.1b the two arrival angles are the minima, which correspond conceptually to array nulls. These are not perfect nulls but minima, because of measurement errors; the reciprocal can be plotted to provide sharp cusps at the arrival angles, as in Figure 2.5.1c. Although the last form is favoured, the diagrams showing sharp cusps are not scientifically valid means of assessing the resolution of a spectral estimator, as pointed out by Kay and Demeure (1984). In fact, Figures 2.5.1a, b and c may be simple transforms of each other, with exactly the same resolution. Once the directions of arrival of the rays have been determined, the amplitude of each ray is calculated according to the SRDF algorithm in use.

A difficult case for many algorithms is when perfect coherence between two arriving signals occurs. Two signals are said to be coherent when the fluctuations with time of their amplitudes and phases are perfectly correlated. In this situation the two rays may not be resolved, particularly if they are closely spaced in arrival angle. Local reflections at the receiving site can produce a condition of coherence. When coherence occurs, the spectral estimator is likely to produce a peak at an intermediate angle between the two real angles of arrival. The peak will result as a weighted mean determined by the amplitude of the two rays (Gething, 1991).

As previously said, the SRDF algorithms have the common ability of resolving two signals, which are separated of an angle smaller than the natural beam of the antenna array. Although the resolution limit of these SRDF algorithms depends on several factors, among which the antenna array geometry and size, the direction of arrival, signal-to-noise ratio and relative power of the incoming signals, experimental results suggest that they generally show a resolving power which is a factor of 3-4 higher than the natural beam forming algorithm. However, the accuracy of the bearing estimates can be affected when two signals are close to one another in direction of arrival (see Chapter 6).

2.6 Superresolution direction finding (SRDF) algorithms.

One of the main limitations of conventional beamforming is the inability to resolve two closely separated incoming signals (the separation angle varying according to the antenna array and the DF algorithm in use), which is the consequence of the fact that the signal-to-noise ratio (SNR) is a limiting factor

for the resolution of conventional arrays (Manikas *et al*, 1997). In fact, the lower the SNR, the higher the chances that the weaker signal is lost in the noise. This inability gave stimulus to the development of a new class of techniques, i.e. the superresolution algorithms (Mermoz, 1981).

The most important of these techniques can be grouped in three classes:

- the Maximum Likelihood Methods (MLMs),
- the Maximum Entropy Methods (MEMs),
- the eigenanalysis methods

2.6.1 The Maximum Likelihood Methods.

The MLMs are based on the assumption that a wave-field consists of a single wanted incoming signal plus distributed noise. The bearing estimates are obtained by maximising the MLM power spectrum (Johnson and Miner, 1986), that can be expressed as

$$P_{MLM} = \frac{1}{a^H R^{-1} a} \quad (2.6.1.1)$$

where a is the steering vector, R is the sample covariance matrix (see Section 2.4), the superscript H denotes the Hermitian transpose (this notation will be used throughout this section).

An Iterative Scanning Null Beam method was developed by Brandwood (1989), based on positioning nulls on the strongest ray at each iterative scanning of the sensor array in the whole region containing incoming signals. The position of the nulls could be adjusted subsequently. Ziskind and Wax (1988) introduced an iterative algorithm (called Alternating Projection, AP) applicable to multiple coherent incoming signals. The AP technique solves first for a single incoming signal, which leads to an update equation for the projection matrix, used to solve for the second strongest signal, and so on. A convergence to stable values can be obtained after about 4 or 5 iterations (Ziskind and Wax, 1988). For a generic array of p sensors with arbitrary locations and directional characteristics, and for q incoming sources from locations $\theta_1 \dots \theta_q$, the estimate of location θ_i at the $(k+1)$ -th iteration, can be expressed as

$$\hat{\theta}_i^{k+1} = \arg \max_{\theta} \frac{a^H(\theta)(I - PA_i)^H R(I - PA_i)a(\theta)}{a(\theta)^H (I - PA_i)^H (I - PA_i)a(\theta)} \quad (2.6.1.2)$$

where $a(\theta)$ is the steering vector corresponding to location θ ; I is the identity matrix and A_i is a $p \times (q-1)$ matrix given by

$$A_i = \left[a \begin{pmatrix} \hat{\theta}_1^k \\ \vdots \\ \hat{\theta}_{i-1}^k \end{pmatrix}, a \begin{pmatrix} \hat{\theta}_{i+1}^k \\ \vdots \\ \hat{\theta}_q^k \end{pmatrix} \right]$$

and PA_i is a projection operator onto the column space of A_i .

Oh and Un (1989) suggested an iterative method, which reduced the computational burden of the AP algorithm, based on transforming Hermitian forms into inner vector products. By defining $\Gamma_i = (I - PA_i)$, with some manipulations Oh and Un obtained the following estimator

$$\hat{\theta}_i^{k+1} = \arg \max_{\theta} \frac{a^H(\theta) \Gamma_i R \Gamma_i a(\theta)}{a(\theta)^H \Gamma_i a(\theta)} \quad (2.6.1.3)$$

The algorithm proposed by Oh and Un reduces the computational burden by a factor of about $p/2$ at the price of a slight memory increase by a factor of $1/p$. Thus the efficiency of this algorithm with respect to AP becomes more significant as p increases.

Another variant of MLM is the Adaptive Angular Response (AAR) algorithm, developed independently by Borgiotti and Kaplan (1979) and Lagunas-Hernandez and Gasull-Llampallas (1984), which measured the power density instead of the power in the spectrum. The output power of the array is normalised to a thermal noise term:

$$P_{AAR} = \frac{a^H R^{-1} a}{a^H R^{-2} a} \quad (2.6.1.4)$$

Borgiotti and Kaplan demonstrated the superresolution capabilities of AAR, although they did not compare its performance to previously existing algorithms. Lagunas-Hernandez and Gasull-Llampallas showed that the algorithm they developed (which they simply named Modified MLM) had a higher resolving power than MLM.

Gabriel (1980) obtained analogous results with his Thermal Noise Algorithm (TNA) to those obtained by Lagunas-Hernandez and Gasull-Llampallas with their Modified MLM. The TNA measures the thermal noise power and its power spectrum (Equation 2.6.1.5) is given by the denominator term of AAR:

$$P_{TNA} = \frac{1}{a^H R^{-2} a} \quad (2.6.1.5)$$

Stoica and Sharman (1990) developed a Method of Direction Estimation (MODE), which combined the performance of MLMs in terms of their ability to detect highly correlated signals with the performance of MUSIC in terms of its lower computational cost. MODE satisfies the ML criterion by adjusting the

taper coefficients obtained from eigenanalysis. The estimates are obtained by minimising the following

$$f_B = \text{Tr}\{(S^* B)(B^* B)^{-1}(B^* S)(\Lambda - \sigma I)\} \quad (2.6.1.6)$$

where S is the matrix of the eigenvectors associated with the incoming signals, σ denotes the Gaussian distribution of the noise, I is the identity matrix, and $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_n\}$, λ_i being the eigenvalues of the covariance matrix R of the incoming signals plus noise. Stoica and Sharman pointed out that, while ML requires, in general, more than two iterations to converge, MODE is not intrinsically iterative, and therefore it has no convergence problems, even though one of the steps of MODE procedure may be iterated. Furthermore MODE does not require multidimensional search (unlike ML), this resulting in the computational advantage mentioned above.

Clarke (1991) and Mather (1991) developed the Incremental Multi Parameter (IMP) algorithm. By iteratively testing the SNR of the estimated number of incident signals, by means of the higher-rank spectrum estimator, IMP aims to obtain near-optimal high-discrimination performance. IMP spectrum estimator is given by

$$\text{SNR}_{M-1}(\theta_M) = [P_{a_{M-1}}^\perp a(\theta_M)]^\# X X^H P_{a_{M-1}}^\perp a(\theta_M) \quad (2.6.1.7)$$

At the first iteration $M=1$ and the orthogonal-projection operator $P_{a_0}^\perp = I_N$ (N -dimensional identity matrix), where N is the number of incoming signals. X is the data matrix, and the generic orthogonal projection operator is defined as

$$P_{a_{M-1}}^\perp = I - a(\theta_{M-1}) a^\#(\theta_{M-1})$$

where $\#$ denotes the pseudo-inverse operator and θ_{M-1} are the bearing estimates at the $(M-1)$ -th iteration step.

Another algorithm based on the ML principles is the Direction Of arrival by Signal Elimination (DOSE), developed by Zatman *et al* (1993), which iteratively attempts to detect one signal while nulling the others. Although DOSE is capable of resolving multi-component wave-fields into separate modes, when two sources are close to one another, it might provide a direction of arrival, which corresponds to an average of the actual DOAs. At each iteration, the DOAs previously estimated are suppressed, and a new estimate is obtained until the peak of the last DOA searched for is below an initially set threshold. Therefore, the DOSE algorithm does not require an a priori knowledge of the number of incoming signals. An Iterative Null Steering (INS) algorithm

(Warrington, 1995), which is based on DOSE and IMP (see Section 2.7 for full mathematical details), and also includes a steering vector tuning procedure has been utilised for this research. Another SRDF algorithm used for this research is a development of Capon's Minimum Variance Estimator, by Featherstone *et al* (1997), which they called Loaded Capon (also detailed in Section 2.7).

All above mentioned algorithms model the incoming signal as a point source. However, the received signal can be highly spread in both azimuth and elevation, which is a usual occurrence at high latitudes (see Section 1.5). Goldberg and Messer (1998) addressed the problem of localising a single coherently scattered source, whose received signal components are fully correlated at each direction over which the scattering extends (see also Valaee *et al*, 1995). Goldberg and Messer assumed a Gaussian distributed source with additive Gaussian noise. They used the Cramér-Rao bound (CRB) to study the inherent limitations, in terms of lower accuracy, in the estimation of the mean bearing of a coherently scattered source when compared to the estimate of a point source at the same mean DOA.

Trump and Ottersen (1996) attempted to address the problem of estimating the DOA and the angular spread of a source surrounded by a large number of scatterers. They proposed two computationally less complex estimators than the ML, based on least-squares fits of the sample data covariance matrix to the theoretical covariance matrix derived from the assumed model. The modelled signal assumes a scattering situation by a reasonable spatial distribution function (they used the Gaussian) and the parameters of this distribution are estimated. In particular, Trump and Ottersen showed that the proposed optimally weighted least-squares criterion is asymptotically efficient, i.e. it achieves the minimum estimation error variance with increasing number of snapshots.

If we call σ_ϕ^2 the standard deviation of the distribution, the estimates of nominal DOA and angular spread for the weighted least-squares estimator are obtained as

$$[\theta, \sigma_\phi^2] = \arg \min \text{Trace}[(SR + \sigma^2 I)W - I]^2 \quad (2.6.1.8)$$

where S is the received signal power, R is the model covariance matrix, σ is the noise variance and W is a positive-definite weighting matrix, which can be replaced by a consistent estimate

$$\hat{W} = \hat{R}_N^{-1}$$

where \hat{R}_N is the sample covariance matrix, obtained with N samples, see Trump and Ottersen (1996) for full mathematical justification. This replacement would not affect the asymptotic properties of the estimator, i.e. when N tends to infinity. Trump and Ottersen also proposed a less complex (non-weighted) form of the criterion function by choosing $W = I$, but also stated that a degradation in performance was to be expected. However, Trump and Ottersen made a number of assumptions for their model, among them that the angular spread φ is small so that $\cos(\varphi)$ and $\sin(\varphi)$ can be approximated by the first term in the Taylor series expansions. As we will see in the next chapters, the angular spread can be of several tens of degrees, in which case the approximation cannot be considered valid, and the model proposed by Trump and Ottersen would fail.

Raich *et al* (2000) introduced the concept of “partial coherence” to account for temporal correlation (as well as spatial correlation) for propagation channels formed between a source and an antenna array. They generalised the previously proposed distributed source models for DOA estimation to a parametric spatial-temporal model for what they called partially coherent distributed (PCD) sources. This model lies between the two extreme cases of incoherently distributed (ICD) sources (zero temporal channel correlation) and fully coherently distributed (FCD) sources (full temporal channel correlation). The PCD model is more likely to represent a realistic channel scenario. The inherent accuracy limitations for this model become more severe as temporal channel correlation increases. For the PCD channel model, the channel vector $c(k)$, where k is the discrete time index, can be expressed as a superposition of weighted point source contributions over a continuum of directions

$$c(k) = \int_{-\pi}^{\pi} f(\theta, k) a(\theta) d\theta \quad (2.6.1.9)$$

where $f(\theta, k)$ is defined as the angulo-temporal weighting function, which describes the temporal evolution of the scattered component arriving from each θ .

In order to address the problem of the computational complexity that characterises the ML estimators (caused by multidimensional searches over the parameter search space), Raich *et al* (2000) presented a concentrated ML estimator for PCD sources. The cost function for bearing estimates, involving a search over four parameters, is given by

$$\hat{\Psi}_{PCD}^{ML} = \arg \min_{\theta_0, \Delta, \rho, \alpha} \left\{ \sum_{k=1}^K \log \left| \frac{R_k}{\sigma_n^2} \right| + KN \log(\sigma_n^{2ML}) \right\} \quad (2.6.1.10)$$

where θ_0 and Δ denote the angular mean and angular standard deviation parameters respectively, $\rho = \sigma_s^2 / \sigma_n^2$ represents the signal power to noise power ratio, α denotes the temporal correlation between two adjacent channel vector samples, K and N are the number of snapshot measurements and antenna array elements, and

$$R_x = \sigma_s^2 \lambda_t^{(k)} R_c + \sigma_n^2 I_N \quad (2.6.1.11)$$

$\lambda_t^{(k)}$ being a weighting factor for the temporal cross-correlation matrix R_c between the channel vectors. The described concentrated ML estimator is said to be optimum in that its performance asymptotically achieves that predicted by the CRB. Raich *et al* stated that work was being undertaken focusing on the development of more computationally efficient, suboptimal alternatives to the ML estimators.

A Spread Maximum Likelihood (SML), which models the signal as coming from a distributed source, was available at the times of this research and has been used for part of the results presented in Chapter 4. The SML code was written by Read (1999). SML has been chosen as representative of those superresolution algorithms which assume the incoming signal to be generated from a spread source. However, SML was still under development (Read, 1999) at the times of this research, and its operational characteristics are detailed in Section 2.7.

2.6.2 The Maximum Entropy Methods.

The MEMs, based on the work of Burg (1972, 1975), attempt to maximise (subject to some limitations) the entropy of the output from the processing algorithm. The reader is referred to Childers (1978) for a review of the early key papers. Unlike conventional filters, which aim to remove as much noise as possible from the input signal, the MEM algorithm attempts to make its output spectrum as noise-like and structureless as possible. In order to do so, it places nulls on all the signals in the input spectrum, and the transfer function of the filter is used to determine the characteristics of the removed signals. The MEM criterion is preferable when little or nothing is known about the signal characteristics, whereas more is known about the noise features (Johnson and Sherril, 1982). Several solutions have been proposed for a multidimensional analysis. In the solution provided in the papers by McClellan (1982) and Lang and McClellan (1982), the MEM estimates are given by the MEM spectrum, P_{MEM} , that maximises the entropy $H(P)$

$$H(P) = \int_K \ln \left(\frac{1}{P_{MEM}} \right) dk \quad (2.6.2.1)$$

P_{MEM} is expressed as a positive polynomial in the frequency-wavenumber domain K over which power is assumed to be present (the above mentioned papers provide full mathematical details).

Nickel (1987) provided another calculation of the MEM spectral estimator. Nickel explained that the first and last columns of the covariance matrix, corresponding to the largest space interval between the sensors of an antenna array, give full benefit of its aperture. Nickel proposed an MEM spectral estimator by using the first column of the inverse of an estimate of lower order of the covariance matrix:

$$P_{MEM} = \frac{\left(\tilde{R}^{-1} \right)_{11}}{\left| a^H \left(\tilde{R}^{-1} \right)_{.1} \right|^2} \quad (2.6.2.2)$$

where

$\left(\tilde{R}^{-1} \right)_{11}$ is the (1,1) element of \tilde{R}^{-1} and

$\left(\tilde{R}^{-1} \right)_{.1}$ is the first column of \tilde{R}^{-1} .

Maximum Entropy Methods have shown lower angular resolution than other estimators, and therefore no MEM has been utilised for this research.

2.6.3 The eigenanalysis methods.

The eigenanalysis methods (Schmidt, 1981) are based on the principle of searching for the best linear relationships among the sensor signals by decomposing the covariance matrix into its eigenvectors and eigenvalues. Such methods are sensitive to sampling and modelling errors and uncertainties. In fact, they are based on a precise knowledge of the array characteristics (geometry, sensor gain and phase, mutual coupling between elements, etc.). If these parameters are incorrectly applied, the eigenanalysis methods give poor results; therefore calibration with respect to array parameters is crucial in signal-subspace techniques (Fistas and Manikas, 1994).

Eigenanalysis methods have been developed for the study of adaptive arrays (Gabriel, 1976) from the 70s. Pisarenko (1972) developed a spectral estimator form from various weighted means of the eigenvalues of the covariance matrix and studied its characteristics, and also suggested (Pisarenko, 1973) a solution method based on the smallest eigenvalue, λ_{\min} , which gives the noise power. Pisarenko's method, applied to DOA estimates in the case of noise-free

propagation environment and r incoming rays, produces a matrix of $(r+1) \times (r+1)$ elements, one eigenvector being orthogonal to the r ray vectors. The corresponding eigenvalue is zero in a noise-free environment and σ^2 with noise, where σ^2 is the variance of the noise distribution. The DOAs can be determined from the roots of a polynomial, whose coefficients are the elements of the eigenvector corresponding to λ_{\min} , but with applicability limited to λ_{\min} having multiplicity one. However, a generalisation of Pisarenko's method was presented by Lang (1981) for λ_{\min} with multiplicity greater than one.

The importance of eigenanalysis for radio direction finding was recognised by Gething (1978), who highlighted the importance of distinguishing between signal and noise eigenvalues, and later by Ellis (1980), who gave a complete geometrical interpretation. Van Blaricum and Mitra (1978) discussed applications of the eigenanalysis method and, by using the known standard deviation of the noise, provided the basis for the determination of the maximum number of DOAs. Bienvenu and Kopp (1980, 1981) investigated some properties of the eigenanalysis solution (such as the distinction between signal and noise eigenvectors). Whereas Reddi (1979) used the signal subspace for the DOAs estimates, the Multiple Signal Classification (MUSIC) algorithm (Schmidt, 1979, 1981, 1986) uses the noise subspace (see Equation 2.7.1 for the mathematical expression of the MUSIC spectrum). MUSIC is the most well-known and widely-used eigenanalysis method for direction finding. Although many other eigenanalysis methods have shown interesting features, MUSIC can be regarded as the classical eigenanalysis algorithm, and as such it has been utilised for this research as a reference SRDF algorithm (detailed in Section 2.7).

As MUSIC performs poorly in case of fully correlated sources, a number of techniques have been developed to overcome this problem. However, some of them have the limitation to be applicable only to linear arrays of uniformly-spaced identical sensors, such as the technique called 'spatial smoothing' (see, for example, Pillai, 1989). However in this thesis we focus on algorithms applicable to any array geometry. A development from MUSIC capable of working with a generic array geometry is the Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT) (Paulraj *et al*, 1986; Roy and Kailath, 1989). ESPRIT has significant advantages in computer speed with respect to MUSIC, and it can work without array calibration. However, to achieve this ESPRIT imposes a constraint on the structure of the antenna array, i.e. the N elements have to be displaced in $d = N/2$ doublets. The elements of each doublet must have identical sensitivity patterns, and are translationally separated by a displacement vector Δ , which sets the reference direction. Thus

this array corresponds to having two identical sub-arrays translated by d , which is the magnitude of Δ .

Let us consider $n \leq d$ incoming sources centred at frequency f and with speed c . Let us call

$$\Phi = \text{diag}\{e^{j\varphi_1}, \dots, e^{j\varphi_n}\} \quad (2.6.3.1)$$

the diagonal matrix of the phase delays between the doublet elements for the n wavefronts, where

$$\varphi_k = \frac{2\pi f d}{c} \sin \theta_k \quad (2.6.3.2)$$

θ_k being the DOA of the k -th signal. The diagonal elements of Φ are equal to the eigenvalues ($\gamma_1, \dots, \gamma_n$) of the ESPRIT covariance matrix for the signal subspace (see Roy and Kailath, 1989, for full mathematical justification). The estimate of the DOA of the k -th signal is then given by

$$\hat{\theta}_k = \sin^{-1} \left\{ c \arg \left(\frac{\hat{\gamma}_k}{2\pi f_0 d} \right) \right\} \quad (2.6.3.3)$$

However, since the maximum number of incoming signals that can be detected is half the number of antennas, ESPRIT has the disadvantage of a much reduced maximum number of rays that can be resolved with respect to MUSIC, using the same array geometry and dimensions. Furthermore, ESPRIT (like MUSIC) shows difficulties in resolving coherent signals.

Manikas and Turner (1991) developed an Adaptive Signal Parameter Estimation and Classification Technique (ASPECT), which is capable of resolving correlated signals and can also be applied to an arbitrary array geometry. This algorithm uses an initial guess as to the number and the DOAs of incoming signals, since the actual number of the sources present in the signal environment cannot be known *a priori*. The number and DOAs of the incident signals are then estimated during the processing. However, any initial information about the incoming signals will speed up the process of convergence. As with MUSIC, a wrong estimate of the number of signals with ASPECT can lead to wrong estimates of the DOAs. The bearing estimates are obtained by minimising one of the three following ASPECT cost functions

$$\Psi = \text{trace}(P_a P_E^\perp P_a P_E) \quad (2.6.3.4)$$

$$\Psi = \prod_i \left[\frac{1}{\sqrt{E_i^H P_a E_i}} \right]^r \quad (2.6.3.5)$$

$$\Psi = \text{trace}(P_a^\perp P_E) \quad (2.6.3.6)$$

where E is the matrix whose columns are the eigenvector belonging to the signal subspace, P and P^\perp are the projection and the orthogonal-projection operators respectively, and n is the number of signals.

To address the problem of estimating the DOA of a spread source (see sub-Section 2.6.1), Lee S.R. *et al* (1996) proposed a parametric distributed source model, and Lee Y.U. *et al* (1997), proposed both a parametric and a non-parametric distributed source model. Lee Y.U. *et al* considered signals from different sources to be either correlated or uncorrelated, and signals within a source to be correlated. The shape of the distributed source is assumed to be known, and is represented in terms of the spatial harmonics. The output of an array is obtained by integrating the effect of the incoming plane wave signal over all directions weighted by the distribution source density. The parametric source model must satisfy the condition that the distributed source density of two closely-spaced combined sources located at any point between their centres is less than the source density of either one at its centre. The non-parametric source is assumed to be a general function defined by any azimuth-of-arrival ($0 \leq \theta < 2\pi$). The aim with the non-parametric model is generalised to estimating the distributed source as opposed to the conventional DF problem of estimating the number and DOAs of point-like sources. Lee Y.U. *et al* used an eigenstructure-based method for the DOA estimation. The estimates of the DOA θ_i and distribution parameter ρ_i for the i -th distributed source are obtained by minimising the cost function

$$\left(\hat{\theta}_i, \hat{\rho}_i \right) = \arg \min \left\{ a_{p/np}^H(\theta, \rho) \hat{G} \hat{G}^H a_{p/np}(\theta, \rho) \right\} \quad (2.6.3.7)$$

where \hat{G} is an estimate of the noise subspace matrix and $a_{p/np}$ is the steering vector for the parametric/non-parametric model. Lee Y.U. *et al* (1997) displayed their results for a uniform linear array for sake of simplicity, and stated that these are also valid for a generic array geometry. However, they did not specify in what circumstances their model would fail for a generic array.

In the next section the SRDF algorithms utilised for this research will be described in detail.

2.7 MUSIC, Iterative Null Steering (INS), Loaded Capon and Spread Maximum Likelihood (SML) algorithms.

In this section the SRDF algorithms used in this research are introduced. These are the Loaded Capon, INS, MUSIC and SML algorithms. MUSIC and Capon

(Capon, 1969, 1970) are commonly used in operational environments, which is the reason why MUSIC and Loaded Capon chose for this research, MUSIC being a classic algorithm and Loaded Capon a recent development of the original MVE (Minimum Variance Estimator) Capon algorithm. INS has been utilised in order to test its validity at high latitudes in comparison to the other two SRDF algorithms, since it had been used successfully at mid-latitudes (Warrington, 1995) (see Section 2.8). SML, the most recent, has been utilised because it models the high latitude propagation environment more accurately than the other algorithms mentioned above, even though it has some limitations. The overall choice of these algorithms is driven by the fact that they are representative of different types of algorithms. It follows that a comparison of their performance is very useful since it allows us to infer the characteristics of the propagation environment, and therefore the behaviour of the algorithms in such an environment. The results should also be of great interest to DF operators at high latitudes, as we will see in the following chapters. The MATLAB programming language has been used to implement the five algorithms utilised in this research (Beam and the four superresolution ones named above).

Multiple Signal Classification (MUSIC)

Hill (1990) and Gething (1991) provided an introduction to this algorithm, whereas for a full description the reader is referred to Schmidt (1986). It has been widely used for direction finding of uncorrelated signals within the HF and VHF bands.

MUSIC is based on the formation of a covariance matrix (Equation 2.4.5) from samples of the signal amplitude and phase at a number of spaced antennas followed by eigen-decomposition of the covariance matrix into terms associated with the incident signals and other terms associated with the noise components. The various signal components (propagation modes, co-channel signals and noise) must be uncorrelated over the interval when the data are collected. In fact, in the case of correlation between signal components, the accuracy is greatly reduced and the algorithm is unable to resolve them.

The Equation 2.7.1 shows the MUSIC function

$$f_M(\alpha, \epsilon) = \frac{1}{a(\alpha, \epsilon)^H R_N a(\alpha, \epsilon)} \quad (2.7.1)$$

where

$$R_N = E_N E_N^H$$

and

E_N is a matrix whose columns are the eigenvectors assumed to be associated with the noise components.

In Equation 2.7.1 the directions of arrival correspond to minima in the denominator resulting in sharp peaks in the overall function.

The number of incident signals is unknown in all practical applications of the algorithm. It is therefore necessary to estimate this parameter before R_N can be calculated. In simulation, the eigenvalues tend to belong to two separate groups, one with large values belonging to the signal components, the other with small values belonging to the noise components. The number of signal components can be therefore estimated from the number of the large eigenvalues. With real data, the division between the two groups becomes more blurred and, furthermore, the noise may swamp an eigenvalue associated with a small signal.

It is important to note that the eigenvalues are not a measure of the power of the individual signal components. Therefore a strong signal component may be associated with a relatively small eigenvalue. The estimates of the directions of arrival are likely to be erroneous if the number of signal components is wrongly estimated. Gething (1991) showed that large errors occur in the case of under-estimation, reduced errors in case of mild over-estimation, whereas larger over-estimation leads to additional peaks which may wrongly be interpreted as directions of arrival of a signal.

When the MUSIC algorithm is implemented, it is therefore necessary to set a threshold between the eigenvalues associated with the signal components and the eigenvalues associated with the noise components. The other parameter to set is the number of signals estimated. The position of the peaks in $f_M(\alpha, \epsilon)$ is then determined. The number of the largest peaks is the same as the number of signals estimated. The positions of the largest peaks are associated with the true directions of arrival.

Iterative Null Steering (INS)

The INS algorithm (Warrington, 1995), which is based on IMP and DOSE, introduced in Section 2.6, operates iteratively by first estimating the direction of arrival of one signal from the peak of a scanned (in software) beam whilst suppressing all the other signal estimates through the projection of nulls. Initially, a single signal component is assumed and its direction of arrival is estimated as the direction of the maximum received power from the scanned beam. A null is then positioned in this direction and an estimate made of the direction of arrival of any second component. The second component is then nulled and a revised estimate made of the original direction and the process

repeated until the estimated directions of both signal components remain stable. Both of these signal components are then nulled and if there is still significant energy being detected in the scanned beam, the number of signals searched for is increased by one and the iteration process repeated with all but one component being nulled at any time. The search terminates when the peak of the scanned beam (with the nulls imposed) is less than a pre-defined threshold or when a new estimate is identical to one of the nulled directions.

In order to compensate for system errors and the effect of noise, which can result in the failure to correctly null out incoming energy from a specified direction, a steering vector tuning procedure has been incorporated (Moyle and Warrington, 1997). Once the peak has been found by the beam scanning process, the associated steering vector is then adjusted ("tweaked") by successively modifying each element in turn by a small amount ($\pm 0.1\text{dB}$ and then $\pm 0.1^\circ$) and recalculating the value of the peak with the modified steering vectors. The modified steering vector giving the largest peak is then used in further stages of the algorithm. The tweaking process is repeated until no further increases in the peak value occur (i.e. the algorithm adopts a "hill climbing" approach). The final modified steering vectors are then used in the nulling procedure to form deeper nulls when searching for other signal components.

A conditioning matrix, Q (given by Equation 2.7.2), based on a Gram-Schmidt projection can be applied to the signal data (Equation 2.7.3) to form nulls in specified directions (Hudson, 1981). Note that the signal-conditioning matrix is Hermitian (i.e. $Q = Q^H$) and that $Q = Q^2$.

$$Q = I - A(A^H A)^{-1} A^H \quad (2.7.2)$$

$$y = Qx \quad \text{or} \quad Y = QX \quad (2.7.3)$$

where

A is a matrix whose columns are the steering vectors corresponding to the required null directions, and

I is the identity matrix.

The functions on which the IMP and DOSE algorithms are based may be written as

$$p_N(\alpha, \epsilon) = \frac{a(\alpha, \epsilon)^H y y^H a(\alpha, \epsilon)}{a(\alpha, \epsilon)^H Q a(\alpha, \epsilon)} \quad (2.7.4)$$

for snapshot data, or for more than one sample set by

$$p_N(\alpha, \epsilon) = \frac{a(\alpha, \epsilon)^H R_C a(\alpha, \epsilon)}{a(\alpha, \epsilon)^H Q a(\alpha, \epsilon)} \quad (2.7.5)$$

where

R_C is the conditioned covariance matrix given by

$$R_C = \frac{1}{N} YY^H \quad (2.7.6)$$

$p_N(\alpha, \epsilon)$ represents the power received in the specified direction after nulls have been imposed on the directional sensitivity pattern of the array. The term in the denominator of Equations 2.7.4 and 2.7.5 represents the gain of the array in the specified look direction after the nulls have been imposed. It is included in these equations to correct for the influence of the signal conditioning at directions other than the nulled ones.

This algorithm may be regarded as an approximate maximum likelihood estimator in which the detected power (given by the Equations 2.7.4 or 2.7.5) is minimised by a suitable choice of steering vectors included in matrix A (see Equation 2.7.2). This differs from the true Maximum Likelihood Estimator (MLE) in which the function given in Equation 2.7.7 is minimised (the trace function is a matrix operator which returns the sum of the elements on the leading diagonal of the argument).

$$f = \text{Trace}(QR) \quad (2.7.7)$$

Loaded Capon

The original algorithm, named after its inventor Capon (1969), is based on a Minimum Variance Estimator (MVE) method. Capon created a high resolution algorithm which marked a decisive step forward with respect to the conventional algorithms, whose resolution was determined essentially by the natural beam pattern of the array of sensors. His experimental results showed an improvement in resolving power of about a factor of four. In this original algorithm the steering weights for a specific look direction were adjusted to maintain a constant processing gain in that direction but to minimise the total output power from the array, thus minimising contributions to the output power arising from noise and from signals arriving at other directions.

The covariance matrix of the Capon/MVE algorithm is given by

$$R = \frac{1}{K} \sum_{i=1}^K x_i x_i^H \quad (2.7.8)$$

where K is the number of samples and x_i is the i^{th} complex data vector recorded by the receiving array.

The output power of the antenna array is given by

$$P_{MVE}(\alpha, \varepsilon) = \frac{1}{a(\alpha, \varepsilon)^H R^{-1} a(\alpha, \varepsilon)} \quad (2.7.9)$$

An enhanced version of the Capon's original estimator has been used in the work of this thesis, which is named 'Loaded Capon' (Featherstone *et al*, 1997), which overcomes the possible errors in the estimate of the covariance matrix, which occur when using the MVE algorithm, in particular when relatively few data samples are available. A non-infinite sample support causes the spread of the eigenvalues of the noise subspace of the covariance method, thus degrading the performance of the MVE algorithm. This problem can be overcome by adding a constant to each eigenvalue, this constant being given by a fraction of the total power contained within the covariance matrix. The leading diagonal of the covariance matrix is then loaded by this constant and the amount of the noise subspace corrupting the output will thereby be lowered.

The possibility of successfully estimating the direction of arrival of a signal by means of only a few samples of the data set is a highly desirable feature, in particular in conditions of disturbed propagation. In fact, the possible sudden change in propagation conditions may affect the signal in a very short amount of time (see Section 4.3 for an investigation on the effect of the integration time on the observed bearing spread). The Loaded Capon algorithm operates successfully when using only a few samples of the data set to form a time averaged covariance matrix and this allows the signal to be tracked in conditions of disturbed propagation.

The covariance matrix of the Loaded Capon is

$$R = V\Lambda V^H \quad (2.7.10)$$

where

$$\Lambda = \text{diag}(\lambda_1 \lambda_2 \dots \lambda_N)$$

and

$$V = [v_1 v_2 \dots v_N]$$

$\lambda_1, \lambda_2, \dots, \lambda_N$ are the eigenvalues of R and

v_1, v_2, \dots, v_N are the corresponding eigenvectors.

The Matrix Inversion Lemma allows the covariance matrix to be rewritten, by a scaling factor of $1/\lambda_{\min}$, as follows

$$R^{-1} = I - \sum_{i=1}^N v_i \frac{\lambda_i - \lambda_{\min}}{\lambda_i} v_i^H \quad (2.7.11)$$

where λ_{\min} is the smallest eigenvalue.

The Minimum Variance Estimator algorithm in Equation 2.7.9 can now be rewritten as

$$PLC(\alpha, \varepsilon) = \frac{1}{a(\alpha, \varepsilon)^H a(\alpha, \varepsilon) - \sum_{i=1}^N \frac{\lambda_i - \lambda_{\min}}{\lambda_i} |a(\alpha, \varepsilon)^H v_i|^2} \quad (2.7.12)$$

Each eigenvector will then contribute to the output in the proportion to $(\lambda_i - \lambda_{\min})/\lambda_i$. When the number of samples used to form the covariance matrix is infinite, every eigenvector has an eigenvalue of λ_{\min} , in the noise subspace. The noise subspace will therefore make no contribution to the output, because $(\lambda_i - \lambda_{\min})/\lambda_i = 0$.

In the case when a finite number of samples are used, the eigenvalues are spread, i.e. $\lambda_{m+1} > \lambda_{m+2} > \dots > \lambda_{\min}$, where m is the rank of the signal subspace. Besides, if the noise in data samples is not statistically independent, the noise eigenvalues may be smaller. In this case we have $(\lambda_{m+i} - \lambda_{\min})/\lambda_{m+i} \neq 0$, because $\lambda_{m+i} \neq \lambda_{\min}$. The output is therefore corrupted by the noise subspace.

The Loaded Capon method reduces the term $(\lambda_{m+i} - \lambda_{\min})/\lambda_{m+i}$ by adding a constant to each eigenvalue, thus diagonally loading the covariance matrix. The practical effect is then to reduce how much the noise subspace corrupts the output. The effect on the signal subspace's contribution to the output is minimal, providing that the signal subspace eigenvalues are much larger than the loading level. Furthermore, the diagonal loading will help in the inversion of the covariance matrix, in the case that the spread of the eigenvalues originates an ill-conditioned covariance matrix. The modified covariance matrix R_L can be defined as

$$R_L = R + \alpha I \quad (2.7.13)$$

where I is the identity matrix and α is the loading factor.

Spread Maximum Likelihood (SML)

Many DF algorithms assume that the incoming signal is generated by a point source, an assumption which would not provide optimum bearing estimates when the received signal is spread over a range of azimuth and elevation angles, which is a usual occurrence at high latitudes (see Chapter 3). This has given the stimulus to researchers, in the recent years, to develop new algorithms which can model the signal as arriving from a distributed source, as

well as a point source. Read (1999) developed the SML algorithm to improve the sensitivity of DF systems to weak signals in the presence of stronger spread signals. Therefore, SML would be more suitable to the high latitudes, since it would model more realistically the propagation encountered there, which is the main property compared to the other algorithms utilised for this research.

SML assumes both signal and noise to be random processes. This is because, even though the signal is transmitted at known amplitude and phase, the often highly variable ionospheric environment tends to randomise both amplitude (eg. fading) and phases (eg. Doppler shifting), see Section 1.3. By assuming Gaussian random processes, the associated probability density function, valid for HF measurements (no dc component) is given by

$$f(x_0, x_1, \dots, x_{K-1}) = \frac{1}{(\pi^N \det C)^K} e^{-\text{trace}(X^H C^{-1} X)} \quad (2.7.14)$$

where x_k (for $k=0, \dots, K-1$) is the sample data vector measured by N antennas at time instant t_k , X is the $N \times K$ data matrix whose columns are the data vector, and C is the $N \times N$ model covariance matrix describing the spatial correlation between the antennas. The estimates are obtained by choosing the covariance matrix C that maximises $f(x_0, x_1, \dots, x_{K-1})$, which is equivalent to maximising the cost function

$$L = -\ln(\det C) - \text{trace}(RC^{-1}) \quad (2.7.15)$$

where R is the data covariance matrix given by

$$R = \frac{1}{K} XX^H$$

However, SML has the limitation to fail for angular spreads higher than 15° (see next section). Nevertheless, SML was available at the times of this research and has been utilised in order to compare its results (see Chapter 4) with those obtained by means of the other algorithms described in this section, which are substantially different.

2.8 Comparison between the SRDF algorithms utilised for this research.

In the last three decades, superresolution direction finding algorithms have been considerably improved, and new ones have been developed (e.g. Read, 1999). Many researchers undertook comparisons between the performances of these algorithms applied to both real and simulated data, for direction finding of signals propagating at different latitudes and therefore in different

ionospheric environments. For the reasons discussed in the previous two sections, the SRDF algorithms utilised for this research are MUSIC, INS (based on IMP and DOSE), Loaded Capon and SML (detailed in Section 2.7), and therefore, for the scope of this thesis, this section focuses on the comparison between them. It is worth bearing in mind that, of these algorithms, MUSIC and Capon are the 'classical ones', whereas INS and Loaded Capon are more recent developments of previously existing algorithms.

An important parameter in evaluating the performance of a direction finding algorithm is its ability to identify a target at the lowest possible value of signal-to-noise ratio. From this aspect, MUSIC was outperformed by IMP (Mather, 1991). Making use of simulated data, IMP was shown to be capable of resolving two correlated signals at a signal to noise ratio 10-12dB lower than that required by MUSIC to resolve uncorrelated signals. Furthermore, Mather demonstrated the higher robustness of the IMP algorithm with respect to the MUSIC, since the results obtained with IMP were less affected by antenna calibration errors.

Zatman and Strangeways developed the DOSE algorithm, reported by Zatman *et al* (1993). They undertook analysis of data collected over three mid-latitude paths by means of a multichannel receiver, and compared the performances of DOSE and MUSIC. The results showed that DOSE was characterised by a much higher resolving power and robustness than MUSIC. The results obtained with DOSE were also compared to those obtained with interferometry techniques (see work reported earlier by Warrington and Jones (1991)), which were incapable of resolving multimoded signals in the experiments reported by Zatman (see Gething, 1991, for interferometry techniques in multisignal/multimoded environments). In case of single-moded propagation the results were similar, except that with DOSE the spread of the estimated directions of arrival was much smaller.

Manikas *et al* (1997) compared the performance of the IMP and DOSE algorithms (among others) operating in a coherent signal environment with simulated signals, using a planar 6-element circular array of 25 metres radius with a natural beamwidth of 23°. They used two coherent sources separated by various angles and additive noise taken to be 20dB and 10dB below the power of the two signals. The results were similar with both values of signal-to-noise ratio. IMP and DOSE failed for angular separations of 25° (or smaller) and 15° (or smaller) respectively.

IMP and DOSE have been taken as the basis for developing the INS algorithm by Warrington (1995). INS does not assume that the incident signals are uncorrelated, it does not require a predefined number of estimates, and it includes a steering vector tuning procedure (see Section 2.7). INS and MUSIC,

applied to experimental data from three mid-latitude paths, were compared by Warrington (1995). The data were analysed in one snapshot and for the following integration times: 1s, 2s, 4s, 8s, 16s and 32s. INS was significantly better than MUSIC at resolving the various propagation modes in terms of success rate and bearing standard deviation, particularly for short signal analysis intervals. Warrington concluded that the poor performance of the MUSIC algorithm was mainly due to its characteristic of assuming that the incident signals were uncorrelated, whereas they can be well correlated, especially in the case of multimoded propagation and over short periods.

The performances of MUSIC in terms of variances of estimated DOAs were compared to the Capon and the Loaded Capon algorithms by Featherstone *et al* (1997). As we have seen in Section 2.7, Featherstone *et al* (1997) improved the Capon's MVE algorithm (1969) into an algorithm which they named Loaded Capon. They used both simulated data and experimental data for a mid-latitude path (recorded with a multichannel HF DF system) to emphasise the advantages of their new Loaded Capon algorithm with respect to the classic Capon's algorithm (1969). For a more comprehensive comparison, they processed the same data by means of MUSIC. Figures 2.8.1a&b illustrate the variances of the estimated DOAs versus number of samples for simulated data, for signal to noise ratios of 9dB and 12dB respectively, using all three algorithms. In both cases, Loaded Capon showed lower variances than Capon, this difference being very remarkable for low numbers of samples. While MUSIC consistently exhibited lower variances than Loaded Capon regardless of the number of samples, Loaded Capon has two main advantages with respect to MUSIC: a) it requires a much lower computation burden and b) it does not require *a priori* knowledge of the number of signals being received. The errors which occur when this number is incorrectly estimated are therefore avoided, errors which are more likely to occur in multicomponent wavefields.

A multicomponent wavefield is generally produced by ionospheric multipath or co-channel interference from other transmitted signals. This can cause two or more signals to arrive at the receiver with different angles of arrival and different amplitudes. Johnson and Miner (1986) undertook simulations that compared the performance of several superresolution algorithms in resolving multiple signals. In particular, the comparison between the MUSIC and the Capon algorithms, in terms of angular resolution and co-channel interference, is of special interest for this thesis. For the simulations, the antenna array (Figure 2.8.2) consisted of five antennas deployed in a crossed array configuration in the X-Y plane, separated by 30 m. Two additional antennas were located vertically on the Z axis, over the centre of the crossed array. Signals at 4 MHz

were assumed to be incident on the composite 7-antenna array, which implied a spacing between antennas of 0.4λ . Johnson and Miner simulated two unspread and mutually incoherent signals arriving at the receiver with elevation angle separations of 10° , 5° and 4° , having the same power and the same signal-to-noise ratio (Figure 2.8.3). MUSIC resolved the two signals in all cases, whereas Capon only resolved the two signals when they were separated by 10° , whilst in the other two cases it showed one broad peak. A co-channel interference simulation was also carried out to reproduce the HF propagation environment perturbed by an interfering signal. In this simulation, Johnson and Miner assumed three unspread and mutually incoherent signals with relative signal amplitudes of 1.0, 0.5 and 0.1, elevation angle separation of 10° to each other, same signal to noise ratio (Figure 2.8.4). The MUSIC algorithm resolved the three signals for every case considered, while the Capon algorithm never resolved the weakest signal, but did resolve the two strongest signals for angular separation of 10° or more. The research reported by Johnson and Miner is limited to the case of unspread incident signals, which is often an unrealistic scenario, especially at high latitudes (see Chapter 6).

The ability of the DF algorithms to detect more than one signal is obviously a crucial one. Investigations on the sensitivity of some algorithms were carried out by Dumas (1997), who compared the performance of several algorithms, including MUSIC, INS and SML (Read, 1999) discussed later in this section, applied to an experimental data interval. Figure 2.8.5 shows that the gross structure of the traces detected by means of the MUSIC and the INS algorithms was similar, both in the case of one signal detected and when evidence of a secondary signal showed. Analysis of the same portion of data, undertaken for this research with INS, showed that the secondary signal could be detected for a longer time interval (see Figure 4.2.6), possibly due to a different termination threshold (see Section 2.7), which was not specified by Dumas. It is evident in Figure 2.8.5 that SML was remarkably more sensitive to the secondary signal than the other algorithms.

Read (1999) compared some results obtained with SML and MUSIC using experimental data. The traces detected by SML were cleaner, i.e. less scattered, than those for MUSIC, and this can be seen in Figure 2.8.5 for the analysis undertaken by Dumas. Using simulated data, SML performed better than MUSIC, especially when the problem of estimating the bearing of a weak point-source signal in the presence of a stronger distributed-source signal was considered (see Section 2.7). However, the modelling failed when the azimuth and/or elevation spreading of a distributed signal was assumed to be larger than 15° . This is an important limitation, as incoming signals often arrive at the

receiver spread in azimuth for more than 15° (see Chapters 3 and 4). Furthermore, SML required a much higher computational burden than MUSIC. Recently, Read has improved the latter deficiency with an algorithm, which was named Spatial Incoherent Region Estimator (SPIRE) (Read, 1999). Even though the results obtained by means of SPIRE appeared to be promising, Read admitted that further efforts were needed to investigate the full capabilities and limitations of this new algorithm. Since it has been developed so recently, it has not been possible to utilise SPIRE in this research. However, the assumption that a signal propagating at high latitudes has to be modelled as generated by a spread-like source is certainly supported by the results presented in this thesis.

2.9 Visibility of multiple signals.

Different antenna array geometry and SRDF algorithms will generally perform differently, especially when multiple signals are present. The presence of both single-reflection (point source) and multiple reflection/scattering (extended source) ionospheric radio propagation is typical of the observed high latitude night-time conditions (Jenkins, 1997). The extended source can be caused by a reflection from moving features in the F-region (see Section 1.5). The DOAs tend to cover the solid angles subtended by these features. At the same time, the occasional appearance of a sporadic E-layer can give rise to signal propagation along the Great Circle Path. The received signal may be weaker and narrower than the spread signal reflected from the F-region.

Jenkins modelled such an environment, in which to compare the performances of four array patterns, shown in Figure 2.9.1, and two algorithms, MUSIC and deterministic Maximum Likelihood. All four arrays consisted of 12 antennas. The Vortex and the log-spiral array were based on previously implemented experimental arrays, while the centred circle and the star array were selected on the basis of observed good DF performance in low signal-to-noise situations. The analysis was undertaken over a wide range of aperture sizes and frequencies of operation. The arrays were scaled in size so as to have the same mean aperture (over all azimuths) at any frequency. This allowed performance to be evaluated separately in terms of array geometry and aperture in wavelength. The modelled data were processed by means of MUSIC and deterministic Maximum Likelihood (ML) (Jenkins, 1994), both requiring an estimate of the number of signal directions.

The performances of the four array patterns and the two algorithms were characterised in terms of visibility of the weaker signal, i.e. how well a weaker point source was observed in the presence of a stronger extended source. Figure 2.9.2 shows several azimuth-time plots obtained with the ML algorithm, for

different point-source/spread-source power ratios. The point source is set at 180° azimuth East of North, the spread source is scattered between 140° and 165° E of N. The trace of the point source is clearly visible when the point source power is -10dB relative to the spread source. The direction of arrival of the point source becomes harder to estimate when the power ratio decreases to -15dB , and even more at -20dB . At -25dB point-source/spread-source power ratio, the point source trace cannot be seen as such. The visibility of two spread signals is presented later in this thesis (see Section 6.3, including a comparison with Jenkins' results).

The effect of array geometry on DF performance is illustrated in Figure 2.9.3, where the curves of point-source visibility as a function of aperture in wavelength are shown, for the four array geometries. The three plots correspond to three point-source/spread-source separations: 5° , 10° and 15° azimuth. The general effects for all four array patterns is an improvement in performance as the aperture is increased up to an optimum size (which depends on the point-source/spread-source separation), followed by a subsequent deterioration in performance as the aperture is increased over the optimum size. This was explained as follows: a) when the aperture is too small, a DF array cannot resolve two closely separated sources, because its beam is too broad; b) when the size of the array is relatively increased, the beam is relatively broad, and thus few steered directions are needed to cover a spread source, which leaves more directions to be used to detect a weaker point source; c) when the aperture is too large, the beam is so narrow that all available directions are needed to detect the amount of energy of the stronger signal, leaving none for the weaker point source. According to these criteria, the results of the four array geometries were similar for small apertures. But for optimum and large apertures, the star array (consisting of three arms with its smallest spacing at its extremities) performed much better than the other three array geometries, in terms of visibility of the secondary energy, for all three values of azimuth separation. The comparison between the MUSIC and the ML algorithms showed that the latter performed better, at the cost of a heavier computational burden.

2.10 SRDF for non-zero bandwidth and finite angular spread signals.

A classical definition of narrowband, when an adaptive antenna array is employed for signal detection, is that 'there is no decorrelation between signals received on opposite ends of the array' (Compton, 1988). Zatman (1998)

provided a more analytical expression for the classification of a signal environment as narrowband.

For sake of simplicity, Zatman considered a uniform linear array of N antennas, with half-wavelength spacing at the operating frequency f_0 . However, the results can be extended to any irregular array geometry and any operating frequency. The steering vector representing the transfer function between the angle of arrival θ from the array broadside and the output of the array is given by

$$a(\theta) = [1, e^{j\varphi}, \dots, e^{j(N-1)\varphi}]^T \quad (2.10.1)$$

where T represents the transpose operator and the inter-element phase φ is given by

$$\varphi = \frac{2\pi df}{c} \sin \theta \quad (2.10.2)$$

d being the inter-element spacing, f the frequency of the received signal and c the propagation velocity. As φ depends on both the angle of arrival and the frequency of the incoming signal, for a linear array, a non-zero bandwidth signal appears as an extended angular source, whereas a zero bandwidth signal appears to be received from a discrete source (studies of the spread of incoming signals are presented in Chapters 3 and 4 for experimental data and in Chapters 5 and 6 for modelled data).

A covariance matrix R (see Section 2.7) can be expressed as

$$R = V\Lambda V^H \quad (2.10.3)$$

where V is the matrix of eigenvectors and Λ is the diagonal matrix of eigenvalues. If we partition the eigenvectors into a signal and a noise subspace, we obtain

$$R = V_s \Lambda_s V_s^H + V_n \Lambda_n V_n^H \quad (2.10.4)$$

where the subscripts s and n refer to the signal and noise subspaces respectively. The matrix V_s is the minimum-rank orthonormal basis for the received signals. In the zero bandwidth case, the rank of the signal subspace is equal to the number of signals present, which means that each signal present corresponds to a rank-one representation in signal subspace. Zatman (1998) defined the 'effective rank' of a signal-only covariance matrix as the number of signal subspace eigenvalues greater than 1 (i.e. 0dB since no noise is included). If the bandwidth of a signal is sufficiently wide, the rank of the signal subspace

can be larger than the number of signals present, because one or more signals can have a rank representation greater than one. This means that one or more signals will be represented by two (or more) eigenvalues. In such cases the received signal (or signals) cannot be assumed to be zero-bandwidth. Zatman (1998) defined the notion of narrowband signal as follows: 'If the bandwidth of a signal is such that the second eigenvalue of the signal's noise-free covariance matrix is larger than the noise level in the signal-plus-noise covariance matrix, then that signal may not be described as narrowband'. Therefore, as a signal's bandwidth increases, eigenvalues pop up from the noise floor one at a time. In order to define a signal as non-zero-bandwidth, an expression of the two larger eigenvalues is sufficient. Methods for the calculation of the eigenvalues of a covariance matrix have been proposed by Hudson (1981) and Lee H.B. (1992). The derivation used by Zatman (1998) is a special case of the general results by Hudson (1981).

Let us consider the case of two equal power uncorrelated zero-bandwidth signals. The two eigenvalues of the corresponding covariance matrix (see Zatman, 1998, for the full mathematical derivation) are given by

$$\lambda_{1,2} = N \frac{s_m}{2} (1 \pm |\psi|) \quad (2.10.5)$$

where $s_m/2$ is the power of each of the two signals and ψ is the cosine of the angle between the two eigenvectors in the N -dimensional space.

For the purposes of this discussion, we are interested in the smaller eigenvalue, which can be obtained by using the '-' sign in Equation 2.10.5. Now we need to apply the formulation of the obtained second eigenvalue of the covariance matrix to the wideband problem, for which we need to calculate $|\psi|$.

Because a signal ceases to be narrowband when its effective rank increases from one to two, we can affirm that two discrete uncorrelated sources adequately model a non-zero bandwidth signal. The two discrete sources will have mean and variance of their instantaneous frequency spectrum corresponding to the mean and variance values of the corresponding non-zero bandwidth signal they model. For a rectangular power spectrum with bandwidth b and centre frequency f_c the variance of the spectrum is given by

$$\text{var}(f) = \frac{1}{b} \int_{f_c - \frac{b}{2}}^{f_c + \frac{b}{2}} (f - f_c)^2 df = \frac{b^2}{12} \quad (2.10.6)$$

The signal can be represented as two delta functions each separated by k from the centre frequency. The variance of such representation is equal to k^2 ,

therefore, as the variances of the model and the real non-zero bandwidth signal must be the same, we have

$$k^2 = \frac{b^2}{12} \quad (2.10.7)$$

and the separation between the two delta functions is given by

$$2k = \frac{b}{\sqrt{3}} \quad (2.10.8)$$

Now we need to obtain an expression relating the bearings of the two discrete signals originated by the signal model to the bandwidth of the real signal. For an arbitrary array, $|\psi|$ is given by

$$|\psi| = \frac{|a(\theta, \bar{f} + k)^H a(\theta, \bar{f} - k)|}{\sqrt{a(\theta, \bar{f} + k)^H a(\theta, \bar{f} + k) a(\theta, \bar{f} - k)^H a(\theta, \bar{f} - k)}} \quad (2.10.9)$$

where \bar{f} is the mean of the received signal's power spectral density. For a uniform linear array, let us consider the two values of the inter-element phase φ_{\pm} (see Equation 2.10.2) for the two deltas as given by

$$\varphi_{\pm} = \frac{2\pi d(f \pm k)}{c} \sin \theta \quad (2.10.10)$$

Because $|\psi|$ depends on both the frequency and angle of arrival of the incoming signal, the model of Equation 2.10.8 (in a few steps from Equation 2.10.10) corresponds to having two incoming signals from bearings equal to

$$\sin(\theta) \left(1 \pm \frac{b_f}{2\sqrt{3}} \right) \quad (2.10.11)$$

where $b_f = b/f_0$ is defined as the fractional bandwidth. The two signals are spaced in $\sin(\theta)$ by

$$\varepsilon = \frac{\sin(\theta) b_f}{\sqrt{3}} \quad (2.10.12)$$

For a uniform linear array $|\psi|$ is given by

$$|\psi| = \left\{ \frac{\sin\left(\frac{N\pi\varepsilon}{2}\right)}{N \sin\left(\frac{\pi\varepsilon}{2}\right)} \right\} \quad (2.10.13)$$

Given $|\psi|$ from Equation 2.10.13, the smallest eigenvalue in Equation 2.10.5 can be expressed as

$$\lambda_2 = \frac{N_{s_m}}{2}(1-|\psi|) \quad (2.10.14)$$

The calculation of λ_2 does not include the noise, whereas if noise is included λ_2 is given by Equation 2.10.14 plus 1 (Zatman, 1998). Therefore, a signal can be regarded as narrowband, i.e. its effective rank is equal to 1, if λ_2 is smaller than 1 (i.e. $\lambda_2 < 2$ or $< 3\text{dB}$ if noise is included). As a result, the narrowband condition is expressed by

$$\frac{N_{s_m}}{2}(1-|\psi|) < 1 \quad (2.10.15)$$

For a linear array, the apparent spreading of a signal caused by the non-zero bandwidth effect, is analogous to spreading due to array rotation (Zatman, 1995). A similar result to Equation 2.10.15 for the case of a rotating array was derived by Hayward (1997). In fact, array motion affects the structure of the resulting covariance and therefore the formation of the relevant eigenvalues. Therefore, for systems involving arrays mounted on rapidly moving platforms, the assumption of approximately stationary signal environment may be violated. Such systems, mainly used for military purposes, aim to adaptively suppress unwanted sources of interference, such as jamming. However, when the location of jamming sources is not known *a priori*, the waveforms of the jammers need to be calculated from the data. In order to minimise the sum of squared errors between the beamformer output and the wanted signal, a set of weights for some finite set of data is usually computed. This set of weights is regarded as the optimal Wiener solution (approached asymptotically as the size of the data set increases), in a wide-sense stationary signal environment. The best linear mean-square (MS) estimates (regardless of the statistics of the inputs or when the inputs are gaussian) can be obtained by means of the Wiener solution (Van Trees, 1968), which therefore maximises the signal-to-noise-plus-interference ratio (SNIR) at the beamformer output. The least-squares approach is also taken with the Sample Matrix Inverse (SMI) algorithm (Reed *et al*, 1974) and the Recursive Least-Squares (RLS) algorithm (Haykin, 1986).

In a non-stationary signal environment, the weight solution needs to be time-varying. If the data are input into the beamformer at a fixed rate, the effective rate at which the weight solution is updated is increased by either decreasing the number of samples used to calculate the weights in the SMI algorithm or the memory length of the RLS algorithm. As a consequence, the convergence of the solution at each update is reduced. Van Veen (1991) and Krolik and Swingler (1994) exploited the convergence properties of the SMI algorithm in a wide-sense stationary signal environment. Observations from simulations (Balance and Miller, 1991) showed that spatially distributed sources can cause significant

degradation in beamformer performance. The results by Hayward (1997), reported below, confirmed these observations, as they showed that the weight solution is sensitive to very small variations in the apparent angle of arrival of incoming signals, during the time interval in which sufficient data are acquired to update the weights.

Let us consider an irregularly spaced antenna array. Given the polar angles θ and Φ describing a point in the array far field (Figure 2.10.1), $a(\theta, \Phi)$ is the generic steering vector. The locus of a as θ and Φ vary is referred to as the array manifold. In the time domain, and in the presence of a single narrowband source, the data vector at the array output is given by

$$y(t) = s(t)a(t) + v(t) \quad (2.10.16)$$

where $s(t)$ is the transmitting waveform and $v(t)$ is the white thermal noise process, uncorrelated with the $v(t)$, both $s(t)$ and $v(t)$ being wide-sense stationary. In the case of $s(t)$ being an interferer or jammer, $y(t)$ is a noise-only data vector. In order to obtain a noise-only covariance matrix, following the SMI algorithm, we obtain a set of such vectors and average the n rank-one estimate, thus having

$$\hat{R} = \frac{1}{n} \sum_{i=1}^n y_i y_i^H \quad (2.10.17)$$

Goodman N.R. (1963) showed that, if a were fixed, Equation 2.10.17 would represent the maximum likelihood estimate of the true covariance. However, as a is time-varying, the estimate is that of the covariance matrix

$$R = E\{\hat{R}\} = \frac{1}{n} \sum_{i=1}^n R_i^s + R^v = R^s + R^v \quad (2.10.18)$$

where R_i^s is the rank-one covariance matrix for the source in the i th position and R^v is the noise covariance. The covariance matrix R is equivalent to that which would be obtained from the presence of n uncorrelated sources, whose spatial distribution varies according to the shape of the array manifold and the sampling rate. In fact, the ordering of the data vectors in the averaging process is arbitrary, therefore any information related to the motion is lost.

If R is known, the weight vector of an ideal beamformer is given by

$$w_0 = \frac{R^{-1} a_t}{a_t^H R^{-1} a_t} \quad (2.10.19)$$

where a_t is the steering vector corresponding to the direction of the target. In order to understand how the structure of the covariance matrix can affect the

beamformer performance, we need to consider an expression of the SNIR in case of uncorrelated target signal and noise as input data. We assume that there is no loss caused by mismatch between a_t and the true target steering vector. Even though the true target steering vector is time-varying, the assumption is justified by the very small angles involved. Monzingo and Miller (1980) obtained the following expression:

$$SNIR_0 = P_t a_t^H R^{-1} a_t \quad (2.10.20)$$

where P_t is the target signal power. The covariance matrix can also be expressed as in Equation 2.10.3 ($R = V\Lambda V^H$). The inverse of the covariance matrix is given by

$$R^{-1} = V\Lambda^{-1}V^H \quad (2.10.21)$$

and substituting in Equation 2.10.20 we have

$$SNIR_0 = P_t \sum_{i=0}^{k-1} \frac{|a_t^H V_i|^2}{\lambda_i} \quad (2.10.22)$$

Assuming the noise process spatially white, R takes the form of $\sigma^2 I$ and each of its eigenvalues is given by the sum of the corresponding eigenvalues of R^s and R^v . From Equation 2.10.22 we can see that an increase of the i th eigenvalue of R^s causes a loss in the beamformer performance proportional to the value of the numerator. In case of single strong point source present (a discrete scatterer or a jammer), R^s has a single non-zero eigenvalue with no motion, whereas with motion R^s has additional eigenvalues, whose value depends on the effective displacement of the source. The maximum size of the source displacement that can be regarded as tolerable (no considerable deterioration in the beamformer performance caused), can be obtained by setting an upper limit of σ^2 on the largest additional eigenvalue. When this limit is reached, the second term of the summation in the Equation 2.10.22 will have increased of 3dB. Therefore, the effect of the source displacement on the beamformer performance depends on the value of the numerator, which is function of the eigenvectors of R^s . A situation in which the value of this factor becomes significant can be, for example, when the array is rotated. In this jamming scenario the beamformer performance can deteriorate significantly, and the jammer suppression can be ineffective.

Assuming that m sources with equal power P_s/m are present and that the platform is in uniform motion, being

$$r_i = x_i \cos \Phi + y_i \sin \Phi \quad \text{and} \quad \hat{r}_i = y_i \cos \Phi - x_i \sin \Phi$$

we obtain (for a full mathematical derivation, the reader is referred to Hayward, 1997)

$$\lambda_1 \cong \pi^2 \frac{P_s}{3} k (\cos \theta)^2 \overline{r^2} \delta \theta^2 \quad (2.10.23)$$

for small uniform motion in θ and

$$\lambda_1 \cong \pi^2 \frac{P_s}{3} k (\sin \theta)^2 \overline{r^2} \delta \Phi^2 \quad (2.10.24)$$

for small uniform motion in Φ . In both cases $\lambda_0 \cong P_s k$. It is now possible to set a limit to the maximum effective source displacement that can be tolerated during the estimation of the covariance matrix. Although this discussion is restricted to uniform linear arrays, the analysis can be applied to arbitrary array geometries with similar results. From Equation 2.10.23, for motion in elevation, we have

$$\delta \theta_{\max}^2 \cong \frac{3\sigma^2}{k\pi^2 P_s (\cos \theta)^2 \overline{r^2}} \quad (2.10.25)$$

In the case of the source representing the mainbeam jammer, we can assume that $\cos \theta \cong 1$. For a k sensor linear array with half-wavelength sensor spacing, $\overline{r^2} \cong k^2/48$, $k \gg 1$. The beamwidth of the array (peak to first null) is equal to $2k$, thus in a few steps, in beamwidths, for large k we obtain

$$\delta \theta_{\max} \cong \frac{6}{\pi \sqrt{JNR}} \quad (2.10.26)$$

where $kP_s/\sigma^2 = JNR$ (the quiescent pattern output jammer-to-noise ratio). As a numerical example, from Equation 2.10.26 we can calculate that, in the presence of a mainbeam jammer with $JNR=10000$ (i.e. 40dB), a linear array will be sensitive to rotations as small as 0.02 beamwidths during the time taken to estimate the sample covariance matrix.

As a conclusion, Hayward (1997) and Zatman (1998) obtained (with a different approach) the common result that a single source can give rise to an estimated covariance matrix equivalent to that due to two or more uncorrelated spatially distributed sources. This equivalence was exploited by both Zatman and Hayward by calculating the second largest eigenvalue of the covariance matrix, which, in the above mentioned scenarios, can originate from the same single source of the first largest eigenvalue, instead of originating from a different source. The practical consequence is that the observations obtained by means of direction finding systems can show the presence of multiple traces, whereas only a single broad source is present in the signal environment. Therefore, a

non-zero bandwidth signal and/or a non-stationary signal environment originated by array motion can be analogous situations to that of spread incoming signals discussed in Chapters 3 to 6. For a theory of two-signal representation of one signal spread in azimuth and a comparison with simulation results the reader is referred to Sections 5.7 and 5.8 respectively.

2.11 Summary.

Fixed land-based stations for Direction Finding are classified in three main groups: Narrow Aperture DFs (NADFs), Wide Aperture DFs (WADFs) and Very Wide Aperture DFs (VWADFs). NADFs present limited cost, but poor resolving power. VWADFs have good resolving power, but are very expensive. In between is the compromise of using WADFs, which have become the most popular since the advent of multichannel receivers and digital computers. This has given the operators the ability to utilise modern signal processing techniques.

Although conventional interferometry techniques have the ability of detecting more than one direction of arrival in a multisignal or multimoded environment, superresolution direction finding algorithms are necessary to resolve multisignal or multimoded environments more accurately, without the use of very large arrays. In fact, superresolution direction finding algorithms are characterised by the ability to resolve two signals which are separated in azimuth less than the natural beamwidth of the antenna array at the receiving site. Therefore in the last decades researchers have focused their efforts in the development of SRDF algorithms and have compared their performances with both observed and simulated data, with respect to a large number of factors (see Sections 2.8 and 2.9). The SRDF algorithms used for the work detailed in this thesis are MUSIC, INS, Loaded Capon and SML.

Recent studies have shown that a non-zero bandwidth signal (Zatman, 1998) or a non-stationary signal environment originated by array motion (Hayward, 1997) can give rise to an estimated covariance matrix equivalent to that caused by two or more uncorrelated spatially distributed sources. Such scenario can cause direction finding systems to show the presence of multiple traces, whereas only a single spread signal is present in the signal environment.

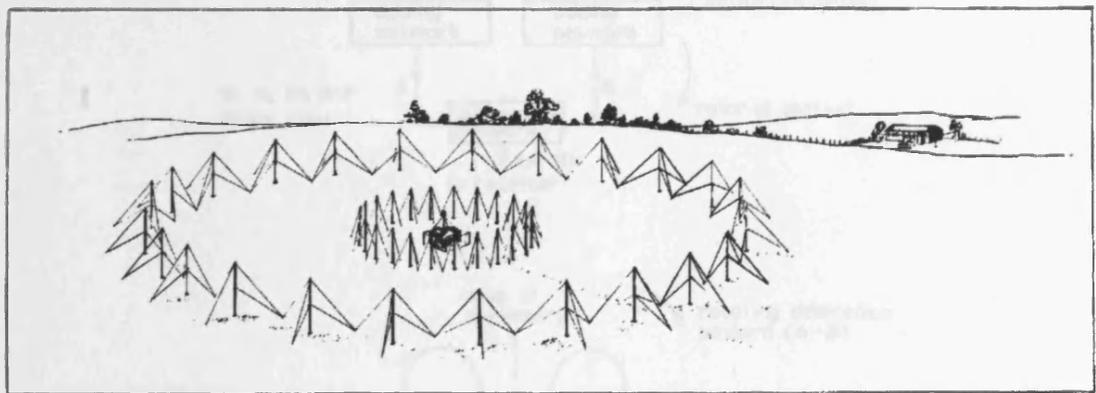


Fig. 2.2.1 Formation of rotating subarrays radiation pattern with
Two-ring DF array (Gething, 1991).

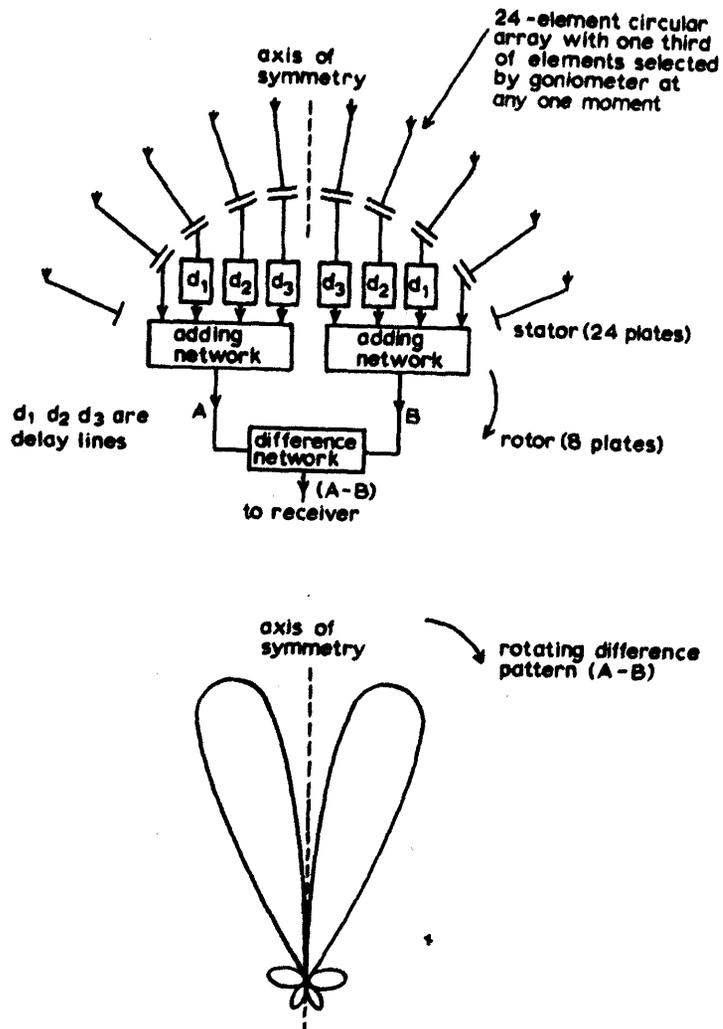


Fig. 2.2.2 Formation of rotating 'rabbit ears' radiation pattern with goniometer (Gething, 1991).

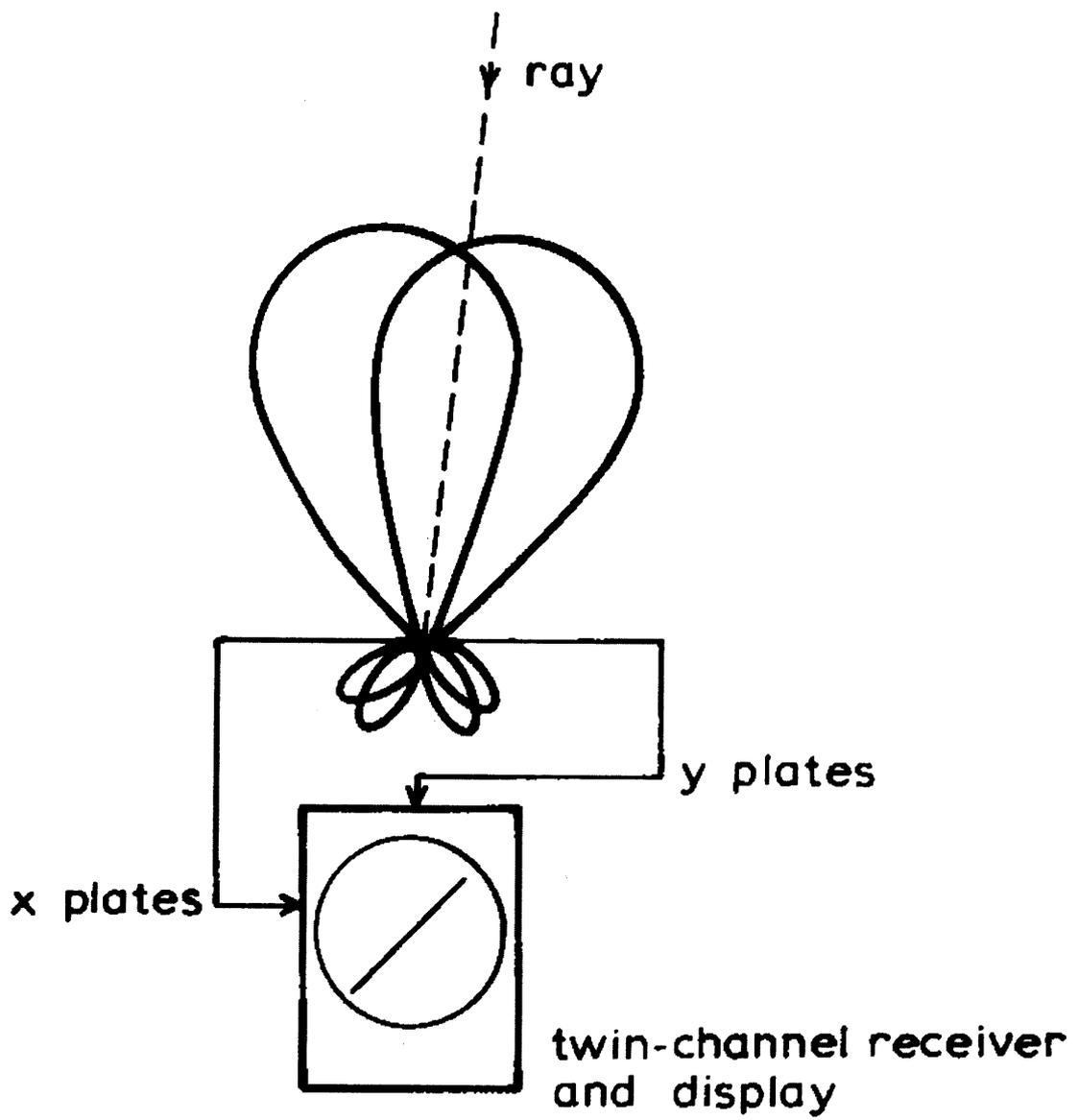
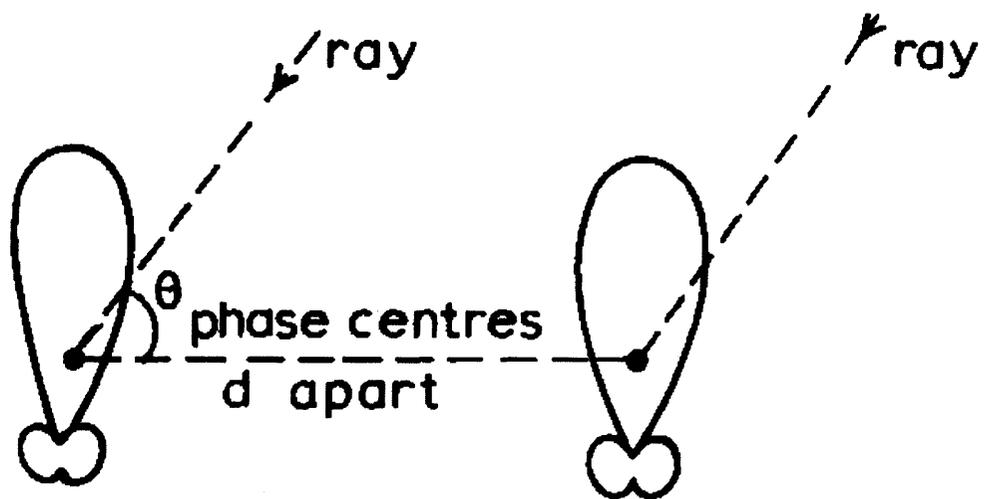
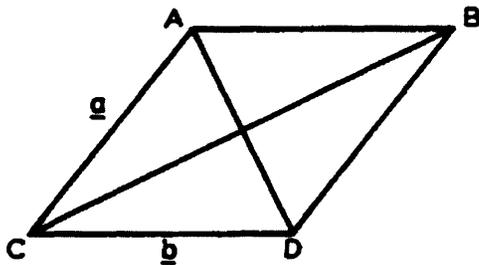


Fig. 2.3.1 Amplitude-comparison DF system (Gething, 1991).



phase difference $\phi = \frac{2\pi d}{\lambda} \cos\theta$

Fig. 2.3.2 Phase-difference DF system (Gething, 1991).



Sum-and-difference phase meter
 a and b are of equal magnitude and $\angle ACD = \varphi$
 $CB \perp AD$
 sum $CB = 2CD \cos(\varphi/2)$
 difference $AD = 2CD \sin(\varphi/2)$
 $\therefore \frac{AD}{CB} = \tan \frac{\varphi}{2}$

Fig. 2.3.3 Sum-and-difference phase meter (Gething, 1991).

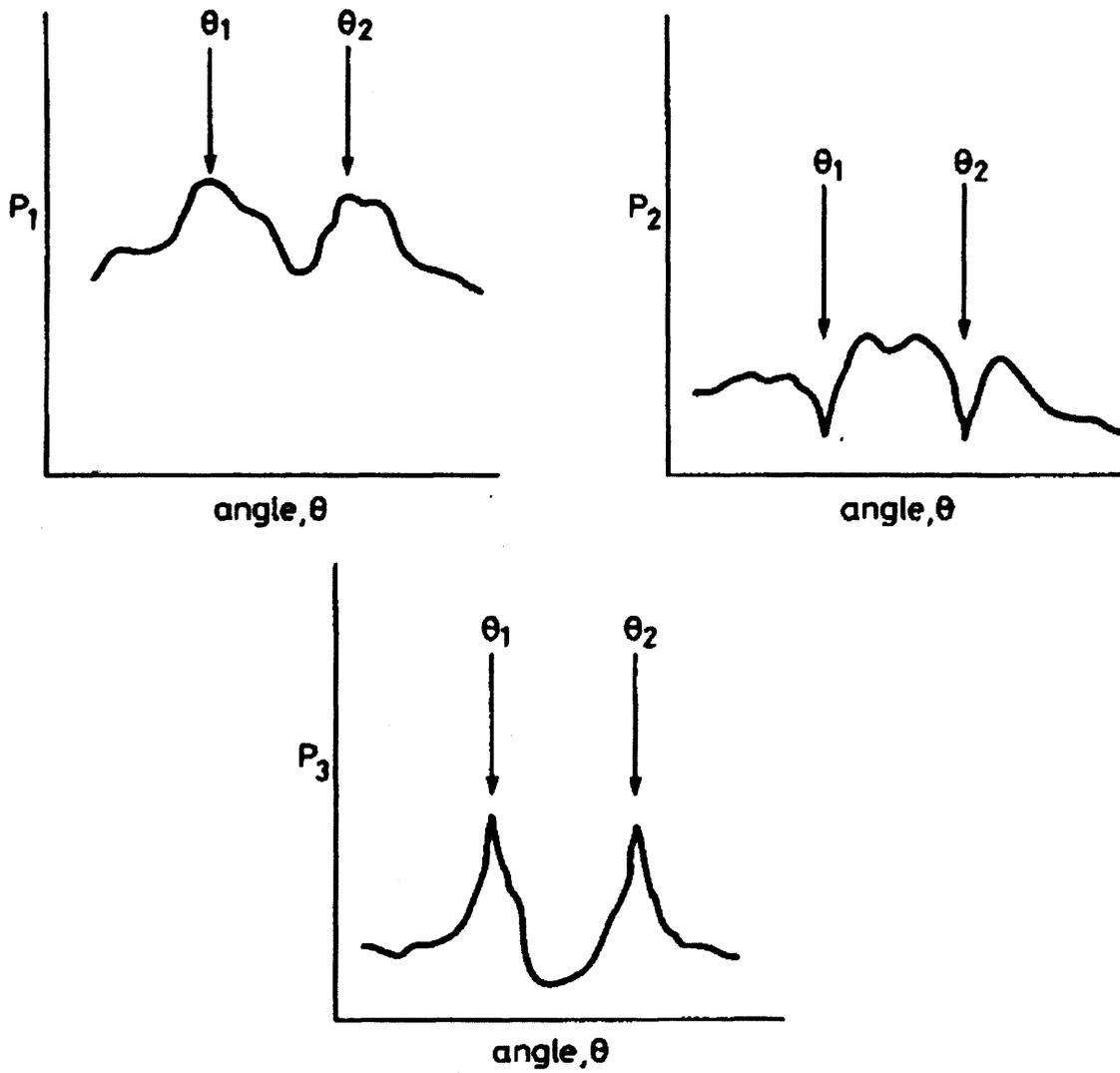


Fig. 2.5.1 Three types of spectral estimator P for two-ray problem
a) use of maxima; *b)* use of minima; *c)* use of inverted minima (cusps) (Gething, 1991).

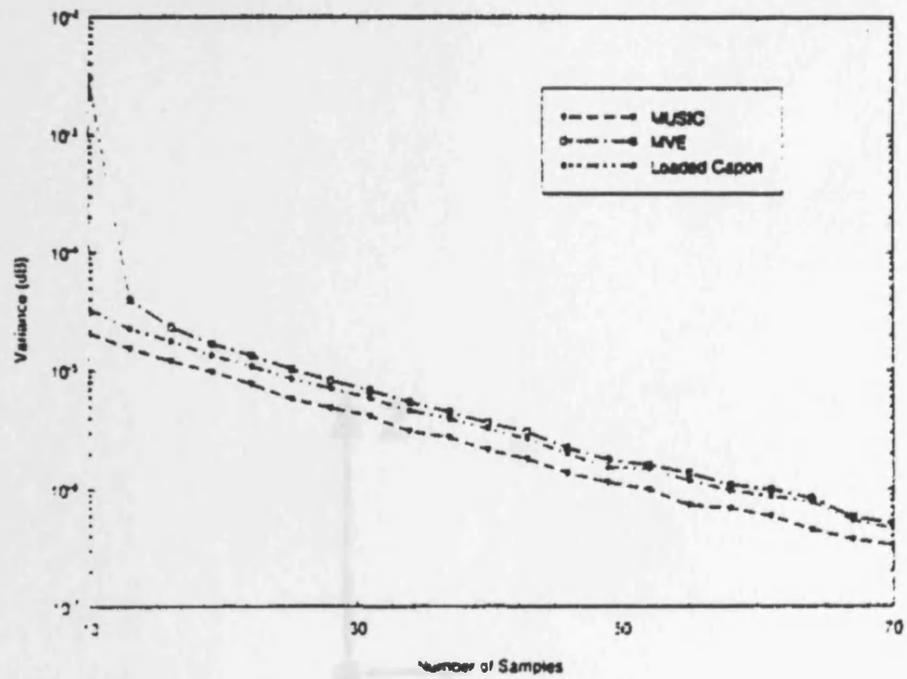


Fig. 2.8.1a Variance of the estimated angles of arrival made by Capon's Minimum Variance Estimator, Loaded Capon and MUSIC, for SNR=9dB (Featherstone *et al*, 1997).

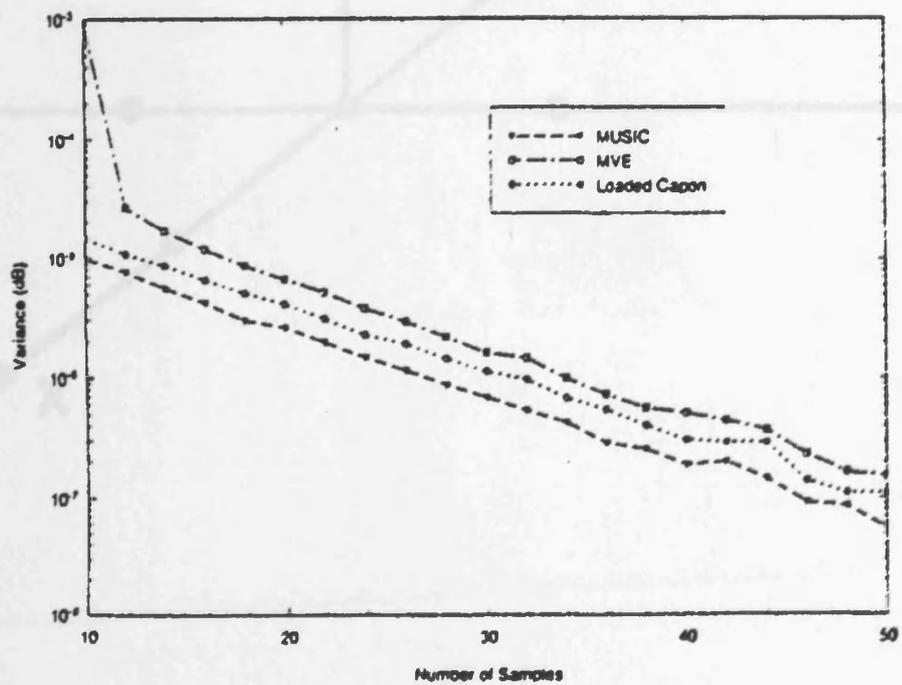


Fig. 2.8.1b Variance of the estimated angles of arrival made by Capon's Minimum Variance Estimator, Loaded Capon and MUSIC, for SNR=12dB (Featherstone *et al*, 1997).

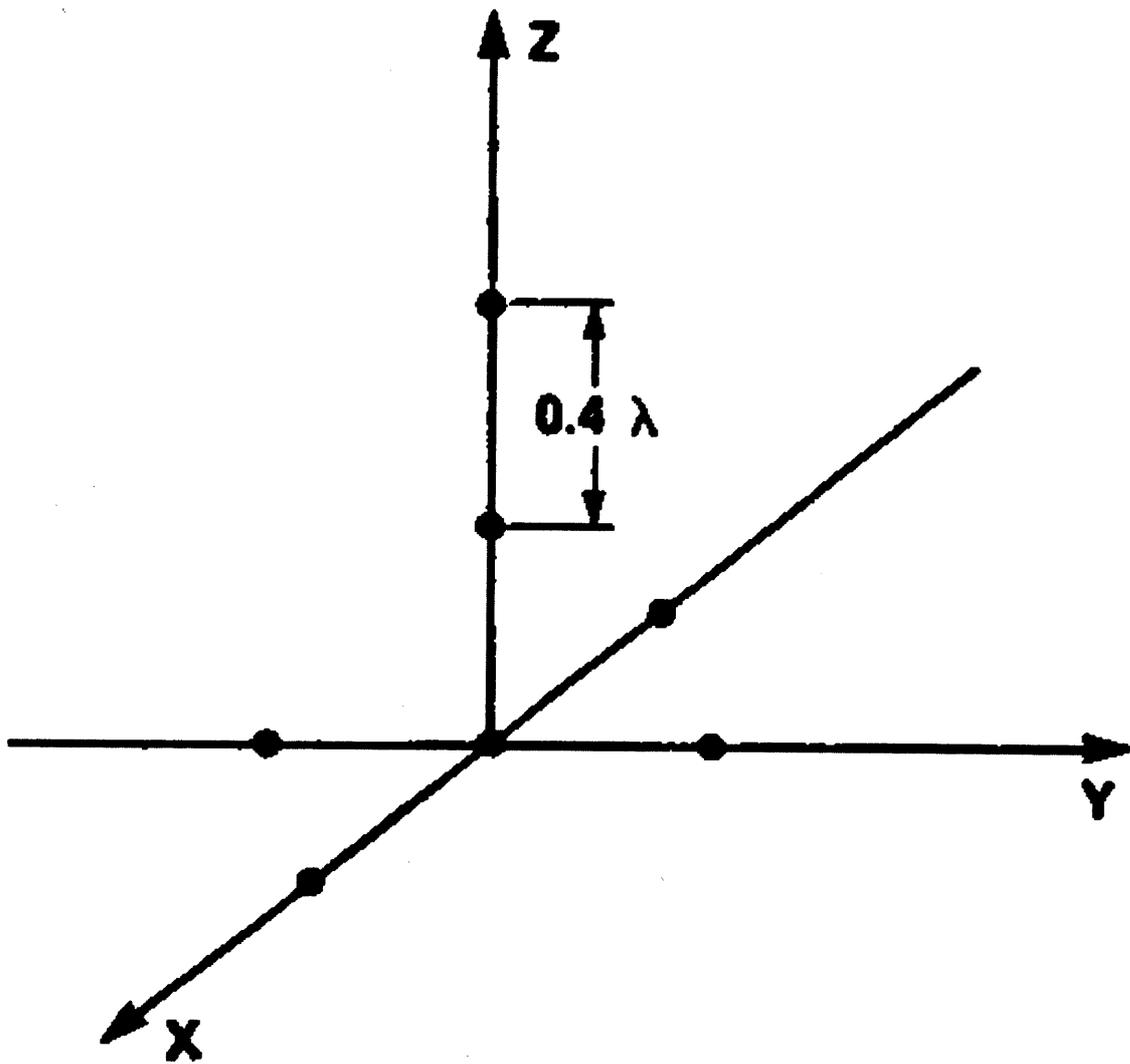


Fig. 2.8.2 Seven antenna array (Johnson and Miner, 1986).

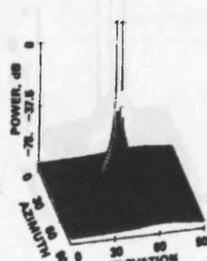
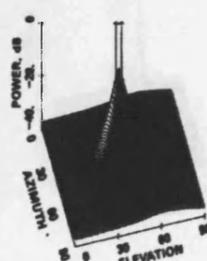
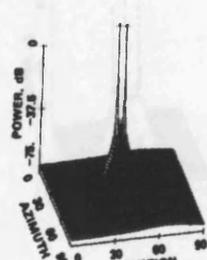
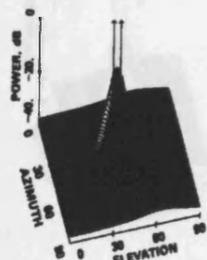
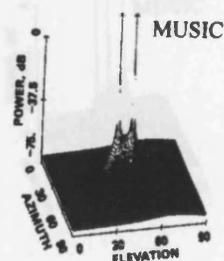
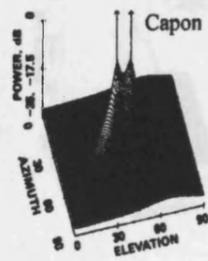


Fig. 2.8.3 Signal separation simulation. Equal amplitudes, mutually incoherent, SNR=30dB, Az=45°, Δ El=10°, 5°, 4° (Johnson and Miner, 1986).

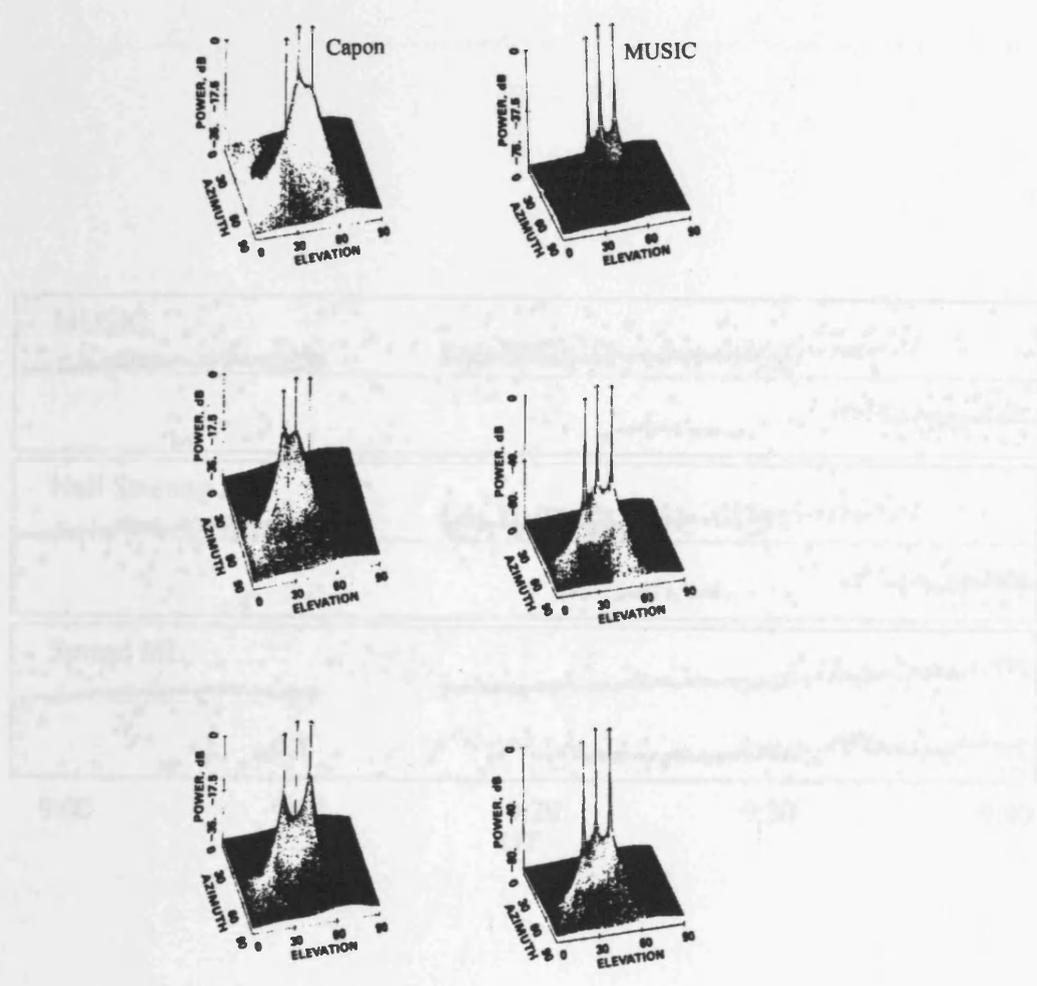


Fig. 2.8.5 The difference between the observed and true azimuths, $\Delta\theta_{obs}$, for

Fig. 2.8.4 Co-channel interference simulation. SNR=30dB, mutually incoherent, (Az,El)=(45°,40°), (45°,40°), (50°,30°), relative amplitude (top) 1.0, 0.5, 0.1, (middle) 0.5, 0.1, 1.0, (bottom) 0.1, 1.0, 0.5 (Johnson and Miner, 1986).

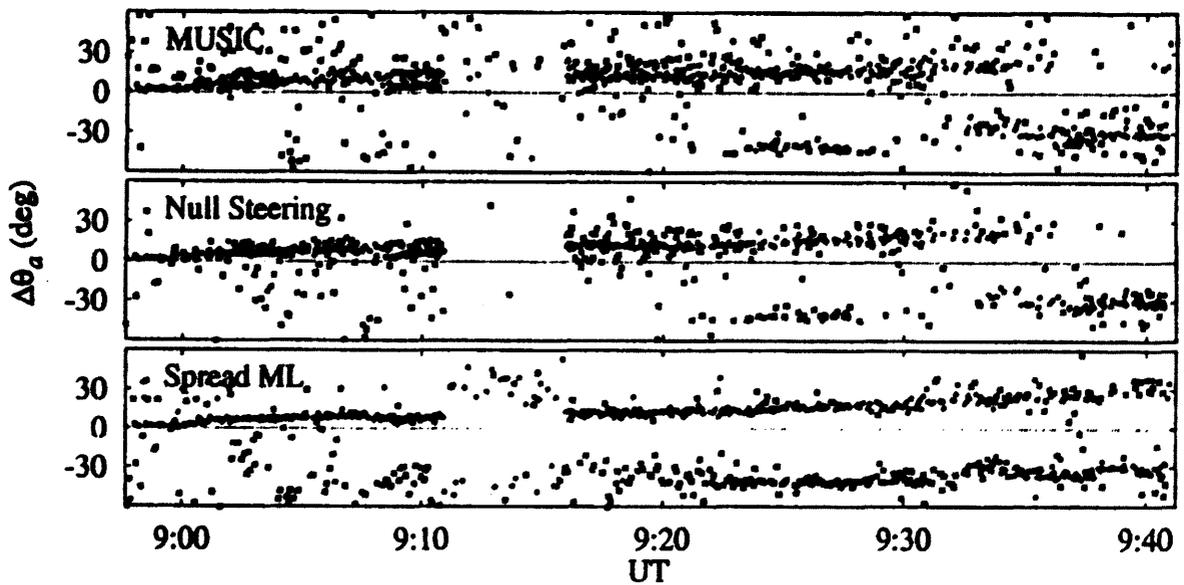


Fig. 2.8.5 The difference between the observed and true azimuths, $\Delta\theta_a$, for three DF algorithms indicated in the top left hand corner of each frame. The Spread Maximum Likelihood algorithm (bottom frame) clearly outperforms the other algorithms in detecting the two signals (Dumas, 1997).

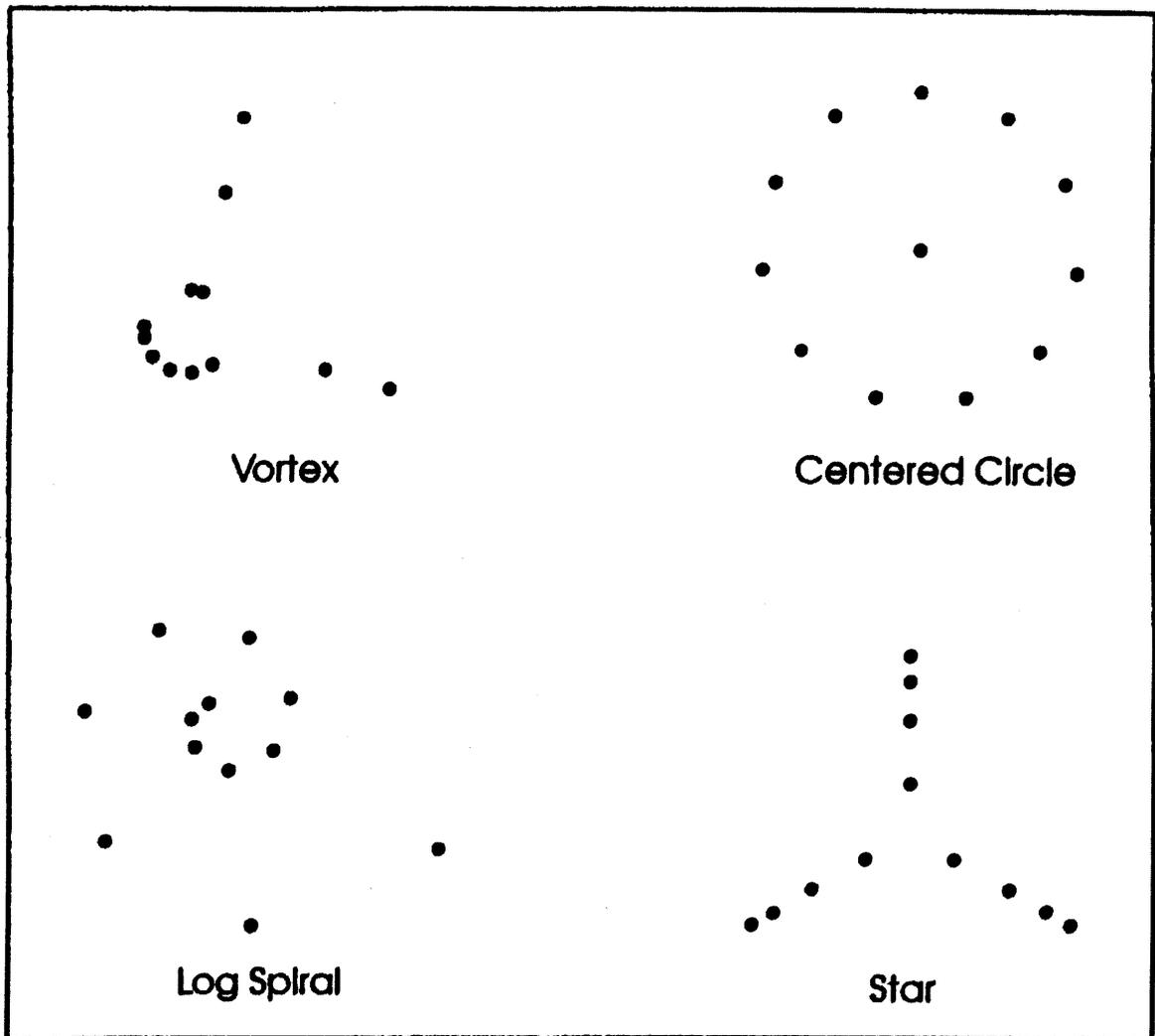


Fig. 2.9.1 Plan views of simulated antenna arrays (Jenkins, 1997).

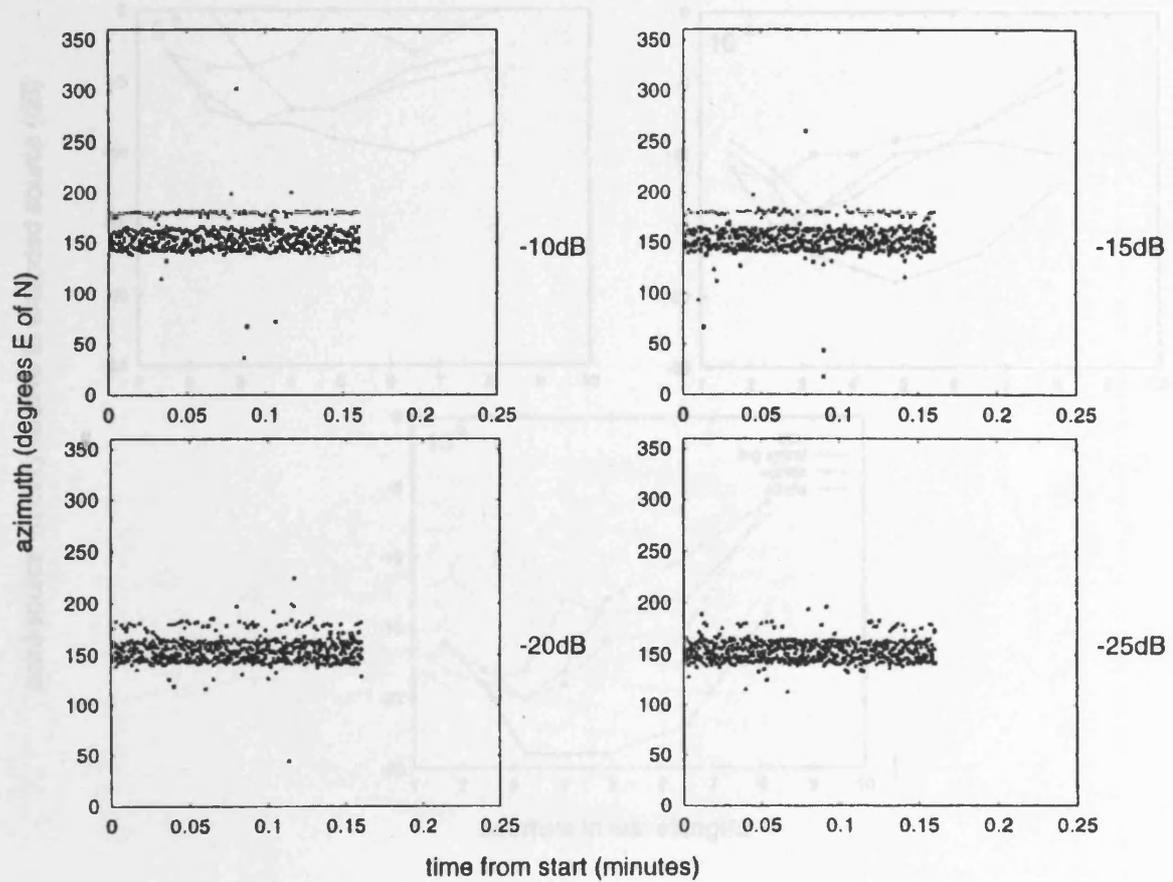


Fig. 2.9.2 Azimuth-time displays for the Vortex array, 4.05 wavelength aperture, receiving a point-source signal from 180° E of N and a stronger signal from an extended source 140° - 165° E of N, for various point-source powers relative to the extended source, with a Maximum Likelihood algorithm (Jenkins, 1997).

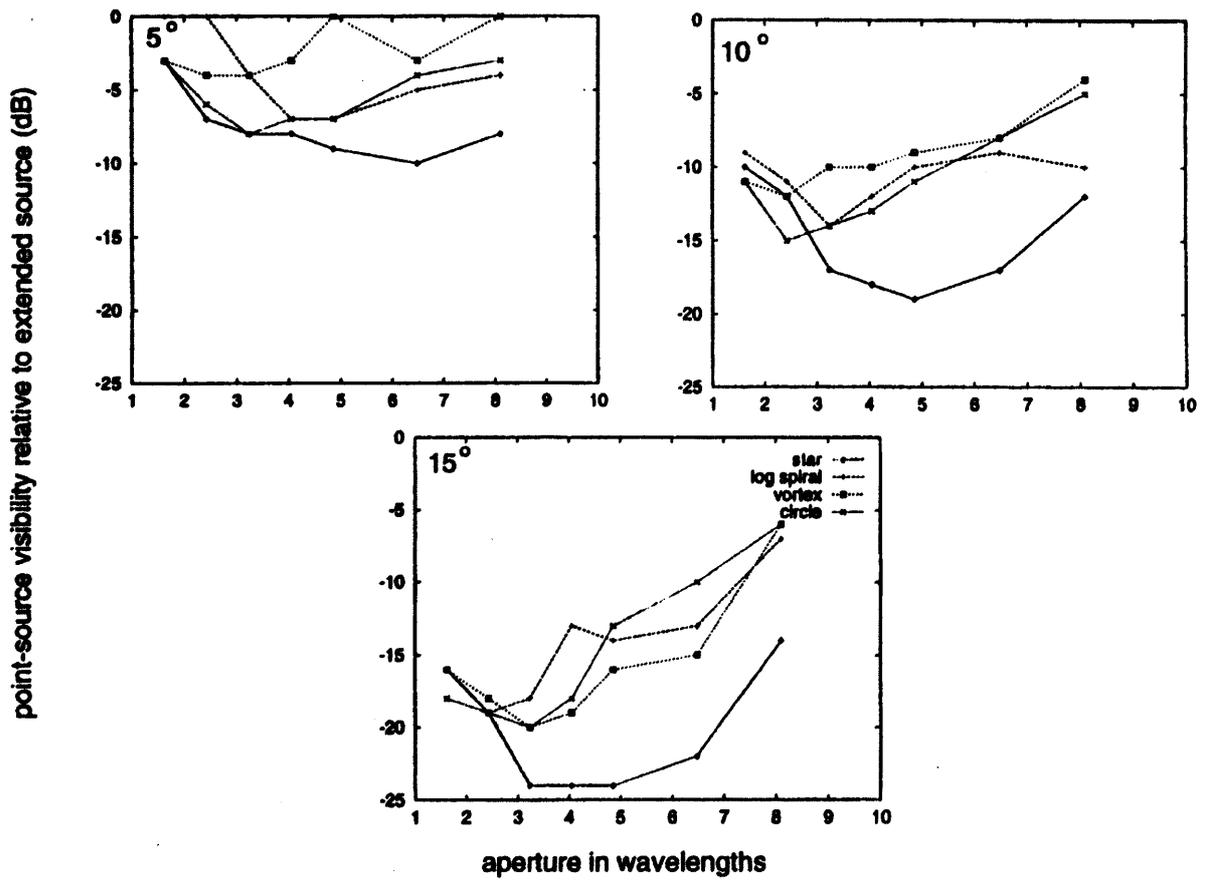


Fig. 2.9.3 Point-source visibility as a function of array aperture for the four array geometries and three point-source/spread-source azimuth separations (Jenkins, 1997).

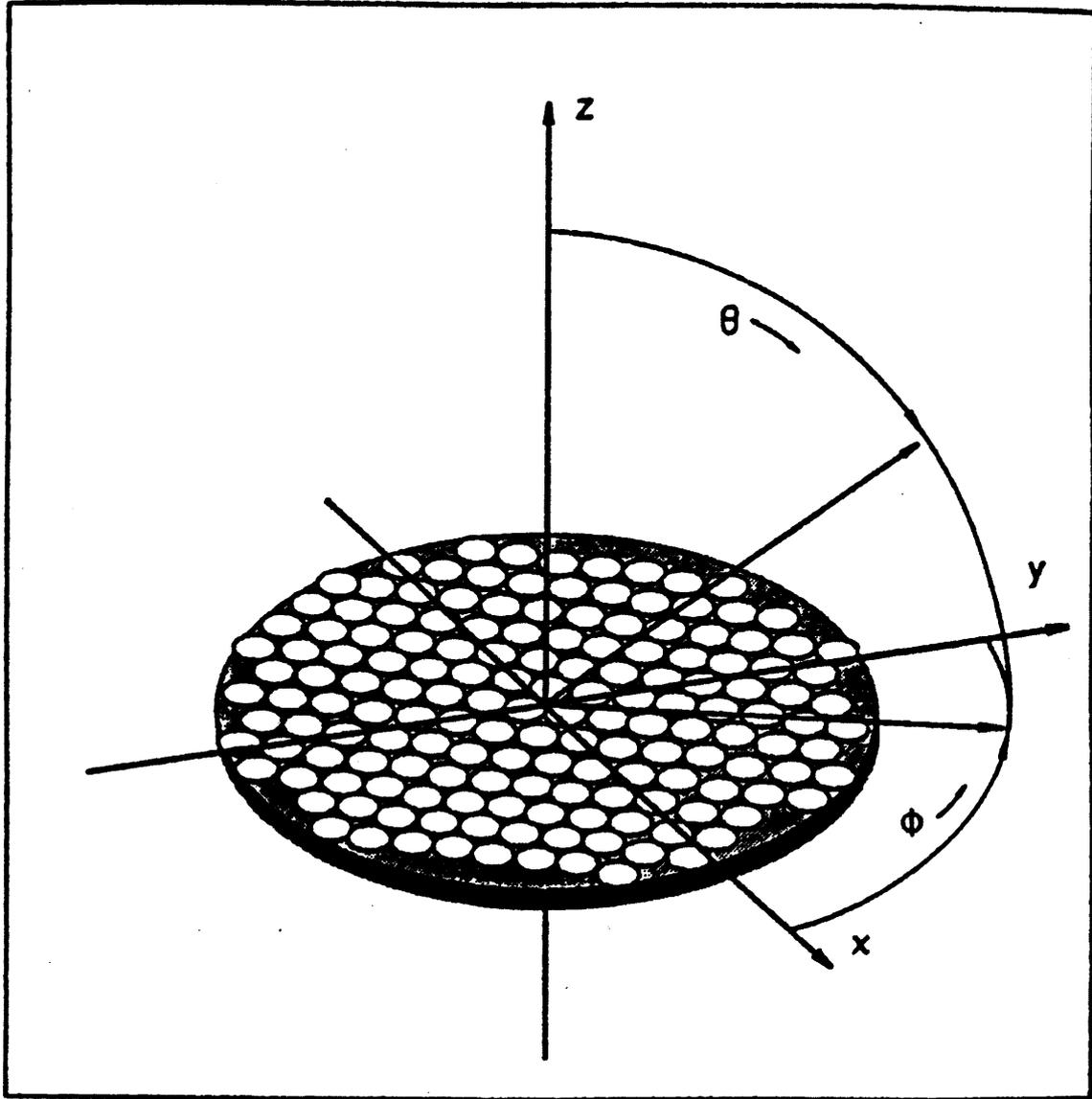


Fig. 2.10.1 Array and far-field co-ordinate system (Hayward, 1997).

3. DATA COLLECTION METHOD AND OVERVIEW OF THE MEASUREMENTS

3.1 Introduction.

The data acquisition system (known as Vortex) used to collect the data analysed in this work was deployed at Alert, in the Canadian Arctic. The data were collected in wintertime by the Canadian Communications Research Centre, when the receiver was recording a signal from Iqaluit, in the North West Territories, 2100 km South of Alert. The locations of the transmitter and receiver make this a polar cap path and therefore the performance of the Vortex DF system is limited by this challenging propagation environment.

Direction-of-arrival observations over three days during the Arctic winter (22nd-24th January 1996) are presented. During these days several ionospheric disturbances occurred which affected the propagation environment, making the signal arrive at the receiver at angles far removed from the Great Circle Path. At times, the direction of arrival of the signal showed very rapid changes (see Section 3.6). The oblique ionograms taken during the data interval over the I-A path are shown.

Bearing measurements for data recorded over the same path have been undertaken by Warrington *et al* (1997a and b) with a goniometric system for a period between January and April 1994 (see Section 1.5). Bearing measurements for the 23rd January 1996 (see Section 1.5) were also investigated by Dumas (1997) and for the 24th January 1996 (see this and next chapter) by Warrington *et al* (1997a and 1999).

The data in this chapter have all been obtained by means of the Iterative Null Steering (INS) superresolution algorithm, with the search for peaks of secondary energy set to terminate when the peak of the scanned beam was less than 6dB below the strongest peak (see Section 2.7). The 6dB threshold was chosen in order to filter noise from the data without excluding important information about the incoming signal. Since major features are generally observed by all algorithms (see Chapter 4) the INS algorithm was arbitrarily chosen to present an overview of the measurements. All plots in this chapter were obtained with the full array configuration of 12 antennas.

Table 3.1.1 (see below) shows the 3-hourly values of the geomagnetic index K_p on the 22nd, 23rd and 24th January 1996 respectively. The value of K_p was never greater than 3, and these three days can be regarded as days of quiet geomagnetic activity. However, some of the observed bearings show a strongly disturbed propagation environment, as we will see in the following sections.

This makes the data very interesting for the purpose of investigating how the superresolution algorithms perform in situations more extreme than typically found at mid-latitudes. The time intervals over these three days when the data were available are shown in Table 3.1.2, in total 7h32' on the 22nd, 8h27' on the 23rd and 9h33' on the 24th January 1996.

Simulations of a drifting mono-dimensional over-dense plasma structure have been undertaken, the results of which are consistent with the observed change in DOA.

Table 3.1.1 Values of K_p on 22nd, 23rd and 24th January 1996 (US Department of Commerce, National Geophysical Data Center FTP).

UT	00-03	03-06	06-09	09-12	12-15	15-18	18-21	21-24
K_p on 22/01/1996	3-	1	1-	1-	2	2+	2	2
K_p on 23/01/1996	3-	1	1-	2-	1+	1-	1	3
K_p on 24/01/1996	0+	1	0+	1-	1	3	3-	2+

Table 3.1.2 Time intervals (UT) over the three days in January 1996, when the collected data were available for processing.

22/01/1996				13:58-16:38		19:03-23:55
23/01/1996	00:25-04:59	05:09-10:02				
24/01/1996			09:36-14:22		15:53-20:40	

3.2 Data acquisition system.

The Vortex DF system, which was used to collect the data, is installed at Alert, located on the northern tip of Ellesmere Island in the Canadian Arctic (82.60° N, 62.35° W). The analysed data were collected during a period when the system was recording a signal from the CZD transmitter, located in Iqaluit (63.45° N, 68.30° W), also in Canada, about 2100 km south of Alert. The I-A Great Circle Path is therefore within the polar cap (see Figure 1.4.1). Both in Iqaluit and Alert the Local Time (LT) corresponds to the Universal Time (UT) – 4 hours. The transmitter in Iqaluit provided a fixed target, allowing deviations from the true

bearing to be used as estimates of the errors in bearing-determination that could be expected in the operational field.

The direction finding antenna array deployed at Alert is illustrated in Figure 3.2.1a, where the full squares represent the positions of the antennas used for the Vortex array. The Vortex array consists of 12 irregularly spaced antennas. The antennas are vertical elements, 8 of which are deployed in a circle measuring 25-m radius, which is part of a 2-ring DF system. The 2 rings are Pusher arrays (circular arrays with 24 antennas, spaced at 15° intervals). Elements 9 and 10 are at a distance of 88 m from the centre of the circle; elements 11, and 12 are at 125 m from the centre of the circle. Elements 9 and 11 are at azimuth $\theta_a = 15^\circ$, measured clockwise of North, elements 10 and 12 are at $\theta_a = -75^\circ$. All 12 elements are 6-m high and 5-cm diameter aluminium poles with ground radials. The 4 outlying antennas (9 to 12) are used to increase the array aperture, thus increasing the resolving power of the Vortex DF system. The maximum dimension of the antenna array is 176 m.

The 12-antenna Vortex array is repeated for better clarity in Figure 3.2.1b. Each antenna is connected to an individual receiver. Figure 3.2.2 shows the Vortex data acquisition system that received a signal transmitted from Iqaluit at 9.292 MHz (wavelength $\lambda=32.3$ m) which alternated between a 15-sec tone and a 15-sec interval during which the Morse callsign was sent twice. This 30-sec cycle was repeated continuously for a 25-min period, followed by a 5-min sounder slot during which the 9.292 MHz signal was switched off while an oblique ionogram lasting 4 min and 40 sec was taken along the I-A path. The ionogram consisted of a sweep in frequency from 2 to 30 MHz at a rate of 100 kHz/sec. The entire 30-min sequence was then repeated throughout the experimental period under consideration. At the receiver, the 12 signals (one at each antenna) were mixed with a common reference signal to downconvert the input signal from HF to 2.5 kHz, bandpass filtered between 1-4 kHz and then sampled and digitised at a rate of 10 kHz. The bandwidth of the received CW signal was typically 10 Hz or less. However the bandwidth was about 30 Hz when the Morse callsign was transmitted. The raw data were written to 8-mm tapes for subsequent processing and analysis. Each tape recorded about 5.7 hours of continuous data. The array response in amplitude and phase were calibrated for each tape. The calibration constants were incorporated into the data (read in blocks, see later in this section) before forming the covariance matrix. The data were filtered using an additional bandpass filter to reduce the amount of noise in the processed signal. A width of 200 Hz was found to represent a useful compromise for achieving noise reduction without affecting the DOA measurements.

The ionograms were collected to allow the mode structure of the I-A path to be determined. In this way, the propagation conditions could be well-assessed. The traces obtained by means of the oblique ionograms gave important information about the usable range of frequencies of operation, as well as information about the disturbances in the propagation environment. Furthermore, the evolution of the ionospheric conditions along the I-A path could be inferred by interpreting series of consecutive ionograms.

The collected data have been analysed with a simple Beam algorithm and five SRDF algorithms (see Sections 2.4 and 2.7 respectively). These algorithms have been implemented in the MATLAB-computing language. The software allows complete flexibility in the number and the position of the antennas to be used for the processing. This enables the effective beamwidth of the antenna array as well as the angular resolution to be varied. It is important to note that the runtime increases with the number of antennas in use.

The data have been processed with two sets of elements of the Vortex array, the full array of 12 antennas and a sub-array of the 8 inner antennas in circle, numbered 1 to 8 in Figures 3.2.1a&b. The two arrays have two different apertures and shapes. In the case where the full array is in use, the maximum dimension is 176 m, which corresponds to 5.45λ , whereas the maximum dimension of the sub-array is given by the diameter of the inner ring of the DF system, that is 50 m (1.55λ). The corresponding beam patterns will then have different shapes, with as expected the main lobe obtained from the full array of 12 antennas (Figure 3.2.3) being markedly narrower than that obtained from the sub-array of 8 antennas (Figure 3.2.4). In the case of 12 antennas there are two strong side lobes close to the main lobe (at about 170° and 200°), whereas in the case of 8 antennas the main lobe is broader, covering the range of values between 170° and 200° . Furthermore, a strong sidelobe exists at around 250° . Sidelobes occur in directions other than the current main lobe direction for which a similar phase distribution occurs across the array. In the presence of noise, calibration errors, diffuse signals, etc. it is possible for the measured phase distribution to be more like that for the sidelobe direction than the actual signal direction. This sidelobe excitation would cause erroneous DOA estimates.

Table 3.2.1 shows the expected elevation angles of arrival for the signal reflected from the E or F layer, in case of 1-hop or 2-hop propagation. The E region reflection height is assumed at 100 km and the F region reflection height at 250 km. Figures 3.2.5a, b, c and d show the beam patterns formed by the 12-antenna Vortex array, in the look direction of the true bearing (188.5°) for various elevation angles. The main lobe of the array is highly spread in

elevation, i.e. it is sensitive to a wide range of values, and therefore the elevation angle measurements are not likely to be very accurate.

Table 3.2.1 Approximate elevation angles of arrival and propagation times for a signal propagating through the ionosphere from Iqaluit to Alert. The two cases of the signal reflected from the E or the F layer, in case of 1 hop (no reflections from the ground) or 2 hops (one intermediate reflection from the ground) are illustrated. The reflection heights are assumed 100 km for the E layer and 250 km for the F layer.

TYPE OF PROPAGATION	E LAYER (100 km)	F LAYER (250 km)
1-HOP (El - Prop. time)	0.6° - 7.09 ms	8.3° - 7.33 ms
2-HOP (El - Prop. time)	8.1° - 7.11 ms	22.9° - 7.81 ms

The raw data were read in blocks of 0.8 sec, which correspond to 8000 samples for each of the 12 antennas, the received signals having been sampled at a rate of 10 kHz. To generate the data covariance estimates, an FFT was performed on each block of 12 by 8000 data points and only the data within the 200 Hz frequency band around 2.5 kHz retained. The interference due to noise and other unintended HF signals was thereby reduced. The data covariance matrix was formed directly from the frequency domain data. The data were then input to the direction finding algorithm at the estimation rate of 1 direction of arrival each 0.8 sec. This value was used because it was regarded as a good compromise between the need of accurate DOAs estimates, low bearing standard deviations and reasonable computational time (see Section 4.3). This process then produced a predefined number of estimates of azimuth, elevation and amplitude of the received signal each 8000 samples. In the plots shown throughout this thesis primary and secondary estimates are included (red and green dots respectively). The whole data set will be displayed in the next sections, along with the relevant oblique ionograms.

3.3 Observations on 22nd January 1996.

The bearings and amplitude measured between 13:58 and 16:38 UT on 22nd January 1996 are illustrated in Figure 3.3.1. In all figures a number of gaps in the trace of the signal from Iqaluit are apparent. These correspond to the 5-minute periods ending at 16 and 46 past each hour, during which time the transmitter was off and the sounder was on for the collection of oblique ionograms, shown in Figure 3.3.2, on the same I-A path (see Section 3.2). The

Maximum Usable Frequency (MUF) was well above the frequency of operation of 9.292 MHz (vertical line in the plots). This implies that the signal from Iqaluit was expected to be received at Alert. Because the transmitter and receiver were not synchronised, the delay of the ionogram traces is an indication of the relative propagation time. The azimuth vs. time plot in the upper panel of in Figure 3.3.1 exhibits a narrow trace around the true bearing of 188.5° from 14:00 to 15:00. The relative amplitude (bottom panel) is higher until 15:00 than afterwards, when it decreases by 15-20dB. After 15:00 the azimuth trace becomes spread. There is evidence of two weaker signals coming from about 60° and 275° after 15:00, when the decrease in the signal strength of the wanted signal from Iqaluit allows these secondary signals to emerge, as the gap between the strongest peak and the other peaks decreases (see Section 2.7).

The elevation angle of arrival (middle panel) is below 10° until 15:00. This value is typical of 1-hop F mode propagation along the I-A Great Circle Path (see Table 3.2.1 for the values of the expected elevation angles for the I-A path). Even though this elevation angle value is also typical of 2-hop E mode propagation, the ionograms at these times are suggestive of F region reflection. After 15:00 the elevation angle is slightly higher on average, and in particular between 16:05 and 16:25. The 16:15 ionogram suggests a two-moded propagation (also suggested by the higher spread in the elevation angle trace), possibly a 1-F mode and a 2-F mode. During the whole time interval the elevation angle varies periodically. Since the azimuth of arrival is constantly fairly close to the true bearing, this would possibly suggest either the occurrence of vertical movements of the reflection point or the presence of periodical ionospheric tilts (possibly generated by TIDs, which typically have periods from a few minutes to more than an hour, or other ionospheric disturbances) or the occurrence of moving patches along the I-A path causing the reflection point to move nearer, or otherwise, to the receiver. The elevation shows oscillation also in other intervals of the data set, particularly apparent in Figures 3.3.3 and even more in Figure 3.5.3.

The time plots and the relevant ionograms in the interval between 19:03 and 23:55 are shown in Figures 3.3.3 and 3.3.4 respectively. The MUF was greater than the 9.292 MHz operational frequency at all times. All ionograms show a thin trace, at times more apparent than others, possibly representing a sporadic-E reflection. This is very evident in the 22:15, 22:45 and 23:15 ionograms, at 9.292 MHz. At all other times there is still evidence of the presence of a thin trace, supporting frequencies of operation higher and/or lower than 9.292 MHz. We can infer that a sporadic-E layer was constantly present on the 22nd January 1996 between 19:15 and 23:45 along the I-A path, supporting different

frequency ranges at different times. The azimuth vs. time trace of Figure 3.3.3 is quite narrow and 'clear' until 22:00 and the amplitude is relatively high. After 22:00 the scenario changes: the bearing becomes spread and there is evidence of a second spread trace, the amplitude drops by about 20dB and the elevation angle of arrival also becomes highly variable. In fact, the ionograms after 22:00 show evidence of two-moded propagation. For example, the 22:45 ionogram shows two thin traces at the operational frequency, suggestive of 1-hop and 2-hop sporadic-E, whereas the following ionogram exhibits a thin and a spread trace, possibly sporadic-E and F mode propagation.

3.4 Observations on 23rd January 1996.

The data observed between 00:25 and 04:59 UT on 23rd January 1996 are shown in Figure 3.4.1 and the ionograms taken at these times in Figure 3.4.2. Unlike those recorded on the 22nd January 1996, some of these ionograms show that the MUF was below the frequency of the wanted signal, e.g. at 00:45 and 03:15. The azimuth panel in Figure 3.4.1 confirms that at these times the signal from Iqaluit does not show a clear trace of primary bearings. The energy of the received signal was therefore lower than at other times due to the conditions of the propagation environment, because penetration often occurred, although clearly some generally weak reflection did occur. The occurrence of sporadic-E reflection (thin trace) is generally present in Figure 3.4.2. The amplitude is higher at around 02:30 and from about 04:00 (Figure 3.4.1), when the corresponding ionograms show a thin trace well above the operational frequency, and therefore the signal strength would appear to depend on this sporadic-E propagation mode.

The azimuth panel of Figure 3.4.1 indicates that multiple signals were present. Several traces are evident during the whole interval, at various azimuths contemporarily. The trace that corresponds to the signal from Iqaluit, at around 188.5°, is fairly constant in azimuth between 00:30 and 02:45, when it starts to increase and becomes more spread until 04:45. Between about 00:30 and 01:30 secondary bearings are representative of the wanted signal, which means that a stronger energy coming from unwanted signals or via other modes was received. There are in fact several other traces, the most visible of all being a very spread energy arriving between around 350° and 40° azimuth. This is evidence of a strong signal coming from N-NE. In fact, it is important to note that the amplitude panel shows energy being received during the sounder slot 5-minute intervals, when the transmitter was off. This indicates that one or more signals, other than the signal from Iqaluit, were received during the time interval of Figure 3.4.1. Other secondary traces are also evident. A long trace at

70-80° between 00:30 and 04:00 with a gap of half of an hour around 01:20, which probably represents an interfering signal, because it is not interrupted when the transmitter is off. Two traces from 00:30 to 01:40 and after 03:50, both at 270-280° and a quite strong, swinging and spread trace between 220° and 260° from 00:50 to 01:40 are also evident, which are likely to represent other modes of the signal from Iqaluit, as they disappear when the transmitter goes off.

The middle panel shows a wide range of elevation angles of arrival during the whole period, which is not surprising when we consider that a number of spread signals were received at Alert. The values range from a few to about 20 degrees at most times. Due to the presence of multi-signal and/or multi-moded propagation, and to the lack of accuracy of the elevation measurements, the elevation panel does not provide precise information about the signal environment. However, at around 02:30 we can see elevation measurements below 10°, and a single-moded propagation from the azimuth panel, this indicating that possibly a 2-hop E mode was present, which is in agreement with the thin and straight trace shown by the ionograms around that time. At around 01:05 and 02:50, two 10-minute periods stand out immediately, in which the elevation angles ranged between 20° and 30°. Since in both cases there is a break in this signal when the transmitter is off, the trace is likely to represent a propagation mode from Iqaluit with the reflection point closer to the receiver than the transmitter, possibly caused by ionospheric tilts or by patches of enhanced ionisation.

The observations between 05:09 and 10:02 are shown in Figure 3.4.3 and the corresponding ionograms in Figure 3.4.4. The same time interval was analysed by Dumas (1997) (see Section 1.5). The azimuth vs. time plot exhibits very pronounced swings throughout the whole time interval. The largest variations in bearings over the 23rd January 1996 occur between 05:30 and 09:30, when the value of the geomagnetic index $K_p=1-$ was the lowest in the data interval, which is consistent with Table 1.5.1. These bearing swings are often regarded as caused by reflections from large structures of enhanced ionisation (patches and arcs, see Section 1.5).

The ionograms exhibit a highly variable structure of the ionosphere, this variability being clearly confirmed by the DOA and amplitude measurements. The ionograms often show multipath propagation, and in particular most of those in the right hand column (07:45 to 09:45) exhibit one or two patches around 9.292 MHz. These patches are detached and at different delay values with respect to the main trace, and are consistent with the presence of large

drifting structures of enhanced ionisation. Furthermore the roughly oscillating elevation angle indicated the presence of ionospheric tilts, as in other intervals.

In detail, a number of swings from low to high azimuth can be seen between 05:20 and 07:10. Each of these swings lasted for about 20-30 minutes, with the azimuth increasing by about 20-30°. When each of traces ends, the next one follows starting from a lower bearing, but ends at a value of azimuth greater than the final value of the previous one. The steepest swing occurs between 07:20 and 07:40, from 120° to 170° azimuth, i.e. 50° azimuth swing in 20 minutes. After this swing, the azimuth drops suddenly to about 130° at 07:50, then goes up to 210° at 09:35 (80° in 1h45'). A strong interfering signal is present 0° for most of the time interval. A weaker signal (which is an interferer as it does not disappear when the transmitter is off) is detected at about 260° at around 07:00, between 07:30 and 08:00, and between 09:10 and 09:35.

This interval is clearly suggestive of a periodic phenomenon of large density structures drifting towards increasing values of azimuth, that is westwards for the I-A path. This periodicity was also seen in the work carried out by Dumas (1997) and Warrington *et al* (1997b), also for other propagation paths. In particular the ionograms after 07:30 show detached features at the operational frequency. In the three ionograms between 08:45 and 09:45 we can see that a patch is present alternatively in the first and third panels, again indicating a recurrent series of features. In general, the ionograms of Figure 3.4.4 provide evidence for a multipath propagation environment. Even though the precise identification of these swings is often difficult, there are periods when traces corresponding to two swings are present at the same time, for example at about 09:30, with a stronger trace at around 200° and a weaker trace at around 150° (also seen by Dumas, 1997). Even though relating this occurrence to the ionograms is not obvious, it is possible that the swing terminating at 200° is caused by the patch visible on the 08:45 ionogram (this swing starts at around 08:30), whereas the swing starting at 150° is caused by the patch visible on the 09:45 ionogram.

Rogers *et al* (2001) stated that the ionograms at these times indicate the reflection of structures that drift transverse to the propagation path. They also stated that the occurrence of these bearing swings could, possibly, result from reflections from multiple sun-aligned arcs (see Section 1.4) over the propagation path. Arcs drift from dawn to dusk (see Section 1.4), which at about 07:45 (when patches are visible for the first time in the ionograms) corresponds to a direction diagonally across the I-A path. Whereas at about 09:45 (when patches are visible for the last time) the arcs would drift in a direction along the I-A path. Furthermore and more importantly, the elevation angle would be expected to

increase gradually (slower towards the beginning of this 2-hour interval and faster towards the end) and then drop down suddenly when the signal is reflected from the following arc, then gradually increase again, and so on. As these circumstances do not occur, the bearing swings are unlikely to have been caused by arcs. However, the bearing swings are consistent with those of a multiple patches simulation (see Section 3.6) drifting across the propagation paths with the convection flow (see Section 1.4).

3.5 Observations on 24th January 1996.

The DOA and amplitude observations from 09:36 to 14:22 UT on 24th January 1996 are illustrated in Figure 3.5.1 and the ionograms in Figure 3.5.2. The signal from Iqaluit is very variable in azimuth in the first 2.5 hours of this time interval and other traces can be seen representing interfering signals incoming from various directions. The wanted signal has a number of swings with increasing values of azimuth between 09:40 and 11:50, the steepest occurring between 09:35 and 10:00, when the bearing goes from about 145° to 210°, 65° in 25 minutes. At the same time a peak in amplitude occurs. In the 10 minutes around 09:50 the elevation is above 10°, making more likely the signal path to be a 1-hop F rather than a 2-hop E propagation. The relevant 09:45 ionogram (Figure 3.5.2, top left panel) shows a patch at 9.292 MHz, detached from the main trace. At that time, the ionosphere did not support a Great Circle Path (GCP) mode; thus a patch might 'offer' a reflection point displaced from the GCP. The following ionograms (up to 11:45) show detached patches at various relative delay differences from the main trace. As we have already seen in other examples in the previous section, a patch moving transversally to the propagation path can cause a swing in the direction of arrival of the signal (see next section). The elevation angle of the signal is generally below 15° when the azimuth swings were present, and the amplitude is relatively low, with the exception of two further peaks at about 10:30 and 11:25.

After 12:00 the bearing of the signal tends to stabilise around the GCP. There is a slow positive change in azimuth, which starts at 12:00 and ends at 14:00 from about 180° to 195°, i.e. 15° in 2 hours. At these times the amplitude of the signal is much higher. The relevant ionograms (right hand panels) no longer show evidence of a detached feature and the MUF is constantly well above the frequency of operation, unlike in the previous ionograms (left hand panels), where the frequency of operation is generally close to the MUF. There appears to be some periodicity in the elevation with peaks at around 20°, at 12:30, 12:45 and 13:15. The interpretation of the sharp trend of the elevation between 12:00 and 13:30 in Figure 3.5.1 is made more difficult by the 5 min breaks when the

ionograms were taken between 12:15 and 13:15. However, an expanded view of the observations at these times (not shown in the thesis) would suggest a varying elevation angle, with growth faster than the decay. This, given the fairly constant bearing, would then suggest a series of features drifting along the propagation path from Alert to Iqaluit, i.e. from dusk to dawn, which would exclude the presence of arcs, which drift in the opposite direction (see Section 1.4). However, these elevation swings could be caused by patches. Although there is no clear evidence of patches in the ionograms, the trace due to signal being reflected by them could be hidden within the rather spread main trace. In fact, the difference in the propagation delay for two paths through the F layer, one via a patch, would be between 0 and about 0.5 ms, and therefore both traces could be part of the ionogram spread. In such circumstance, the series of patches would not be in agreement with the theoretical convection flow (see Sections 1.4), which, however, is estimated on a statistical basis and therefore the occurrence of patches having caused these elevation swings cannot be dismissed. Even though this interesting portion of data is not easy to interpret, a simulation presented in the next section has shown consistency with a patch drifting along the GCP in the A-I direction.

The time plots and the ionograms from 15:53 to 20:40 are illustrated in Figures 3.5.3 and 3.5.4 respectively. During this period the azimuth of the received signal is around the true bearing all the time. The single trace is narrow and clear, rarely surrounded by secondary bearings, probably as a result of fairly high amplitude. In the middle panel the elevation angle of arrival oscillates, being on average below 10° (either 2-hop E or 1-hop F mode). This apparent oscillating elevation angle coupled with a fairly constant azimuth of arrival (phenomenon also shown towards the end of the previous time interval in Figure 3.5.1, as well as in the Figures of Section 3.3), is once again likely to be caused by periodical ionospheric tilts, with subsequent movements of the reflection point closer to the receiver (higher angle of arrival) or otherwise (lower angle of arrival). All these ionograms show a well-defined main trace, the MUF being well above the 9.292 MHz operational frequency. Some of the ionograms show that a multipath propagation was likely to occur. In particular the 19:15 ionogram exhibits 2-hop F and both low and high angle propagation, which would explain the slightly higher elevation angle at around 19:15 than elsewhere (excluding when the transmitter was off). Multimoded propagation then occurred, possibly over several paths on the GCP, since the bearing is fairly constant at this time.

Further interpretation of the results is left to the next chapter. Expanded views of some portions of the time plots will be shown, supported by the relevant

ionograms, in order to assess the effect of different propagation conditions on direction finding systems. The results obtained with other algorithms (Loaded Capon, MUSIC, SML and Beam) will be compared. Furthermore, the results obtained with two different array patterns will be compared in order to investigate the effect of the array geometry.

3.6 Simulation of a drifting over-dense plasma structure.

Large bearing deviations in the azimuth of the received signal at high latitudes are often present (Warrington *et al*, 1997c). These swings are usually attributed to the movement of over-dense plasma structures, such as patches and arcs, over the polar cap. When the MUF is exceeded by the operational frequency, paths other than the GCP can become more apparent, and these structures are likely to reflect the signal. The reflection point will then move as these structures move, resulting in bearing deviations, which are more prevalent during magnetically quiet periods (see Table 1.5.1).

Simulations of a drifting mono-dimensional reflector have been undertaken. Firstly given the geographical co-ordinates (latitude and longitude) of two points, the GCP distance between them, the azimuth and elevation (for a given height) of one point with respect to the other are calculated assuming a perfectly spherical Earth. The trajectory of the patch is then simulated by simply varying the co-ordinates of the reflector with the desired speed, and recalculating the new azimuth and elevation (one estimate per second). The purpose of these simulations is to show that the large bearing deviations in the observations can be simulated with a simple model consisting of a reflecting moving structure, taking care in the assumption of realistic trajectories (see below).

Two examples are illustrated in Figure 3.6.1 of intervals where observed and simulated bearings are increasing. The observations have been obtained with the INS algorithm using the 8-antenna sub-array (see Section 3.2). The top left hand panel exhibits a swing from about 145° to 210° azimuth, between 05:37 and 06:01 Local Time (LT) on 24th January 1996. The elevation angle varies between 5-10°, consistent with a 1-hop F mode. In the top right hand panel the azimuth rises from about 195° to 205°, between 05:01 and 05:25 LT on 23rd January 1996. The elevation panel below exhibits a slightly decreasing angle, between 5-10°. An example of experimental observations and relevant simulations in which the elevation angle varied but the azimuth was fairly constant is presented later in this section.

The third and bottom panels show the results of two bearing vs. time and elevation vs. time simulations respectively. A point reflector is assumed to

move westwards, at a height of 240 km and 210 km in the left and right hand side panels respectively, with trajectories shown in Figure 3.6.2. These reflection heights are typical of the F region, where large-scale electron density structures can occur. 24 minutes of data were simulated. In the case on the left hand side in Figures 3.6.1 and 3.6.2, the reflector would cover 1200 km (from 800 km to the east to 400 km to the west of the GCP) at a speed of 840 m/s. This velocity is more typical for a patch of enhanced ionisation than an arc (see Section 1.4). The signal is assumed to arrive at the receiver at Alert after a reflection from a point moving westwards (according to the convection flow pattern, see later in this section), crossing the GCP at a distance of 1000 km from Alert (about half distance between transmitter and receiver). It is important to bear in mind that this combination of parameters does not form a unique solution, and is given as an example (see also later in this section). The azimuth of arrival of the signal increases, becoming greater than the true bearing (188.5°) after the point crosses the GCP. In the right hand panels, the reflector is also assumed to move westwards, covering 200 km, starting at 100 km to the west of the GCP (also at the level of 1000 km from Alert), and arriving at 300 km to the west of the GCP at the end of the simulation. The speed of this reflector is assumed to be about 140 m/s. Even though this speed is lower than the speed typical of patches, the direction of this reflector would follow the convection flow pattern, and as such it is more likely to be a patch than an arc (see below). Both the bearing and the elevation panels of this simulation (third and bottom in Figure 3.6.1) show good agreement with the experimental data (top and second panels in Figure 3.6.1). However, it is worth bearing in mind that the elevation angle estimates are much less accurate than the bearing measurements (see Figures 3.2.3, 3.2.4 and 3.2.5).

It is important to calculate the minimum height at which a point reflector can be seen at a given distance from a receiver, to make sure that the assumed distances of the reflector from the receiver are realistic. Figure 3.6.3 shows that when a point is about 1100 km from the receiver, which is about the maximum distance of the point from Alert in either simulation, the minimum height for it to be visible from the receiver is between 90 km and 100 km. A 90-km height corresponds to the bottom of the E region, whereas these structures drift across the F region. This implies that the drifting features have not gone below the horizon. Therefore all the parameters assumed in the simulations are realistic.

It is also important to note that very often the bearing swing traces overlap one another. For example, in the top left panel of Figure 3.6.1 we can see the end of a trace just after 05:37 LT at about 190° , when the main trace of the panel starts at 145° . The main trace ends at about 06:00, when another trace starts at 170°

(see Figure 3.5.1 for a more comprehensive view). This indicates that the main trace stopped being visible when another more intense feature crossed the propagation environment, and not that, for example, the previous feature disappeared below the horizon. Two traces are present at the same time around the first and last two minutes of the time interval, whereas in between the main trace only is visible. This happens either because the gap in strength between the main signal and other possible signals was too big (so the weaker signal was not detected), or simply because the main trace represents the only signal that was present at that time. A combination of the two above mentioned situations across the time interval is likely to have occurred. Furthermore, the top left hand panel trace of Figure 3.6.1 represents one of the most dramatic scenarios of the whole data set, as it starts at about 145° , i.e. 45° below the true bearing, and ends at about 20° above the true bearing. Therefore, if the feature was visible at 45° below the true bearing, it was very likely to also be visible at beyond 20° above the true bearing. Furthermore, at the end of the simulation the point reflector is at about 1100 km from Alert while a point reflector which is at 230 km of height can be seen from a distance of about 1700 km (see Figure 3.6.3), and therefore it should be above the horizon.

Simulations have been undertaken to reproduce the sharp elevation angle swings (with fairly constant bearing along the GCP) occurring on 24th January 1996 between 12:30 and 13:30 UT (see Figure 3.5.1). Figure 3.6.4 shows an example of particularly sudden decrease in the elevation, from about 15° to 5° in 4 minutes. The observations have been obtained with INS and 8 antennas (left hand plots). Given the fairly constant bearing at about the direction of the GCP, a point reflector drifting along the GCP, in the direction from Alert to Iqaluit (decreasing elevation angles) has been simulated (right hand panels). This point reflector is not likely to represent an arc, since they move from dawn to dusk (see Section 1.4), i.e. the opposite direction. However, it could represent a patch (see previous section) moving in a direction which does not agree with the theoretical convection flow. The good agreement of Figure 3.6.4 is obtained with a reflection height of 220 km, a speed of 2200 m/s (typical of a patch, but not an arc), starting and ending points at 670 and 1180 km from Alert respectively.

Simulations of all the bearing swings which occurred during the data set have been undertaken. Although the identification of the number and precise trend of the swings is often far from obvious, at least 12 swings can be identified, eight on the 23rd and four on the 24th January 1996. Table 3.6.1 shows the time interval (third column) and the azimuth range (fourth column) of these swings, numbered in chronological order in the first column. The fifth and sixth

columns show the direction and the speed respectively of a simulated moving point reflector, which would cause the corresponding measured bearing swings. The results in Table 3.6.1 are illustrated in Figure 3.6.5. The straight arrows, numbered 1 to 12 (corresponding to the numbers in Table 3.6.1), represent the direction of the moving features simulated by point reflectors. The points representing the drifting patches are reflected at heights comprised between 200 km and 250 km, therefore in the F region. The length of the arrows is approximately proportional to the speed of the moving features (see Table 3.6.1). The direction and speed of the reflectors match both azimuth and elevation measurements in all cases. The simulation was able to obtain a range of reasonably good matches for the reflector crossing the GCP between about 800 and 1200 km from Alert, and by varying the values of direction (generally about 250-300°), speed (100 m/s to a few km/s) and reflection height (see above) depending on each swing. The values in Table 3.6.1 were calculated for a patch crossing the GCP at 1000 km from Alert.

Table 3.6.1 Bearing swings occurred on 23rd January 1996 (numbered 1 to 8) and on 24th January 1996 (9 to 12). In the last two columns are the direction and speed of the simulated moving reflector point, for which trajectories cross the GCP at 1000 km from Alert.

Swing No.	Date	Time interval (Local Time)	Azimuth range	Direction, E of N	Speed, m/s
1	23/01/96	01:20-01:45 (25 min)	155°-175° (20°)	290°	730
2	23/01/96	01:45-02:10 (25 min)	160°-185° (25°)	250°	270
3	23/01/96	02:15-02:40 (25 min)	165°-190° (25°)	280°	320
4	23/01/96	02:40-03:10 (30 min)	175°-200° (25°)	290°	270
5	23/01/96	03:20-03:40 (20 min)	130°-165° (35°)	260°	670
6	23/01/96	03:50-04:10 (20 min)	130°-160° (30°)	265°	500
7	23/01/96	04:15-04:35 (20 min)	155°-180° (25°)	270°	390
8	23/01/96	04:55-05:35 (40 min)	185°-210° (25°)	270°	140
9	24/01/96	05:35-06:00 (25 min)	145°-210° (65°)	270°	840
10	24/01/96	05:55-06:25 (30 min)	180°-235° (55°)	280°	610
11	24/01/96	06:50-07:10 (20 min)	155°-195° (40°)	255°	570
12	24/01/96	07:20-07:40 (20 min)	160°-200° (40°)	270°	580

The curved dashed arrow in Figure 3.6.5 represents an indication of the direction of the convection flow during both the examined periods of time, i.e. between about 01:20 and 05:35 LT on 23rd January 1996, and between 05:35 and 07:40 LT on 24th January 1996. The likely trend of the convection flow during these two time intervals has been obtained by considering the average values of the B_y and B_z components of the IMF (NASA Coordinated Data Analysis Web Site). As $B_y > 0$ most of the time during the two time intervals both on the 23rd and on the 24th January 1996, we can infer from Figure 1.4.4 (both top and bottom right hand plots) that the convection flow had about the direction given by the curved dashed arrow in Figure 3.6.5 (for both $B_z < 0$ and $B_z > 0$). Bearing in mind that the convection flow is obtained on a statistical basis, the directions of the point reflectors clearly follow the direction of the convection flow (see Section 1.4) in all 12 analysed cases of bearing swings, which is consistent with the swings occurring on the 23rd and 24th January 1996 having been caused by trains of patches crossing the polar cap at about half way between Iqaluit and Alert.

In conclusion, the results of these simulations show that the measured bearing and elevation swings are consistent with those caused by structures of over-dense plasma (patches in these cases) drifting across the polar cap. However, it is not suggested that a point reflector is a good model for over-dense structures, which have a diffuse nature.

3.7 Summary.

The data acquisition system used to collect the data analysed in this work is situated at Alert, in the Canadian Arctic. The data were collected when the receiver was recording a signal transmitted from Iqaluit, 2100 km south of Alert, which propagated within the polar cap in wintertime.

The observations showed that the performance of HF direction finding systems can be highly affected and severely degraded by the propagation environment at high latitudes, as expected from previous work (see Section 3.1). In fact, the DOA of the signal can be well displaced from the GCP when the signal is reflected by structures of over-dense plasma far from the GCP.

In this chapter the whole data set from the 22nd to the 24th January 1996 for the transmission path from Iqaluit to Alert has been presented. The propagation of the signal varied dramatically from one day to another and also within the same day. Large and rapid bearing swings often occurred, in particular when the relevant ionograms showed detached features from the main trace, which is often likely to be evidence of patches or arcs of enhanced ionisation.

Several simulations of a mono-dimensional moving reflector have been undertaken which show that the signal bearing and elevation swings can be caused by trains of patches of over-dense plasma drifting across the polar cap. The simulated direction of the point reflector closely followed the direction of the convection flow in the case of the bearing swings.

The next chapter will focus on the interpretation of the results obtained by processing the data with various algorithms and with different antenna geometries.

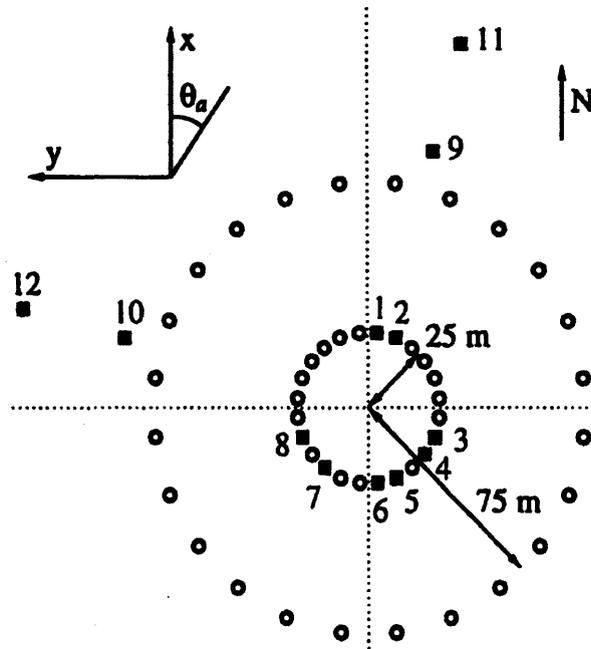


Fig. 3.2.1a Alert Vortex DF system. The full squares represent the antenna elements in the Vortex antenna array, and the circles are unused antennas which are part of a 2-ring DF system (Dumas, 1997).

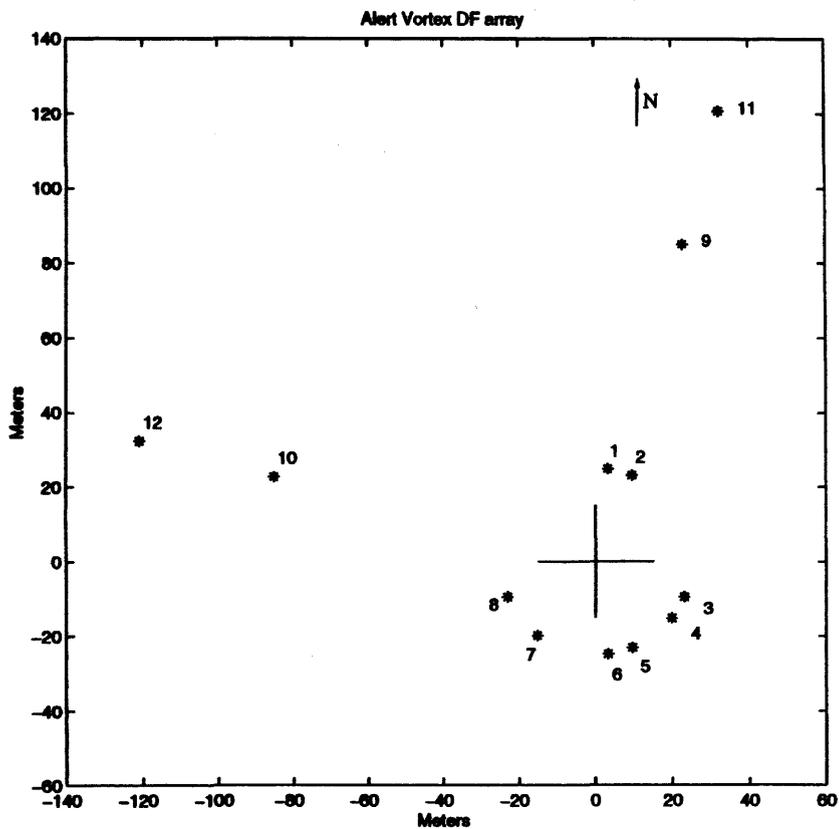


Fig. 3.2.1b Positions of the antennas in the Vortex array.

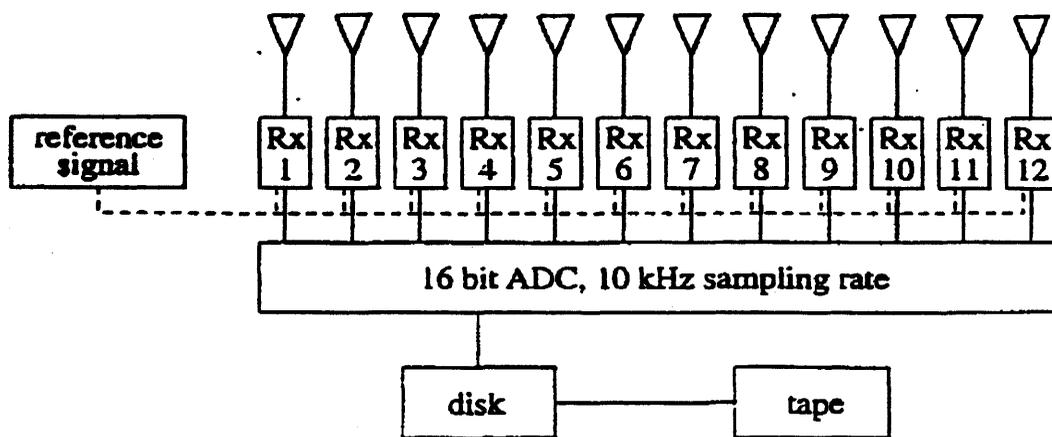


Fig. 3.2.2 Vortex data acquisition system (Dumas, 1997).

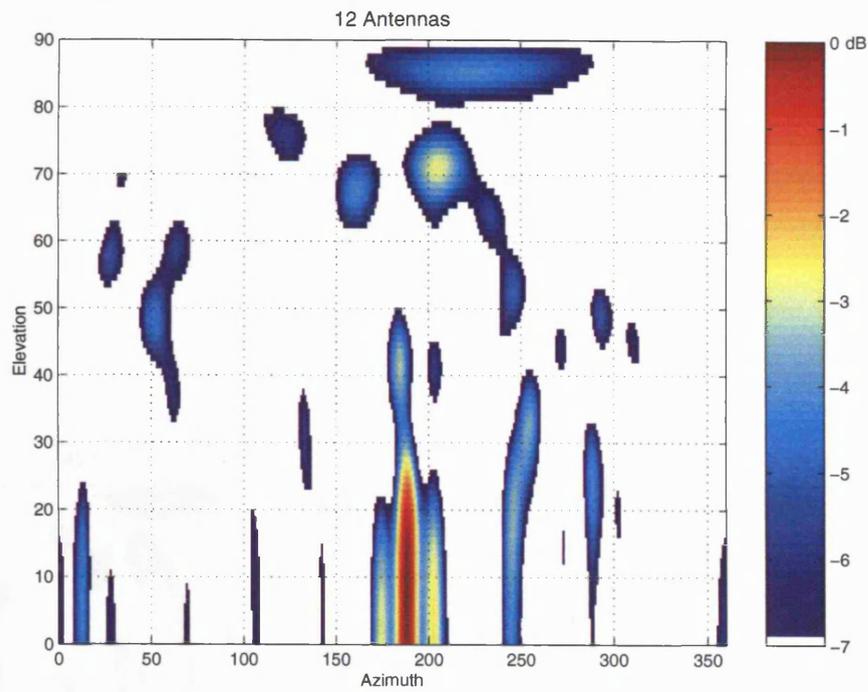


Fig. 3.2.3 Beam pattern formed by the full 12 antenna Vortex array. Look direction: $az=188.5^\circ$, $el=8.3^\circ$ (GCP direction, 1 hop F region reflection at 250 km height).

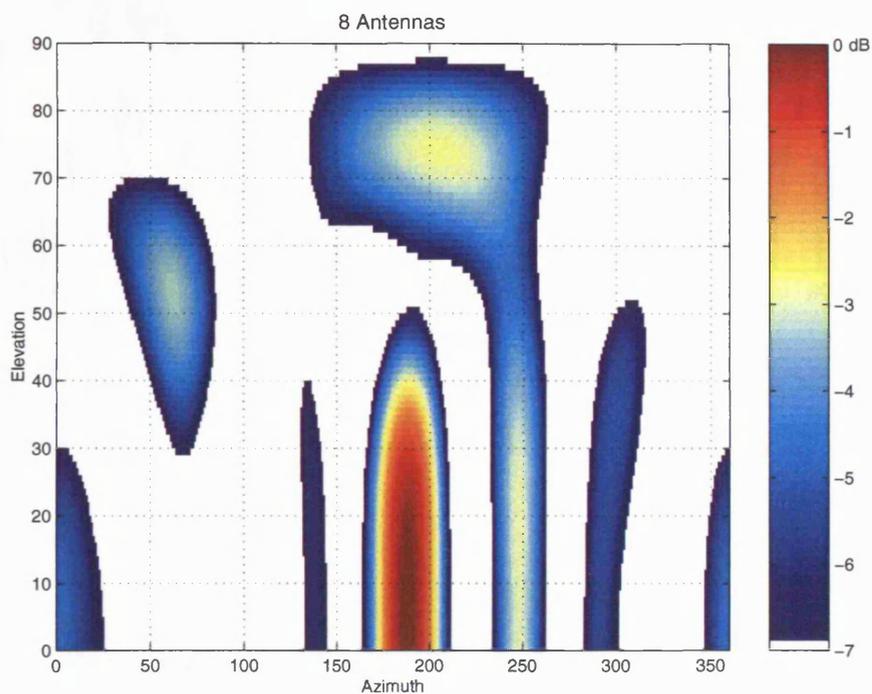


Fig. 3.2.4 Beam pattern formed by the Vortex sub-array of antennas 1 to 8. Look direction: $az=188.5^\circ$, $el=8.3^\circ$ (GCP direction, 1 hop F region reflection at 250 km height).

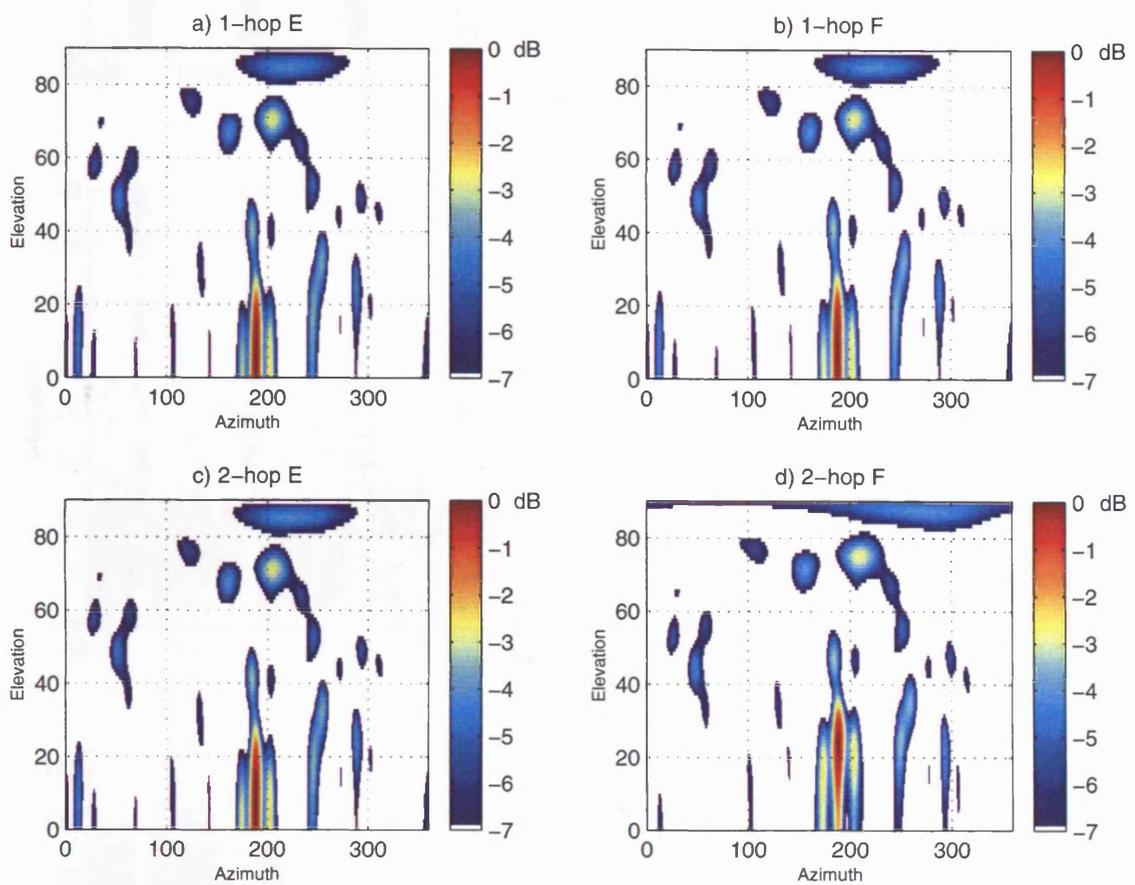


Fig. 3.2.5 Beam patterns formed by the full 12-antenna Vortex array. Look direction: $az=188.5^\circ$; *a*) $el=0.6^\circ$ (1-hop E), *b*) $el=8.3^\circ$ (1-hop F), *c*) $el=8.1^\circ$ (2-hop E) and *d*) $el=22.9^\circ$ (2-hop F).

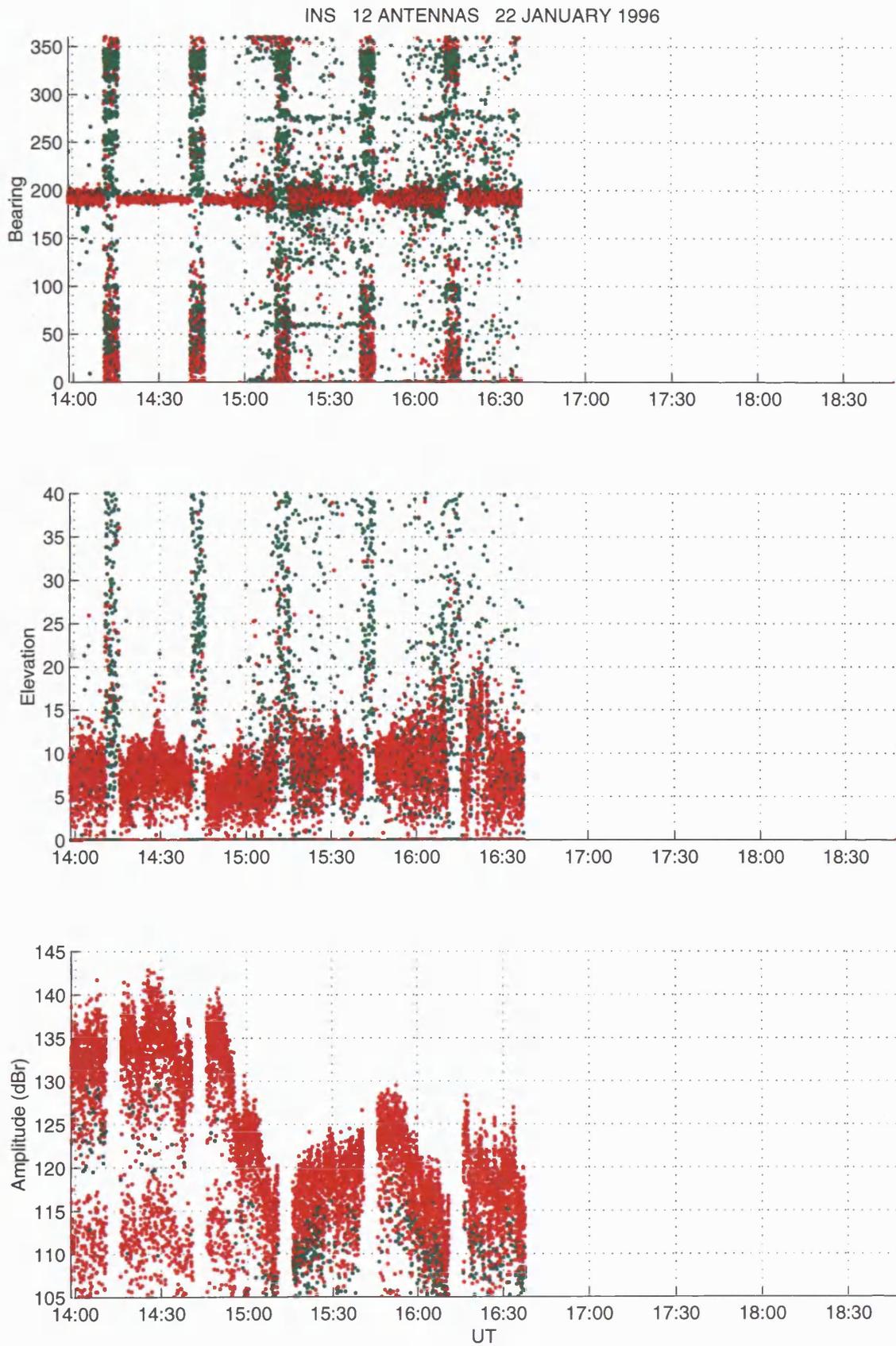
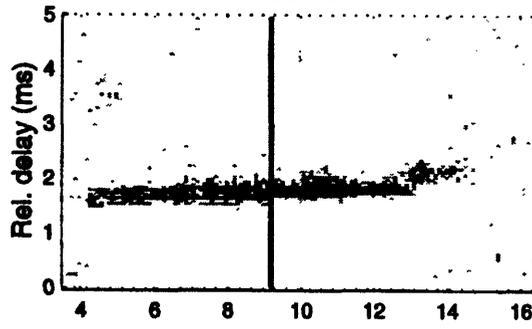


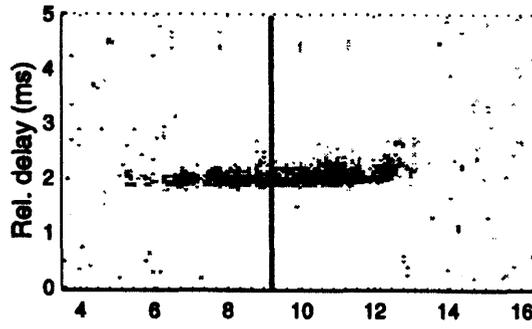
Fig. 3.3.1 Iqaluit-Alert azimuth, elevation and relative amplitude vs. time plots for the 22nd January 1996 (13:58 - 16:38 UT). The red and green dots represent the primary and secondary bearing estimates respectively.

22 JANUARY 1996

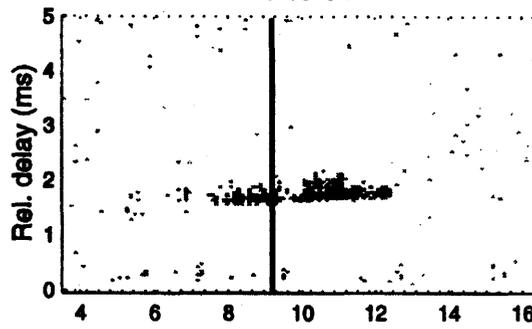
14:15 UT



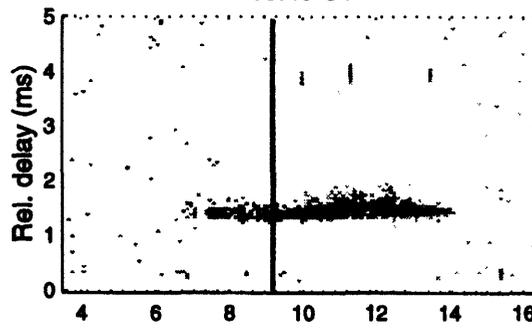
14:45 UT



15:15 UT



15:45 UT



16:15 UT

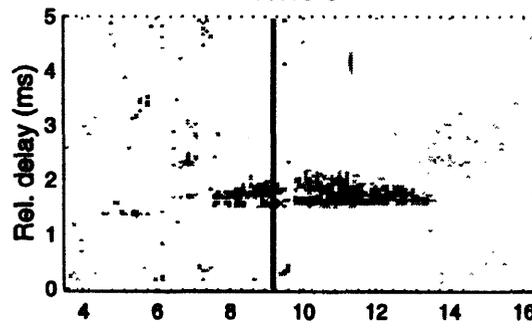


Fig. 3.3.2 Oblique ionograms for the Iqaluit-Alert path (22nd January 1996).

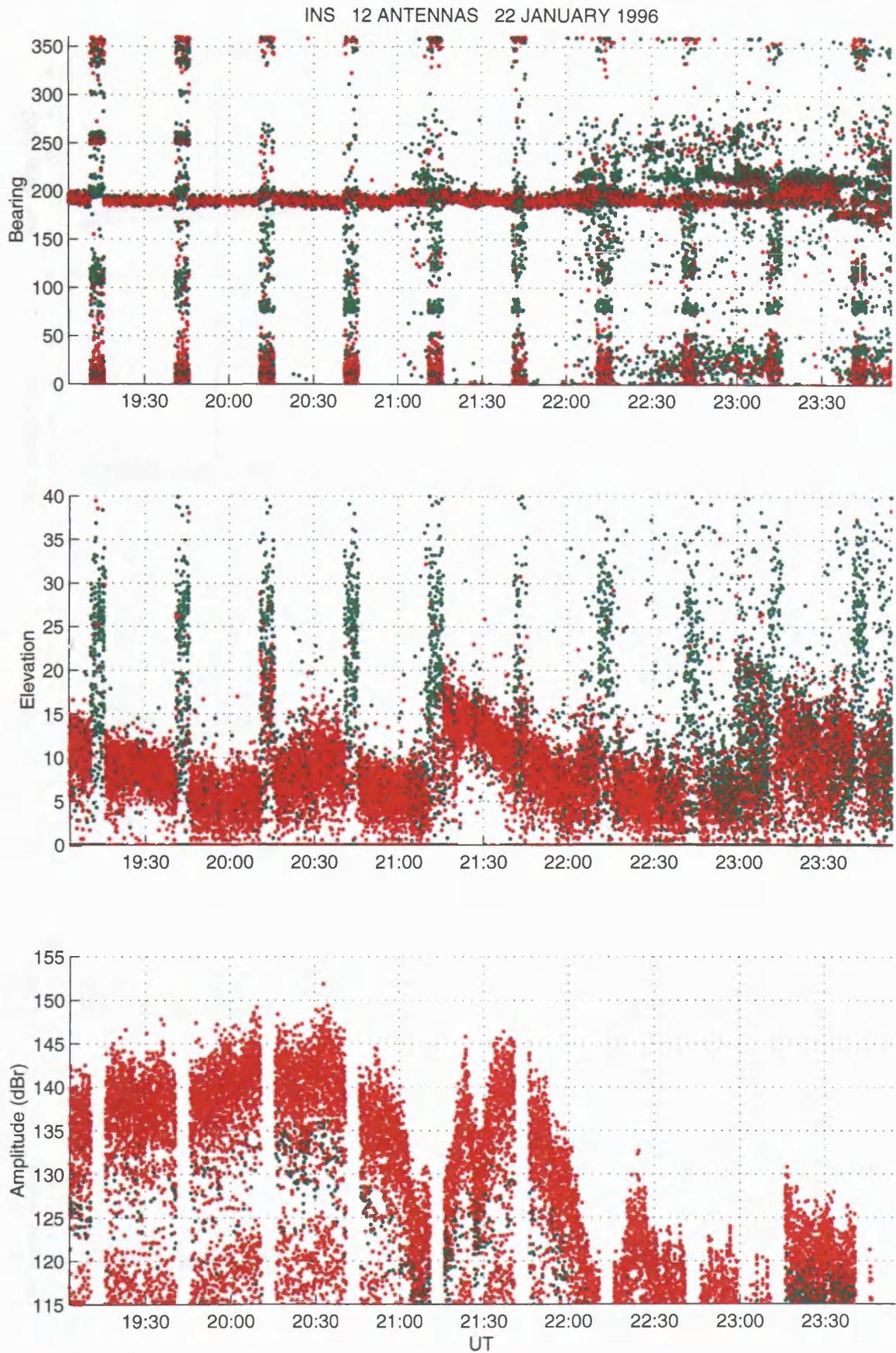


Fig. 3.3.3 Iqaluit-Alert azimuth, elevation and relative amplitude vs. time plots for the 22nd January 1996 (19:03 - 23:55 UT).

INS 12 ANTENNAS 23 JANUARY 1996

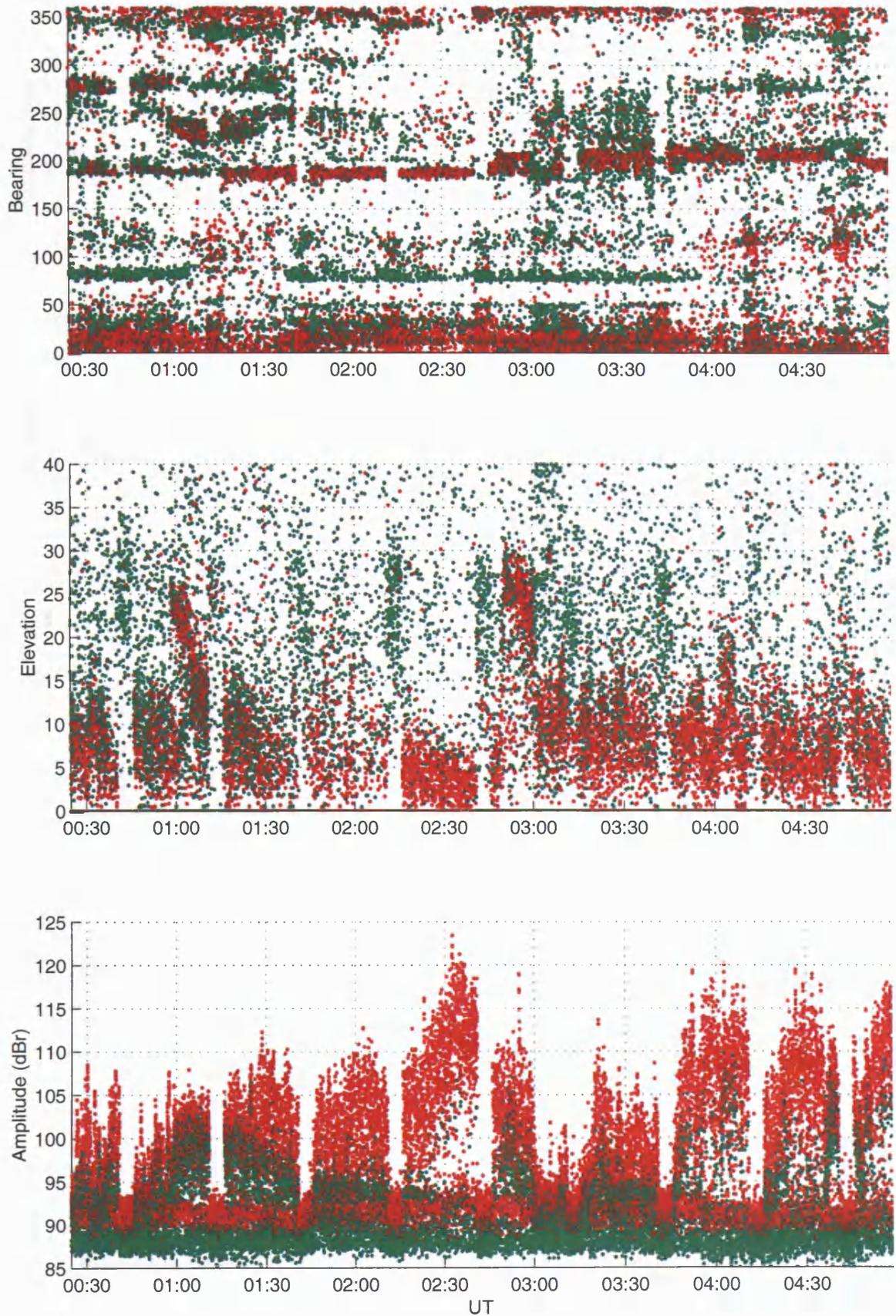


Fig. 3.4.1 Iqaluit-Alert azimuth, elevation and relative amplitude vs. time plots for the 23rd January 1996 (00:25: - 04:59 UT).

INS 12 ANTENNAS 23 JANUARY 1996

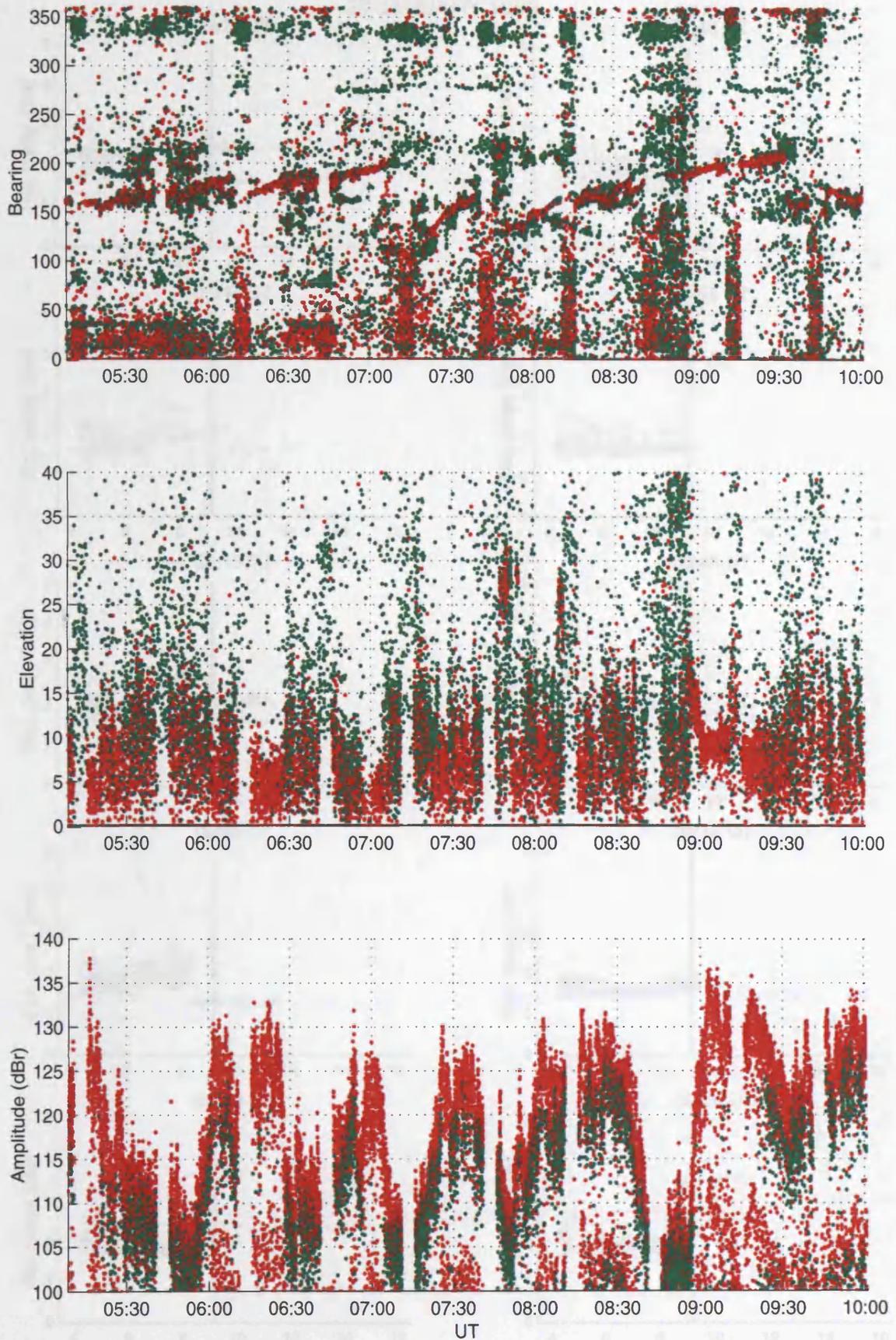


Fig. 3.4.3 Iqaluit-Alert azimuth, elevation and relative amplitude vs. time plots for the 23rd January 1996 (05:09 - 10:02 UT).

23 JANUARY 1996

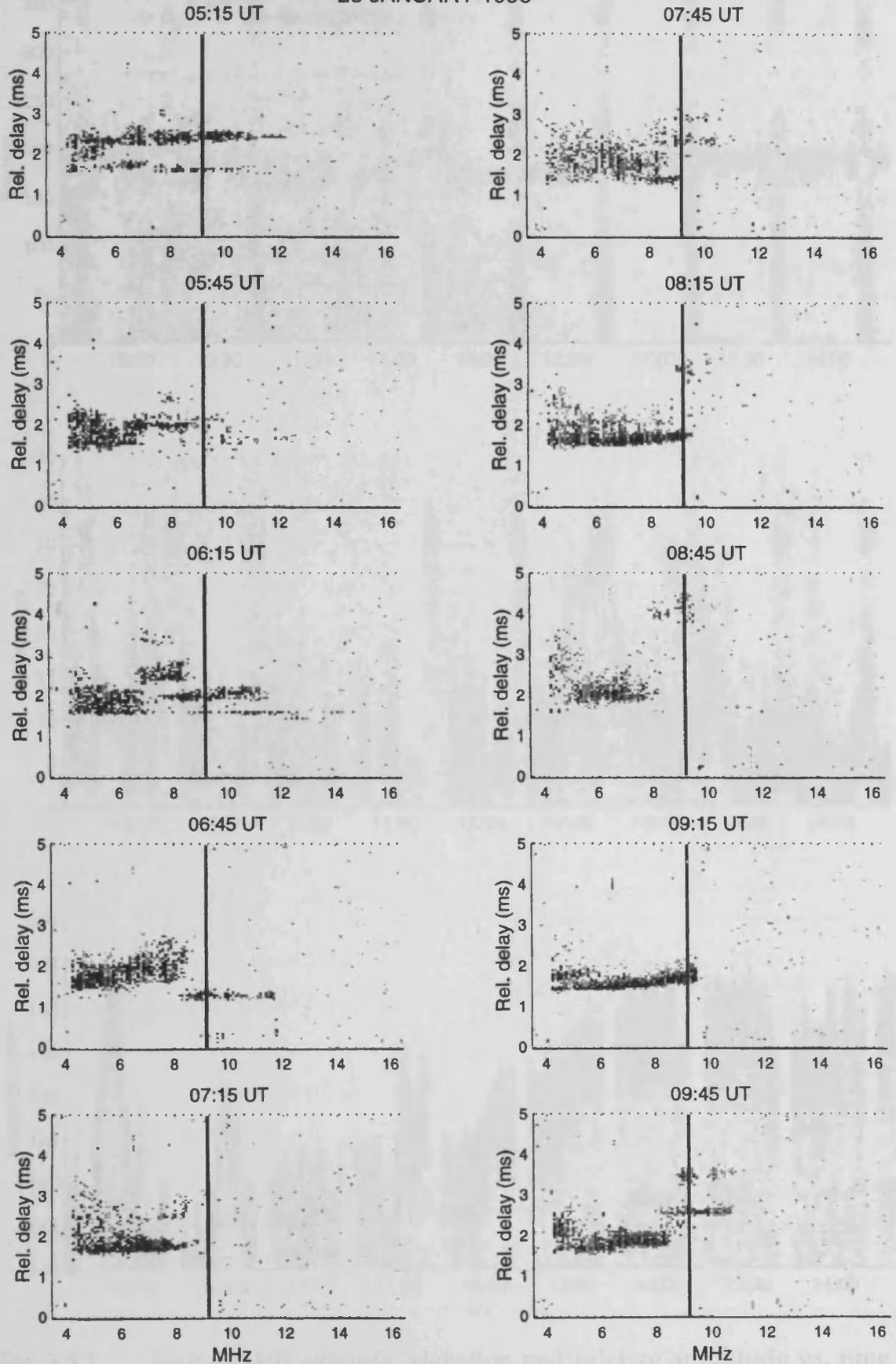


Fig. 3.4.4 Oblique ionograms for the Iqaluit-Alert path (23rd January 1996).

INS 12 ANTENNAS 24 JANUARY 1996

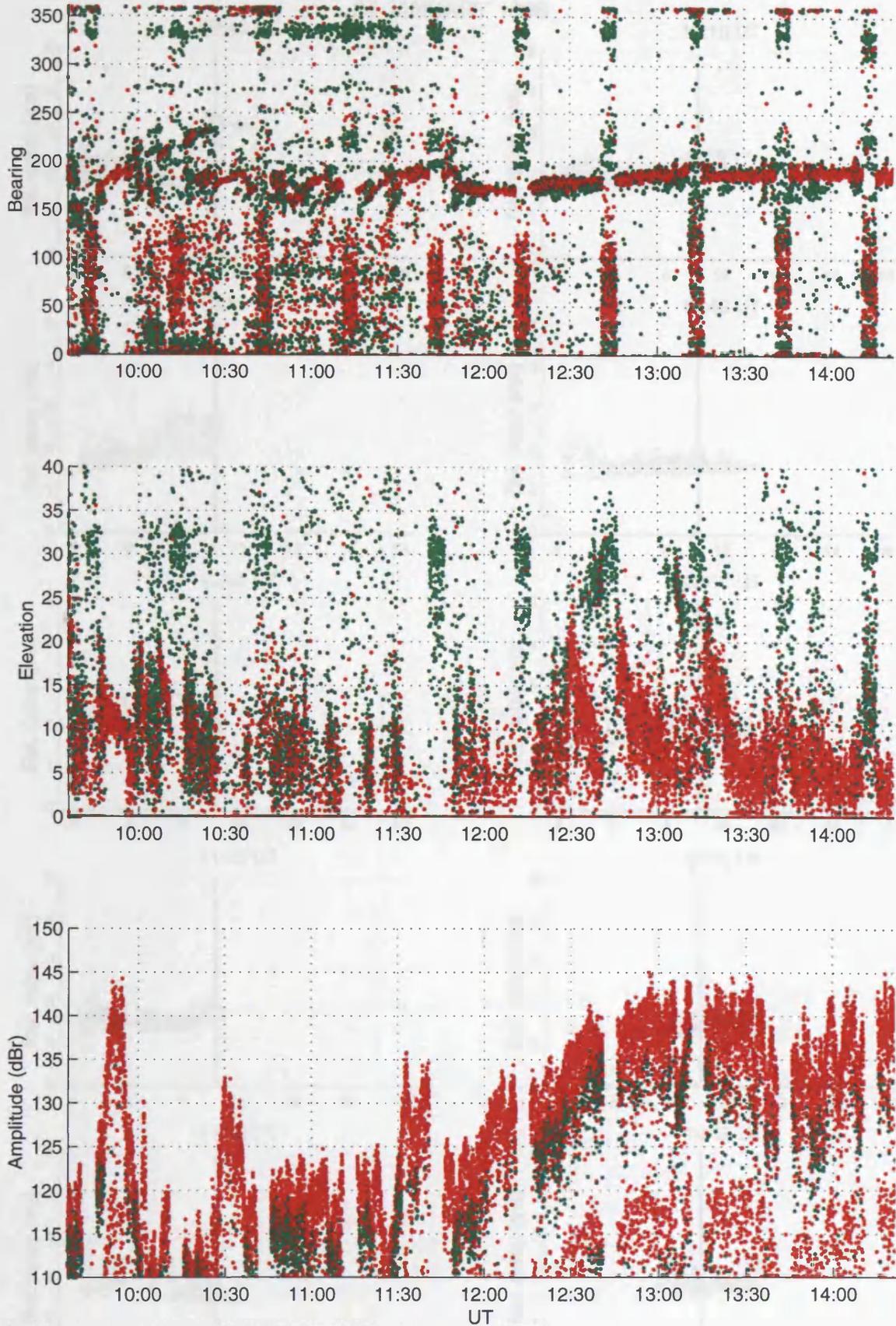


Fig. 3.5.1 Iqaluit-Alert azimuth, elevation and relative amplitude vs. time plots for the 24th January 1996 (09:36 - 14:22 UT).

Fig. 3.5.2 Oblique ionograms for the Iqaluit-Alert path (24th January 1996).

24 JANUARY 1996

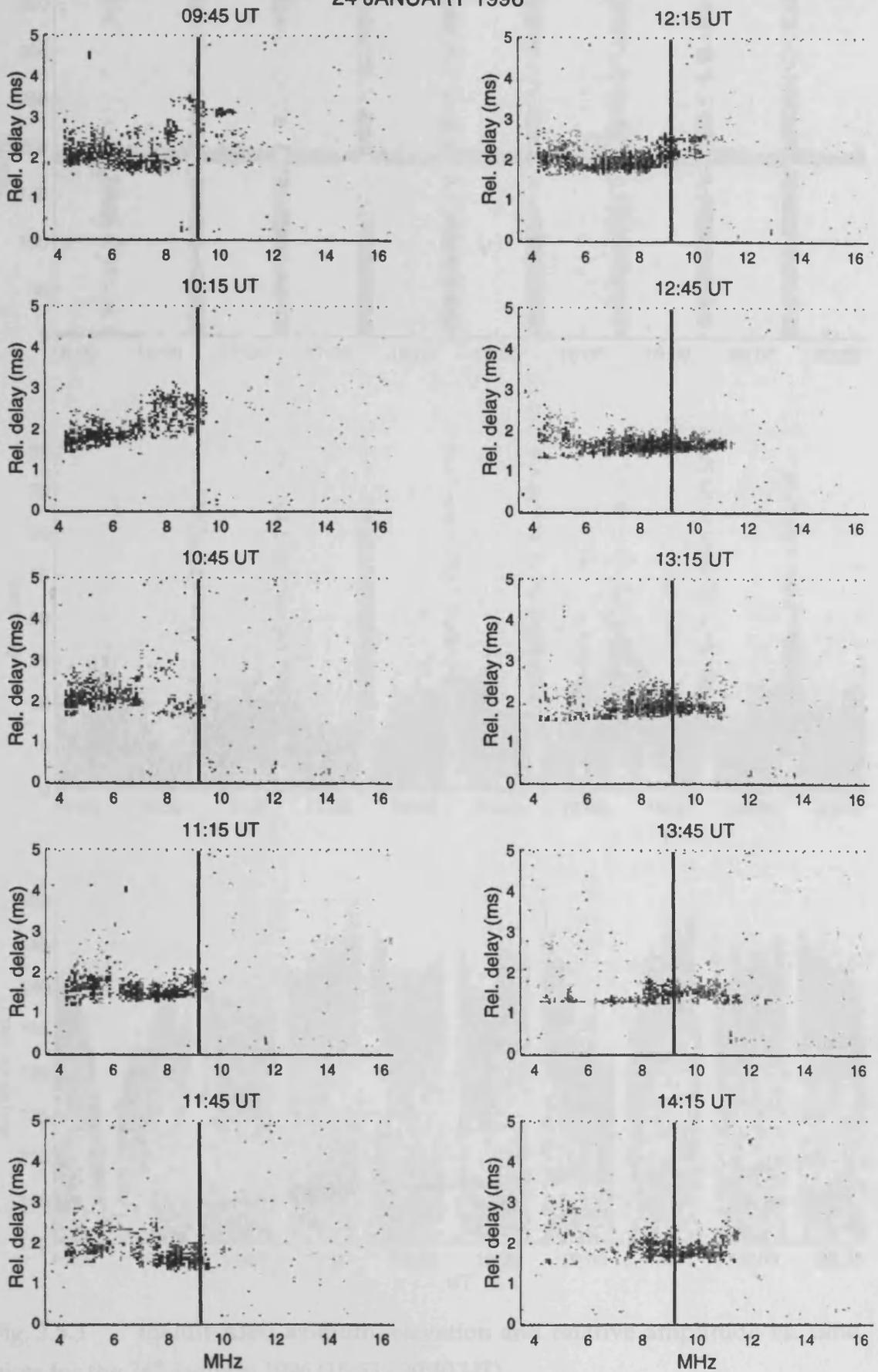


Fig. 3.5.2 Oblique ionograms for the Iqaluit-Alert path (24th January 1996).

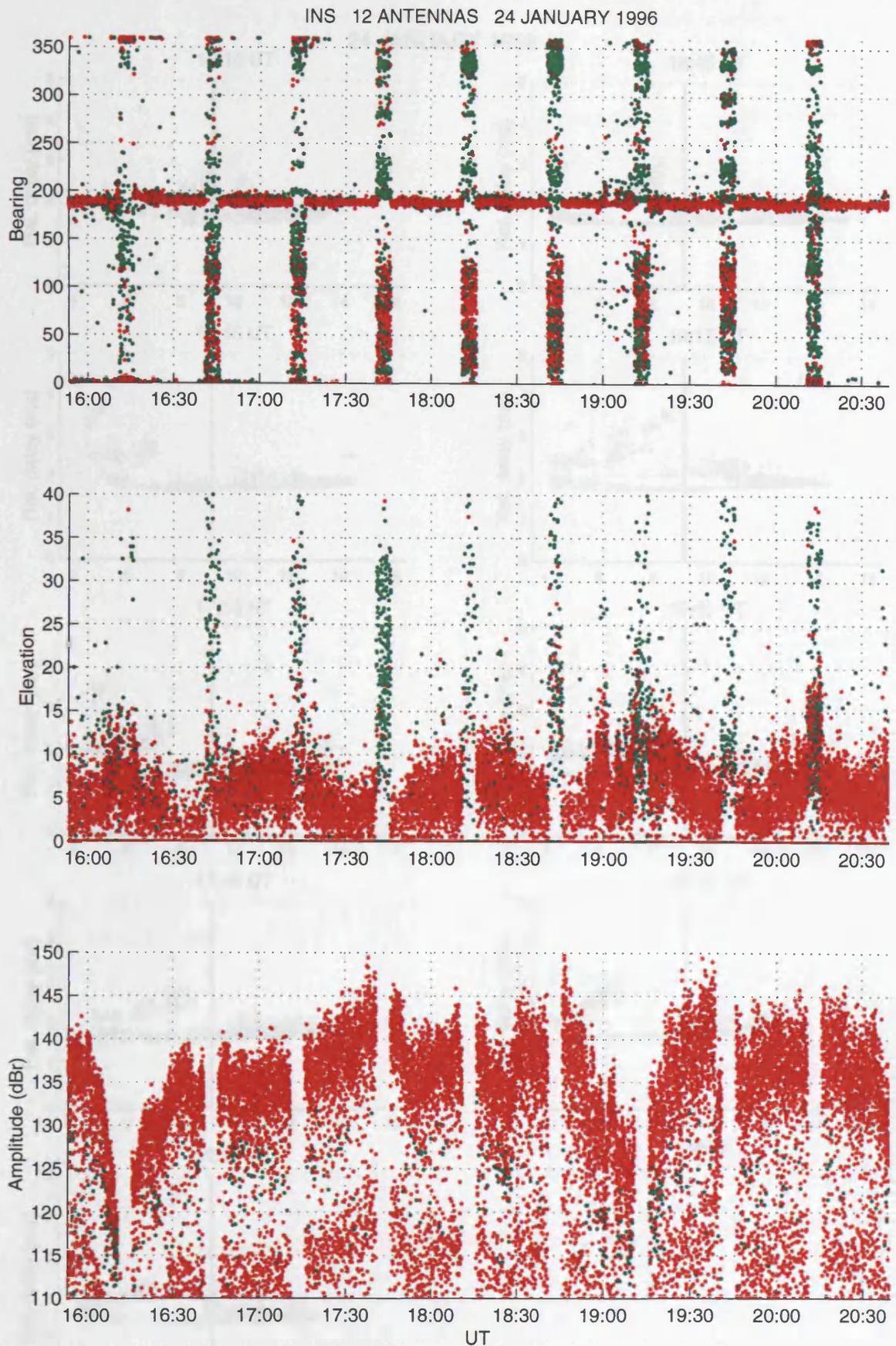


Fig. 3.5.3 Iqaluit-Alert azimuth, elevation and relative amplitude vs. time plots for the 24th January 1996 (15:53 - 20:40 UT).

Oblique ionograms for the Iqaluit-Alert path (24th January 1996)

24 JANUARY 1996

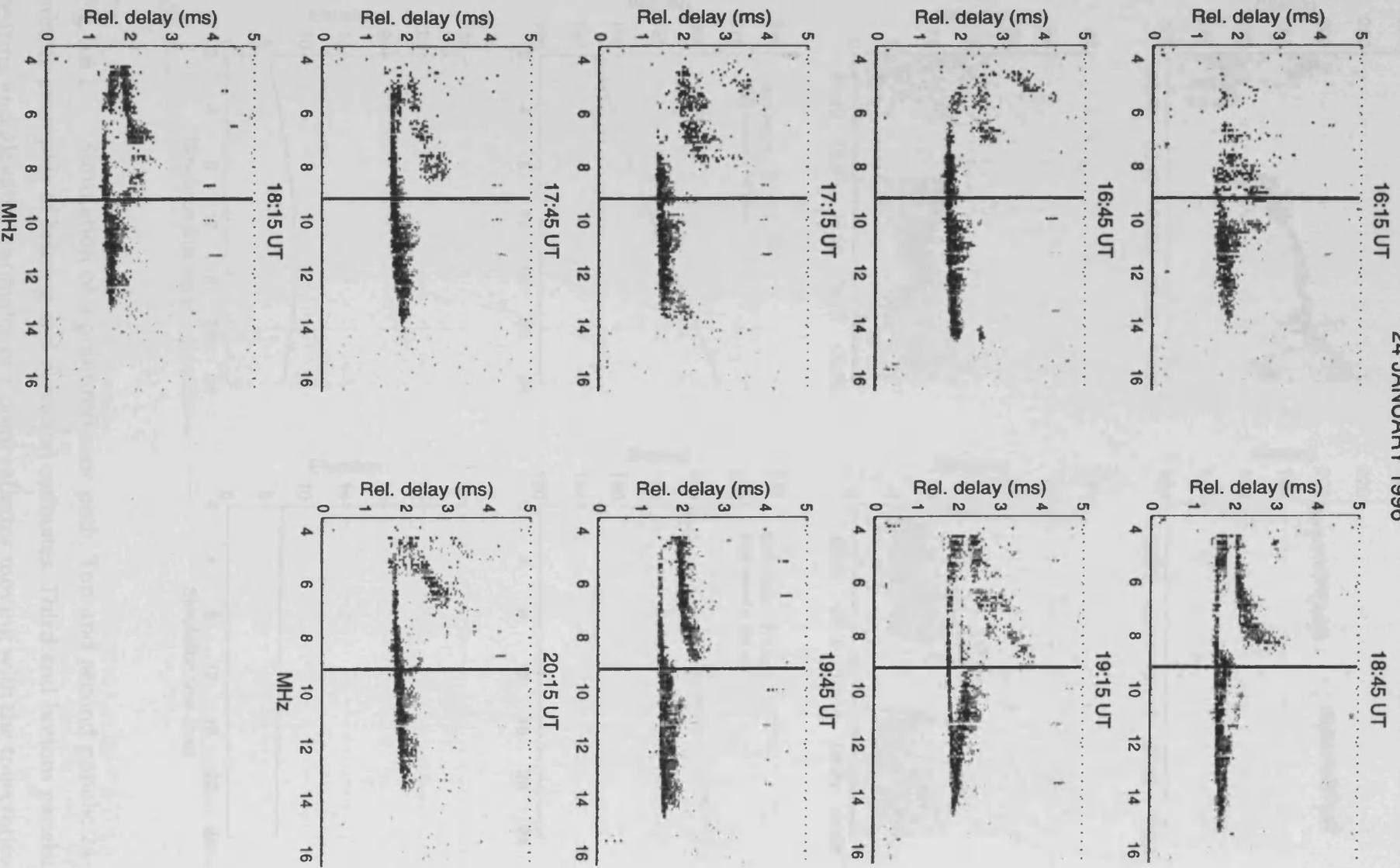


Fig. 3.5.4 Oblique ionograms for the Iqaluit-Alert path (24th January 1996).

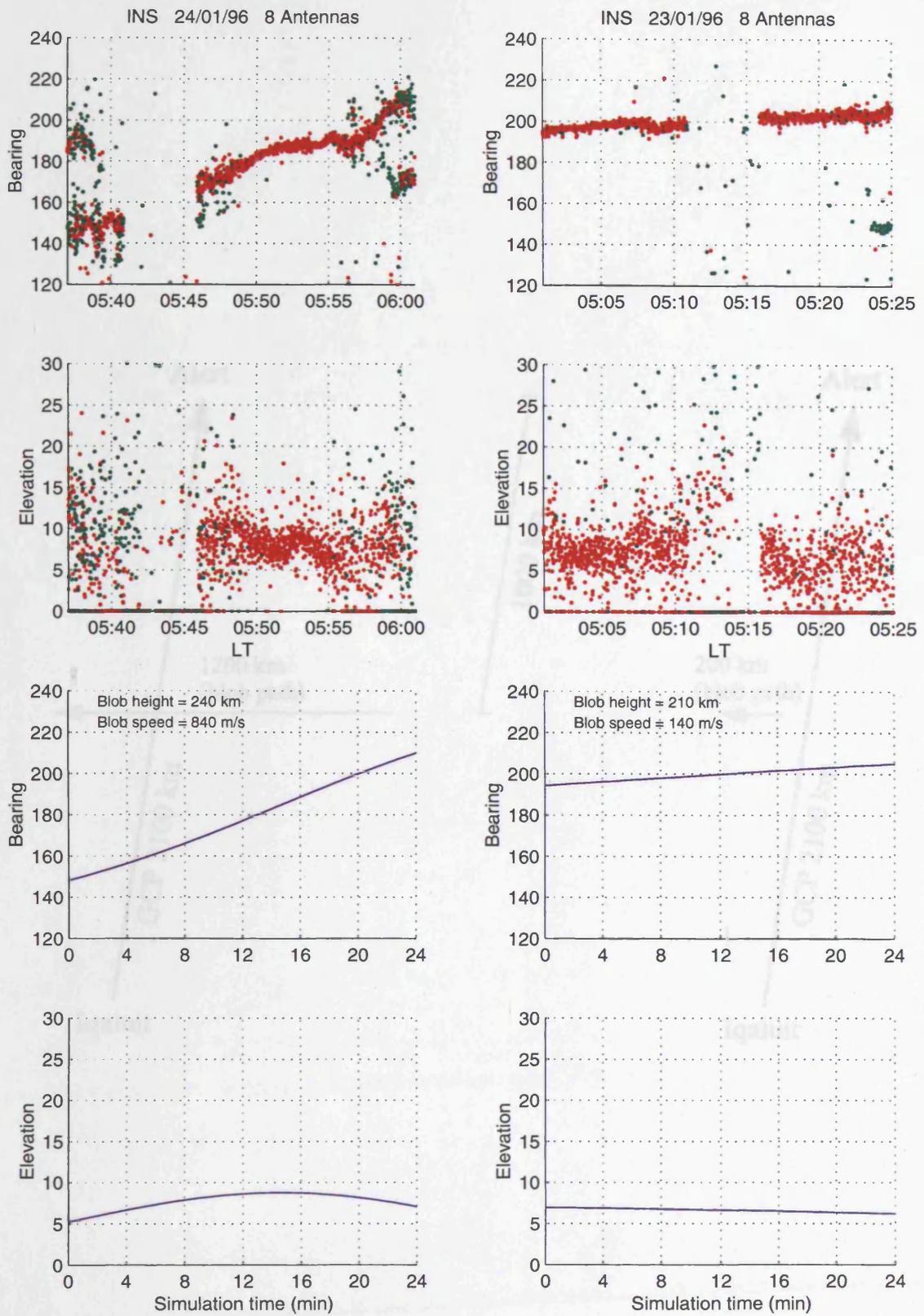


Fig. 3.6.1 Simulation of a point reflector path. Top and second panels: 24-minute examples of bearing and elevation estimates. Third and bottom panels: bearing and elevation estimates of a point reflector moving with the trajectories shown in Figure 3.6.2.

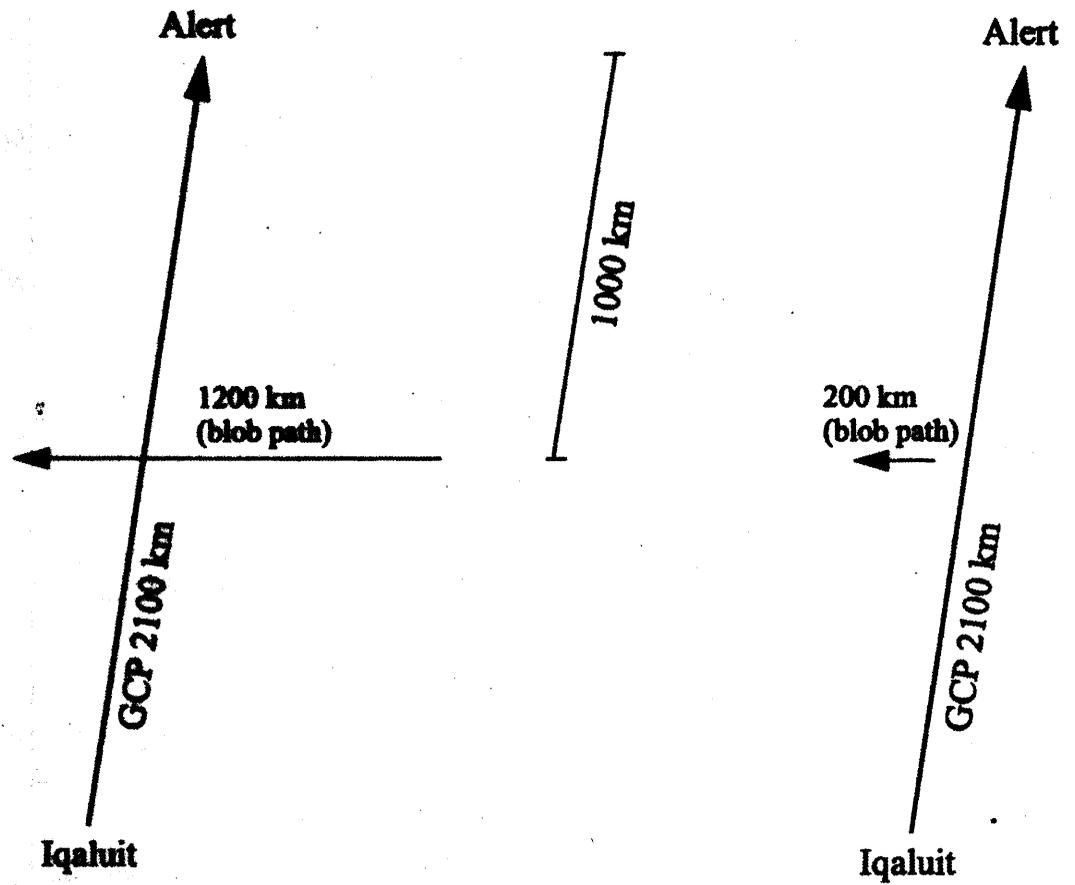


Fig. 3.6.2 Trajectories of two moving point reflectors (1200 km and 200 km paths), corresponding to the simulations in Figure 3.6.1 (third and bottom left panels for the left hand trajectory, and third and bottom right panels for the right hand trajectory).

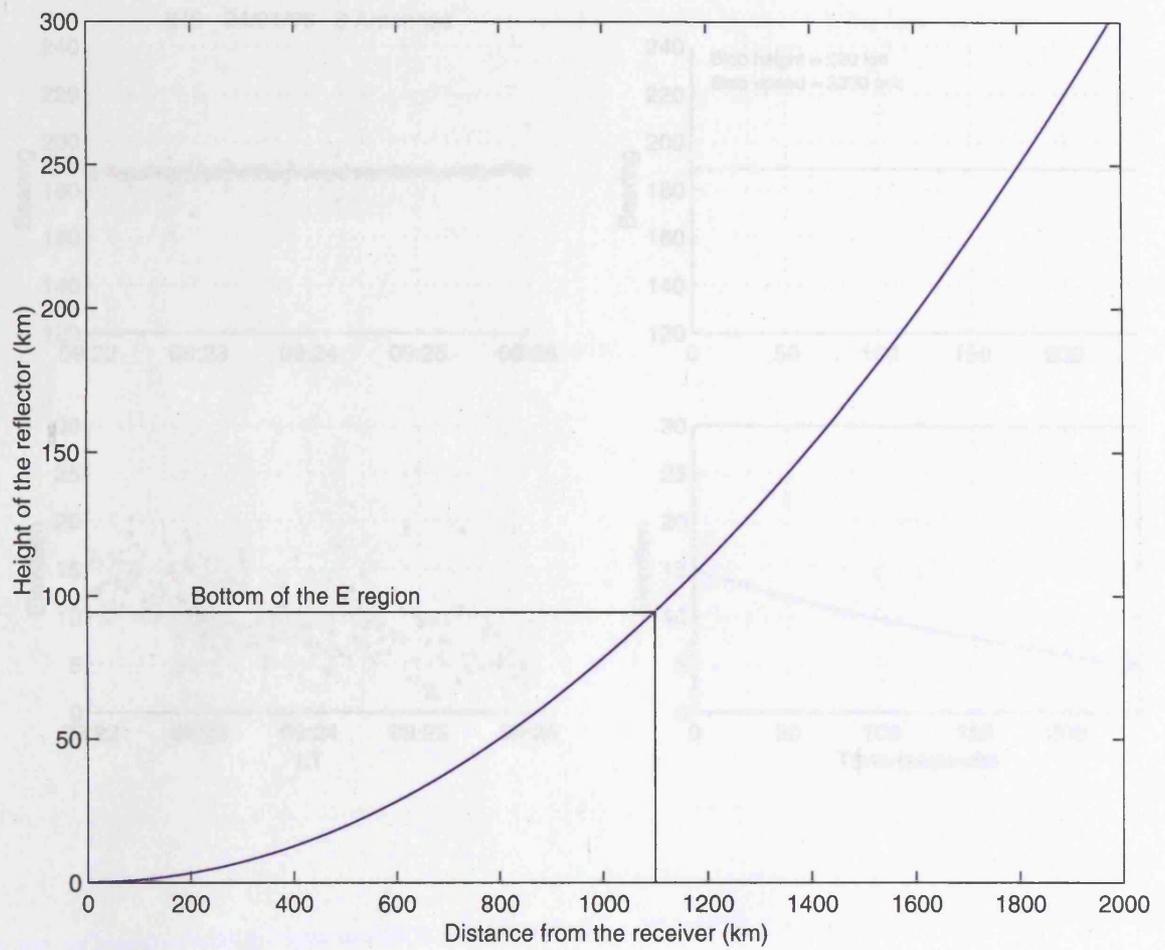


Fig. 3.6.4 Simulation of a point reflector path. Left hand panel shows example of bearing and elevation observations, with a 1000 km range.

Fig. 3.6.3 Minimum height at which a point reflector can be seen from a receiver for distances up to 2000 km.

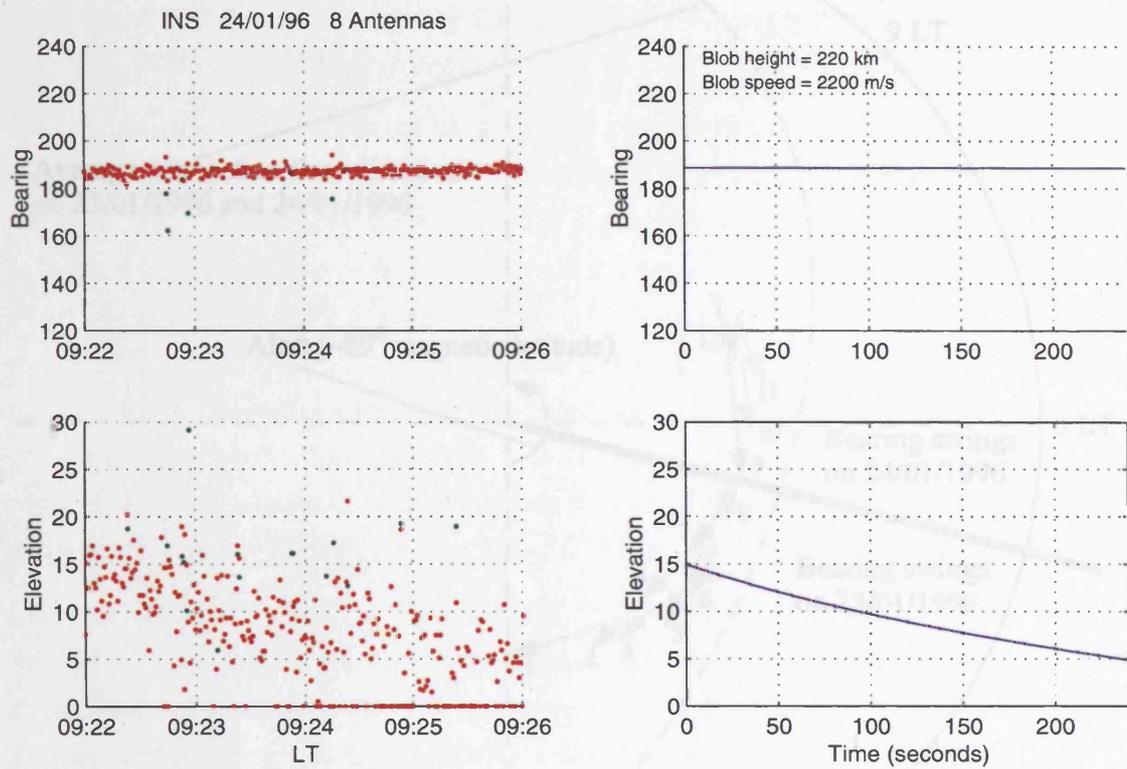


Fig. 3.6.4 Simulation of a point reflector path. Left hand panels: 4-minute example of bearing and elevation observations, with a strong elevation swing. Right hand panels: bearing and elevation estimates of a point reflector moving drifting along the GCP in the Alert-to-Iqaluit direction.

4. Behaviour of the flow

4.1. Introduction

In the previous chapter, we have seen that the flow is generally directed towards the south-east. This is due to the fact that the wind is generally from the north-west. The flow is also affected by the Coriolis force, which causes it to curve to the right of its direction of motion.

The average convection flow direction on 23/01/1996 and 24/01/1996 is shown in the diagram. The flow is generally directed towards the south-east.

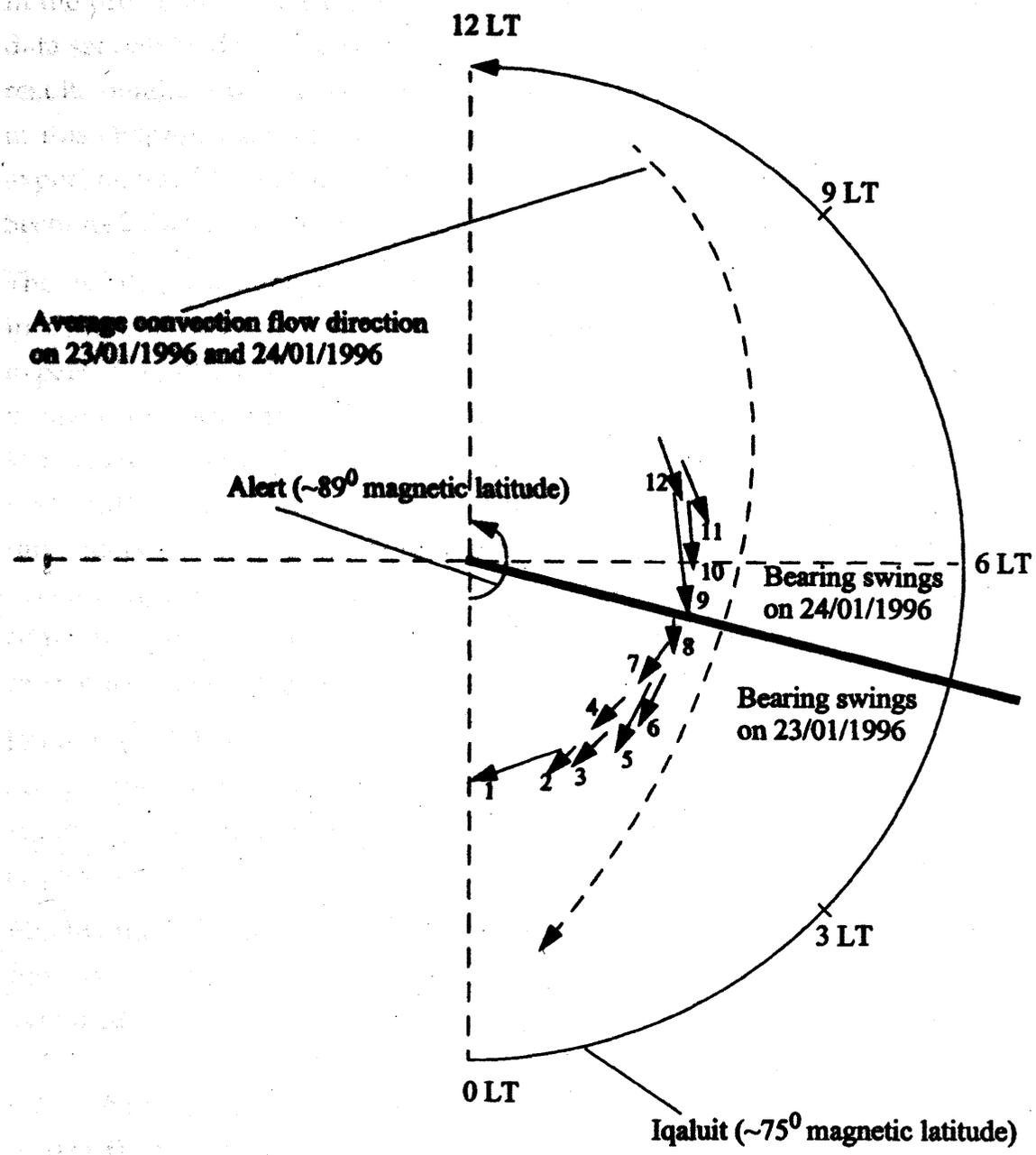


Fig. 3.6.5 Simulation of 12 point reflectors paths (straight arrows numbered 1 to 8 for the 23rd January 1996 and 9 to 12 for the 24th January 1996). The length of the arrows is proportional to the speed of the reflectors (see Table 3.6.1).

4. Behaviour of the INS, Loaded Capon, MUSIC and SML algorithms

4.1 Introduction.

In the previous chapter we have seen the results of the processing of the whole data set obtained by means of the Iterative Null Steering algorithm. Some of the results obtained with Loaded Capon, INS, MUSIC and Beam will be compared in this chapter. Examples of bearing measurements obtained by means of the experimental Spread Maximum Likelihood (SML) algorithm (Read, 1999) (see Sections 2.7 and 2.8) are also reported.

The results show dependence both on the algorithm and on the array geometry, in different ways according to the various propagation conditions. This is expected, as others have compared DF algorithms, underlining the differences in terms of performance (see Sections 2.8 and 2.9 for a review). In particular, Warrington *et al* (1997c) and Dumas (1997) have compared results obtained with different algorithms for the same propagation path and for some of the time intervals investigated here. Warrington *et al* investigated the behaviour of several algorithms in different propagation conditions, whereas Dumas compared performances of algorithms in terms of sensitivity with respect to a secondary weaker signal.

Jenkins (1997) undertook a comparison between the performances of four arrays, (in terms of minimum angular separation at which a weaker point source signal would be detected in presence of a stronger spread-source signal, see Section 2.9), including the Vortex antenna array (see Section 3.2).

Finally, the behaviour of the standard deviation in bearing is investigated, as a function of integration time and the time over which bearing estimates are averaged.

4.2 Algorithm and array geometry dependence of the measured azimuth of arrival.

Bearing estimates have shown dependence on both the algorithm and the array geometry in use throughout the whole data set. Bearing observations during three 1-hour intervals and the corresponding ionograms, obtained with Loaded Capon and all 12 antennas in use, are illustrated in Figure 4.2.1 as an introduction to this investigation. Although interesting features have been seen throughout the whole three-day data collection, these three 1-hour intervals have been selected because they are representative of three different propagation conditions, occurring during the same day, i.e. on the 24th January

1996 (results presented by Warrington *et al*, 1997c). In particular the Maximum Usable Frequency (MUF) changed dramatically during the day.

During the first period, from 09:45 to 10:45, the signal showed several swings of bearing, varying between about 160° to 235° (top right panel). In particular, we can see a very rapid bearing deviation of the primary trace from around 170° at 09:45 to around 210° at 10:00, possibly extending on to about 235° at 10:25. These types of bearing swing are often attributed to reflection from patches or arcs of enhanced ionospheric electron density convecting over the polar cap (see Sections 1.5 and 3.6). It is also interesting to note that at this time the signal frequency was close to the MUF and that the ionogram trace is very spread, this indicating the presence of irregularities which will affect the propagation. Some two and a half hours later, from 12:15 to 13:15, the MUF increased to around 11.5 MHz and the ionogram trace is much less spread than in the earlier example (middle left panel). At this time, a swing in the measured bearing is still evident but not as pronounced as during the earlier period. In the final period, between 18:45 and 19:45, the ionogram traces are much less spread than in the previous examples and the main features are readily identifiable (bottom left panel), with a 2-hop F and both low and high angle 1-hop F propagation. The direction of arrival of the signal was fairly constant and relatively close to the Great Circle direction of 188.5° (bottom right panel).

The whole data set has been analysed with Loaded Capon, INS (6dB termination threshold, see Section 2.7), MUSIC and Beam, and with both the full 12-antenna Vortex array and a sub-array consisting of the 8 inner antennas in circle of Figure 3.2.1b, numbered 1 to 8, to give a second beam pattern. The gross structure of the results was consistent for all four algorithms and both array geometries. However, three 24-minute intervals of data around the times of Figure 4.2.1 are shown in Figures 4.2.2 to 4.2.4 in order to compare the detailed structure of the bearing measurements for each algorithm and array geometry. Examples of bearing measurements with SML are also shown later in this section, and merits and drawbacks of this algorithm discussed.

A preliminary inspection of Figures 4.2.2 to 4.2.4 shows that with both 12 antennas (left hand columns) and with 8 antennas (right hand columns), INS generally exhibited the narrowest traces, followed by Loaded Capon. MUSIC generally showed very spread traces, with primary bearings often intermixed with secondary bearings. Beam was often capable of detecting a relatively narrow main trace, but also usually showed evidence of secondary traces well detached from the primary one (this is particularly evident in Figure 4.2.4). This is not surprising, because Beam is not a superresolution algorithm, and as such it cannot resolve two signals separated by less than the natural beamwidth of

the antenna array. Therefore, in order for the Beam algorithm to detect two peaks of energy, these must be separated by more than the beamwidth. Otherwise the spread energy arriving at the receiver from one signal may be split in well-separated traces. These traces will then not represent two different signals, but portions of energy from the same signal (confirmed by the break when the transmitter was off), and might arise from the presence of sidelobes (see Figures 3.2.3 to 3.2.5). All the above features were often seen in the whole data set. It is then important to point out that while the gross structure of the results was consistent for all algorithms, differences arose in the detailed structure.

Multiple traces which appear to closely track one another were often seen with SRDF algorithms (this is particularly apparent in Figures 4.2.2 and 4.2.3, to a greater or lesser extent depending on the algorithm). Since SRDF algorithms, unlike Beam, can resolve two signals separated by less than the natural beamwidth of the antenna array, these closely separated traces must be different in origin to the well separated ones exhibited by Beam. By comparing the results for SRDF algorithms with 12 and 8 antennas, whether these traces represent real signals or arise as an artefact of the algorithm in use can be determined (wider beam patterns result from smaller apertures and vice versa, see Section 3.2). If the multiple traces represent real signals, their structure would not be substantially affected (provided that the signals were resolvable with both array geometries), whereas considerable changes in the structure of the traces are expected if multiple traces result as artefacts of the algorithm.

Generally there are more multiple traces when the full 12-antenna array is in use. For example, in Figure 4.2.3 Loaded Capon and MUSIC produced three traces with the full array, but only two when the sub-array was in use. INS showed two traces with 12 antennas and only one with 8 antennas. This difference in the number of closely separated traces with different arrays indicates that these traces are likely to be an artefact of the algorithm, representing portions of the energy of the same signal. In Figures 4.2.2 and 4.2.3 more closely separated traces are generally present when the 12-antenna full array is used rather than with the 8-antenna sub-array. To interpret this behaviour, we have to consider that the secondary bearings found by the algorithms could represent part of the primary signal when signal is relatively spread. The narrower the beam, the more likely this situation will occur. Since the beamwidth is inversely proportional to the array aperture, this explanation is consistent with the closely separated traces observed with the 12-antenna array, but not the 8-antenna array. Closely separated multiple traces have been reproduced by means of modelling studies in which spread signals are

assumed to arrive at the receiver (see Chapters 5 and 6). It is also noteworthy that, in general, the traces appear to be more spread with the full 12-antenna array. However, the differences between the 8- and 12-antenna array are less marked during times with smaller bearing deviations or narrower signals. For example, the differences are less apparent in Figure 4.2.3 (slow changes in azimuth) than in Figure 4.2.2 (very rapid changes in azimuth), and even less in Figure 4.2.4 (fairly constant bearing).

In addition to the four algorithms discussed above, measurements have also been undertaken by means of the Spread Maximum Likelihood (SML) algorithm (Read, 1999) (see Sections 2.7 and 2.8), which assumes that the signal is spread in direction. SML has been used here in order to test its capabilities, since it was available when this research was being carried out (unlike SPIRE, see Section 2.8). However, it is important to bear in mind that SML was still under development and somewhat limited, e.g. it cannot deal with spreads of greater than 15° (Read, 1999).

The performance of SML has been determined by analysing two 20-minute examples of experimental data (Figure 4.2.5). The upper panels (23rd January 1996, between 08:20 and 08:40 UT) refer to a time interval in which the other SRDF algorithms exhibited a spread incoming signal (about $20\text{-}50^\circ$ azimuth, depending on the specific algorithm, see Figure 3.4.3 for INS, observations with Loaded Capon and MUSIC are not shown in the thesis). Loaded Capon, INS and MUSIC exhibit a narrow primary trace which increases from about $150\text{-}160^\circ$ to $180\text{-}190^\circ$, over this 20-minute time interval. Instead, SML shows primary bearings arriving from about 225° in the first 2-3 minutes, and from $350\text{-}360^\circ$ in the remaining portion of the time interval, while multiple secondary traces between $80\text{-}180^\circ$ occur during the whole interval. This happens with both 12 and 8 antennas. This clear difference between SML and the other SRDF algorithms is not surprising, since all the others consistently exhibit a trace more spread than 15° , this consistency suggesting that the signal was more spread than 15° , and therefore SML was not reliable.

The lower panels in Figure 4.2.5 (24th January 1996, between 16:20 and 16:40 UT) refer to a time interval in which the other SRDF algorithms exhibited a narrow incoming signal ($\sim 10^\circ$ depending on the algorithm, see Figure 3.5.3 for INS, observations with Loaded Capon and MUSIC are not shown in the thesis), and therefore sufficiently narrow to expect SML work correctly. In common with the other algorithms, SML does show a very narrow primary bearings trace at about 190° with both 12 and 8 antennas. However, it also exhibits a number of narrow secondary traces, at clearly different directions for the 12-antenna and 8-antenna arrays, which are not found with the other algorithms. Therefore, this

is likely to be an artefact of the algorithm (unlike the other SRDF algorithms, SML showed narrow secondary traces also in the upper panels).

The above results make it clear that SML needs further development, especially for when the DOAs are highly spread. For this reason SML has not been utilised elsewhere in this research.

It is very important to distinguish between multiple traces which represent multiple signals and those which are artefacts of the algorithm. It is worth recalling here that INS (see Section 2.7) terminates the search for secondary bearings when the remaining energy is less than a previously set threshold below the first detected peak (6dB unless otherwise stated). In order to make INS more sensitive to secondary bearings (and/or noise) a threshold of 15dB has been used for the results illustrated in Figure 4.2.6. In the two plots to the left the bearing separation between the primary and secondary trace changes with the number of antennas in use, which suggests that, in this case, the double trace is an artefact of the algorithm. In the right hand panels the bearing separation between the primary (200°) and the secondary trace (150°) does not change with the number of antennas in use, indicating that these traces represent two separate signals.

It is worth comparing the top right hand plot of Figure 4.2.6 to the middle panel of Figure 2.8.5, showing results presented by Dumas (1997) in the same time interval with the INS algorithm (note that Dumas did not specify the termination threshold employed). The results in the top right hand plot of Figure 4.2.6 showed evidence of the interfering signal at about 150° from just after 09:20 to the end of the interval, at 09:35, whereas the corresponding trace of Dumas' plot shows a clear gap between about 09:28 and 09:32. However, it is important to bear in mind that the 15dB threshold used for the results in Figure 4.2.6 increases the sensitivity of the INS algorithm to secondary energy. This important aspect was not emphasised by Dumas when he compared the performance of INS with other algorithms, in terms of sensitivity to the secondary trace at about 150° .

Contour plots of the bearing estimates obtained with 12 antennas vs. those obtained with the 8-antenna sub-array are shown in Figure 4.2.7 for Loaded Capon, INS, MUSIC and Beam. They refer to the period 15:53-20:40 UT on 24th January 1996, an interval for which large bearing swings were not evident and the signal from Iqaluit was the only one present during the whole time interval (see Figure 3.5.3). It is important to consider that only the primary bearing estimates have been included. Of particular note in the four panels is the offset of the distribution from the 'equal result line', the bearings obtained with the 12-antenna array being, on average, different to those obtained with the 8-

antenna array (results presented by Warrington *et al*, 1997c). This offset appears to depend on both array geometry and algorithm. Loaded Capon and MUSIC produced bearing estimates around 1° higher with the 12-antenna array, whereas INS and Beam produced similar offsets, but in the opposite direction. The differences between the bearing estimates obtained with 8 or 12 antennas (i.e. different apertures), and with different algorithms could arise from the diffuse nature of the signals. It is also clear that the spread of the traces was algorithm dependent during this period. The least amount of scatter occurred with the Beam algorithm, the standard deviation of the bearing estimates being equal to 2.2° in the case of full array, and to 2.1° when the 8-antenna sub-array was used. This is not unexpected, since the primary estimates only were taken into account. Second best, from this point of view, was INS (standard deviation equals 2.4° and 2.2° with 12 and 8 antennas respectively). Loaded Capon showed higher values of standard deviation (4° with 12 antennas and 3.5° with 8 antennas). The distribution of the points for the MUSIC algorithm was significantly greater, the standard deviation being equal to 8.4° with both array geometries.

In this section we have seen that, with the exception of SML, each direction finding algorithm used here showed similar gross behaviour, but differences in the detailed structure of the results. Furthermore, the azimuth has frequently exhibited multiple traces. It is of great importance to determine whether these arise as a consequence of multimoded propagation, interfering transmissions, or as an artefact of the algorithm. Modelling studies of one and two distributed signals are presented in Chapters 5 and 6 respectively. These studies are able to reproduce some experimental observations, including multiple traces arising as artefacts of the algorithm (a clear example is illustrated in Figure 6.3.5).

4.3 The effect of integration time on the observed bearing spread.

An important parameter for DOA estimates is the integration time (see, for example, Warrington, 1995), the choice of this parameter having to meet the needs of the DF operator. In this section several data intervals are presented, in which the integration time (time/snapshot) has been varied, and the bearing estimates have been averaged over a range of time values.

Several intervals of about 20-minutes for each day have been investigated. The intervals have been chosen such that the azimuth spreads are different, a low spread of up to about $15\text{-}20^\circ$ and a high spread of greater than this. Within each time interval, the portions of data collected when the transmitter was off (see Section 3.2) have been excluded. A range of integration times from 0.05 to 10 sec

has been used. It is important to consider that integration times higher than 10 sec are unlikely to be of use in an operational environment at high latitudes, where rapid bearing deviations often occur. In fact, the data set utilised in this investigation has presented several occurrences of very rapid azimuth and elevation changes. For example, both the azimuth swing between 05:35 and 06:00 LT on 24th January 1996 (see Table 3.6.1) and the elevation swing between 09:22 and 09:26 LT on the same day (see Figure 3.6.4) have exhibited angular deviations comprised between 0.4° and 0.5° each 10 seconds on average.

For each integration time value, the primary bearing estimates have then been averaged over time intervals greater than or equal to the integration time up to 10 sec, i.e. for an integration time of 0.05 sec, the bearing estimates have been averaged over 0.05, 0.1, 0.25, ..., 10 sec, and so on. The bearing estimates have been detrended in order to prevent azimuth swings affecting the standard deviations. Furthermore, only primary bearings within 30° either side of the mean azimuth of the signal have been included in the calculations of the standard deviations.

As an example, one minute of bearing measurements is illustrated in Figure 4.3.1, in which the integration time is varied between 0.05-10 sec (top to bottom), using Loaded Capon, INS, MUSIC and Beam with the full 12-antenna array. For all integration time values, the traces for INS are the narrowest and those obtained with MUSIC are the most spread. This behaviour was generally seen in all investigated time intervals. The results are summarised in Figures 4.3.2 and 4.3.3. Figure 4.3.2 shows two examples of bearing standard deviations against average time for an integration time of 1 sec and with different spreads (higher in the top plots and lower in the bottom ones). The important aspect of the results in Figure 4.3.2 is that, if the time/snapshot is kept constant, the standard deviation decreases as the average time increases, and this is true for all the situations investigated. This is a very important result, as a low standard deviation is a desirable feature (assuming that the bearing is fixed); however it has to be reconciled with the need of the bearing estimate to be obtained in the shortest possible time. For example, a time/average of 10 sec could be too high for DF operators who need real-time estimates. In the higher spread interval, with 12 antennas the bearing standard deviations are lower for Loaded Capon than for the other algorithms, whereas they are about similar for all algorithms with 8 antennas (the curves crossing one another), thus Loaded Capon performs better with the full array. In the lower spread interval the higher values are shown consistently by MUSIC, followed by INS, then Loaded Capon and Beam, with both array geometries. Overall, among the SRDF algorithms Loaded Capon performed better than the others in both the higher and lower

spread intervals. With the exception of Loaded Capon in the higher azimuth spread interval, the algorithms did not exhibit substantial differences between the two array geometries.

It is also interesting to investigate the behaviour of the algorithms when, keeping constant the time/average, we vary the integration time. Two examples with azimuth spreads higher (top panels) and lower (bottom panels) than $15\text{-}20^\circ$ are shown in Figure 4.3.3, for time/average of 5 sec. In the higher spread interval, Loaded Capon with 12 antennas performs better than the other algorithms. As in Figure 4.3.2 for a fixed integration time, also for a fixed time/average the algorithms did not show substantial differences between 12 and 8 antennas in terms of bearing standard deviations. These did not show any particular trend for all algorithms for an integration time of up to about 1-2 sec, after which they tended to diverge, increasing, decreasing or remaining constant dependent on the algorithm, array geometry and spread. Also in other intervals, not shown in the thesis, the algorithms showed analogous results to this example for integration times below 2 sec. However, the results could be substantially different for higher integration times. A reason for this may be that the higher the integration time, the fewer the bearings taken into account for the calculation of the standard deviation, and therefore the less reliable the results. An analysis on a much longer data set would be needed for a more comprehensive study. However, from this investigation we can conclude that the bearing estimates are generally lower for Loaded Capon, especially in the higher spread intervals and with the full array in use.

The choice for the optimum integration time value depends on the particular situation in which a DF operator has to deal with. In fact, it is important to take into account that an integration time too high would be unsuitable when rapid deviations in bearing occur (in Chapters 2 to 4 we have seen that this often happens at high latitudes), or when signals are only present for short periods. In such situations the DF finding system would need a suitably short integration time value (even though this would not necessarily imply lower bearing standard deviations, as we have seen). However, for very short integration time values of the order of hundredths of a second, the computational burden increases dramatically. This investigation has shown that the larger the time over which the bearing estimates are averaged, the lower their standard deviation. However, once again, a too high average time could not be appropriate in case of rapid azimuth swings and when a real-time estimate is crucial. A value of 0.8 sec integration time was chosen for the results presented throughout this thesis (unless otherwise stated) as a reasonable compromise between the above mentioned factors.

4.4 Summary.

The whole three-day data set has been analysed with three superresolution algorithms (Loaded Capon, INS and MUSIC) and a simple beam forming algorithm (Beam), and with two antenna array geometries. The gross features in the data are largely independent of algorithm and geometry. However, the detailed results differed especially when the signal was fairly spread and/or showed large bearing deviations. Therefore care must be taken in the interpretation of these results since they can depend substantially on the DF algorithm in use at these times.

The comparison of the results of the processing carried out with the full 12-antenna Vortex array and an 8-antenna sub-array has shown several interesting features. The results obtained with the 8-antenna sub-array generally exhibited less secondary energy, resulting in less spread traces. Multiple traces were often present which could be evidence either of a multisignal environment or multimoded propagation, or artefacts of the algorithms. A comparison of the results obtained with two different array geometries was usually able to distinguish between these mechanisms. In particular, the closely separated traces were likely to arise as artefacts of the DF algorithm, since their angular separation was different when using two different array patterns and apertures.

Examples of experimental data obtained by means of the SML algorithm have also been presented. SML attempts to model the signal environment at high latitudes more realistically than Loaded Capon, INS and MUSIC. However, the poor results obtained with SML confirm the need for its further development suggested by Read (1999).

An investigation on the effect of the integration time and the time over which bearing estimates are averaged on the bearing spread has been undertaken, for intervals of both low and high azimuth spread. In both cases, the results (for all algorithms and array geometries) indicate that the bearing standard deviation generally decreases when, the integration time being fixed, the time/average is increased. In general, the lowest bearing standard deviations were exhibited by Loaded Capon (among the SRDF algorithms). The behaviour of the algorithms did not substantially differ between 12 and 8 antennas, with the exception of Loaded Capon, which showed lower bearing standard deviations in the higher azimuth spread intervals. With a fixed time/average, the bearing standard deviation values were fairly independent of integration time up to about 1-2 sec of integration times.

Integration time values set too high would be unsuitable in the case of the occurrence of rapid bearing deviations, which is a frequent occurrence at high

latitudes. However, low integration time values can dramatically increase the computational burden. A high time/average, although it implies lower standard deviation values, can be unsuitable when real-time DF is a crucial factor for DF operators. The information contained in this investigation can be useful for DF operators in order to make the proper parameter choices according to requirements and propagation conditions.

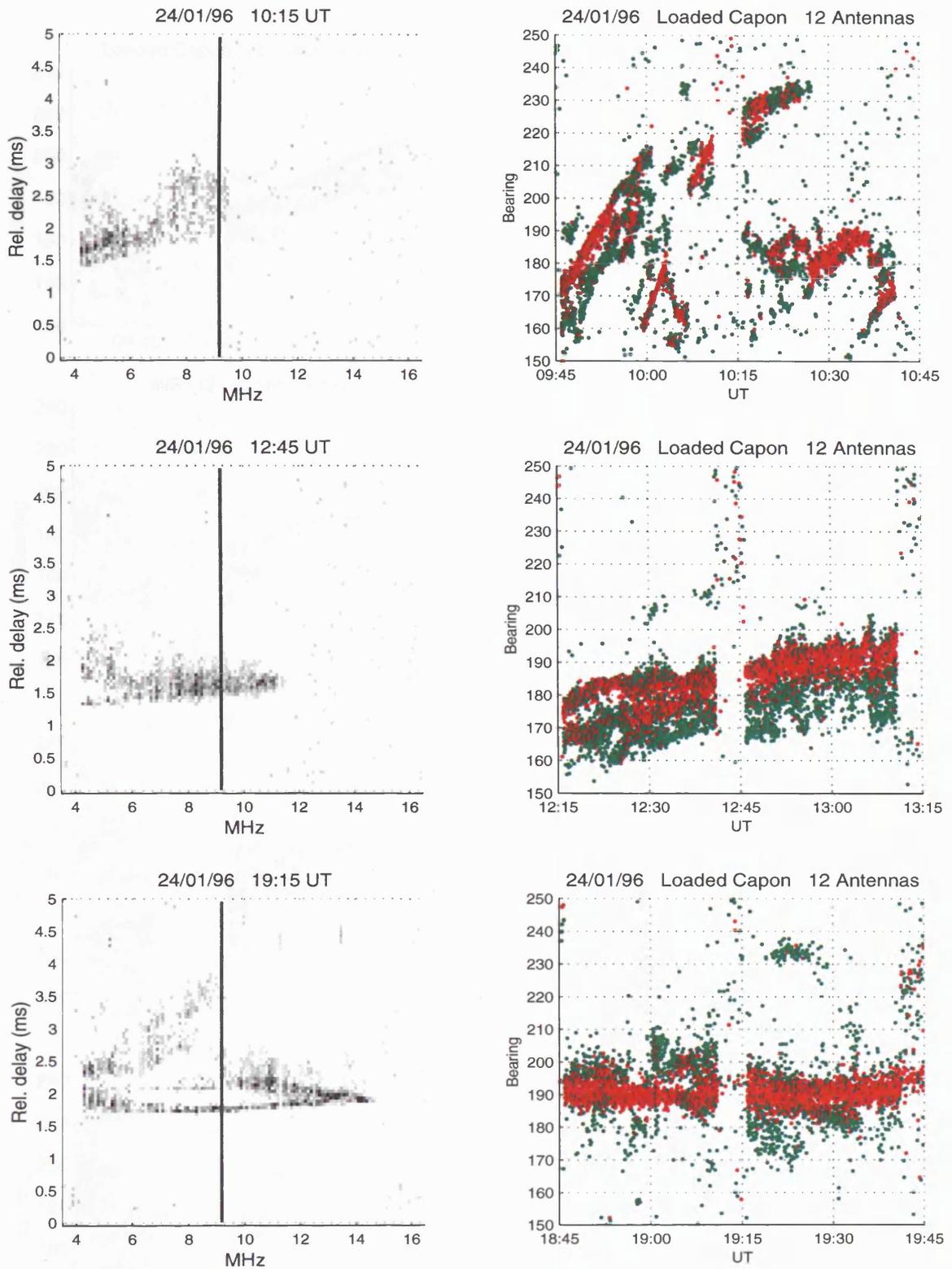


Fig. 4.2.1 The left hand panels show oblique ionograms recorded over the Iqaluit-Alert path at 10:15, 12:45 and 19:15 UT on the 24th January 1996. Bearing estimates obtained using the Loaded Capon algorithm for the period either side of the times of the ionograms are given in the right hand panels (12-antenna array).

24 JANUARY 1996

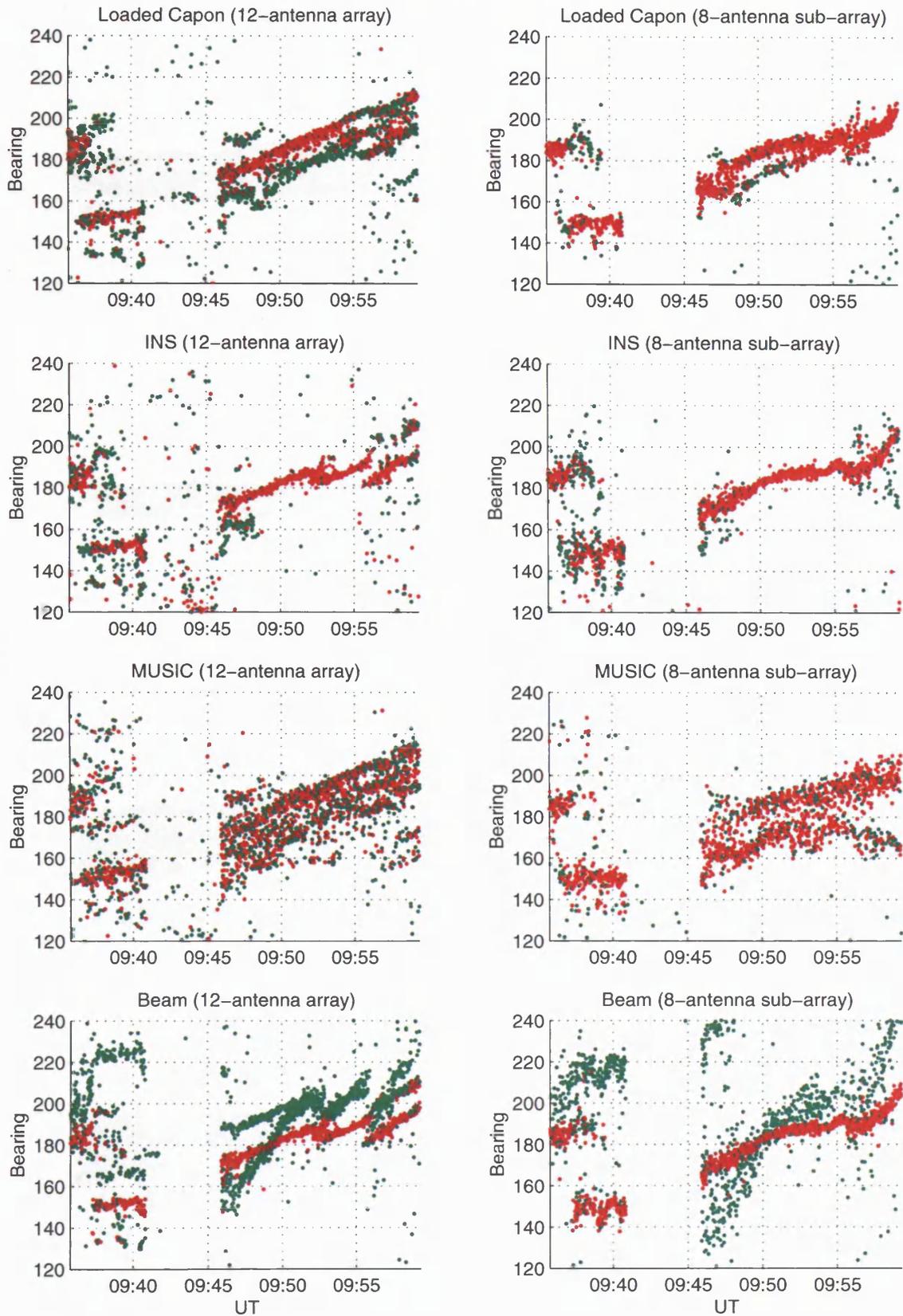


Fig. 4.2.2 Iqaluit-Alert: bearing measurements on the 24th January 1996, 09:36-09:59. Comparison between the algorithms for both the full 12-antenna array (left hand panels) and the 8-antenna sub-array (right hand panels).

24 JANUARY 1996

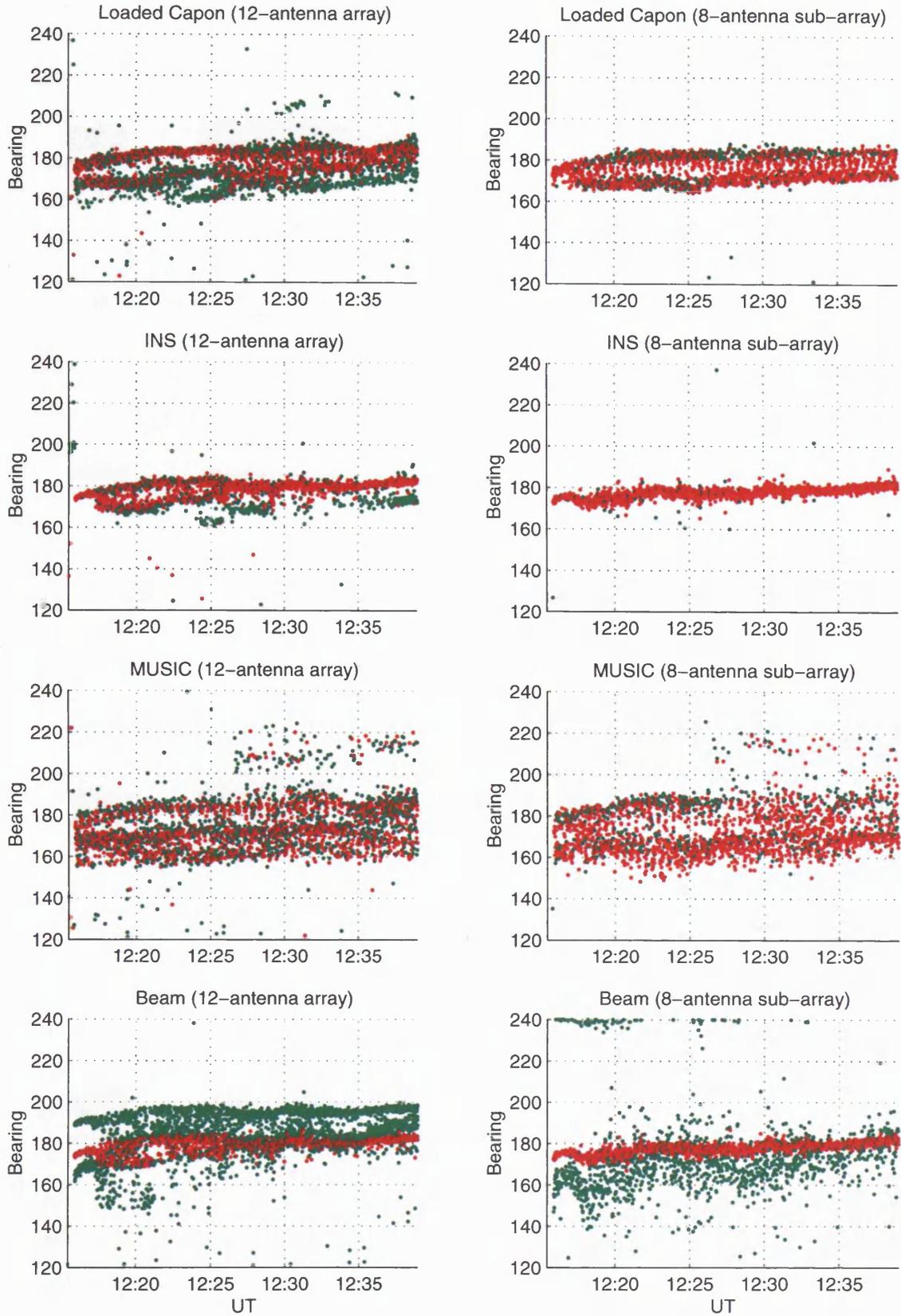


Fig. 4.2.3 Iqaluit-Alert: bearing measurements on the 24th January 1996, 12:16-12:39. Comparison between the algorithms for both the full 12-antenna array (left hand panels) and the 8-antenna sub-array (right hand panels).

24 JANUARY 1996

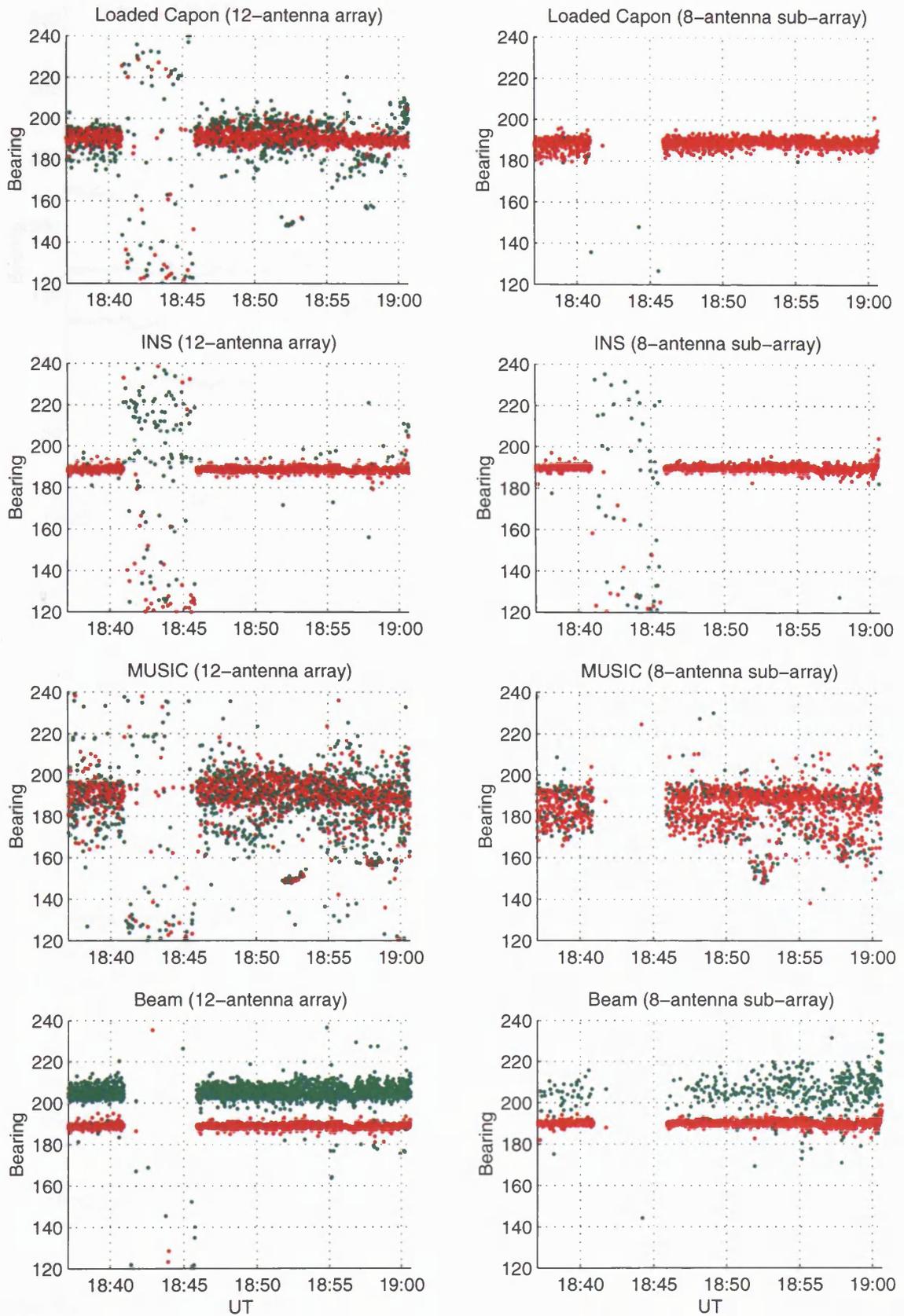


Fig. 4.2.4 Iqaluit-Alert: bearing measurements on the 24th January 1996, 18:37-19:00. Comparison between the algorithms for both the full 12-antenna array (left hand panels) and the 8-antenna sub-array (right hand panels).

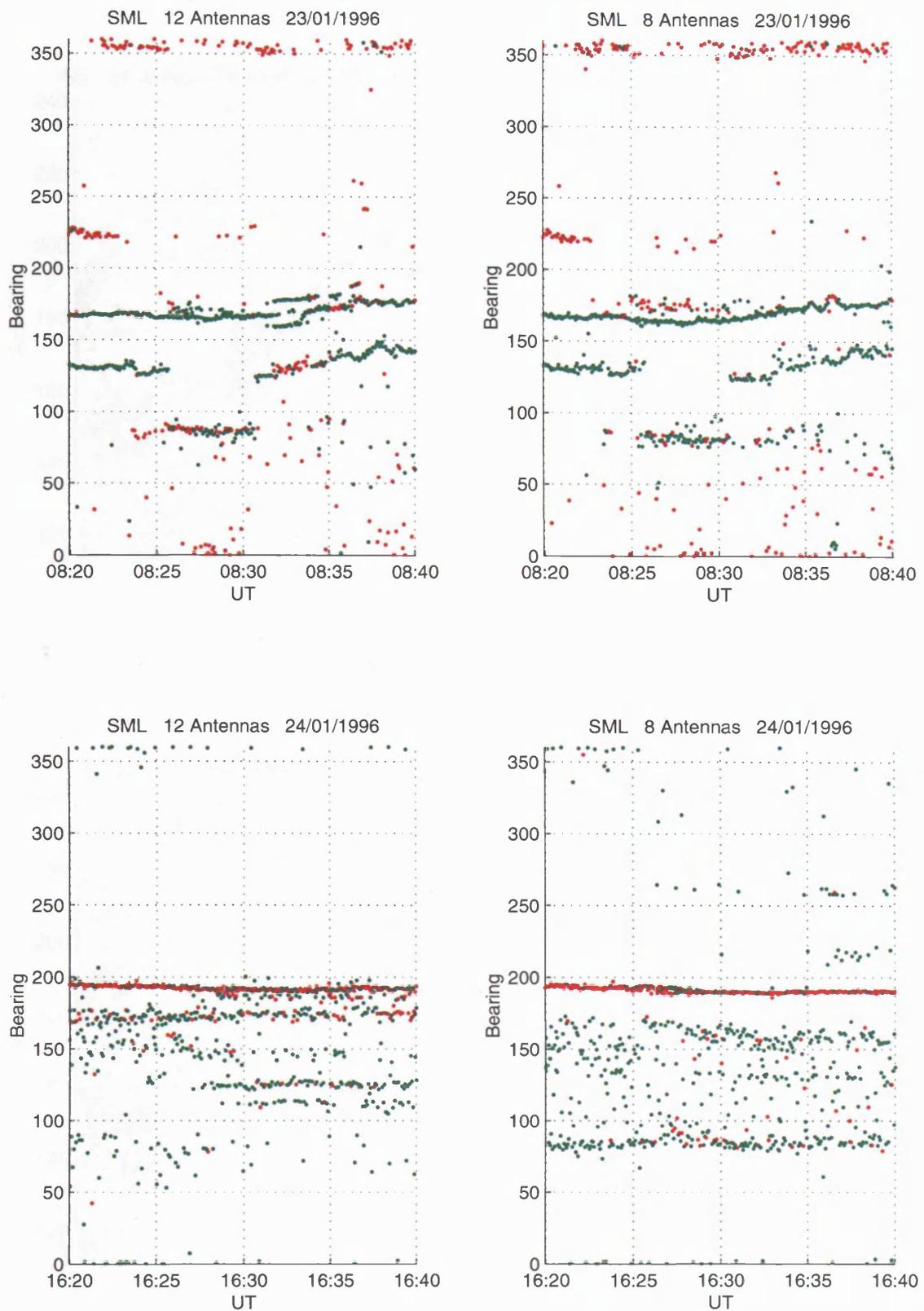


Fig. 4.2.5 Bearing estimates obtained by means of the SML algorithm on 23rd January 1996 (upper panels) and 24th January 1996 (lower panels), with the full 12-antenna array (left hand panels) and the 8-antenna sub-array (right hand panels).

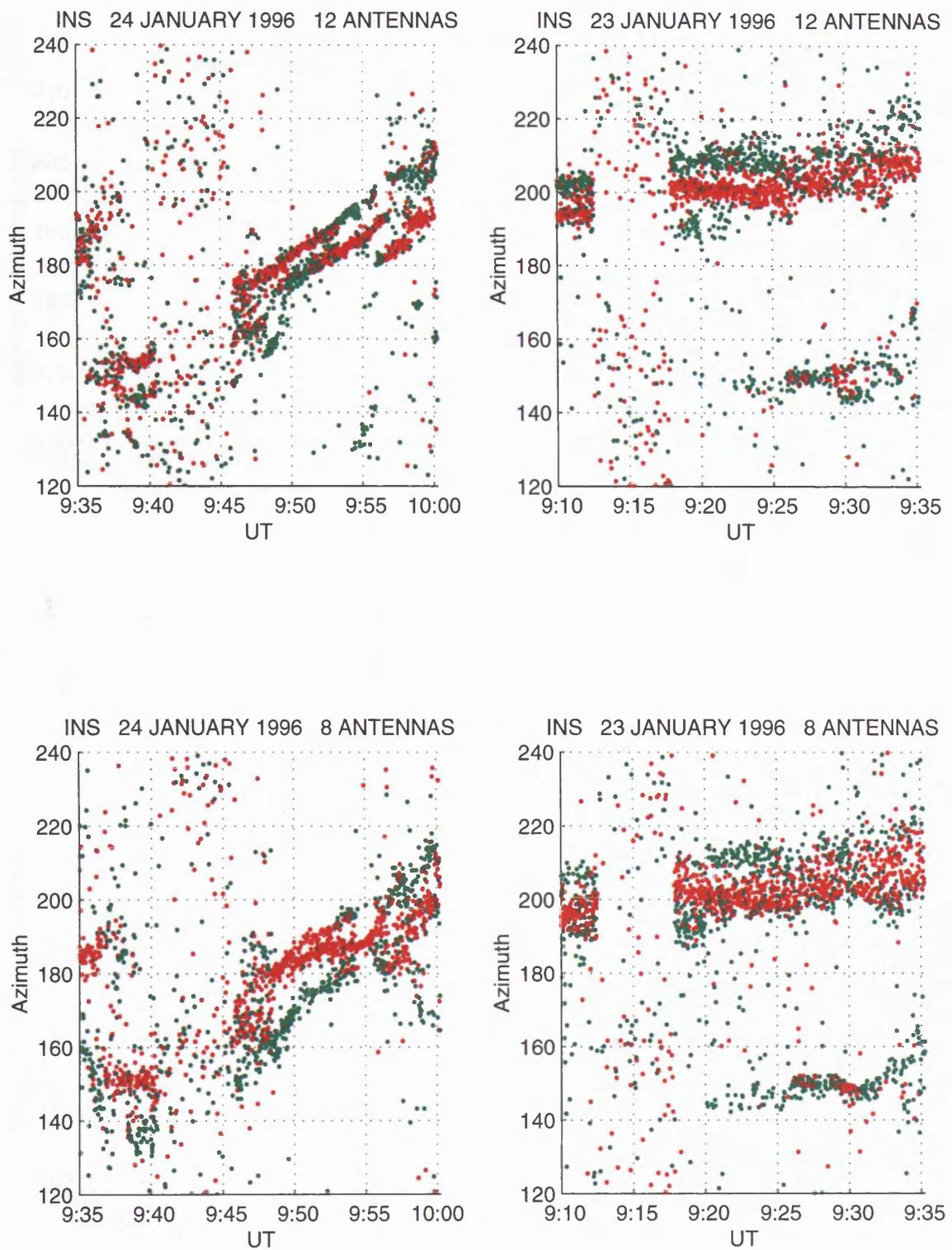


Fig. 4.2.6 Iqaluit-Alert: two examples of bearing measurements obtained with both the full 12-antenna array (top panels) and the 8-antenna sub-array (bottom panels). Algorithm used: INS with 15dB termination threshold.

24 JANUARY 1996 15:53-20:40

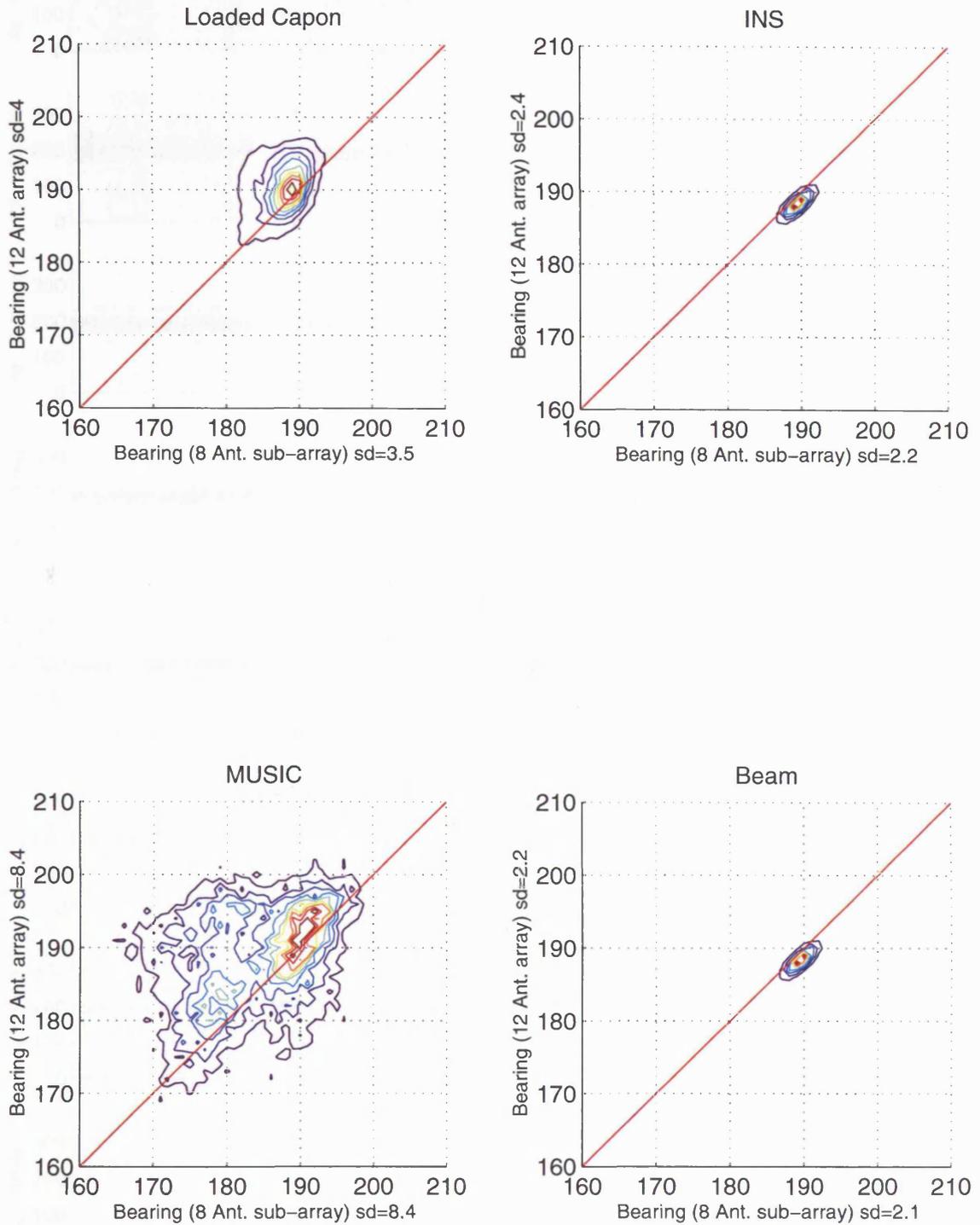


Fig 4.2.7 Iqaluit-Alert: contour plots comparing the bearing estimates obtained with the full 12-antenna array and the 8-antenna sub-array. Only primary estimates are included. Contours are shown at 0.1, 0.2, ..., 0.9 of the peak.

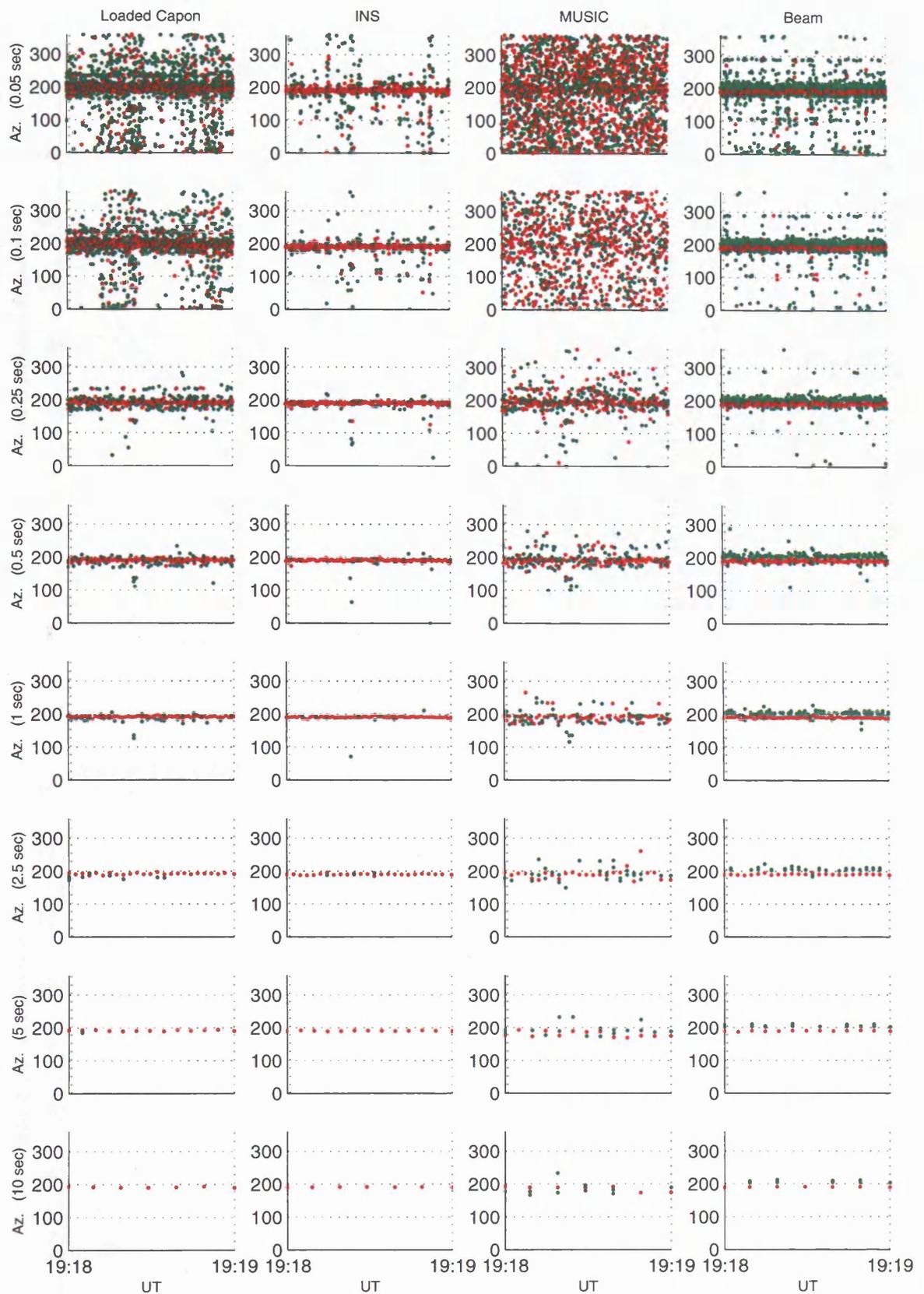


Fig 4.3.1 Bearing measurements taken with integration times between 0.05 sec (top panels) and 10 sec (bottom panels), for Loaded Capon, INS, MUSIC and Beam (left to right) with the full 12-antenna array, on 24th January 1996..

INTEGRATION TIME: 1 SEC

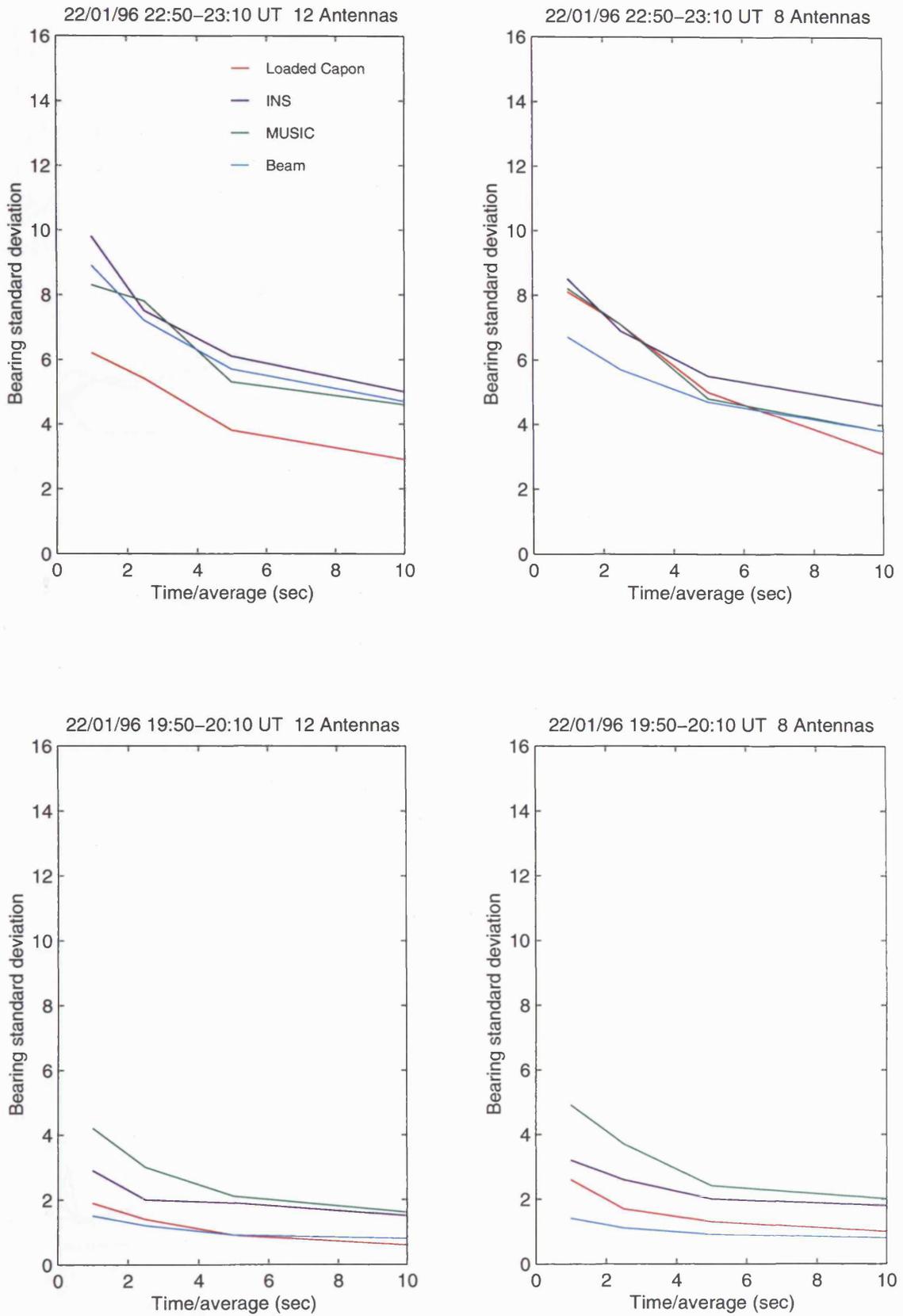


Fig 4.3.2 Standard deviation of experimental bearing estimates as function of the time over which the estimates are averaged, for an integration time of 1 sec (top panels: spread > 15°; bottom panels: spread < 15°).

TIME/AVERAGE: 5 SEC

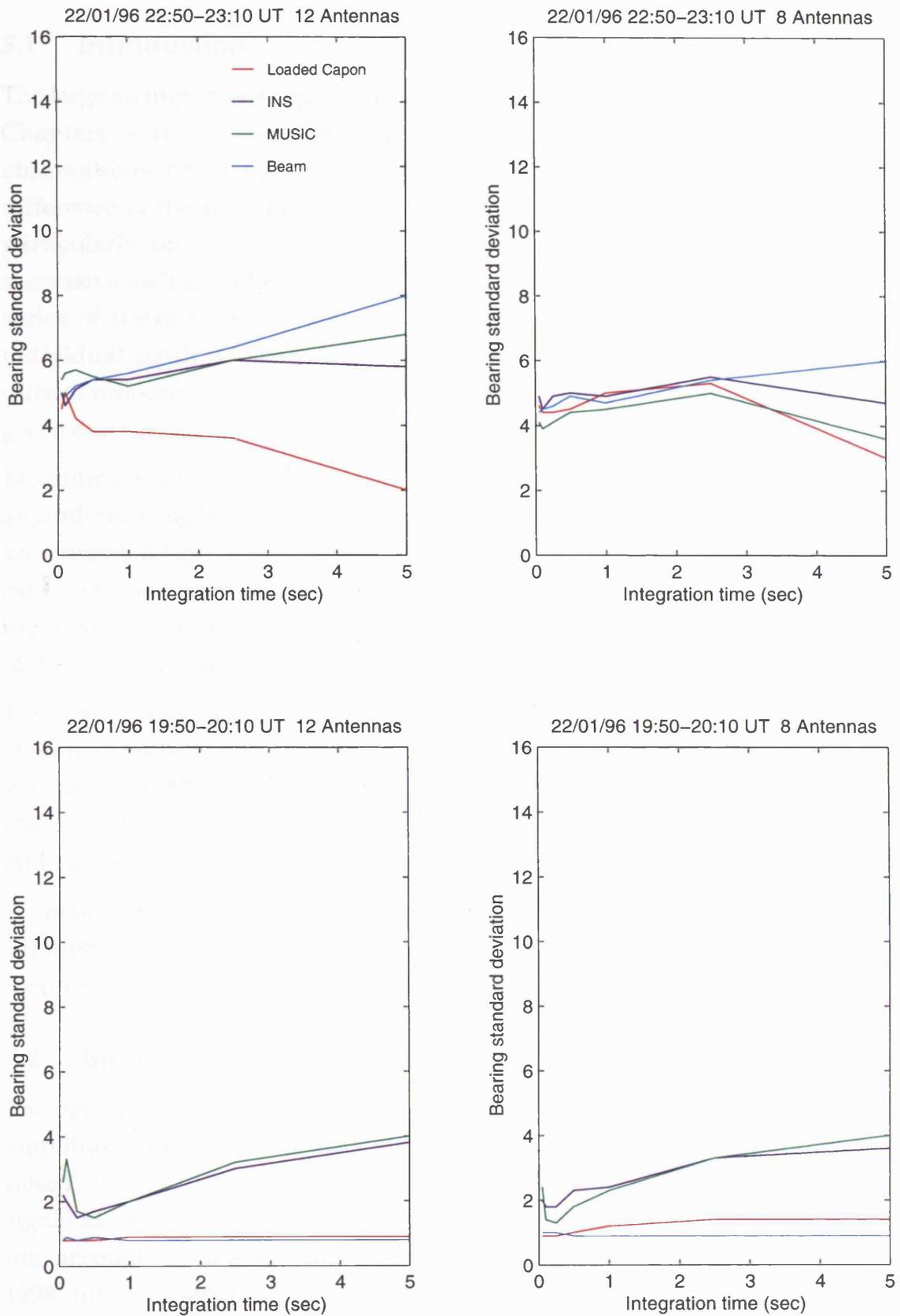


Fig 4.3.3 Standard deviation of experimental bearing estimates as function of integration time, for a time over which the estimates are averaged of 5 sec (top panels: spread > 15°; bottom panels: spread < 15°).

5. MODELLING OF ONE SPREAD SIGNAL

5.1 Introduction.

The large swings in bearing apparent throughout much of the data presented in Chapters 3 and 4 were broadly similar in character to the goniometric observations previously reported (Warrington *et al*, 1997b). The major difference is the presence of multiple traces that closely track one another, particularly on 23rd and 24th January 1996. In considering the presence of successive swings in bearing on 23rd January 1996, Dumas (1997) simulated a series of traces in the bearing time history, each trace corresponding to an individual patch. However, Dumas did not take into account the effect of diffuse reflections from the ionosphere and it is important that this effect is given due consideration.

Modelling studies undertaken to investigate the behaviour of the various algorithms using both antenna array configurations, for diffuse incident energy, are presented here. It is noteworthy that the distribution of energy in practice is expected to be complex and time varying and therefore it was not expected that the modelling exercise would accurately reproduce the time evolving features of the analysed data.

The signal was modelled as a grid of point sources in both azimuth and elevation. Several symmetrical and asymmetrical multiple discrete source distributions (MDSDs), including the raised cosine, were imposed on the grid. Different resolutions were investigated. A resolution of 0.1° in both azimuth and elevation was used for most simulations (see Section 5.2).

A theory of two-signal representation from one incoming signal spread in azimuth is derived and presented here, based on Zatman's theory (1998). The theoretical results are compared with model data results.

5.2 Simulation of a spread signal.

The data have been simulated by modelling an incoming signal whose complex amplitude function $A(\theta, \alpha)$ consists of three parts: the type of distribution (the raised cosine $RC(\theta, \alpha)$ is used here, see Section 5.3 for other types), phase of the signal at each antenna $Ph(\theta, \alpha)$ and Doppler spread $Dop(\theta, \alpha)$. The latter takes into account the time varying nature of the diffuse reflections (see Warrington, 1998, for a presentation of experimental observations which may be used to justify this approach). The amplitude of the signal is then given by

$$A(\theta, \alpha) = RC(\theta, \alpha)Ph(\theta, \alpha)Dop(\theta, \alpha) \quad (5.2.1)$$

where the three contributions (at a given instant and considering one antenna at a time) can be expressed as

$$\begin{aligned}
 RC(\theta, \alpha) &= \frac{1}{2} \left\{ 1 + \cos \left[\frac{2\pi}{\Delta\theta} (\theta - \bar{\theta}) \right] \right\} \left\{ \frac{1}{2} \left[1 + \cos \left[\frac{2\pi}{\Delta\alpha} (\alpha - \bar{\alpha}) \right] \right] \right\}, \\
 Ph(\theta, \alpha) &= \exp \left\{ j \frac{2\pi r_i}{\lambda} \cos(\theta_i - \theta) \cos \alpha \right\}, \\
 Ph(\theta, \alpha) &= \exp \left\{ j \frac{2\pi r_i}{\lambda} \cos(\theta_i - \theta) \cos \alpha \right\}, \\
 Dop(\theta, \alpha) &= \exp \left\{ j 2\pi \left[\frac{\Delta\theta_D}{\Delta\theta} (\theta - \bar{\theta}) + \frac{\Delta\alpha_D}{\Delta\alpha} (\alpha - \bar{\alpha}) \right] \right\},
 \end{aligned}$$

$\bar{\theta}$ and $\bar{\alpha}$ are the mean azimuth and elevation angles of arrival, r_i is the distance of the i -th antenna from a reference point (the centre of the array for the Vortex system), θ_i is the azimuth of the i -th antenna with respect to a reference direction (North), λ is the wavelength of the signal, $\Delta\theta$ and $\Delta\alpha$ are the azimuth and elevation spread of the signal, $\Delta\theta_D$ and $\Delta\alpha_D$ are the azimuth Doppler spread and elevation Doppler spread of the signal, and t is the time.

Analytical solution.

In order to obtain the incoming signal, for each antenna and at a given instant (i.e. the corresponding component of the data vector required for the DF process), the complex amplitude of Equation 5.2.1 needs to be integrated with respect to θ and α within the limits θ_{\min} , θ_{\max} , α_{\min} and α_{\max} . After some manipulation the resulting function of θ and α , for fixed values of all above mentioned parameters, can be expressed as

$$f(\theta, \alpha) = \frac{1}{4} f_1(\alpha) f_2(\theta, \alpha) \quad (5.2.2)$$

where

$$f_1(\alpha) = \int_{\alpha_{\min}}^{\alpha_{\max}} \left\{ 1 + \cos \left[\frac{2\pi}{\Delta\alpha} (\alpha - \bar{\alpha}) \right] \right\} \exp \left\{ j 2\pi \left[\frac{\Delta\alpha_D}{\Delta\alpha} (\alpha - \bar{\alpha}) \right] \right\} d\alpha$$

and

$$f_2(\theta, \alpha) = \int_{\theta_{\min}}^{\theta_{\max}} \left\{ 1 + \cos \left[\frac{2\pi}{\Delta\theta} (\theta - \bar{\theta}) \right] \right\} \exp \left\{ j \frac{2\pi r_i}{\lambda} \cos \alpha \cos(\theta_i - \theta) \right\} \exp \left\{ j 2\pi \left[\frac{\Delta\theta_D}{\Delta\theta} (\theta - \bar{\theta}) \right] \right\} d\theta$$

Focusing on $f_2(\theta, \alpha)$, and taking $\bar{\theta} = 0$ and $\theta_i = 0$ for sake of clarity in the next equation, $f_2(\theta, \alpha)$ would be of the type

$$f(\theta) = \int_{\theta_{\min}}^{\theta_{\max}} [1 + \cos(a\theta)] \exp(jb \cos \theta) \exp(jc\theta) d\theta \quad (5.2.3)$$

where

$$a = \frac{2\pi}{\Delta\theta}; \quad b = \frac{2\pi r_i}{\lambda} \cos\alpha; \quad c = 2\pi t \frac{\Delta\theta_D}{\Delta\theta}$$

For the first snapshot of the simulated data we have $t=0$, thus $c=0$ and $\exp(jc\theta)=1$. In such scenario the solution to Equation 5.2.3, would require the following integral to be solved

$$I = \int_{\theta_{\min}}^{\theta_{\max}} \cos(a\theta) \exp(jb \cos\theta) d\theta \quad (5.2.4)$$

For a frequency of 9.292 MHz, an antenna array such that r_i lies between 25 and 125 m (see Section 3.2) and for elevation angles up to 20° , b would assume values between 4.5 and 24.3. Even though the exponential in the integrand of Equation 5.2.4 could be approximated in an expansion series or in a sum of Bessel functions, for the above range of values for b , the approximated expressions would converge after about 60 and 70 terms respectively, which does not allow a practicable solution. Robinson T.R. (personal communication, 2002) demonstrated that the integral of Equation 5.2.4 could be solved analytically, with good approximation for the conditions $b \gg 1$, $b \gg a$, $b \gg \theta_{\min}$ and $b \gg \theta_{\max}$, and by approximating $\cos\theta \cong 1 - \theta^2/2$. However, the above conditions do not occur in this case, in particular a cannot be smaller than 1, and for typical azimuth spreads up to 60° ($0 < \Delta\theta \leq \pi/3$) we have $3 \leq a < \infty$. Furthermore, for the following snapshots t increases, thus contributing to the increase of a .

Numerical solution.

Since the expression of the amplitude of the incoming signal cannot be integrated analytically, the data simulations presented in this and the next chapter have been obtained by modelling the incoming signal as a grid of point sources separated by a given value in both azimuth and elevation, i.e. a multiple discrete source distribution (MDSD). Resolutions between 0.02° and 2.5° in both azimuth and elevation were used.

The simulated data were analysed using the same DF algorithms and software as for the experimental data. In some of the model data analysis, asymmetries were introduced by specifying different 'half widths' of the distribution in increasing and decreasing angles in both azimuth and elevation (the 'half width' was defined as the angular separation from the peak for the power distribution to fall to zero). No noise was introduced in the modelled data, with rounding errors from the computational process proving negligible. As a consequence the results were identical when, for a given spread (in azimuth

and/or elevation), the amplitude of a signal was varied in order to match the size of the area subtended by the distribution of another signal with a different spread.

For the investigation presented in this sub-section, the MDSD was modelled to produce 1 estimate each 0.1 sec, with 10 snapshots for each estimate. A variety of input azimuth spreads were employed. Simulations have been undertaken with a number of sufficiently small values for the grid size and for a number of azimuth and elevation spreads, over 1 minute of data (i.e. 600 estimates). An example of bearing vs. time plots is shown in Figure 5.2.1, where a raised cosine distribution for the amplitude of the signal has been used (see next section for other distributions). The results are obtained with INS and the full array, for an input azimuth and elevation spread of 20° and 10° respectively (the results for other spreads are analogous). The angular separation of the grid ranges from 0.02° to 2.5° in both azimuth and elevation. Even though the azimuth spread values are consistent for all angular separations, the results clearly show a cyclical trend of the bearing estimates for lower resolutions (right hand panels, particularly apparent for azimuth and elevation separations of 2° and 2.5°). Therefore the MDSD model does not reproduce adequately the experimental data for angular separations greater than 0.5° . In contrast, the results obtained for angular separations between 0.02° and 0.2° (left hand panels), show the typical irregularly scattered estimates that characterise the experimental data. These results are analogous for all algorithms employed and both the 12- and 8-element array. Table 5.2.1 (see below) shows the standard deviations of the bearing estimates obtained with all algorithms and both array geometries. The values are independent of azimuth and elevation separations between 0.02° and 2.5° (only bearings within $\pm 30^\circ$ around the true bearing have been taken into account to avoid possible effects from sidelobe excitations). However, it is worth pointing out that decreasing the separation increases the computational burden.

The cyclical trend of the bearing estimates is caused by the discrete distribution computational routine repeating itself after a number of estimates, which depends on the angular separation of the grid that models the signal. As apparent from the right hand panels of Figure 5.2.1, the lower the resolution the shorter the cycle. Even though the routine would repeat the estimates for any resolution value, we can suggest that this does not affect the validity of the simulations for angular separations of 0.2° and below. In fact, the simulations do not aim to reproduce the experimental observations in detail, but their typical features such as the scattered estimates. Since the MDSD model exhibits similar behaviour for sufficiently small grid sizes, then the numerical solution

can be regarded as a reasonable approximation of the analytical solution for resolution values of 0.2° and below. In the light of these results, a resolution of 0.1° has been used in this thesis, since it provides a reasonable approximation to the integral, with reduced computational burden with respect to higher resolutions.

Table 5.2.1 Standard deviation of the bearing estimates as a function of azimuth and elevation steps for the modelled signal, with input azimuth spread of 20° and elevation spread of 10° .

Az. and El. step (deg.) → Algorithm ↓	0.02	0.05	0.1	0.2	0.5	1	2	2.5
Loaded Capon (12 Antennas)	4.2	3.7	4.0	4.3	3.6	3.9	4.3	3.4
Loaded Capon (8 Antennas)	3.5	3.2	3.6	3.6	3.4	3.9	3.6	3.5
INS (12 Antennas)	3.6	3.6	3.6	4.2	3.4	3.7	4.0	2.9
INS (8 Antennas)	3.7	4.0	4.2	4.2	3.3	3.2	5.0	3.6
MUSIC (12 Antennas)	4.9	5.1	4.6	5.0	5.1	4.9	6.5	5.7
MUSIC (8 Antennas)	5.4	5.7	5.8	5.6	5.5	4.9	5.5	4.3
Beam (12 Antennas)	3.6	3.6	3.6	4.2	3.4	3.7	4.0	3.0
Beam (8 Antennas)	4.1	4.4	4.5	4.5	3.3	3.2	5.3	4.1

5.3 The effect of changing the spread energy distribution model.

In reality, the precise distribution of energy of the incoming signal is not known, and could very well vary with time, location etc. For this reason a number of types of distribution have been used to evaluate how they would affect the bearing estimates, including the gaussian and the square distributions for the amplitude of the signal, and the raised cosine distribution for the amplitude and the power of the signal.

The behaviour of the standard deviation primary bearings as a function of the azimuth spread (given by the width containing 90% of the power) for the four distribution models is presented in Figure 5.3.1. In order for possible sidelobe excitations not to affect the calculation of the bearing standard deviations, only bearings within 30° either side of 188.5° have been included. Whereas MUSIC does not show substantial differences between the various distributions, for Loaded Capon and INS the curves for the raised cosine distributions (both amplitude and power) are fairly close to each other and exhibit standard

deviation values lower than for the gaussian and square distributions. The latter has usually shown the highest bearing standard deviations.

A number of symmetrical and asymmetrical distribution models have been utilised for this investigation. An example is shown in Figures 5.3.2a, b, c and d, where bearings estimates are obtained with all the discrete symmetric distribution models ($\pm 15^\circ$ around the true bearing at 188.5°), for Loaded Capon, INS, MUSIC and Beam respectively, with both the full 12-antenna array (left hand panels) and the 8-antenna sub-array (right hand panels). As expected, the square distribution shows more uniformly scattered bearings estimates than the other distribution models. However, the experimental data have usually shown more irregularly scattered bearings estimates. This difference between the square and the other distributions is more apparent, for example, for Loaded Capon with 8 antennas and INS with 12 antennas. The raised cosine and gaussian distributions show more irregularly scattered estimates, and therefore they better simulate signals propagating at high latitudes. However, in Section 5.6 an investigation is described that shows good quantitative agreement between model (raised cosine distribution) and experimental data in terms of azimuth spread. In this investigation the corresponding ionogram trace was used for an indication of the expected value of the signal azimuth spread. Nevertheless, it cannot be excluded that, in other periods, the gaussian distribution model could also provide good agreement with experimental data and a more comprehensive study is necessary. Results with other values of azimuth spread for the symmetric and asymmetric distributions are analogous to those described above, and as such they have not been included in the thesis. For the reasons, the raised cosine distribution has been chosen for the results in the next sections of this chapter and in the next chapter.

5.4 Behaviour of simulated azimuth as a function of angular spread.

For each model run, 2 minutes of data were simulated for a variety of azimuth spreads with half widths (see Section 5.2) ranging from $2-30^\circ$. The other parameters that varied in these simulations (based on experimental data) were the azimuth Doppler spread (6.4, 8.4, 10.4 and 12.4 Hz), the elevation peak (5° , 10° and 15°), the elevation spread ($\pm 2.5^\circ$, $\pm 5^\circ$ and $\pm 7.5^\circ$) and the elevation Doppler spread (2.4 Hz). The simulations were run with many of the possible combinations of these parameters. It was found that only the value of the azimuth spread had significant effect on the azimuth results and the results for other parameters will not be presented.

Figure 5.4.1 shows the simulated azimuth spread as a function of input azimuth spread. The 'equal values' black dashed line emphasises that the results with MUSIC are often characterised by a similar or slightly higher spread than the input, whereas the opposite is true for Loaded Capon and INS. Particularly apparent is the much lower output azimuth spread with the INS algorithm when only the 8-antenna sub-array is in use, which would result in narrower traces in the bearing vs. time plots similar to those seen experimentally (see Chapter 4).

Examples of bearing distributions for various algorithms are reproduced in Figures 5.4.2a, b, c and d (both for each of the full 12-antenna array and the 8-antenna sub-array). A symmetric azimuth spread with values ranging from 5-30°, which are consistent with values seen experimentally (see Section 5.5), was employed.

The results show that the peaks of primary bearings (red) are always close to the true bearing (188.5°) with all algorithms. The mean bearings produced by all superresolution algorithms were within $\pm 0.1^\circ$ or so of the true bearing, although individual estimates tended to lie within the range of the azimuth energy spread. With Loaded Capon and INS, overall average bearing estimates closer to the peak of input energy were obtained by including both primary and secondary estimates. INS (Figure 5.4.2b) produced lower azimuth spread and less spurious secondary bearings than Loaded Capon (Figure 5.4.2a). MUSIC (Figure 5.4.2c) produced bearing estimates with a much greater spread than those produced by either the Loaded Capon or the INS algorithms. However, the Beam algorithm (Figure 5.4.2d) behaved differently, in that it exhibited secondary bearings well separated from the primary ones (possibly caused by sidelobe excitation), whereas the superresolution algorithms showed secondary bearings surrounding the primary ones. In the time domain this would correspond to well-separated traces for the Beam algorithm, but closely separated traces for the superresolution algorithms. These features were frequently present in experimental observations (see Chapter 4). Differences between the 12 and 8-element array cases are evident. In particular, the bearing distribution splits into multiple peaks at larger spread values with the 8-element array than for the 12-element array.

Simulations were also undertaken with asymmetrical distributions of the azimuth power. The asymmetry was achieved by specifying different values for the 'half widths' of the signal distributions, i.e. different angular separations from the peak for the power distributions to fall to zero. In general, the results were similar to those for the symmetrical case in terms of azimuth spread, whereas the mean bearings produced by all the algorithms were closer to the

mean bearing of the azimuth power distribution rather than the peak. Figure 5.4.3 illustrates one example for all algorithms, obtained with half widths of power distribution equal to -2° to $+30^\circ$, i.e. a portion of the power between 186.5° and the peak at 188.5° and a portion between 188.5° and 218.5° .

For the Loaded Capon and INS algorithms, good average bearing estimates were obtained whilst the bearing estimate distribution was single peaked (i.e. at the lower spreads). In general, the mean bearing was closer to the actual DOA when only the primary bearings were included. However, once multiple peaks in the bearing distribution occurred, better estimates were obtained when averages of both the primary and secondary bearings were taken. With MUSIC, the spread was much larger than with the other superresolution algorithms and the overall average bearing was less accurate. Beam showed even more separated secondary peaks than in the case of symmetric azimuth distribution. It is important to note that asymmetry resulted in small differences (up to about 1°) between the mean of the distributions obtained with the 8 and with the 12-element antenna arrays, which was also observed with in experimental data (see Section 4.2), whereas the symmetric model did not produce this offset (Warrington *et al*, 2000).

5.5 Comparison between modelling and observations.

The results of the modelling described in the previous section often exhibited behaviour similar to that observed experimentally. However, since the model does not fully account for the variability of the propagation environment a precise agreement cannot be expected.

Figures 5.5.1a and b illustrate a typical example of model and experimental bearing estimates for 1 minute of data. In this case an asymmetrical model with half-widths of -28° and $+8^\circ$, i.e. a portion of the power is between 160.5° and the peak at 188.5° and a portion between 188.5° and 196.5° has been employed.

The model data (left hand panels) are generally in good agreement with the experimental data (right hand panels). For example, in Figure 5.5.1a the model data for Loaded Capon with the 12-antenna array (top left panel) exhibit primary and secondary bearings spread over about 25° azimuth very similar to that shown by the corresponding experimental data (top right panel). The agreement is more marked when the 8-antenna sub-array is in use (second panels from the top), where the bearing distribution is more contained (about $15\text{-}20^\circ$), and two traces clearly arise from one spread signal. For INS the azimuth spread is 10° with the full 12-antenna array (third panels) and less than 10° for 8 antennas (bottom panel). Two traces are also apparent in the INS 12-antenna case, for both model and experimental data (third panels). The two

closely separated traces which arise from one spread signal in the model is an important result since two closely separated parallel traces are often seen in the observations (Chapters 3 and 4). A theoretical investigation on this matter and a comparison with model data results are presented in Sections 5.7 and 5.8 respectively.

With both 12 and 8 antennas, the MUSIC algorithm (top four panels, Figure 5.5.1b) exhibits similar levels of azimuth spread for simulated and experimental results. However, there appears to be two traces when using the sub-array with experimental data which are not present in the model – therefore in this case the agreement is not as good as it was for Loaded Capon and INS. The Beam algorithm (third and bottom panels) exhibits good agreement between model and experimental primary bearings (especially with 8 antennas, bottom panels). With the full 12-antenna array (third panels from the top), the secondary bearings are scattered at various azimuth values for model and experimental data. It is important to point out that the secondary bearing traces at around 200° do not arise from the spread incoming signal as an artefact of the Beam algorithm, but are possibly caused by sidelobe excitations (this was commonly seen for Beam results, differently from the SRDF algorithms). In fact, the signal is mainly spread below the true bearing of 188.5° (as we have seen with Loaded Capon with 8 antennas and INS with 12 antennas, one trace is around the true bearing and the other trace around 180°).

A good general agreement between the model and the observed data was shown, in particular with the Loaded Capon and INS algorithms, for a short set of data. It is likely that by an appropriate choice of the model parameters the majority of the observed data (even though with different levels of agreement according to the algorithm in use), could be adequately simulated. The behaviour of the experimental data with different superresolution algorithms (Figures 5.5.1a and b) generally agrees with the behaviour shown during several runs with model data (Figure 5.4.1). In fact, the Loaded Capon and the INS algorithms, used for experimental data, exhibited traces spread in azimuth less than 36° (thus less than the input azimuth spread), particularly INS with the 8-antenna sub-array, whereas MUSIC exhibited traces spread over $30\text{-}40^\circ$. The Beam algorithm confirmed its tendency in experiment of producing secondary traces well separated from the primary bearings (see Section 4.2), with the model data graphs often showing secondary peaks distant from the primary peaks.

5.6 A comparison of measured azimuth spread with that observed from ionogram and simulation results.

Oblique ionospheric sounding between a transmitter and a receiver is frequently used to investigate the propagation mode content of an incoming signal (see Section 1.5). Figure 5.6.1 shows an oblique ionogram obtained over the I-A path on the 22nd January 1996 during the 5-minute interval between 14:11 and 14:16 UT. The delay spread at the frequency of operation of 9.292 MHz is about 0.4-0.5 ms. In Figure 5.6.2 the inferred signal azimuth spread as a function of spread of ionogram traces is plotted, with the assumptions of a flat earth, a symmetrical signal azimuth spread with respect to the GCP and a constant reflection height of 200 km at the mid-path point (F layer, 1 hop, see Section 3.3), thus neglecting any contribution of the elevation spread to the ionogram trace delay spread. These assumptions are appropriate to this investigation because its purpose is to make a broad comparison between azimuth spread from observed and simulated data, and ionogram trace spread. For example, the curve of Figure 5.6.2 does not show any substantial difference for a reflection height of, say, 250 km (plot not shown). Figure 5.6.2 is derived by calculating a number of points as follows. Let us call l_1 the distance between the transmitter T and the reflection point P along the GCP (1068.88 km, see Figure 5.6.3). Let us call P' a generic reflection point of a signal whose azimuth spread is θ , and l_2 the distance between T and P'. Since the signal is assumed symmetrically spread with respect to the GCP, then the angle between the TP and TP' will be equal to $\theta/2$. For a given value of azimuth spread, the corresponding ionogram trace spread can be calculated as follows. Being l_{diff} the difference between the length of the two paths TP'R and TPR, given by $l_{diff} = 2(l_2 - l_1)$, we can easily calculate the difference in time t_{diff} taken by the signal components to reach the receiver through these two paths as $t_{diff} = l_{diff} / c$, where the velocity of the signal is approximated as the light speed c . Since the ionogram of Figure 5.6.1 does not exhibit any group retardation at 9.3MHz, thus the reflection points are well away from the layer peaks, therefore the signal speeds would be similar on each path. Bearing in mind that the time taken by the signal components to traverse the paths TP'R and TP''R is the same, the value of t_{diff} corresponds to the ionogram trace delay spread. The segments l_1 and l_2 are related by $l_1 = l_2 \cos \theta/2$, therefore we have

$$t_{diff} = \frac{2l_1}{c} \left(\frac{1}{\cos(\theta/2)} - 1 \right) \quad (5.6.1)$$

and with a few steps

$$\theta = 2 \arccos \left(\frac{1}{\frac{ct_{diff}}{2l_1} + 1} \right) \quad (5.6.2)$$

By using the Equation 5.6.2 for a number of values of the ionogram trace delay spread t_{diff} , we obtain the corresponding values of azimuth spread θ . Figure 5.6.2 is given by the resulting series of points.

A relative delay spread of about 0.4-0.5 ms corresponds to an azimuth spread of about 40°. By comparing the results from the simulation with the azimuth spread set to $\pm 20^\circ$ to the relevant observations will provide a suitable test.

The modelled (raised cosine) and experimental azimuth distributions from the interval close to that of the ionogram of Figure 5.6.1 are illustrated in Figures 5.6.4a and b. Note that the time interval for the model (left hand panels, 2-minutes) is shorter than that of the experimental data (right hand panels, 15 minutes). The 5-minute ionospheric sounding slot is the data gap which starts at 14:11. Although the model makes a number of assumptions, the results agree quite well with observations. For example, for Loaded Capon (12-antennas, Figure 5.6.4a, top left panel) the azimuth spread is about 30° around the true bearing, 10° less than the input azimuth spread. The experimental data traces are indeed spread by about 30°, with secondary traces surrounding the primary ones. Analogous behaviour occurs with reduced spread, when using the 8-antenna sub-array (second panels), which both show an azimuth spread of about 20°. While INS exhibits much less secondary energy than Loaded Capon, there is still good agreement between model and experiment. The good agreement is largely held for MUSIC (Figure 5.6.4b). However, Beam exhibits primary bearing estimates concentrated around the Great Circle direction with secondary estimates mainly at azimuths higher than the true bearing for both model and experimental data and with both array geometries. With the full 12-antenna array, Beam shows a trace of secondary bearings just above 200°, possibly caused by sidelobe excitation.

In conclusion, for this example the delay spread in the ionogram trace in Figure 5.6.1 gives a good indication of the azimuth spread. This applies more or less for all algorithms and array geometries, although the superresolution ones perform better. This agreement was generally found for other times, when the ionogram trace delay spread was greater than about 0.4 ms. However, there was poor agreement for ionograms of two types. The first type was when the ionogram trace was relatively narrow (typical examples are the ionograms of Figure 3.4.2), which made the identification of the corresponding azimuth spread difficult, since a small change in the ionogram trace delay spread would lead to a large change in the inferred signal azimuth spread. The second type

was when the ionograms show two traces and/or other features at the operational frequency (typical examples are most of the ionograms of Figure 3.4.4). Although the results shown in this section are promising, the limited number of available examples implies that a longer period of data collection, with the corresponding ionograms, is necessary for a more comprehensive study.

5.7 Theory of two-signal representation from one signal spread in azimuth.

Throughout this thesis we have often seen closely separated traces (in bearing), possibly arising from a spread signal, with both experimental and model data. In Section 2.10 we have seen the results presented by Hayward (1997) and Zatman (1999), which relate the occurrence of two traces arising from one signal to the eigenvalues of the covariance matrix. Hayward stated that this situation can occur in case of a rotating array, which would have the same effect, on the DF system, of an incoming spread signal. Hayward related the eigenvalues to the spread of the signal measured in beamwidths, using a linear array of 16 antennas. His results (Figure 5.7.1a) show that the higher the spread, the more the weaker eigenvalues approach the value of the first eigenvalue. Of particular importance is the value of the second eigenvalue, which could become strong enough to give rise to a second bearing. Figure 5.7.1b illustrates analogous results (for the first two eigenvalues) for the full 12-antenna Vortex array (red curves) and the 8-antenna sub-array (blue curves). These have been obtained by simulating a snapshot of data (see Section 5.1), and then by calculating the eigenvalues of the covariance matrix, for an incoming signal (raised cosine distribution) with a number of values of azimuth spreads, symmetrical with respect to the GCP, expressed in beamwidths (0.005 to 1.5). Therefore, for a given azimuth spread value in beamwidths, the absolute azimuth spread value would be different between the full-array and the sub-array. The solid curves represent the first eigenvalue, and the dashed curves the second eigenvalue. We can see a clear agreement between the Hayward's results in Figure 5.7.1a and the results for the Vortex array in Figure 5.7.1b.

The absolute value of the eigenvalues depends on many parameters used to calculate them, such as the number of antennas, the number of samples for each estimate and the azimuth steps for the discrete multiple distributions. Therefore, the important aspect is the ratio between the eigenvalues, not their absolute value. For this reason, the eigenvalues in Figure 5.7.1b have been normalised to the first eigenvalue obtained with 12 antennas for a signal spread of 0.005 beamwidths. The eigenvalues obtained with 8 antennas have been

normalised with the same value as for 12 antennas. The eigenvalues curves obtained with 8 antennas are lower than those with 12 antennas because fewer signal components contribute to the data matrix.

A mathematical theory behind a two-signal representation of one non-zero bandwidth signal, was obtained by Zatman (1998) (see Section 2.10). Zatman also related this occurrence to the first two eigenvalues of the covariance matrix. Given the analogy between Zatman's theory about a signal spread in frequency, and my results about a signal spread in azimuth, it is worth making a comparison between the two situations. A theory for a two-signal representation of one signal spread in azimuth is derived in this section and compared to simulations results in the next section.

Zatman considered a uniform linear array of N antennas, with half-wavelength spacing at the operating frequency f_0 . Zatman's two-signal representation of a non-zero bandwidth signal is valid when the amount of spreading is consistent with two eigenvalues of the covariance matrix being above the noise floor, i.e. when the narrowband condition fails. For a signal-only covariance matrix, this corresponds to one signal being represented by two eigenvalues greater than 1 (i.e. 0dB, see Section 2.10), as expressed by Zatman's Equation 2.10.15, repeated in Equation 5.7.1

$$\lambda_2 = \frac{Ns_m}{2}(1-|\psi|) < 1 \quad (5.7.1)$$

where λ_2 is the second eigenvalue, $s_m/2$ is the power of each of the two signals and $|\psi|$ is the cosine of the angle between the two eigenvectors in the N -dimensional space.

In analogy with Equation 2.10.9, for a signal spread in azimuth only, with fixed elevation angle and in absence of bandwidth, we have

$$|\psi| = \frac{|a(\bar{\theta}+k)^H a(\bar{\theta}-k)|}{\sqrt{a(\bar{\theta}+k)^H a(\bar{\theta}+k) a(\bar{\theta}-k)^H a(\bar{\theta}-k)}} \quad (5.7.2)$$

where $\bar{\theta}$ is the mean azimuth of the incoming spread signal and k is the azimuth separation of each of the two traces representing the spread signal from the centre bearing $\bar{\theta}$.

For an arbitrary array, such as Vortex (see Section 3.2), the steering vector is given by

$$a(\theta) = \{e^{j\varphi_1}, e^{j\varphi_2}, \dots, e^{j\varphi_N}\} \quad (5.7.3)$$

where φ (the inter-element phase) is given by

$$\varphi_i = A_i \cos(\Theta_i - \theta) \quad (5.7.4)$$

and

$$A_i = \frac{2\pi r_i \cos(\alpha)}{\lambda} \quad (5.7.5)$$

(r_i, Θ_i) being the co-ordinates of the i -th antenna with respect to an arbitrary origin and θ and α the azimuth and elevation angle of arrival respectively.

By substituting Equation 5.7.3 into 5.7.2, we obtain

$$|\psi| = \frac{A}{\sqrt{BC}}$$

where

$$A = \left\{ \begin{array}{l} e^{-jA_1 \cos[\Theta_1 - (\bar{\theta} + k)]} \\ \vdots \\ e^{-jA_N \cos[\Theta_N - (\bar{\theta} + k)]} \end{array} \right\} \left\{ e^{+jA_1 \cos[\Theta_1 - (\bar{\theta} - k)]}, \dots, e^{+jA_N \cos[\Theta_N - (\bar{\theta} - k)]} \right\}$$

$$B = \left\{ \begin{array}{l} e^{-jA_1 \cos[\Theta_1 - (\bar{\theta} + k)]} \\ \vdots \\ e^{-jA_N \cos[\Theta_N - (\bar{\theta} + k)]} \end{array} \right\} \left\{ e^{+jA_1 \cos[\Theta_1 - (\bar{\theta} + k)]}, \dots, e^{+jA_N \cos[\Theta_N - (\bar{\theta} + k)]} \right\}$$

$$C = \left\{ \begin{array}{l} e^{-jA_1 \cos[\Theta_1 - (\bar{\theta} - k)]} \\ \vdots \\ e^{-jA_N \cos[\Theta_N - (\bar{\theta} - k)]} \end{array} \right\} \left\{ e^{+jA_1 \cos[\Theta_1 - (\bar{\theta} - k)]}, \dots, e^{+jA_N \cos[\Theta_N - (\bar{\theta} - k)]} \right\}$$

By developing the expression of $|\psi|$ we have

$$|\psi| = \frac{e^{-jA_1 \{\cos[\Theta_1 - (\bar{\theta} + k)] - \cos[\Theta_1 - (\bar{\theta} - k)]\}} + \dots + e^{-jA_N \{\cos[\Theta_N - (\bar{\theta} + k)] - \cos[\Theta_N - (\bar{\theta} - k)]\}}}{\sqrt{(1+1+\dots)_{N \text{ times}} (1+1+\dots)_{N \text{ times}}}}$$

and therefore

$$|\psi| = \frac{1}{N} \left| \sum_{i=1}^N \exp \left\{ -jA_i \left[\cos(\Theta_i - (\bar{\theta} + k)) - \cos(\Theta_i - (\bar{\theta} - k)) \right] \right\} \right| \quad (5.7.6)$$

Equation 5.7.6 for an arbitrary array and a generic incoming signal spread in azimuth is the analogue of Equation 2.10.13 for a uniform linear array and a non-zero bandwidth signal (Zatman, 1998).

By inserting 5.7.6 into 5.7.1, simplifying the trigonometrical function and inverting the direction of equation 5.7.1, we obtain

$$\lambda_2 = \frac{Ns_m}{2} \left(1 - \frac{1}{N} \left| \sum_{i=1}^N \exp\{-j2A_i[\sin(\Theta_i - \bar{\theta})\sin k]\} \right| \right) > 1 \quad (5.7.7)$$

Equation 5.7.7 states, in terms of azimuth separation between two traces, the condition by which the effective rank of a generic incoming signal increases from one to two. This sets the minimum azimuth separation value above which one signal incoming from a given bearing might have a two-signal representation, for a given array geometry and signal-to-noise ratio (SNR).

The plots of Figure 5.7.2 illustrate the curves of the second eigenvalue as a function of the separation between the two traces representing one signal spread in azimuth around the true bearing for the I-A path (188.5°), for four values of SNR and two array geometries. Results for an elevation of 8.3° are shown, it being found that the curves are largely independent of elevation. For a given SNR, as the azimuth separation increases, the second eigenvalue also does; when it equals 1, the two-signal representation of one spread signal could take place. This is the analogous to the narrowband condition as described by Zatman (1998). It is apparent that, as the SNR increases, the second eigenvalue reaches the unitary value for decreasing traces separations, which are lower for the full array than for the sub-array. For a 25dB SNR, the two-signal representation can occur for about 13° with 12 antennas and about 49° with 8 antennas; these values decrease with increasing SNR falling to 2° and 8° at 40dB SNR. This means that the stronger the SNR, the less spread the signal will 'need to be' in order for the two-signal representation to become possible. It is therefore interesting to relate the azimuth separation of the two traces to the azimuth spread of the signal.

In analogy with Zatman's discussion (see Section 2.10) about the variance of a spread signal with a rectangular power spectrum, let us evaluate the variance of a signal whose bearing distribution is raised cosine (see Section 5.3). The expression of the raised cosine distribution of the azimuth of arrival θ of a signal and for a generic azimuth spread $\Delta\theta$ is

$$f(\theta) = 1/2 \left\{ 1 + \cos \left[(\theta - \bar{\theta}) \frac{2\pi}{\Delta\theta} \right] \right\}$$

Since the following calculations are valid for any mean azimuth value, let $\bar{\theta}$ be equal to 0° for sake of simplicity. Thus the variance of the distribution is given by

$$\text{var}(\theta) = \frac{1}{\Delta\theta} \int_{-\frac{\Delta\theta}{2}}^{\frac{\Delta\theta}{2}} \theta^2 f(\theta) d\theta$$

$$\text{var}(\theta) = \frac{1}{2\Delta\theta} \int_{-\frac{\Delta\theta}{2}}^{\frac{\Delta\theta}{2}} \left\{ \theta^2 + \theta^2 \cos\left(\theta \frac{2\pi}{\Delta\theta}\right) \right\} d\theta$$

$$\text{var}(\theta) = \frac{1}{2\Delta\theta} \left\{ \frac{\theta^3}{3} + 2\left(\frac{\Delta\theta}{2\pi}\right)^2 \theta \cos\left(\theta \frac{2\pi}{\Delta\theta}\right) + \left(\frac{\Delta\theta}{2\pi}\right)^2 \left[\theta^2 - 2\left(\frac{\Delta\theta}{2\pi}\right)^2 \right] \sin\left(\theta \frac{2\pi}{\Delta\theta}\right) \right\} \Big|_{-\frac{\Delta\theta}{2}}^{\frac{\Delta\theta}{2}}$$

With a few steps we obtain the variance of a raised cosine distribution for a given azimuth spread as

$$\text{var} = \frac{(\Delta\theta)^2(\pi^2 - 6)}{24\pi^2} \quad (5.7.8)$$

A two-azimuth-of-arrival representation of the signal would then consist of two traces separated by k either side of the mean azimuth of arrival, where k^2 is the variance of such representation (Zatman, 1998 and personal communication, 2001). Therefore, by substituting $\text{var}(\Delta\theta)$ with k^2 in Equation 5.7.8, we can easily obtain the separation $2k$ between the two traces, as

$$2k = \frac{\Delta\theta}{\pi} \sqrt{\frac{(\pi^2 - 6)}{6}} \quad (5.7.9)$$

Figure 5.7.3 shows the azimuth separation between two traces representing one spread incoming signal, as a function of its azimuth spread, from Equation 5.7.9. It is then useful to relate the second eigenvalue to the azimuth spread, for different values of SNR. This is obtained by substituting k from Equation 5.7.9 into Equation 5.7.7. The results are shown in Figure 5.7.4 for SNR values of 25, 30, 35 and 40dB, the incoming signal being spread in azimuth around 188.5° and with an elevation angle of arrival equal to 15° . With 12 antennas (blue curves) an azimuth spread values of about 51° or more could give rise to a two-trace representation, with the lowest SNR in Figure 5.7.4 (25dB, top left panel). Whereas with 8 antennas (red curves) the theoretical azimuth spread needs to be greater than about 190° (a very unlikely occurrence) for two traces to be shown. For a 30dB SNR (top right panel) an azimuth spread of about 28° and 104° with 12 and 8 antennas respectively is sufficient for two traces to possibly arise. For higher SNR values the minimum azimuth spread value further decreases for a two-traces representation to be possible, dropping to about 15° and 58° (35dB SNR) and about 9° and 33° (40 dB SNR) with 12 and 8 antennas respectively. The above theoretical results are based on the calculation of eigenvalues of the covariance matrix, therefore this discussion refers mainly to

eigenanalysis methods such as MUSIC. However, simulation results with MUSIC did not clearly exhibit a two-trace representation of a spread signal. But since this occurred with Loaded Capon and INS, even though this theory does not relate to these two algorithms, it is worth comparing its results to simulation results (see next section).

5.8 Comparison between theoretical and simulation results of two-signal representation from one signal spread in azimuth.

In this section some of the model data results obtained with a raised cosine distribution presented in this and the next chapter are compared to the theoretical discussion of the previous section. The data were modelled with a given input azimuth spread and in some cases the results exhibited two-closely separated traces (we refer to Figures 5.5.1a, 6.3.2 and 6.3.5, and other results not shown in the thesis, which are relevant to a two-signal representation of one signal spread in azimuth). The algorithm and the array geometry heavily affected the results. The results obtained with INS also depended on the threshold below the first peak for the search for other peaks to terminate (see Figure 6.3.5). In general, the results about the visibility of two signals were dependent on the azimuth separation and the amplitude ratio between them (see Section 6.3). Furthermore, a number of parameters were input to the model, which could affect the results (see Section 6.3). It is worth bearing in mind that any change in the value of any of the various parameters of the model is likely to change the simulation results.

Table 5.8.1 compares some values of theoretical azimuth separation between two traces representing one spread signal to simulation results for Loaded Capon, INS and MUSIC. The theory relates to eigenanalysis DF methods such as MUSIC, which however never clearly showed a two-trace representation (together with Beam), and thus the simulations did not suggest consistency with the theory. Nevertheless, it is interesting to compare the theoretical results to simulations for Loaded Capon and INS, since these showed the two-trace representation of one spread signal, in some circumstances. However, it is important to bear in mind that Loaded Capon and INS are not eigenanalysis methods, and as such the theory does not relate to these two algorithms. It follows that the results in Table 5.8.1 are not obvious to interpret even though they are very interesting, and it would be desirable to consider how the presence of the additional eigenvalue due to the source spread in azimuth would affect the operation of Loaded Capon and INS (see Section 7.3).

From Table 5.8.1 we can see that in the case when two traces arose from one signal spread 15° in azimuth, the simulation results with Loaded capon and INS

exhibited an azimuth separation (between 5° and 7°) higher than the theoretical value of 3.8° (notice that INS with 8 antennas exhibited one trace only). For increasing input azimuth spread values (20°, 30° and 36°), Table 5.8.1 shows that, when two traces arose, the azimuth separation between them increased too, and was fairly close to the theoretical expectations, in particular with Loaded Capon and full array (7-8°) for an input azimuth spread of 30°, the theoretical value being 7.7°. Simulations with higher azimuth spread values did not show any clear two-signal representation of one spread signal. However, it is not suggested that this cannot happen in a real signal environment. Furthermore, it is noteworthy that the two-trace representation occurred, or otherwise, with no apparent trend.

Table 5.8.1 Azimuth separation values of two traces representing one spread modelled signal as a function of the input azimuth spread, for Loaded Capon, INS and MUSIC, with both the full 12-antenna array and the 8-antenna sub-array. The theoretical azimuth separation corresponding to the input azimuth spread is in brackets, calculated from Equation 5.7.9. The signal is modelled with a discrete raised cosine distribution (az peak=188.5°, az Doppler spread=10.4Hz, el peak=15°, el spread=5° and el Doppler spread=2.4Hz). When a two-trace representation did not occur, '1 trace' is stated.

Input azimuth spread →	15°	20°	30°	36°
Theoretical azimuth separation →	(3.8°)	(5.1°)	(7.7°)	(9.2°)
SRDF algorithm ↓				
Loaded Capon (12 antennas)	5-6°	1 trace	7-8°	1 trace
Loaded Capon (8 antennas)	5-6°	6-7°	1 trace	11-12°
INS (12 antennas)	6-7°	1 trace	1 trace	10-11°
INS (8 antennas)	1 trace	1 trace	8-9°	1 trace
MUSIC (12 antennas)	1 trace	1 trace	1 trace	1 trace
MUSIC (8 antennas)	1 trace	1 trace	1 trace	1 trace

The SNR worst-case scenario in Table 5.8.1 is for the lowest azimuth spread value (15°) and the 8 antenna sub-array, as we can infer from Figure 5.7.4. In

such scenario, the minimum SNR necessary for the second eigenvalue to be greater than 1 is about 47dB, according to the theoretical calculations (Equation 5.7.7, in which the value of k is obtained from Equation 5.7.9 for a raised cosine distribution).

The occurrence of a two-trace representation of one spread signal is clearly not a desirable feature. The algorithms behaved correctly when only one trace was detected, representing one spread signal. There appears to be no particular trend in the algorithms behaviour, i.e. it is not possible to infer which algorithm (between Loaded Capon and INS) or which array geometry is more opportune to use in order to avoid the occurrence of a two-trace representation. Even though MUSIC does not exhibit this problem, we have seen that it is not suitable for high latitude DF. Furthermore, a high number of parameters, corresponding to often-unpredictable features in a real signal environment, are likely to change the results. The modelling results for MUSIC were not in agreement with the theoretical results, since MUSIC never exhibited two traces arising from one spread signal. However, for those circumstances when a two-signal representation of one spread signal occurred for Loaded Capon and INS, the theoretical values were in fairly good agreement with the simulations, especially for azimuth spread values of around 30°.

5.9 Summary.

Ionospheric propagation modes have been often modelled by others (Sections 1.5 and 2.9) with the over-simplification of assuming a smooth ionosphere of infinite extent upon which were superimposed the localised irregularities leading to the scattered energy. More realistic modelling studies have been reported in this chapter, which consider the effect of diffuse reflections from the ionosphere. This model reproduces much more faithfully very disturbed regions, where signals are more likely to arrive spread at the receiver, as often occurs at the high latitudes.

A multiple discrete source distribution was modelled as a grid of point sources separated in both azimuth and elevation for a number of values between 0.02° and 2.5°. For angular separations of 0.2° and below the results did not vary significantly and were able to reproduce the typically scattered bearing estimates seen with experimental data. Since the computational burden increased with increasing resolutions, a reasonably fine grid of point sources separated by 0.1° was employed for the modelling studies of this and the next chapter.

Several types of distributions for the multiple discrete source model have been compared. These are the raised cosine, gaussian and square distributions for the

amplitude of the simulated signal, and also the raised cosine distribution for the power of the signal. The raised cosine and the gaussian distributions appeared to be able to reproduce the irregularly scattered estimates seen with experimental data (less spread for the raised cosine distribution), whereas the square distribution exhibited more uniformly distributed bearing estimates, and thus it is not recommended.

In some of the model data analysis a symmetric distribution of the signal around the Great Circle Path direction was assumed, whereas in others analysis the distribution was assumed to be asymmetric. The model was able to reproduce at least the gross behaviour of the observations for all the DF algorithms. For example, the spread in the bearing estimates was found to be different for each algorithm for a given input spread. Loaded Capon and INS exhibited traces spread in azimuth less than the input value, especially INS with the 8-antenna sub-array, whereas MUSIC often showed traces spread about the same or slightly more than the input value, especially with the full 12-antenna array. Beam often produced secondary traces well separated from the primary bearings.

A comparison of the model results and the observations has found that the amount of spread in an ionogram trace can give an indication of the expected value of the signal azimuth spread. A good agreement was found when the relative delay of the ionogram traces spread was above about 0.4 ms.

A theory about the two-signal representation of one signal spread in azimuth has been developed, based on the relationship between the second eigenvalue of the covariance matrix and the azimuth spread of a raised cosine distributed signal. The relationship between the two traces separation and the amount of azimuth spread derived from the theory was found to be in good agreement, especially for azimuth spread values of around 30° (the azimuth separation between the two traces increased for increasing azimuth spread values, as expected). For input azimuth spreads of about 40° or more, the two-signal representation of one spread signal did not occur in the simulations. The occurrence of a two-trace representation of one spread signal depended on the algorithm and array geometry in use, the type of signal distribution, and to a lesser extent the elevation angle of arrival, elevation spread and azimuth and elevation Doppler spread. Unfortunately, the results did not show any particular trend in terms of which algorithm and array geometry would be better to use in order to avoid the undesirable two-trace representation of one spread signal. Furthermore, it is important to consider that the theory relates to eigenanalysis methods such as MUSIC, whereas the best agreements with the simulation results were found with Loaded Capon and INS. Therefore, it would

be useful to undertake further investigations for a better understanding of the relationship between this theory and the operation of non-eigenanalysis methods such as Loaded Capon and INS.

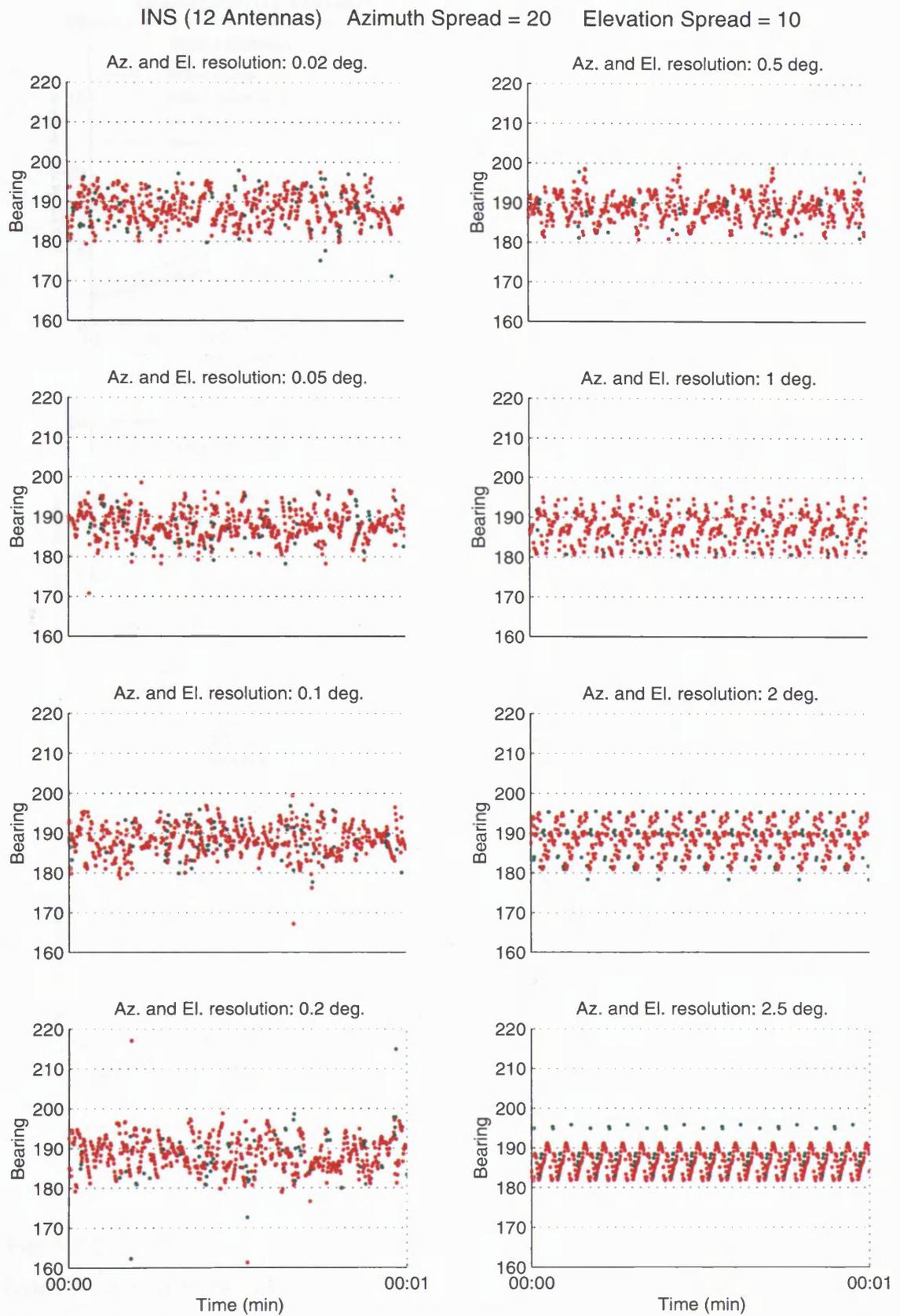


Fig 5.2.1 Bearing estimates obtained from the discrete distribution model (raised cosine), for angular separation values of the grid of point sources from 0.02° to 2.5° in both azimuth and elevation.

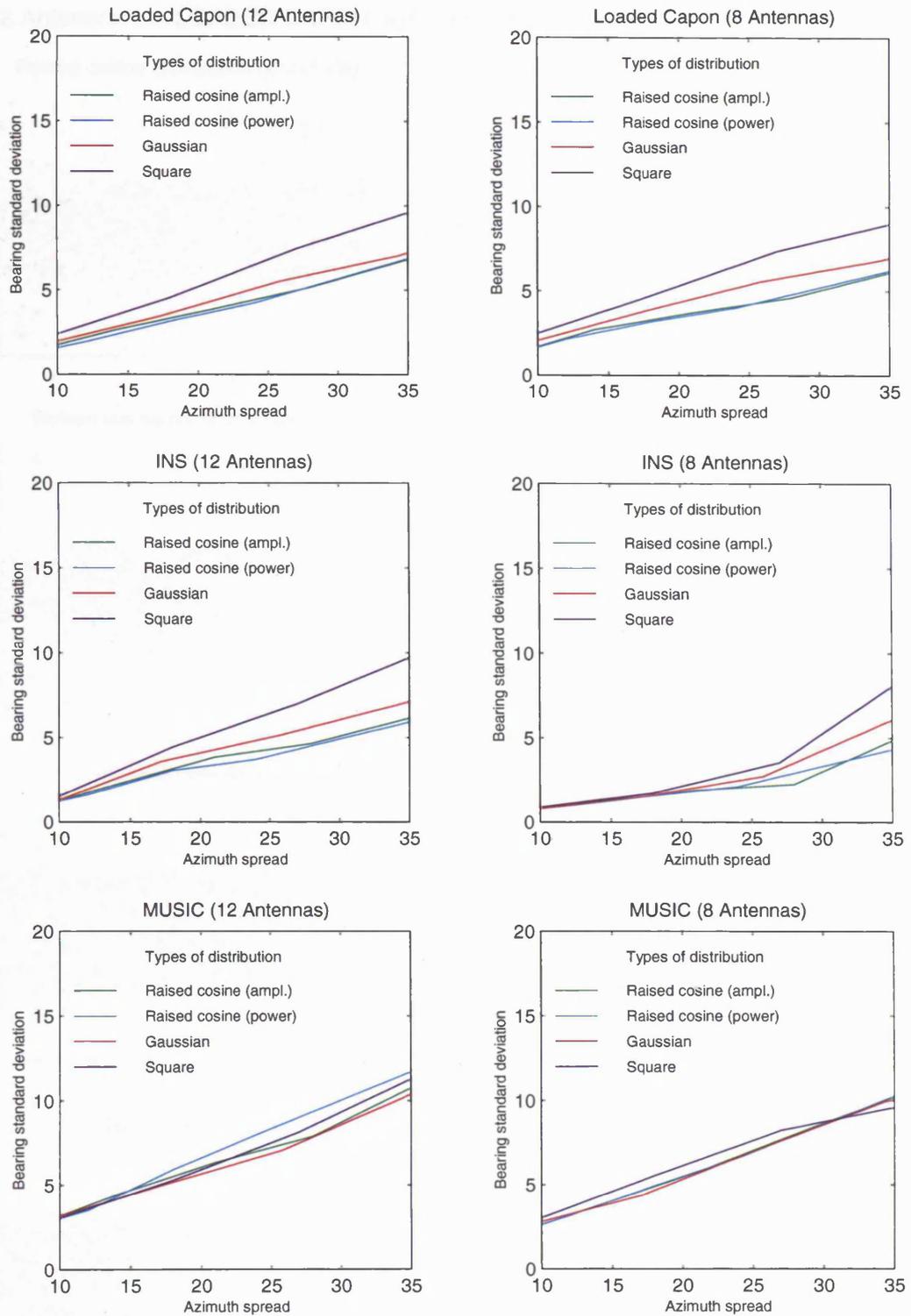


Fig. 5.3.1 Primary bearings estimates standard deviation values obtained from four symmetrical distribution models for the Loaded Capon, INS and MUSIC algorithms (top to bottom), with 12 antennas (left hand panels) and 8 antennas (right hand panels). Input azimuth spread values between 10° and 60° . (az Doppler spread=10.4Hz, el peak=15, el spread= ± 2.5 and el Doppler spread=2.4Hz).

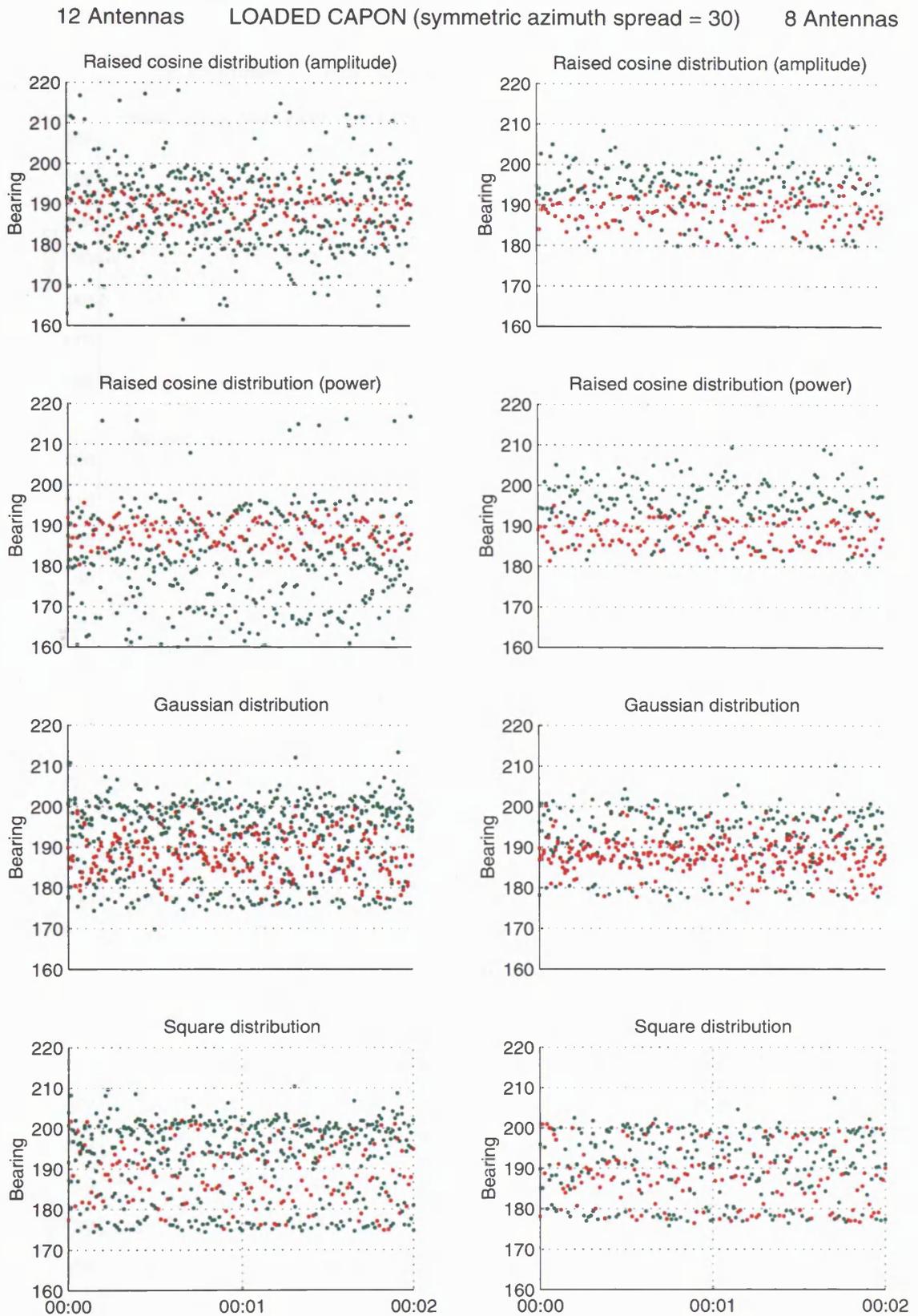


Fig. 5.3.2a Bearing estimates obtained from four symmetrical distribution models, using the Loaded Capon algorithm with 12 antennas (left hand panels) and 8 antennas (right hand panels). The distribution width is equal to $\pm 15^\circ$.

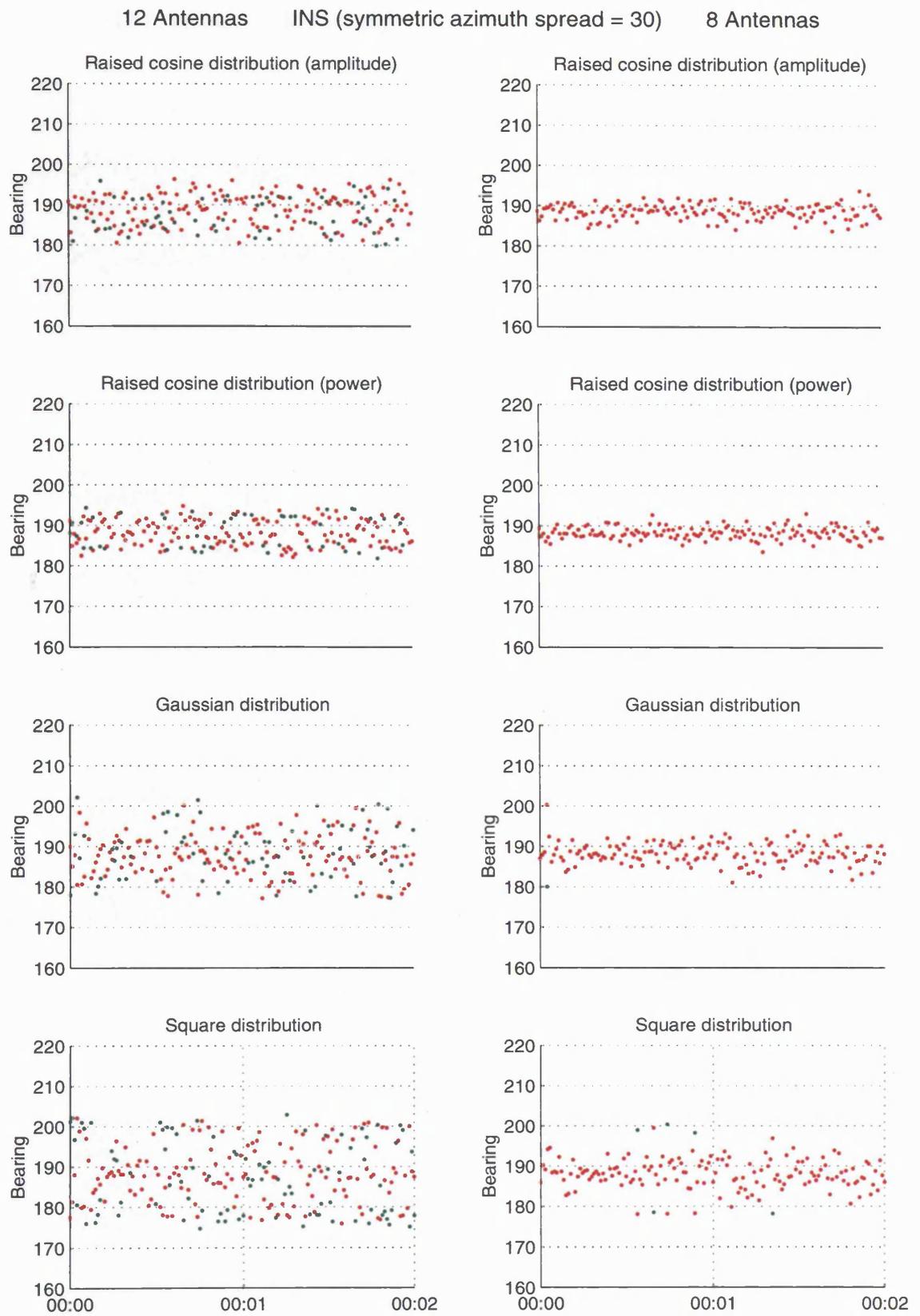


Fig. 5.3.2b As for Fig. 5.3.2a except for the INS algorithm.

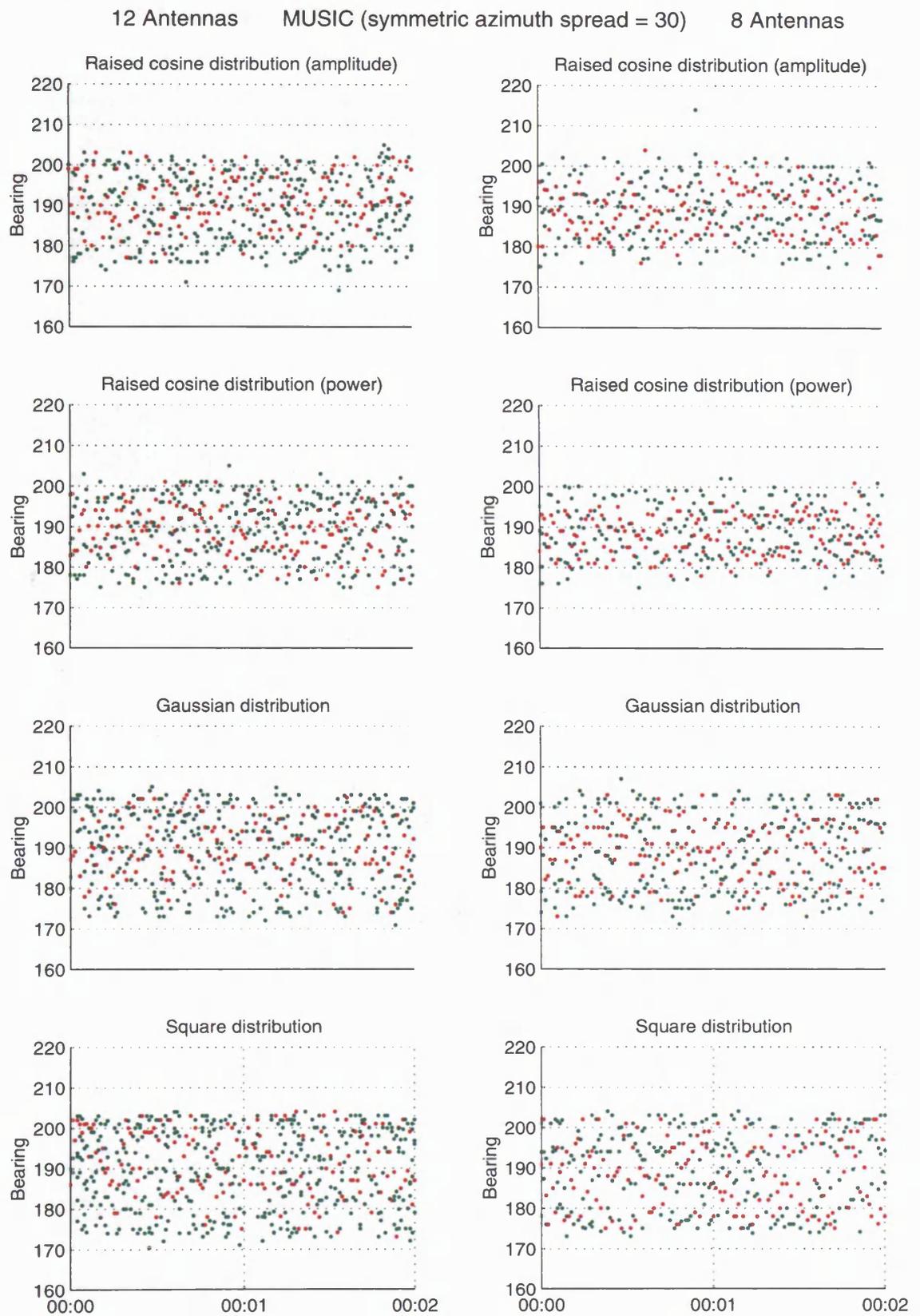


Fig. 5.3.2c As for Fig. 5.3.2a except for the MUSIC algorithm.

12 Antennas BEAM (symmetric azimuth spread = 30) 8 Antennas

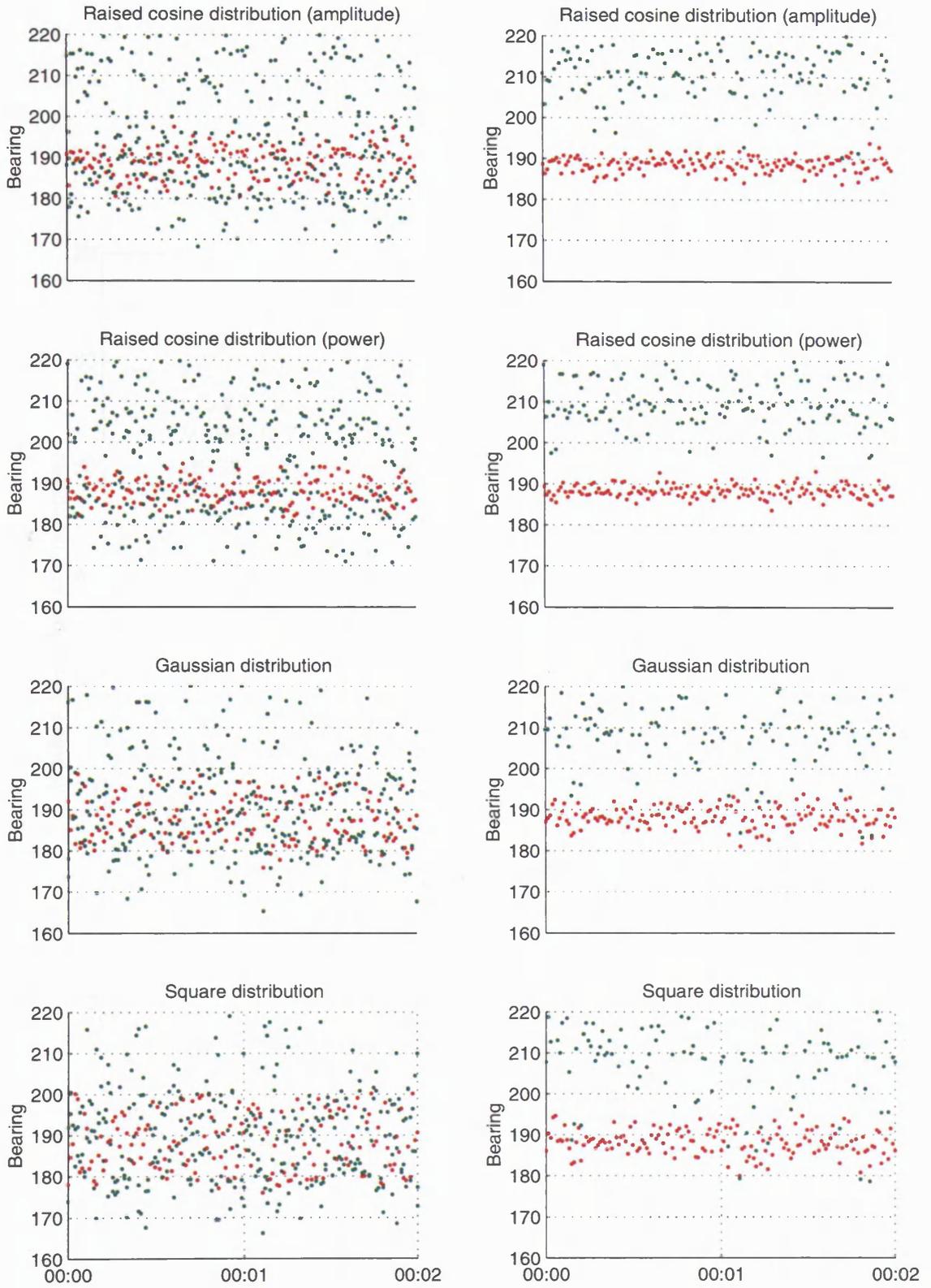


Fig. 5.3.2d As for Fig. 5.3.2a except for the Beam algorithm.

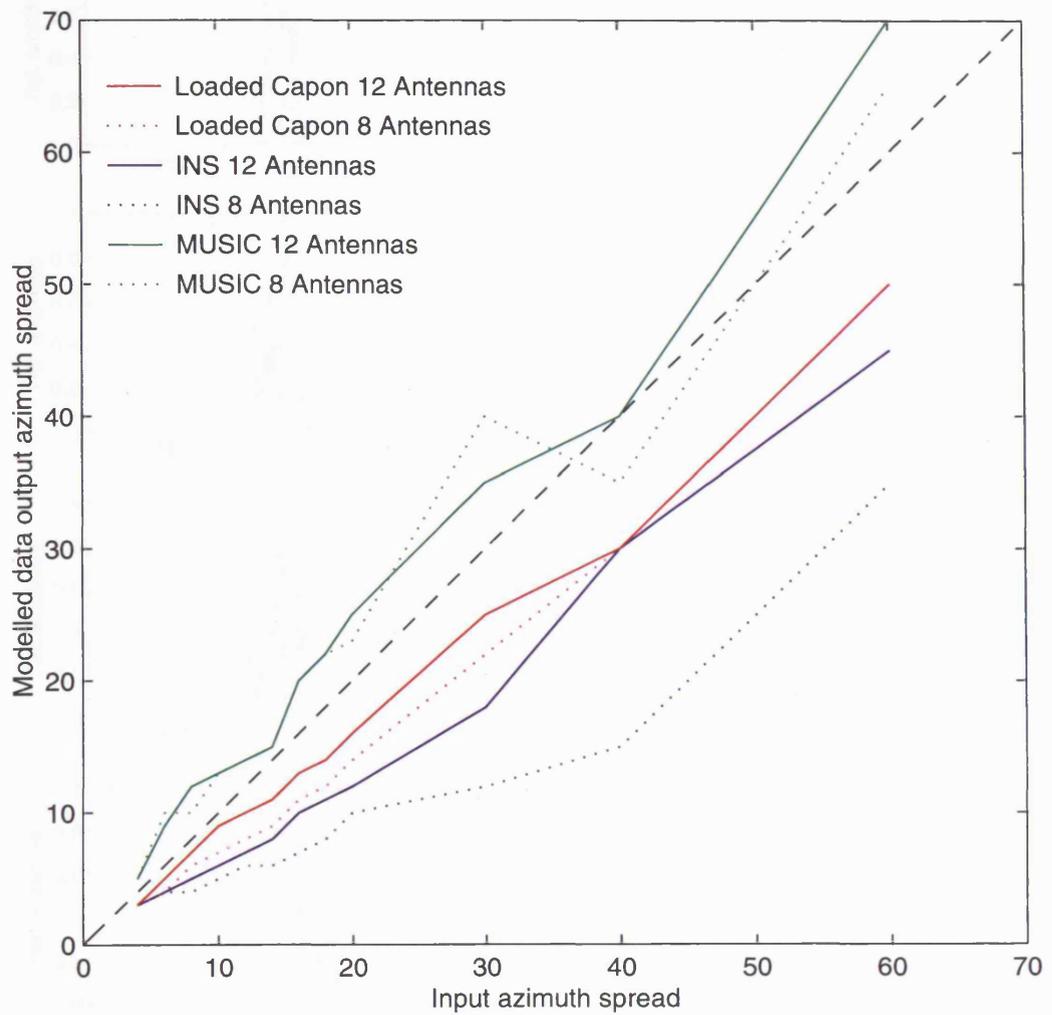


Fig 5.4.1 Simulated azimuth spread vs. input azimuth spread obtained with the superresolution algorithms Loaded Capon, INS and MUSIC (raised cosine distribution).

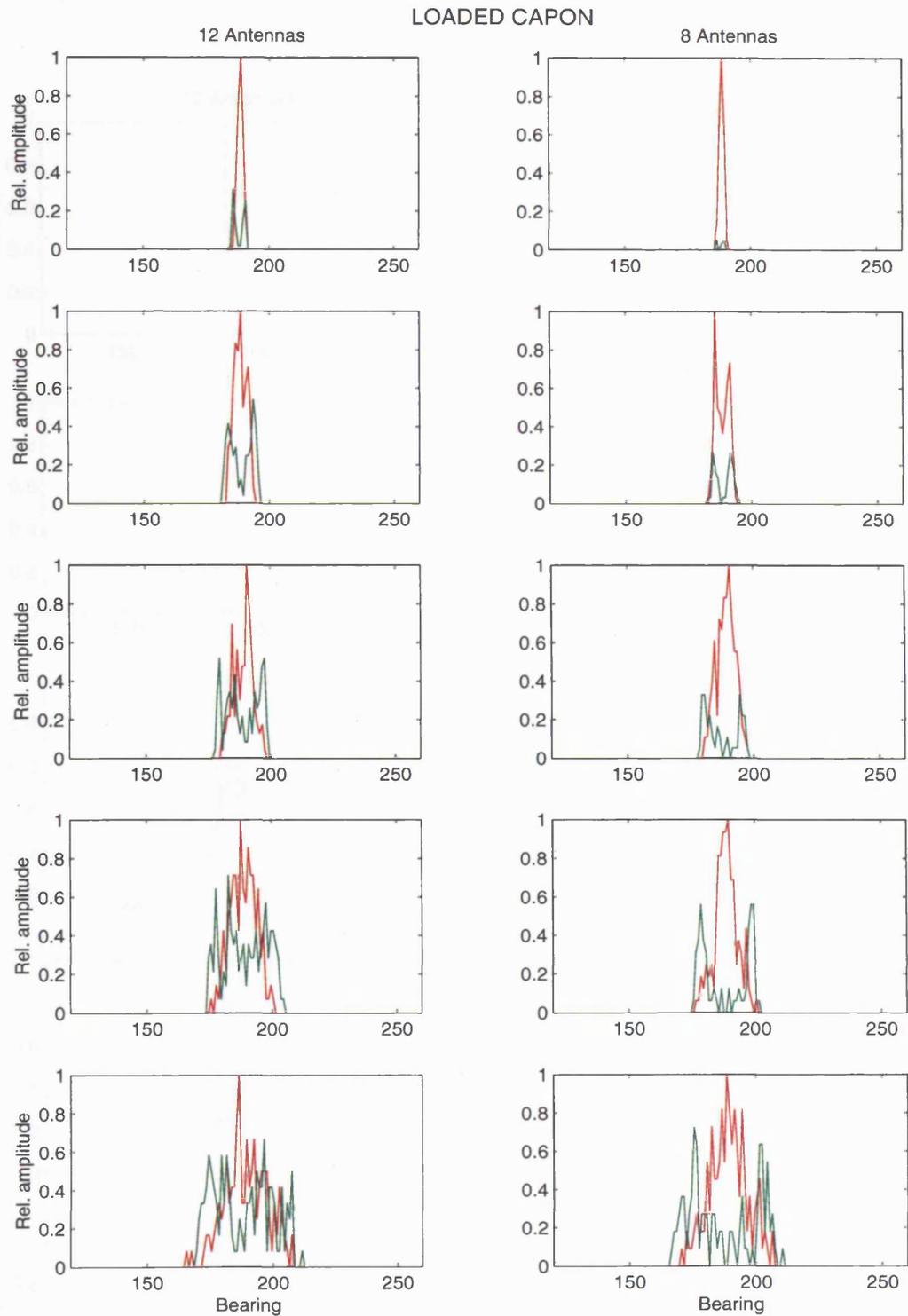


Fig 5.4.2a Bearing estimates distributions obtained from the symmetrical, single raised cosine distribution model, using the Loaded Capon algorithm with 12 antennas (left hand panels) and 8 antennas (right hand panels). The distribution widths in degrees are, top to bottom, $\pm 5^\circ$, $\pm 10^\circ$, $\pm 15^\circ$, $\pm 20^\circ$ and $\pm 30^\circ$. (az peak=188.5°, az Doppler spread=10.4Hz, el peak=15°, el spread= $\pm 2.5^\circ$ and el Doppler spread=2.4Hz). The red curves refer to the primary estimates and the green curves to the secondary estimates.

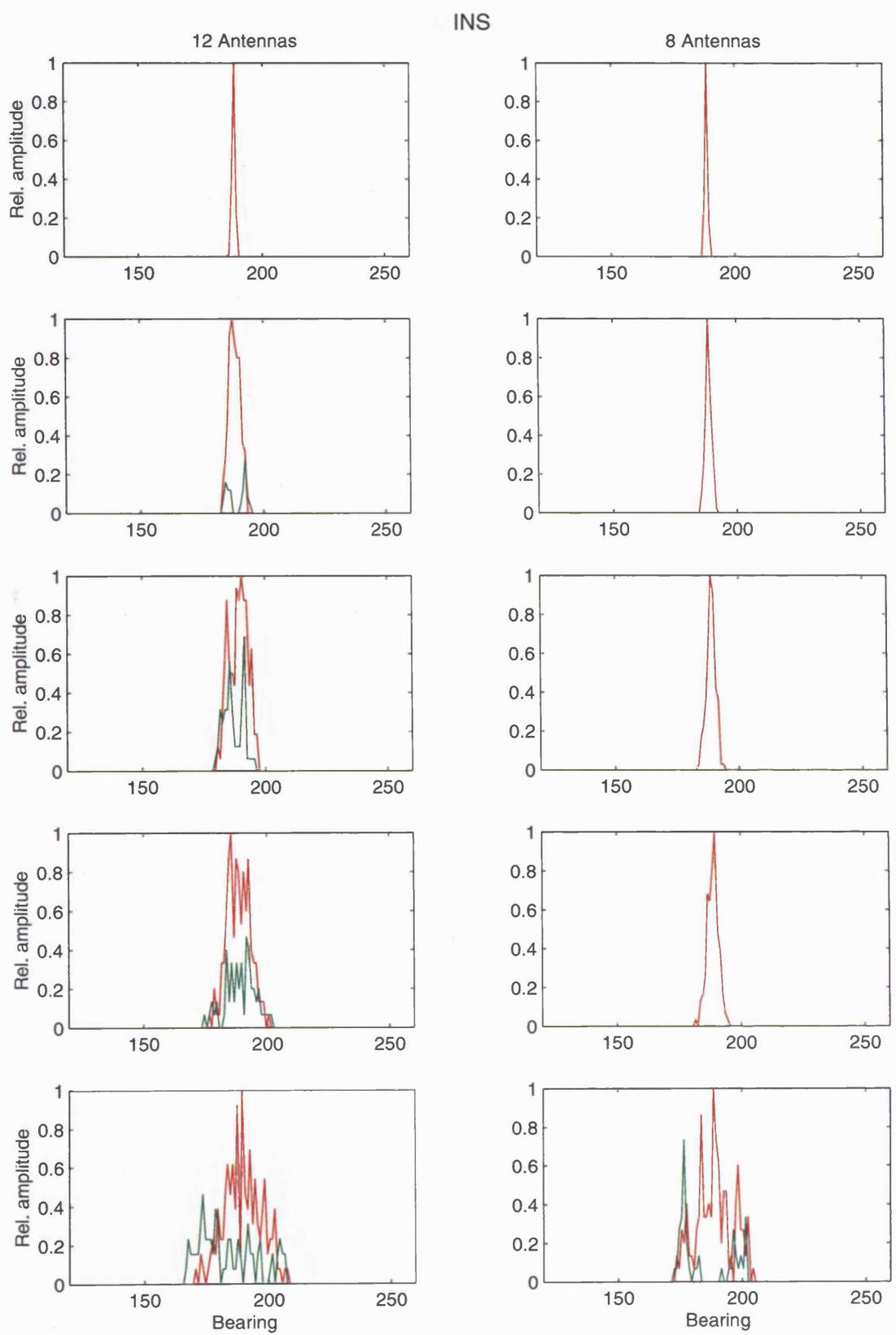


Fig 5.4.2b As for Fig. 5.4.2a except for the INS algorithm.

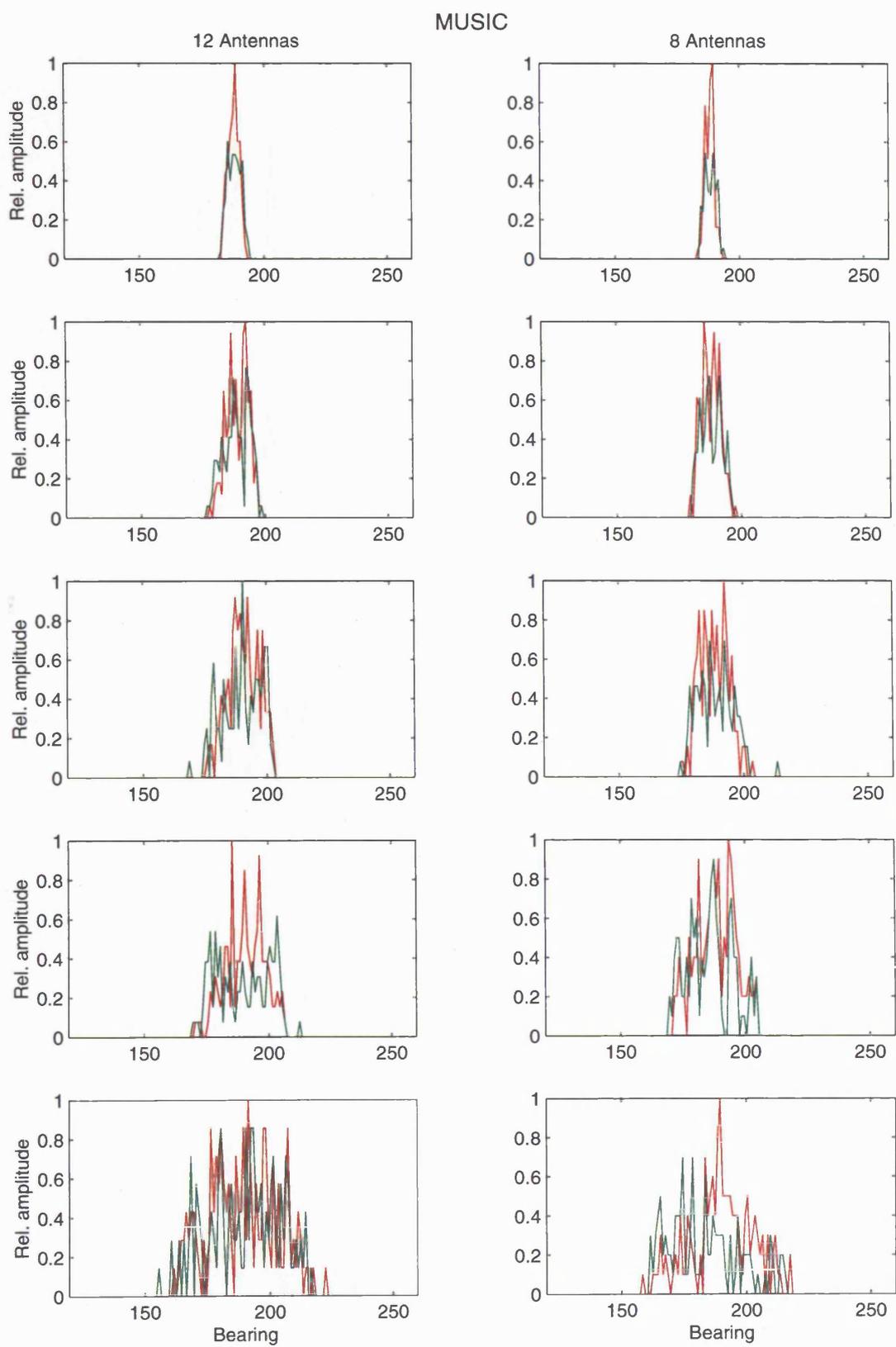


Fig 5.4.2c As for Fig. 5.4.2a except for the MUSIC algorithm.

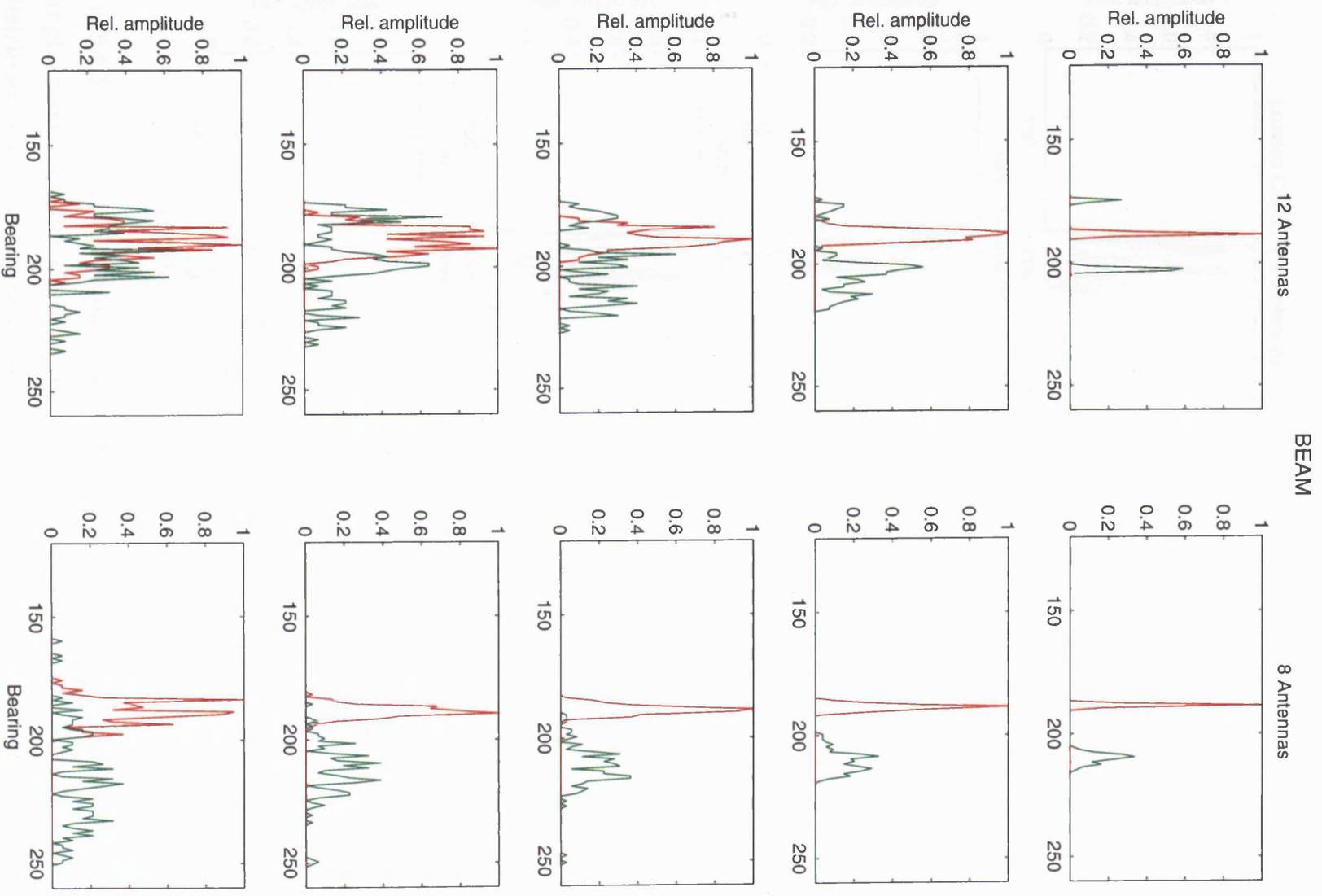


Fig 5.4.2d As for Fig. 5.4.2a except for the Beam algorithm.

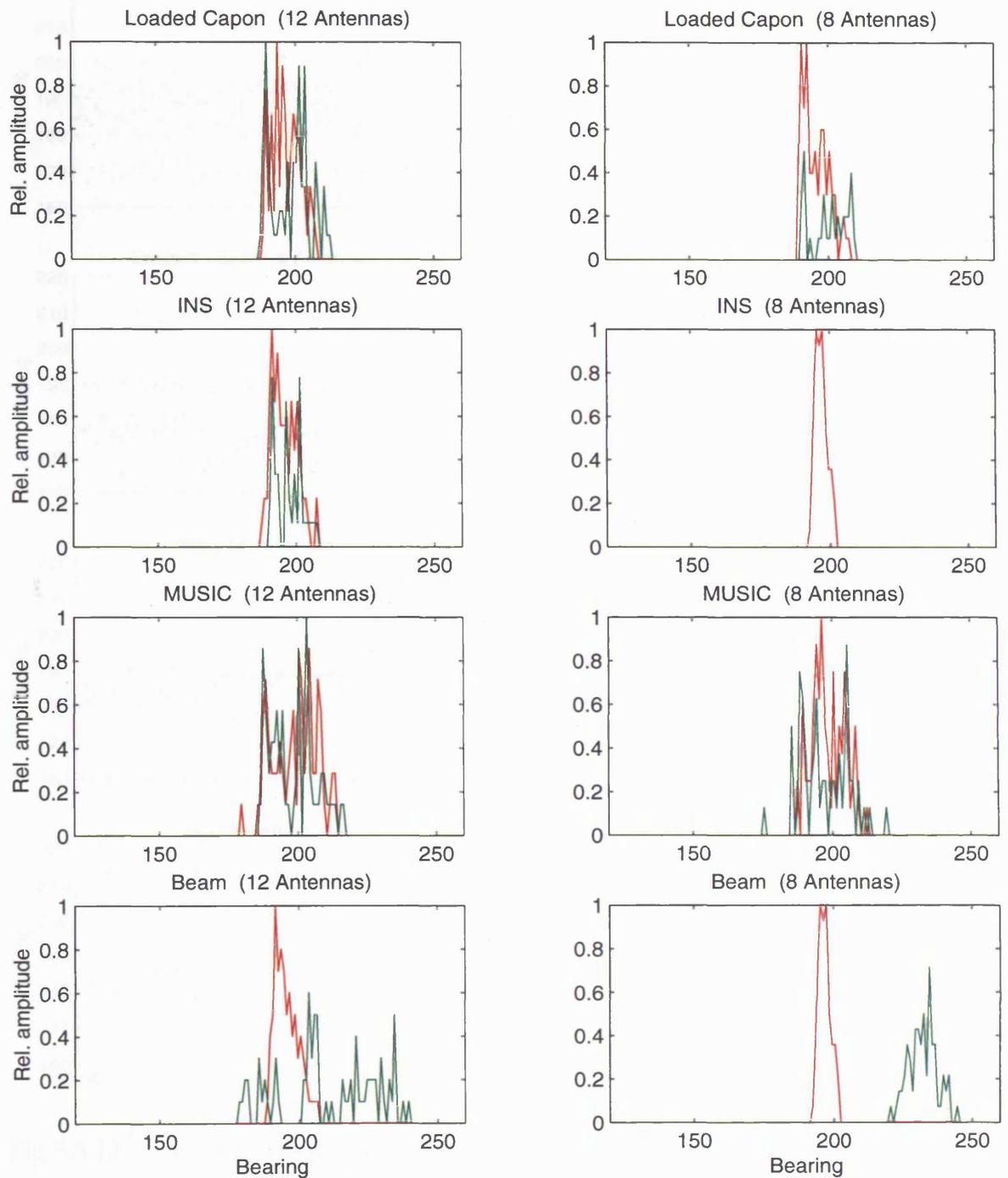


Fig 5.4.3 Bearing estimates distributions obtained from the asymmetrical, single raised cosine distribution model, using all four algorithms. The distribution widths are all equal to -2° to $+30^\circ$, i.e. a portion of the power is between 186.5° and the peak at 188.5° and a portion between 188.5° and 218.5° . (az Doppler spread=10.4Hz, el peak= 10° , el spread= $\pm 7.5^\circ$ and el Doppler spread=2.4Hz). The red curves refer to the primary estimates and the green curves to the secondary estimates.

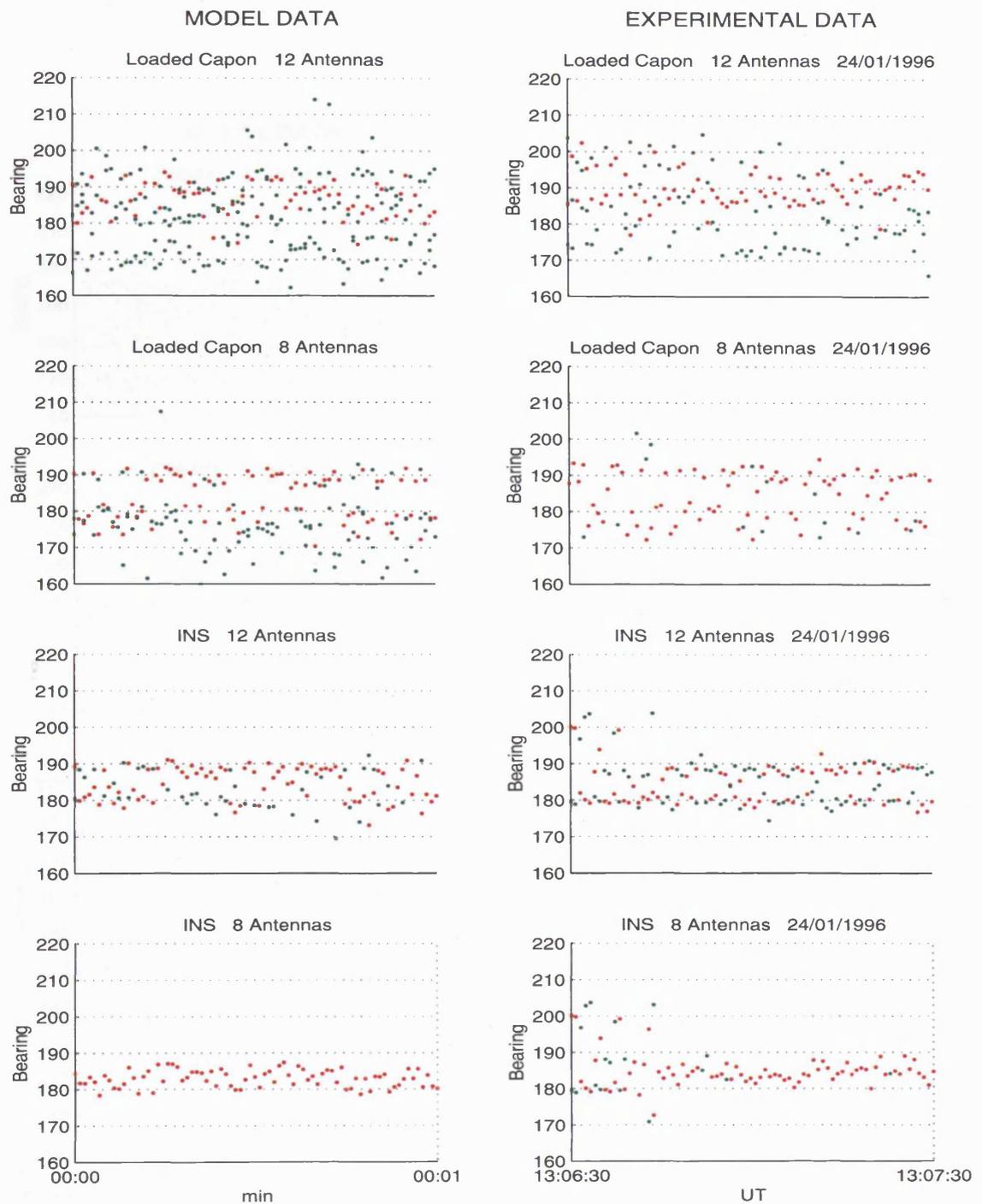


Fig 5.5.1a Comparison between bearing estimates obtained from the asymmetrical, single raised cosine distribution model (left hand panels) and experimental data (right hand panels), obtained over one minute with the Loaded Capon and the INS algorithms. For the model data the distribution half-widths are equal to -28° to $+8^\circ$, i.e. a portion of the power is between 160.5° and the peak at 188.5° and a portion between 188.5° and 196.5° . (az Doppler spread=10.4Hz, el peak= 15° , el spread= $\pm 2.5^\circ$ and el Doppler spread=2.4Hz). The red traces refer to the primary estimates and the green traces to the secondary estimates.

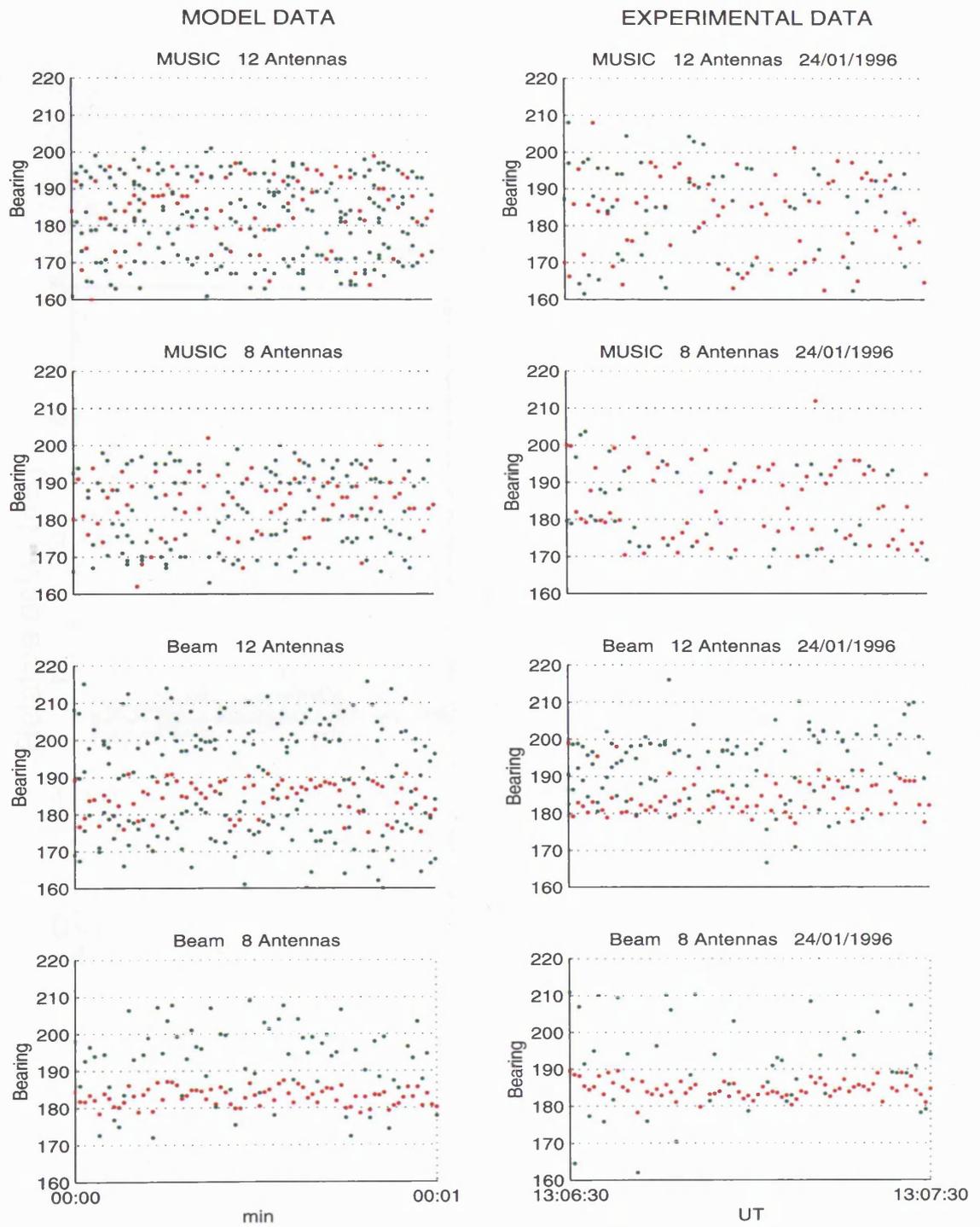


Fig 5.5.1b As for Fig. 5.5.1a except for the MUSIC and Beam algorithms.

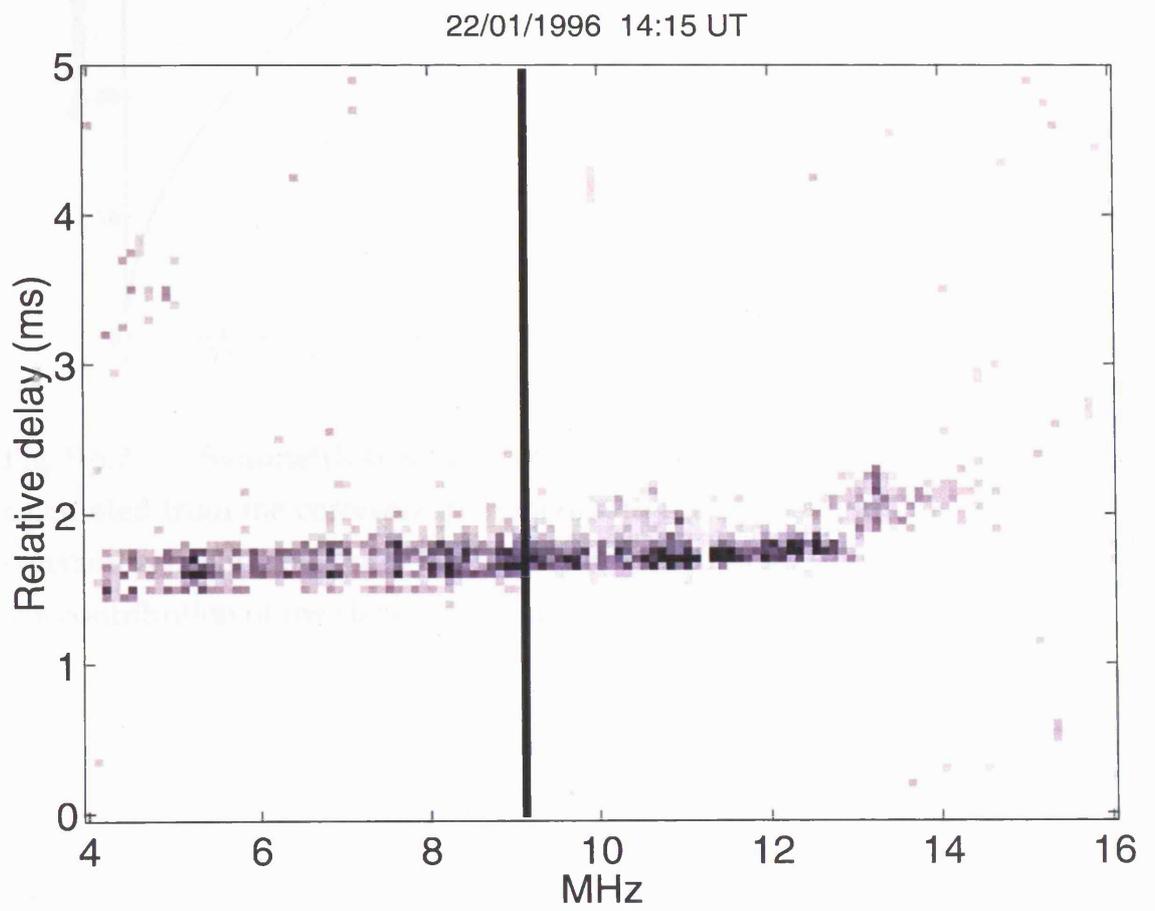


Fig. 5.6.1 Oblique ionogram obtained over the Iqaluit-Alert path.

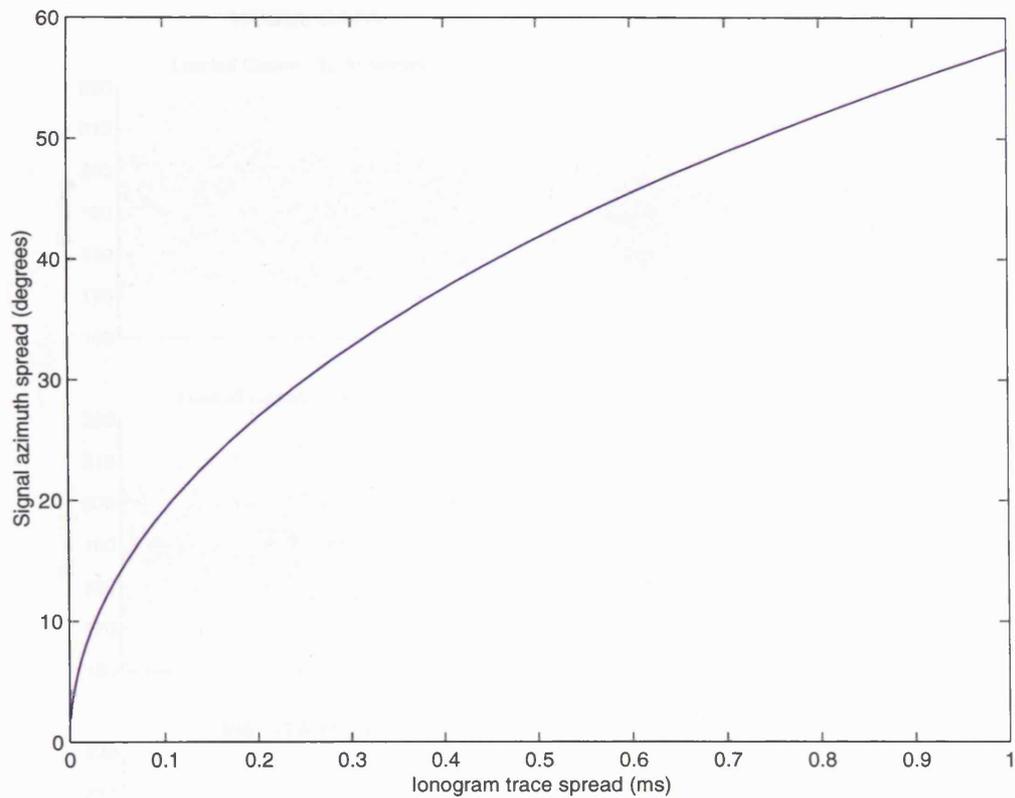


Fig. 5.6.2 Symmetrical signal azimuth spread around the true bearing calculated from the corresponding ionogram trace spread, assuming flat Earth, constant reflection height of 200 km at mid-path point (1-hop F) and neglecting the contribution of the elevation to the ionogram trace spread.

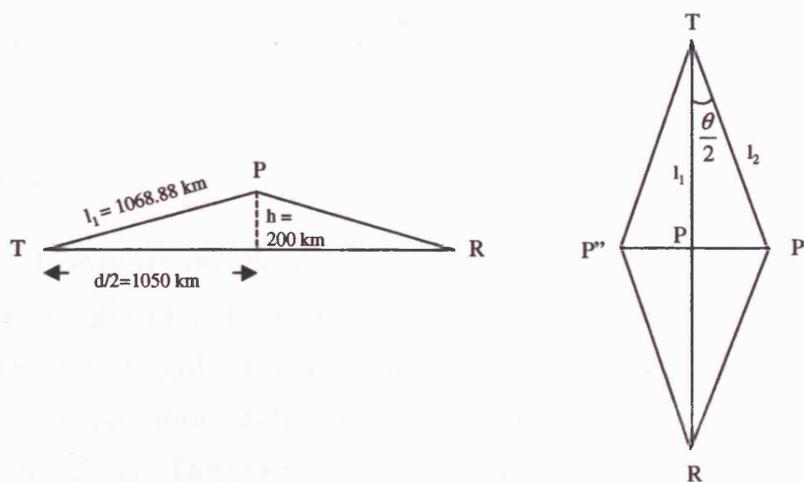


Fig. 5.6.3 Drawing illustrating the method for relating the azimuth spread of an incoming signal to the ionogram trace, with the assumptions stated for Figure 5.6.2.

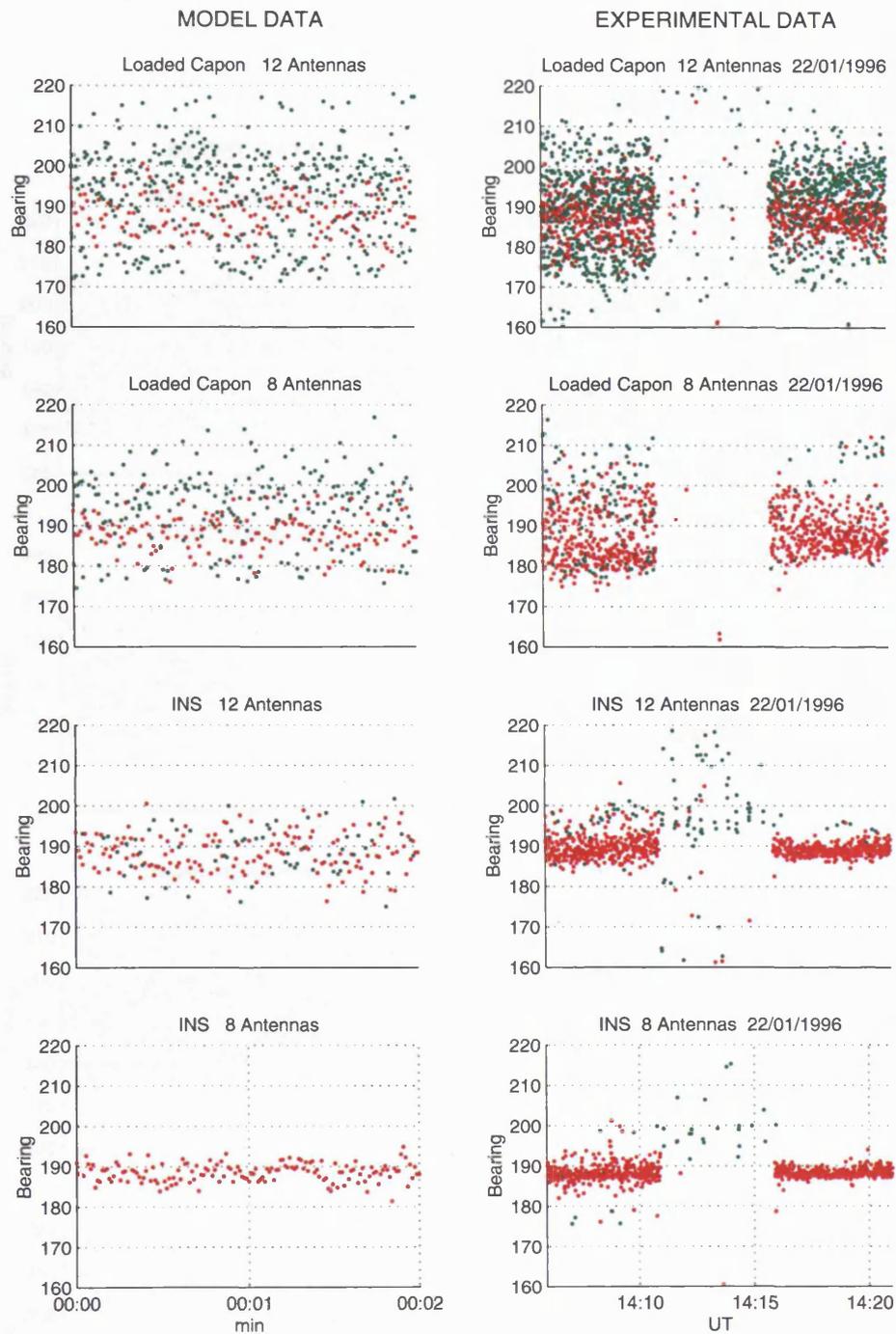


Fig 5.6.4a Comparison between bearing estimates obtained from the symmetrical, single raised cosine distribution model (left hand panels) and experimental data (right hand panels), with the Loaded Capon and the INS algorithms. For the model data the distribution width is equal to $\pm 20^\circ$ around the peak at 188.5° . (az Doppler spread=10.4Hz, el peak=15, el spread= ± 2.5 and el Doppler spread=2.4Hz). The experimental data plots refer to a time interval around the 5-minute sounder slot when the ionogram of Figure 6.5.1 was obtained. The red traces refer to the primary estimates and the green traces to the secondary estimates.

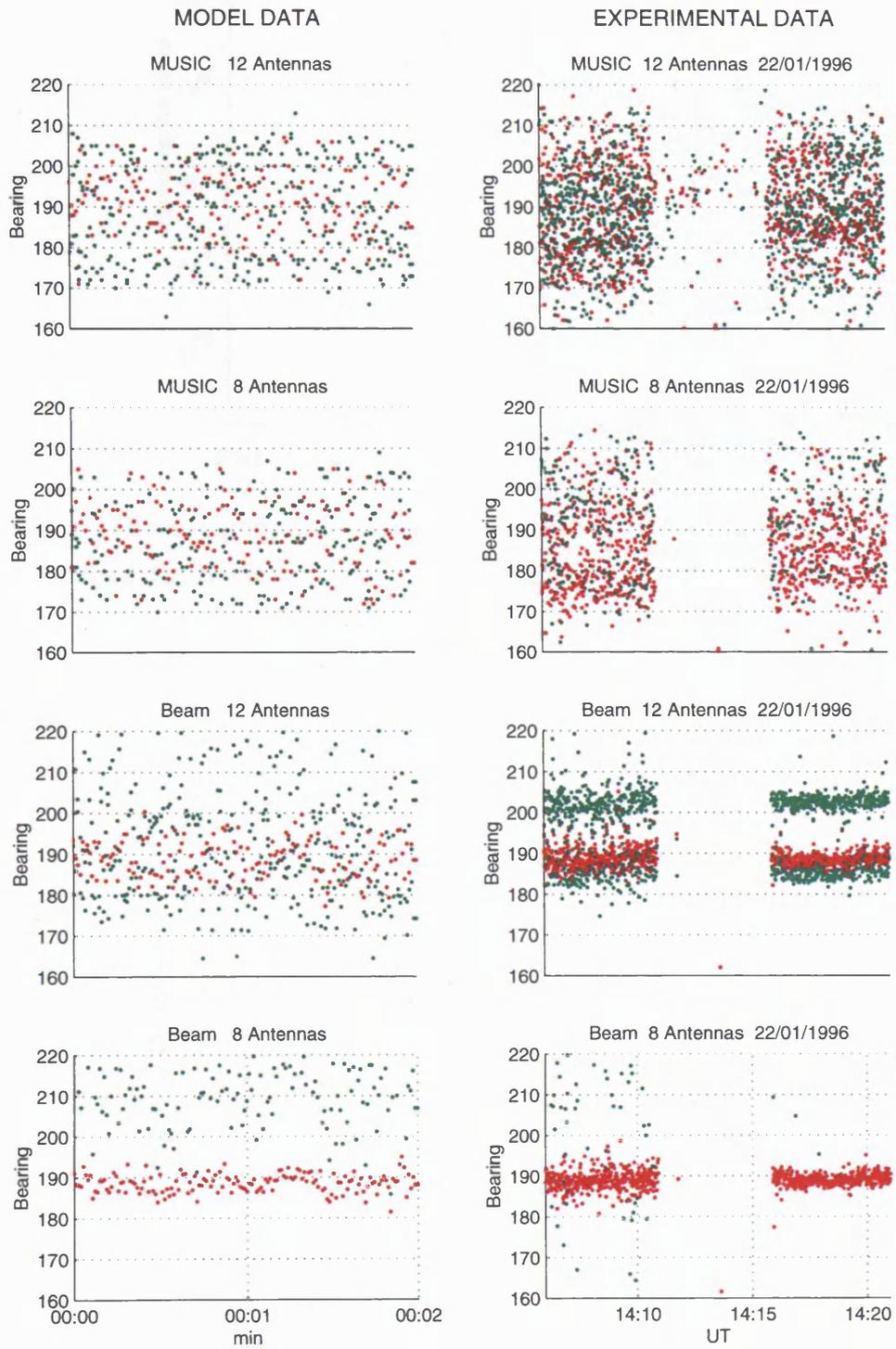


Fig 5.6.4b As for Fig. 5.6.4a except for the MUSIC and Beam algorithms.

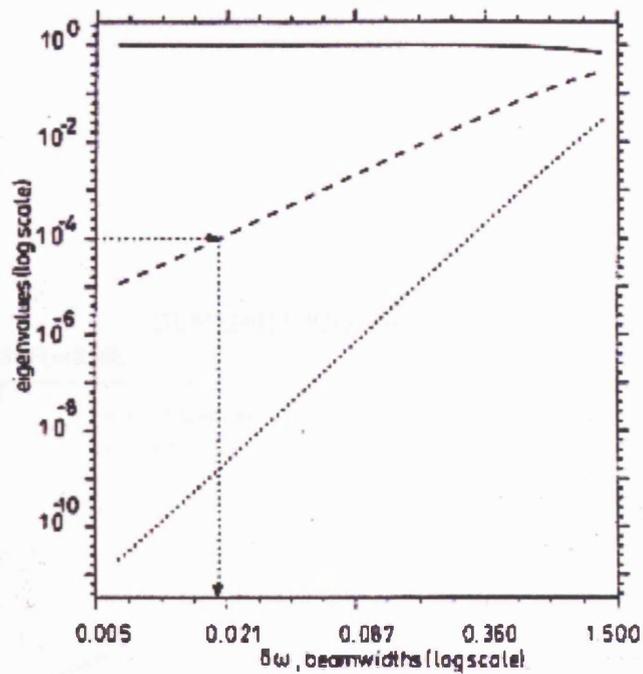


Fig. 5.7.1a Three eigenvalues of the covariance matrix as function of the angular spread (in beamwidths) of a spatially distributed source (Hayward, 1997).

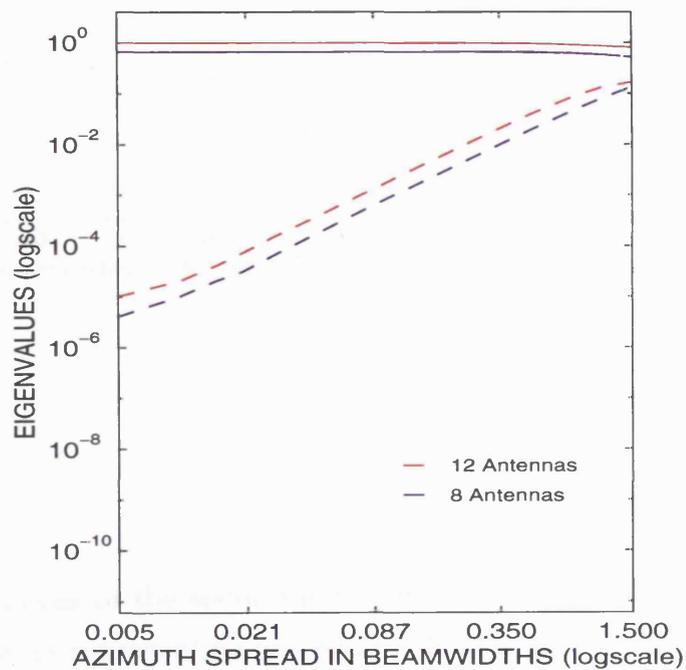


Fig. 5.7.1b First eigenvalue (solid lines) and second eigenvalue (dashed line) of the covariance matrix as function of the angular spread (in beamwidths) of a spatially distributed source for the full 12-antenna array (red curves) and the 8-antenna sub-array (blue curves).

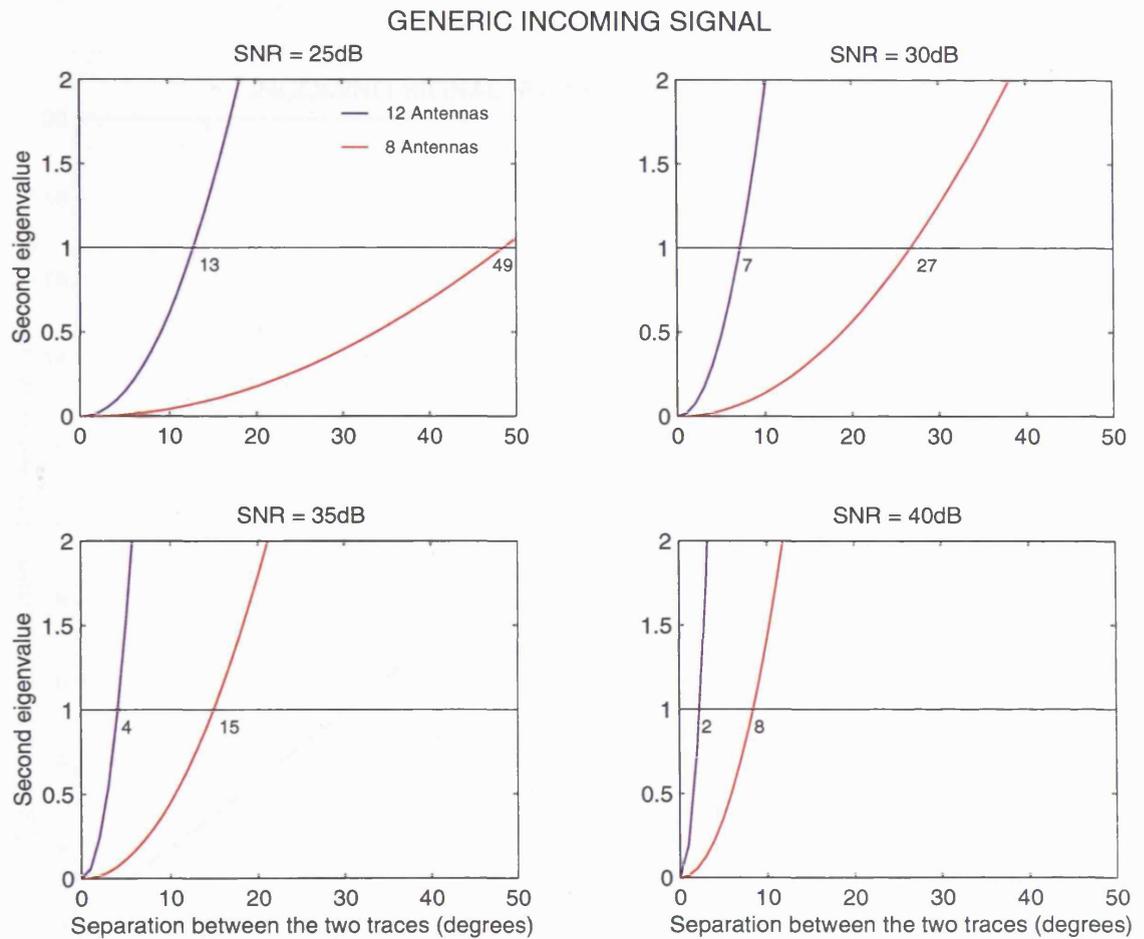


Fig. 5.7.2 Curves of the second eigenvalue as a function of the separation between two traces representing one generic incoming signal spread in azimuth around 188.5° , for four values of SNR (25, 30, 35 and 40dB) and both the full 12-antenna Vortex array (blue curves) and the 8-antenna sub-array (red curves). The elevation angle of arrival is equal to 15° . The numbers below the horizontal line indicate the crossing point, i.e. the minimum azimuth separation values (degrees) at which a two-trace representation of one spread signal could occur.

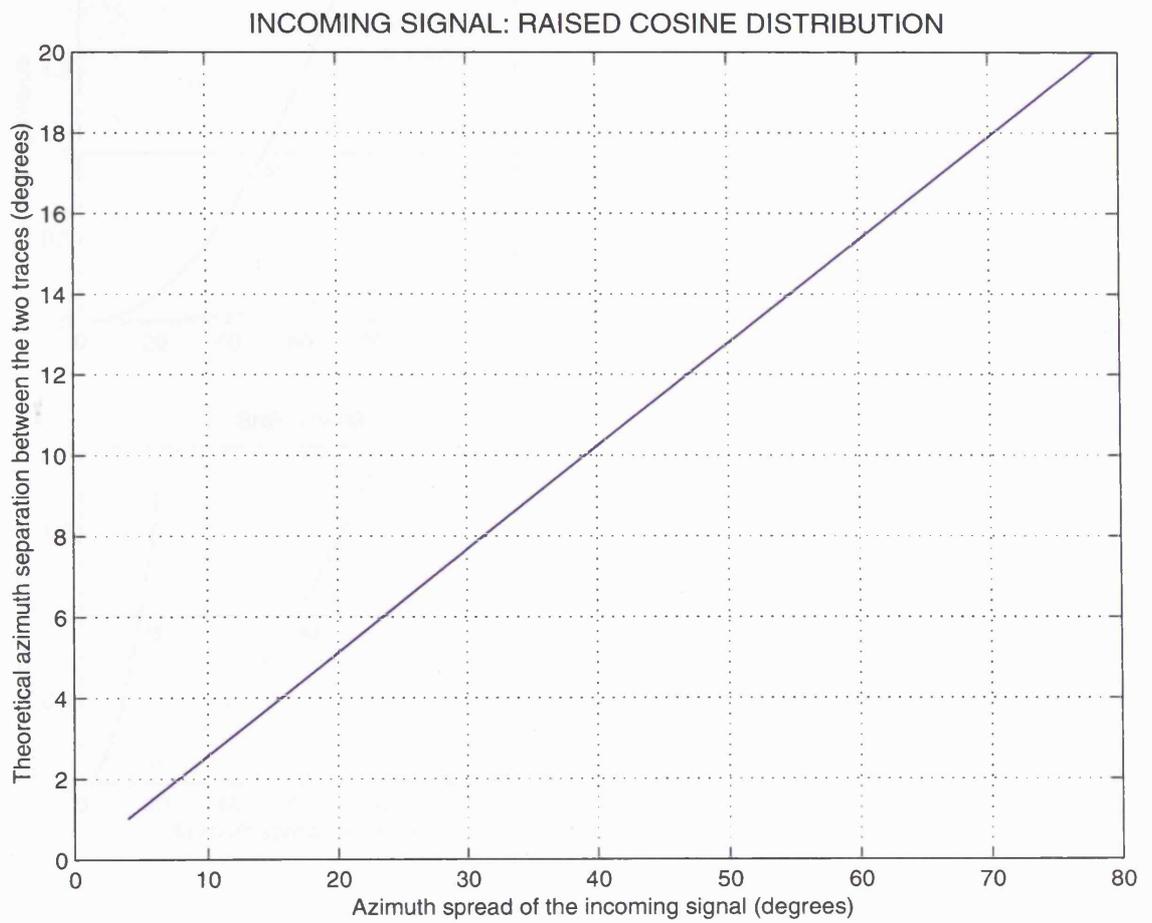


Fig. 5.7.3 Theoretical azimuth separation between two traces representing one spread incoming signal, as a function of its azimuth spread. The signal is assumed to be characterised by a raised cosine distribution.

6. MODELLING OF TWO SPREAD SIGNALS

6.1 Introduction

The observations presented in Chapter 3 and interpreted in Chapter 4 often showed evidence of multiple signals, e.g. interfering signals which arrived at the receiver from different directions.

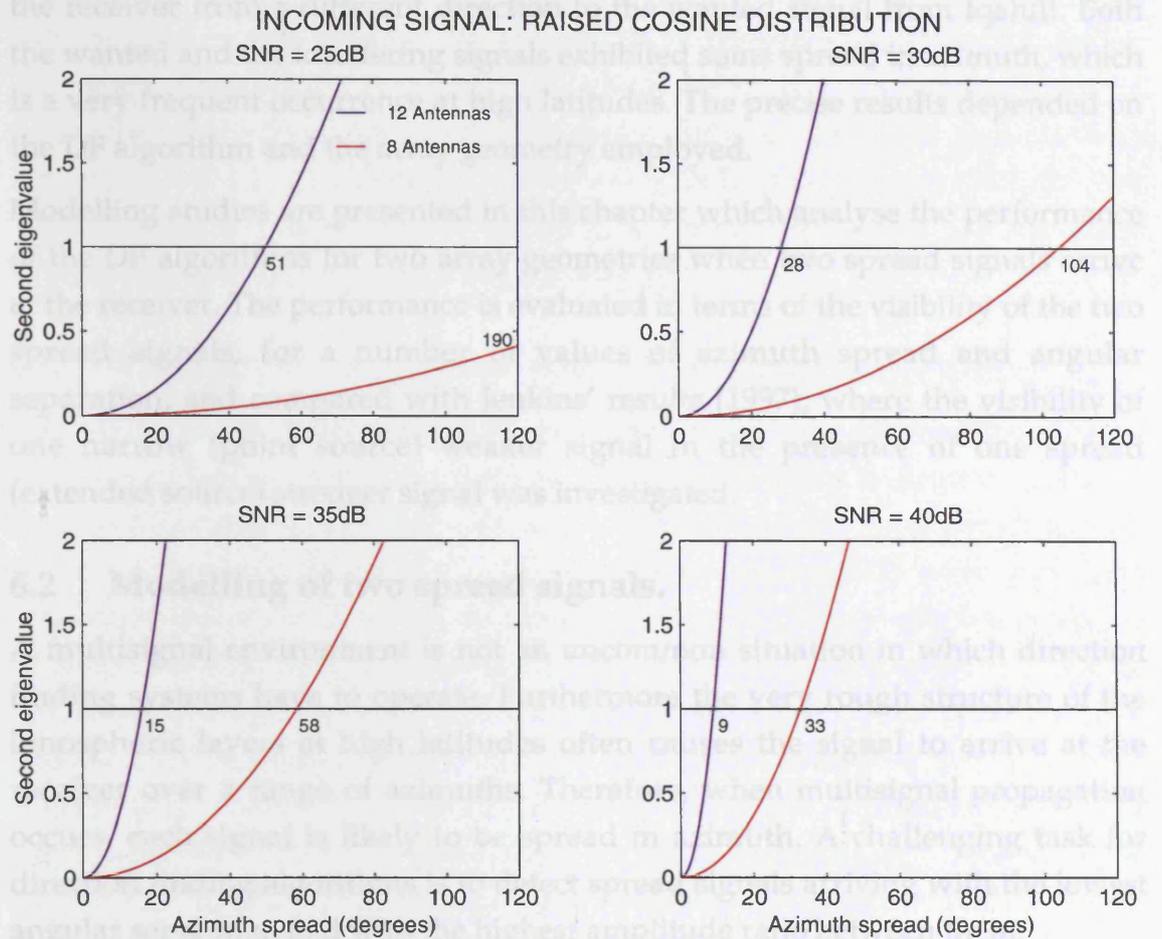


Fig. 5.7.4 Curves of the second eigenvalue as a function of the azimuth spread of an incoming signal (around 188.5°), characterised by a raised cosine distribution, for four values of SNR (25, 30, 35 and 40dB) and both the full 12-antenna Vortex array (blue curves) and the 8-antenna sub-array (red curves). The elevation angle of arrival is equal to 15° . The numbers below the horizontal line indicate the crossing point, i.e. the minimum azimuth spread values (degrees) at which a two-trace representation of one spread signal could occur.

6. MODELLING OF TWO SPREAD SIGNALS

6.1 Introduction.

The observations presented in Chapter 3 and interpreted in Chapter 4 often showed evidence of multiple signals, e.g. interfering signals which arrived at the receiver from a different direction to the wanted signal from Iqaluit. Both the wanted and the interfering signals exhibited some spread in azimuth, which is a very frequent occurrence at high latitudes. The precise results depended on the DF algorithm and the array geometry employed.

Modelling studies are presented in this chapter which analyse the performance of the DF algorithms for two array geometries when two spread signals arrive at the receiver. The performance is evaluated in terms of the visibility of the two spread signals, for a number of values of azimuth spread and angular separation, and compared with Jenkins' results (1997), where the visibility of one narrow (point source) weaker signal in the presence of one spread (extended source) stronger signal was investigated.

6.2 Modelling of two spread signals.

A multisignal environment is not an uncommon situation in which direction finding systems have to operate. Furthermore the very rough structure of the ionospheric layers at high latitudes often causes the signal to arrive at the receiver over a range of azimuths. Therefore, when multisignal propagation occurs, each signal is likely to be spread in azimuth. A challenging task for direction finding algorithms is to detect spread signals arriving with the lowest angular separation and with the highest amplitude ratio between them.

The case where point source and extended source signals occurred was modelled by Jenkins (1997), where the point source would represent a sporadic-E reflection and the extended source would represent an F region reflection caused by a patch of enhanced ionisation (see Section 2.9). Jenkins investigated the visibility of the weaker narrow signal in presence of the stronger spread signal with two algorithms and four array geometries (see Section 2.9). Jenkins assumed that the weaker signal came from a point source. However, at high latitudes this is often not the case, and for this reason the results presented in this chapter were obtained by modelling two signals which both originate from spread sources. Four algorithms (Loaded Capon, INS, MUSIC and Beam) and both the 8-antenna and 12-antenna arrays have been used. In some of the processing runs the two signal distributions overlapped, whereas the minimum angular separation for Jenkins' modelling studies was equal to 5° .

In the simulations whose results are shown in the next section, the same azimuth spread value is given to both signals for each run. The angular separation between the peaks of the two signals has a number of values ranging from 0.2° to 50° . The distributions of the signals, with widths ranging between 0.2° and 30° are all assumed to be symmetrical. The peak of the strongest signal was fixed at 188.5° (i.e. the true bearing). The weaker signal was arbitrarily assumed to arrive at bearings lower than 188.5° . For example, when the angular separation between the two signals was fixed at 20° , the weaker signal would arrive from 168.5° . Four amplitude ratios between the two signals were employed, i.e. 1:1, 2:1, 5:1 and 10:1. The azimuth Doppler spread (10.4 Hz), the elevation (10°), the elevation spread (2° , symmetrical) and the elevation Doppler spread (2.4 Hz) were fixed for both signals.

6.3 Visibility of two signals with varying amplitude ratio.

The results of the simulations presented in this section are different for the four employed algorithms, and for both the 12-element and the 8-element arrays, as expected. Bearing estimates for 10° azimuth spread and 20° angular separation are illustrated in Figure 6.3.1, in which the two signals have the same amplitude. All algorithms are able to detect the two traces, with the exception of Beam with 8 antennas. However, the traces generally exhibit different spreads, as emphasised by the standard deviations values (first decimal precision). Loaded Capon using the sub-array and INS (with a 10dB termination threshold) using the full array show the narrowest traces. The mean azimuth values are generally separated by about $1-2^\circ$ less than the distribution peaks. In Figure 6.3.2 are shown the plots obtained with the same angular separation as in Figure 6.3.1, but much higher azimuth spread, 30° for both distributions instead of 10° . We can immediately notice that INS now performs better than Loaded Capon. The clearest traces are obtained by means of INS when 8 antennas are in use. Loaded Capon shows very spread traces with the full 12-antenna array whereas with the 8-antenna sub-array it exhibits a narrow trace around the true bearing while the other trace around 20° below the true bearing is fairly spread. Furthermore, the latter trace appears to merge with some spread energy coming from $150-160^\circ$ (possibly caused by sidelobe excitation, see Figure 3.2.4), which is an error not occurring with INS. MUSIC fails to detect the two signals with either array geometry. Beam shows a reasonable performance in the detection of the two energies with the full array, but the two mean azimuths are separated by only 15° , i.e. 5° less than the separation between the distribution peaks.

While the separation between the traces is generally slightly lower than the separation between the distribution peaks for relatively low values of azimuth spread of the two distributions (Figure 6.3.1), it is higher when the azimuth spread is set at 30° for both Loaded Capon and INS (Figure 6.3.2). A number of simulations with the incoming signals at various DOAs would be needed for an understanding of this behaviour, and in particular how the beam pattern would influence the results, as its shape and sidelobes will differ according to the look direction, because of the irregularly spaced deployment of the antenna array.

The results of this investigation for two signals with the same amplitude are summarised in Figure 6.3.3, which illustrates the 'visibility curves', i.e. when the two signals are identifiable in the simulated data, for all algorithms with the full 12-antenna array (upper panel) and with the 8-antenna sub-array (lower panel). The stronger signal was always at a bearing of 188.5° (true bearing for the I-A path) for all processing runs, while the weaker signal was offset from this by various amounts. Loaded Capon and MUSIC perform the best, i.e. the angular separation is the lowest, for very low azimuth spreads, with both array geometries. When the azimuth spread of the two distributions increases above 2° , the performance of MUSIC rapidly deteriorates. MUSIC clearly performs well only when the two signals are unspread, which is not expected to be a common situation at high latitudes. In the upper panel (12 antennas), it stands out that for spreads below 15° the performances of Loaded Capon (best), INS (second best) and Beam (third) are fairly constant. With both 12 and 8 antennas Loaded Capon performs the best up to $15\text{-}20^\circ$ of azimuth spread and INS for higher values.

Figure 6.3.4 illustrates the visibility curves for signal amplitude ratios equal to 2:1, 5:1 and 10:1, at various values of azimuth spread (0.2° to 15° , equal for both distributions) and angular separation (1° to 30°). INS has been used with different termination thresholds in the various cases, in order to make possible the detection of the weaker trace. Amplitude ratios of 2:1, 5:1 and 10:1 correspond to 6dB, 14dB and 20dB respectively, thus INS threshold has been set to 10dB, 18dB and 25 dB respectively. Loaded Capon generally performed slightly better than INS with both array geometries. MUSIC showed good performances for very low azimuth spread values (a few degrees) and was generally more sensitive to the weaker trace when the sub-array was in use, rather than the full array. However, MUSIC exhibited more spread traces, as usual. Since Beam never detected the weaker trace, it has been omitted from Figure 6.3.4.

The behaviour of INS for three thresholds (6dB, 14dB and 20dB below the peak: 1st, 2nd and 3rd column respectively), and two amplitude ratios between the two

signals (2:1, 1st and 2nd rows, and 5:1, 3rd and 4th rows) is compared in Figure 6.3.5. The azimuth spread for both signals is equal to 15° and the angular separation is 30°. Amplitude ratios of 2:1 and 5:1 imply that INS would work at its limits with a 6dB and 14dB threshold respectively. The results show that with a 6dB threshold and an amplitude ratio of 2:1 (left panels, 1st and 2nd rows), INS detects a very weak secondary signal (more apparent with 12 antennas), which clearly shows that INS is working at its limits. When the threshold is equal to 14dB (middle panels, 1st and 2nd rows), INS is capable of detecting both traces. However, the stronger signal exhibits two closely separated traces (12 antennas), which appears to be an artefact of the algorithm. The occurrence of this phenomenon is reminiscent of the parallel closely separated traces observed in both experimental (Chapters 3 and 4) and simulated data (Chapter 5) in case of diffuse energy arriving at the receiver. It is important to emphasise that these closely separated traces did not occur for azimuth spread values lower than 15°. The modelled data, once again, show this phenomenon with INS as well as with Loaded Capon (not shown in this section). With a 20dB threshold (right panels, 1st and 2nd rows), the higher amount of detected energy causes the two closely separated traces to occur also with the 8-antenna sub-array for the stronger signal, and there is some evidence of the same phenomenon even for the weaker signal with the full array.

In the 3rd and 4th rows the amplitude ratio is 5:1 (i.e. about 14dB), and therefore INS with a 6dB termination threshold cannot detect the weaker trace (left panels). In the middle panels the threshold is 14dB, therefore INS is again working at its limits and it only detects a few secondary bearings with the full 12-antenna array, whereas with 8 antennas it detects some spread energy at around 170°, probably originated by sidelobe excitation from the signals. If a threshold of 20dB is employed (right panels), the weaker signal becomes visible with 12 antennas and some bearings are also detected with 8 antennas. Two closely separated traces are apparent when 12 antennas are in use with a 14dB and 20dB threshold. The mean values, shown in the panels, are calculated by including both primary and secondary bearing estimates. These mean values are considerably more accurate than those which could be obtained from the primary bearing estimates only. For example, a simple look at the middle and right panels in the 1st row shows that when two closely separated traces are present the primary trace is at about 191-192°, whereas the mean values are 189.1° and 188.8° respectively, i.e. much closer to the distribution peak at 188.5°. The same is true for all other circumstances in which two closely separated traces occur (see other panels).

The importance of the array aperture in terms of visibility of two signals under extreme situations is apparent in Figure 6.3.5 for INS. Likewise for Loaded Capon in Figure 6.3.6, which shows the effect of changing the angular separation from 10-15° for an amplitude ratio of 10:1. In the top left panel we can see the results obtained with full 12-antenna array with two signals with azimuth spread of 10° and separated by 10°. With a 10:1 amplitude ratio Loaded Capon cannot detect the weaker signal, although it shows evidence of some secondary bearings close to the true bearing of the weaker signal (178.5°). Some spread energy is evident between 150° and 170°, which does not correspond to any input energy but may correspond to sidelobe excitation by the two signals.

When the angular separation is increased by one degree to 11° (second left-hand panel), Loaded Capon is able to detect the weaker trace with the full array in use, although not at all times. However, the weaker trace is detected almost continuously when the angular separation is further increased to 12° (third left hand panel). The difference in the performance of Loaded Capon for angular separations of 10° and 12° with the full 12-antenna array is remarkable, whereas with only 8 antennas in use Loaded Capon performed poorly in both cases. However, Loaded Capon with 8 antennas starts exhibiting two traces for 15° of angular separation (bottom right panel), even though the weaker trace is very spread, whereas two clear traces are evident when the full 12-antenna array is in use (bottom left panel). It is evident that for these extreme conditions the array aperture is crucial for determining the visibility of the two signals.

It is interesting to compare the results in this section with those obtained by Jenkins (1997). The third left panel of Figure 6.3.6 is repeated in Figure 6.3.7 (lower panel), and compared with Jenkins' results (upper panel of Figure 6.3.7, repeated from the bottom left panel of Figure 2.9.2). The lower panel shows that Loaded Capon with 12 antennas detects the two signals for a 10:1 amplitude ratio (i.e. 20dB power ratio) when the peaks of these two signals, both spread by 10° in azimuth, are separated by 12° (188.5° and 176.5°). As both signals are characterised by the same values of azimuth and elevation spreads, their peak power ratio equals their integrated power ratio. The stronger signal is spread between 183.5-193.5° and the weaker between 171.5-181.5°, and hence the two distributions are separated by only 2°.

Jenkins' results in Figure 6.3.7 (upper panel) are obtained with the same antenna array (Vortex with all 12 antennas) and with the same power ratio between the two signals (20dB), as in the results of the lower panel of Figure 6.3.7, but with a different algorithm (deterministic Maximum Likelihood) (Jenkins, 1994). It is not possible to clearly distinguish the weaker trace for

power ratios greater than 20dB in either case. However, for Jenkins' case, the point-source signal arriving at 180° and the upper end of the spread-source signal arriving between 140-165° are separated by 15°. Therefore the results obtained in this investigation with Loaded Capon are significantly better than those obtained by Jenkins with the deterministic Maximum Likelihood algorithm, since both signals can be detected at much lower angular separations.

Jenkins also investigated the effect of four array geometries (including the Vortex array) in terms of visibility of the weaker source for different array apertures in wavelengths and for different angular separations (5°, 10° and 15°, see Figure 2.9.3). The values for the Vortex array for an aperture of 5.45λ (corresponding to aperture in wavelengths of the Vortex array for the 9.292 MHz signal sent from Iqaluit, see Section 3.2) have been extracted from Figure 2.9.3 and compared to the results obtained here, as shown in Figure 6.3.8. It is important to consider that, for consistency with Jenkins' results, in this case the azimuth spread separation is compared, i.e. the separation between the limits of the two spread signals, and not the separation between the two peaks of distribution, such as in Figures 6.3.3 and 6.3.4. Also, since Jenkins' investigation refers to the visibility of an unspread source in presence of a spread source having an azimuth width of 25°, it is reasonable to compare his results with those of this investigation with two spread sources having azimuth widths of 12.5° each. Figure 6.3.8 clearly shows that the results for this investigation with Loaded Capon (red curve) are far better than those obtained by Jenkins (blue curve), for the weaker source is detected for a much higher power ratio (about 25dB) with respect to the stronger one.

6.4 Summary.

The presence of multiple incoming signals is a common occurrence, especially at high latitudes, where the signals often arrive considerably spread at the receiver. This applies for both the wanted signal and possible interfering signals or other modes of the wanted signal. The modelling studies presented in this chapter reproduced some of the behaviour of the DF algorithms at high latitudes, when two spread signals arrive at the receiver.

The performance of the algorithms has been evaluated in terms of the visibility of the two signals, for a number of values of azimuth spread and angular separation between them. The DF algorithms have been taken to their limits in terms of the ability of detecting two spread signals arriving at the receiver with the lowest possible angular separation and with the highest possible amplitude ratio. The effect of the amplitude ratio between the two signals has been

simulated. When the signals have the same amplitude, Loaded Capon had the best performance at azimuth spreads of up to about 15-20°, whereas for higher values INS showed the best performance. For amplitude ratios between 2:1 and 10:1 Loaded Capon generally showed slightly better performances than INS, MUSIC (as expected) performed well for unspread signals only and Beam failed for all cases.

The simulations have shown the presence of parallel closely separated traces (often observed in the experimental data, see Chapters 3 and 4), which appear to arise as artefacts of the algorithm, both with the Loaded Capon and the INS algorithms. The estimated direction of arrival of the signal is more accurate when the bearings belonging to both traces are included to obtain the mean value, both for Loaded Capon and INS.

Results obtained for the visibility of two signals for a 10:1 amplitude ratio with Loaded Capon have been compared with Jenkins' results (1997) with the deterministic Maximum Likelihood algorithm. While Jenkins considered a weaker point source signal in the presence of a stronger spread source signal, in the modelling studies presented in this chapter both signals are assumed spread. Since much less angular separation is required for detecting both signals with Loaded Capon, the results of this investigation are significantly better than those obtained by Jenkins with the deterministic Maximum Likelihood algorithm. Furthermore, by using the same 12-antenna array (Vortex), for the same value of angular separation between the two sources, the results of this research by far exceeded Jenkins' results in that the weaker signal was visible for a much higher power ratio with respect to the stronger source.

2 DISTRIBUTIONS: Amplitude ratio 1:1 Azimuth spread=10 Angular separation=20

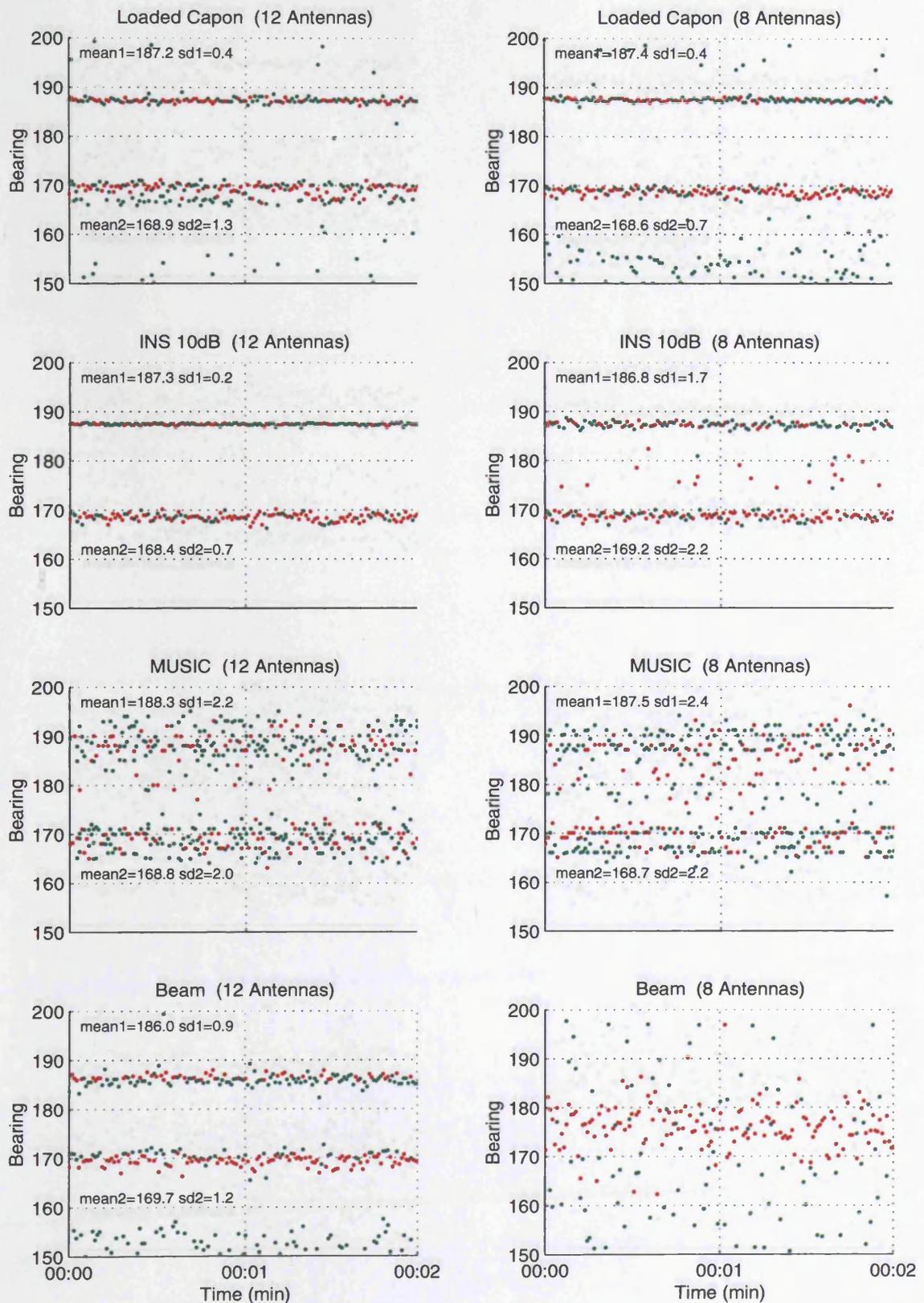


Fig. 6.3.1 Bearing estimates obtained from the double distribution model. Amplitude ratio 1:1. Az. Spread 10°. Distributions peaks at 188.5° and 168.5°.

2 DISTRIBUTIONS: Amplitude ratio 1:1 Azimuth spread=30 Angular separation=20

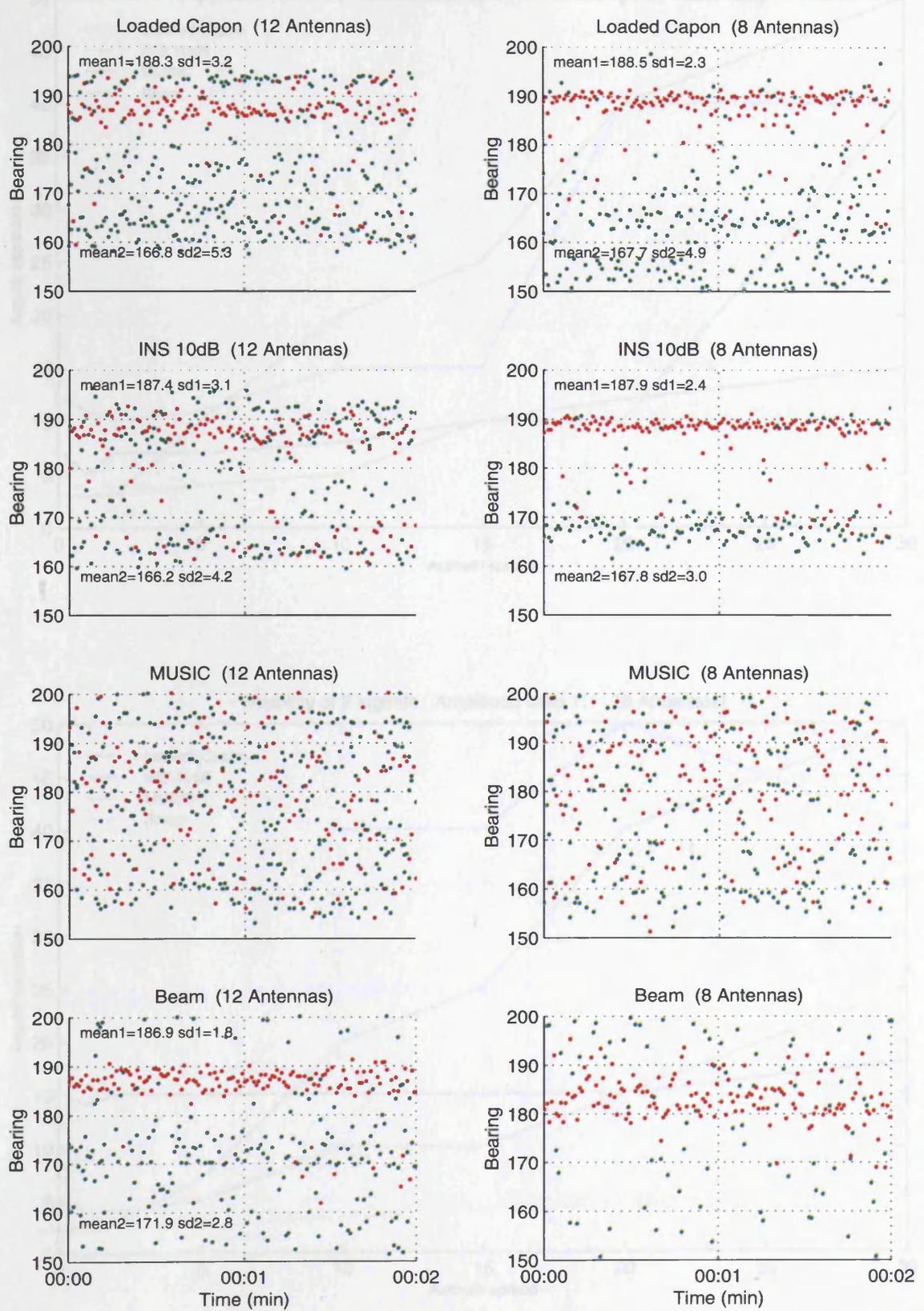


Fig. 6.3.2 Bearing estimates obtained from the double distribution model. Amplitude ratio 1:1. Az. Spread 30°. Distributions peaks at 188.5° and 168.5°.

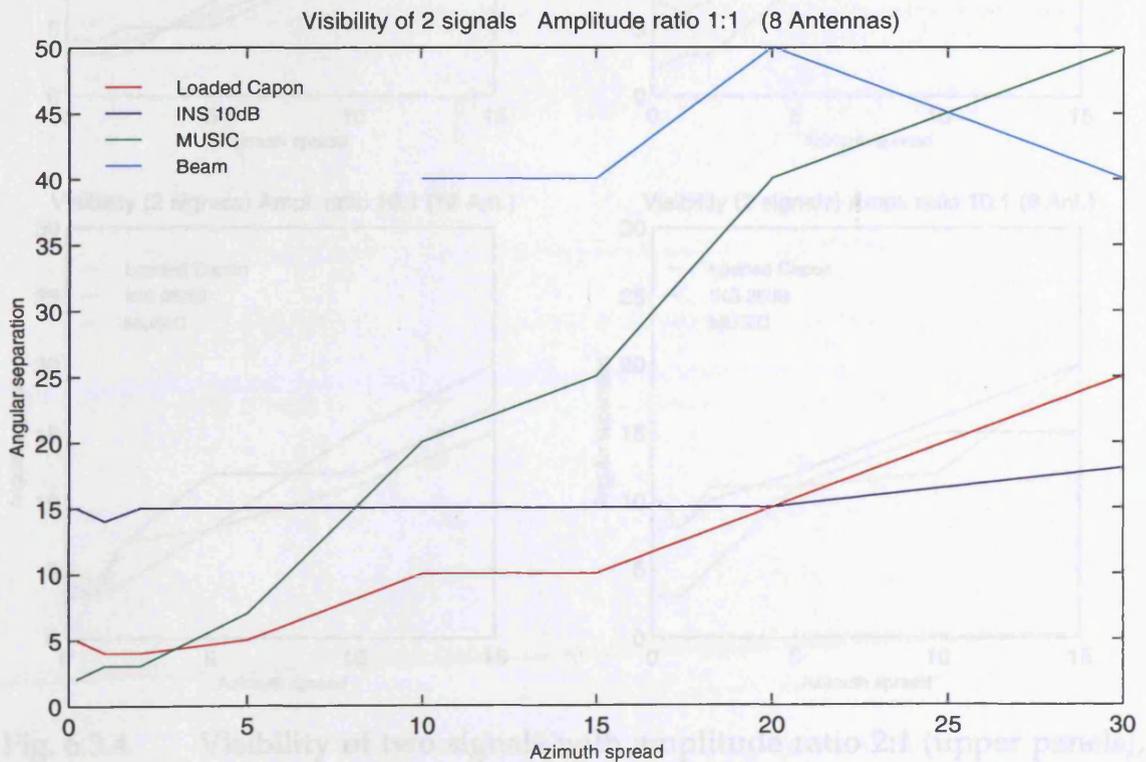
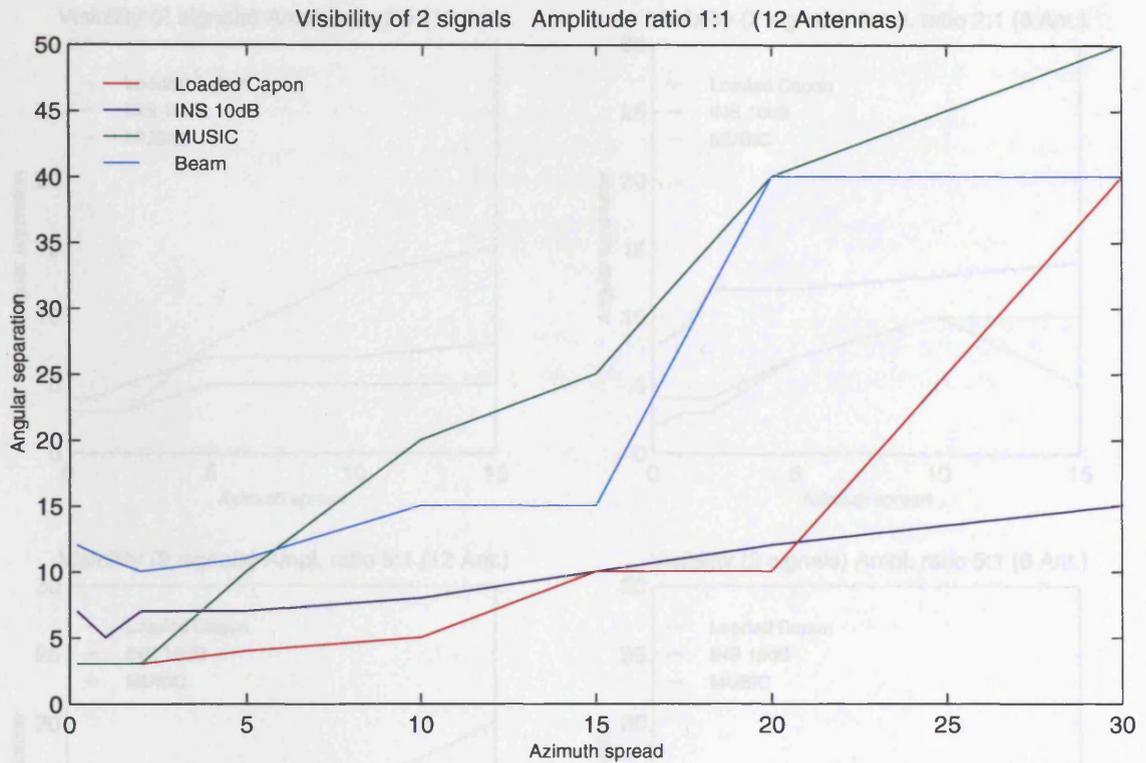


Fig. 6.3.3 Visibility of two signals with amplitude ratio 1:1. The curves represent the minimum angular separation values between the peaks of the two distributions, at which two traces are detected, for different values of azimuth spread (the same for both signals at each processing run).

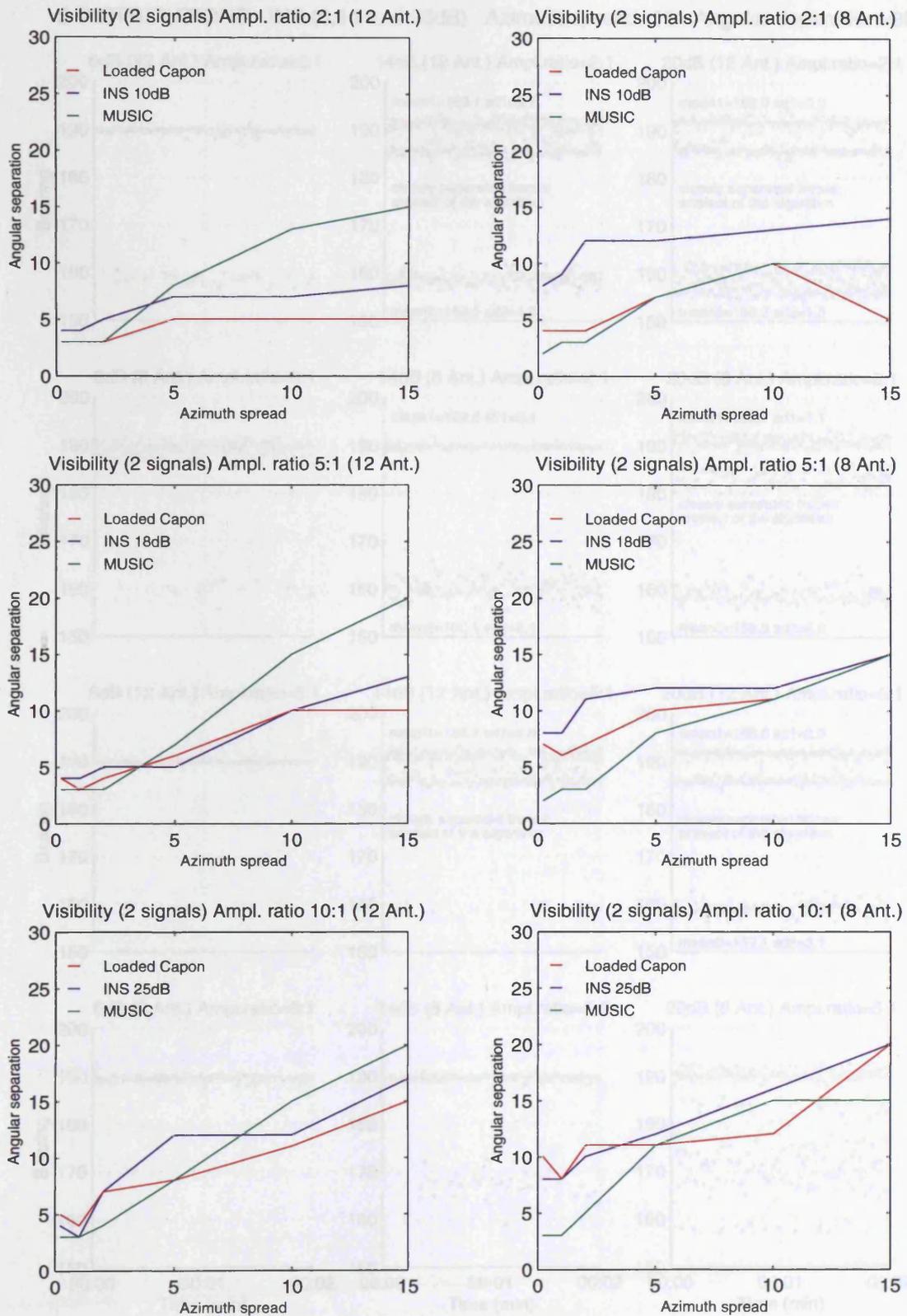


Fig. 6.3.4 Visibility of two signals with amplitude ratio 2:1 (upper panels), 5:1 (middle panels) and 10:1 (lower panels). The curves represent the minimum angular separation values between the peaks of the two distributions, at which two traces are detected, for different values of azimuth spread (the same for both signals at each processing run).

2 DISTRIBUTIONS: INS (6,14 and 20dB) Azimuth spread=15 Angular separation=30

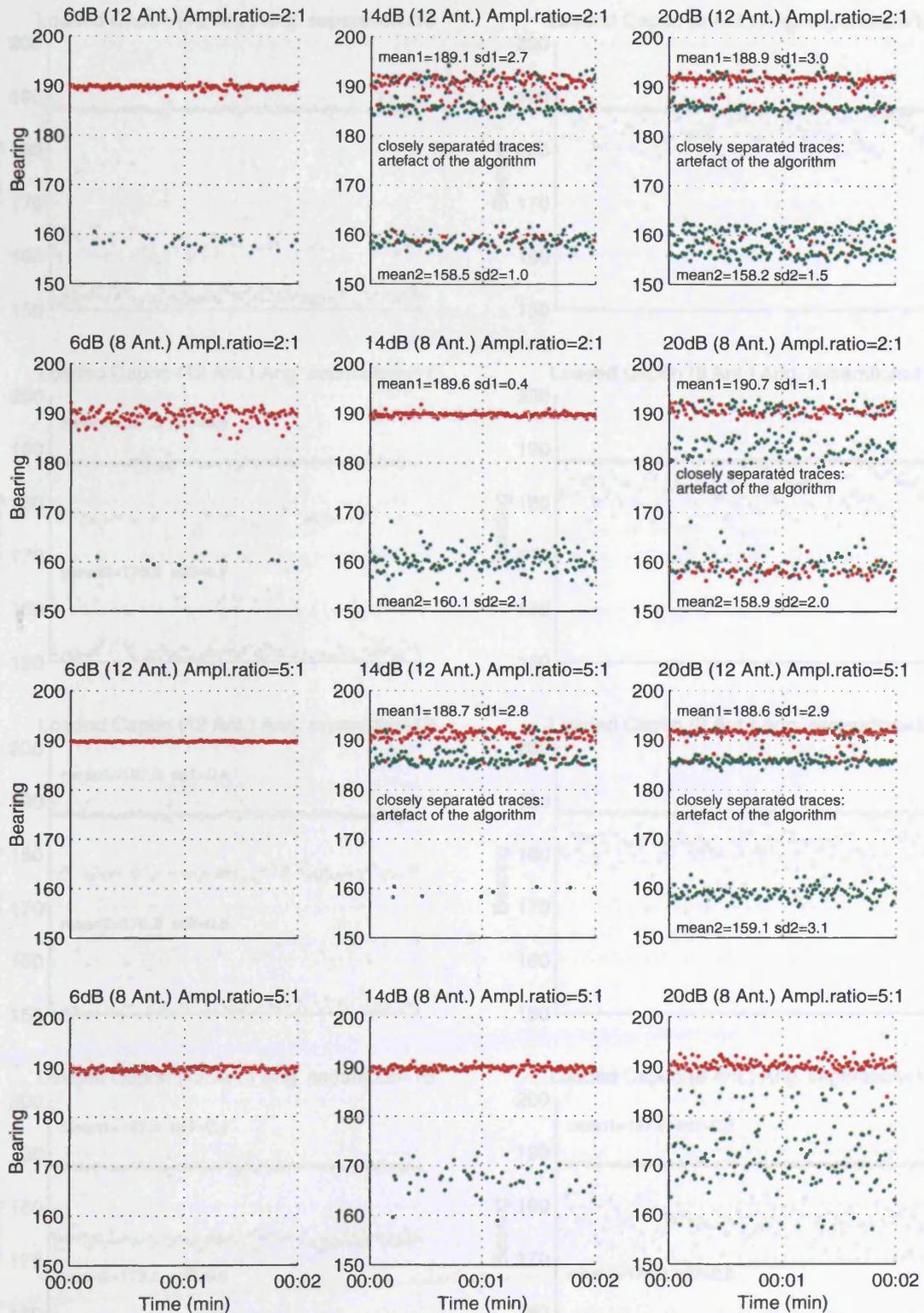


Fig. 6.3.5 Bearing estimates obtained from the double distribution model, with three thresholds for the INS algorithm (6dB, 14dB and 20dB for the first, second and third column respectively). First and third row: 12 antennas, second and fourth row: 8 antennas. Amplitude ratio 2:1 (first and second row) and 5:1 (third and fourth row). Az. Spread 15°. Ang. Separation 30°.

2 DISTRIBUTIONS: Amplitude ratio=10:1 Azimuth spread=10

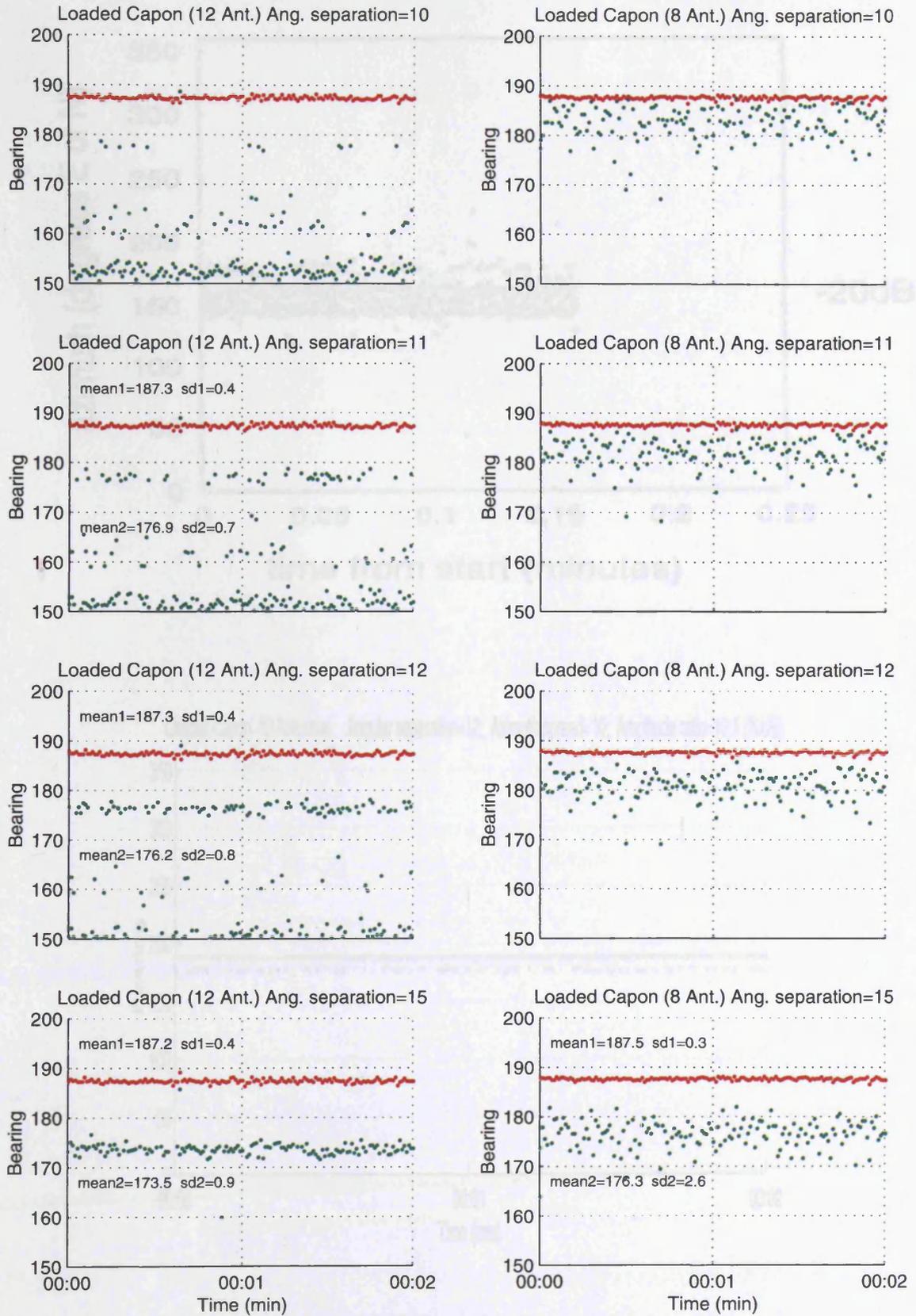


Fig. 6.3.6 Bearing estimates obtained from the double distribution model. Amplitude ratio 10:1. Az. Spread 10°. Ang. Separation 10°-11°-12°-15°. Stronger peak at 188.5°, weaker peak at 178.5°-177.5°-176.5°-173.5° (top to bottom).

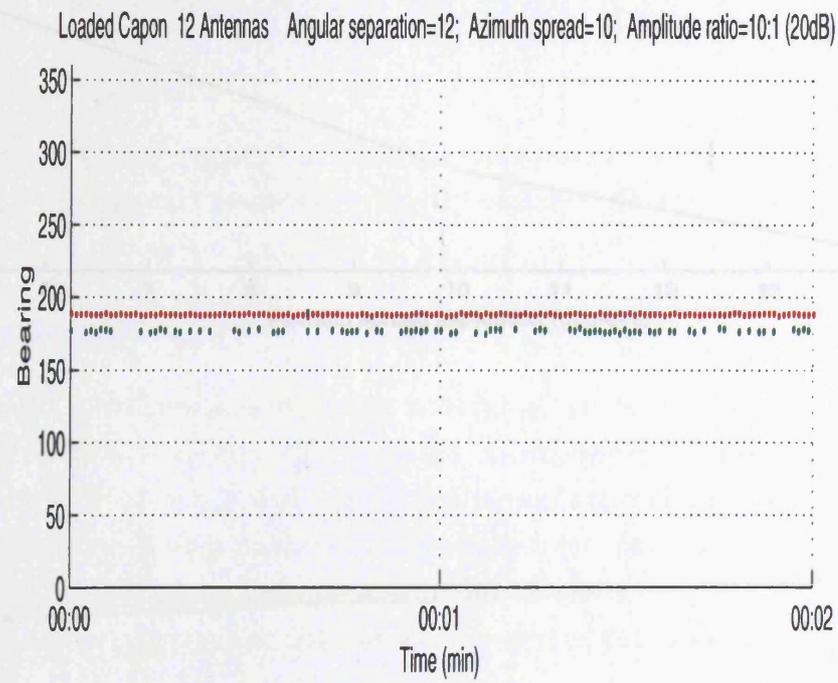
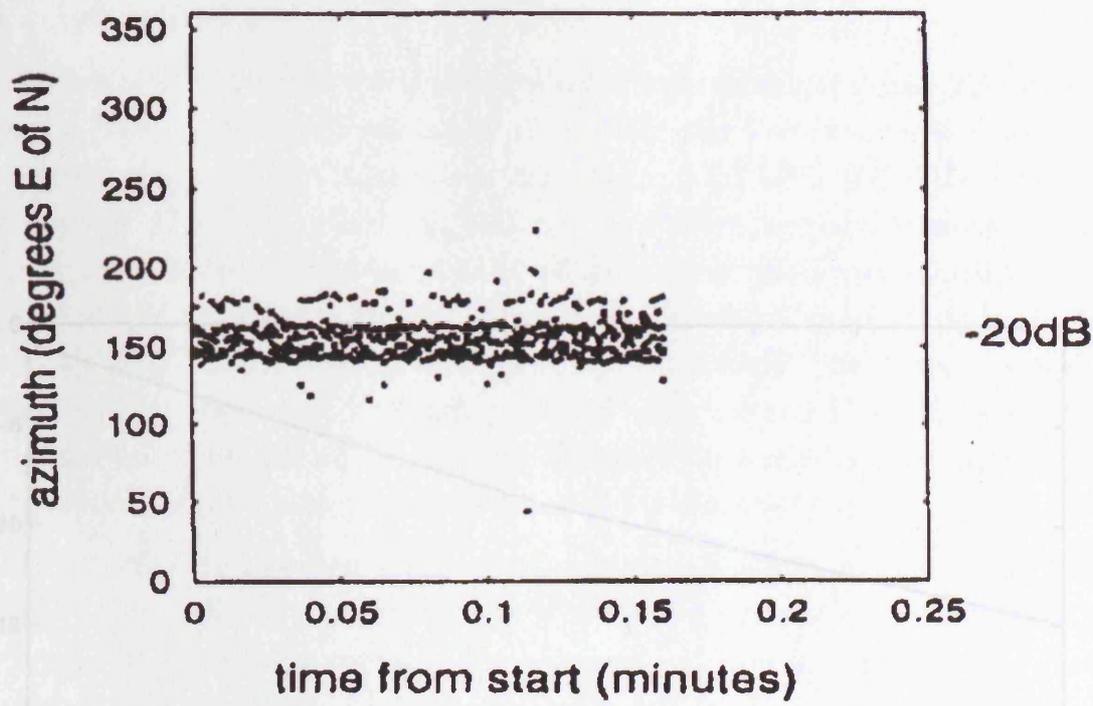


Figure 6.3.7 Visibility of two closely separated distributions. The double distribution model results (lower panel) are compared with those obtained by Jenkins (1997) (upper panel).

7. CONCLUSIONS AND FUTURE WORK

7.1 Introduction.

The data analysed for this work were collected over three days from 22nd to 24th January 1996, in the arctic winter, during which the Vortex direction finding system situated at Alert, Canada, was recording a 9.292 MHz signal transmitted from Iqaluit, 2100 km to the south, making this a polar cap path. Measurements of azimuth, elevation and amplitude of this signal were made using four

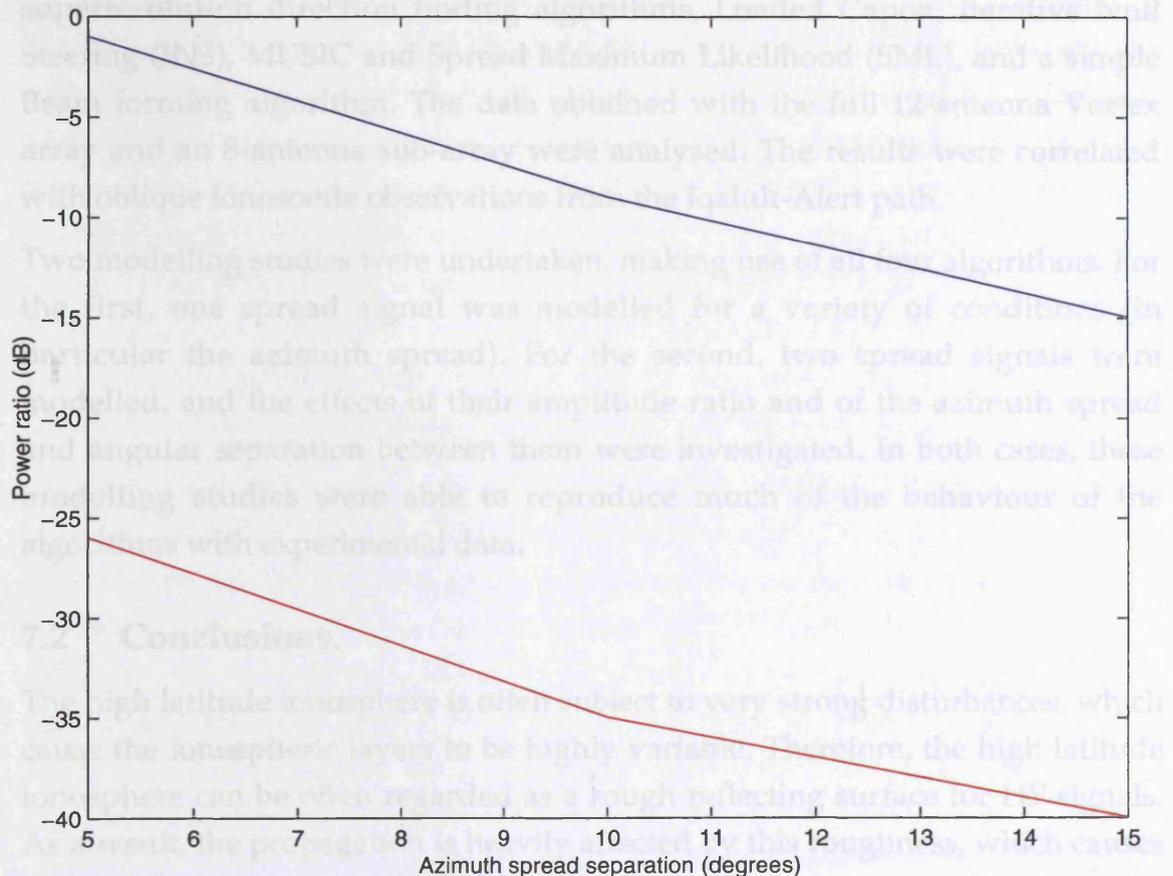


Figure 6.3.8 Power ratio between two signals vs minimum azimuth spread separation for visibility: the results of this investigation obtained with the Loaded Capon algorithm (red curve) are compared to those obtained by Jenkins (1997) with the deterministic ML algorithm (blue curve), for the Vortex array.

7. CONCLUSIONS AND FUTURE WORK

7.1 Introduction.

The data analysed for this work were collected over three days from 22nd to 24th January 1996, in the arctic winter, during which the Vortex direction finding system situated at Alert, Canada, was recording a 9.292 MHz signal transmitted from Iqaluit, 2100 km to the south, making this a polar cap path. Measurements of azimuth, elevation and amplitude of this signal were made using four superresolution direction finding algorithms, Loaded Capon, Iterative Null Steering (INS), MUSIC and Spread Maximum Likelihood (SML), and a simple Beam forming algorithm. The data obtained with the full 12-antenna Vortex array and an 8-antenna sub-array were analysed. The results were correlated with oblique ionosonde observations from the Iqaluit-Alert path.

Two modelling studies were undertaken, making use of all four algorithms. For the first, one spread signal was modelled for a variety of conditions (in particular the azimuth spread). For the second, two spread signals were modelled, and the effects of their amplitude ratio and of the azimuth spread and angular separation between them were investigated. In both cases, these modelling studies were able to reproduce much of the behaviour of the algorithms with experimental data.

7.2 Conclusions.

The high latitude ionosphere is often subject to very strong disturbances, which cause the ionospheric layers to be highly variable. Therefore, the high latitude ionosphere can be often regarded as a rough reflecting surface for HF signals. As a result, the propagation is heavily affected by this roughness, which causes the signal to arrive spread in azimuth and elevation at the receiver.

Ionospheric disturbances at high latitudes, such as patches and arcs of enhanced ionisation, can often cause multimoded propagation. In such circumstance well-correlated signals can reach the receiving site. Therefore those algorithms which assume uncorrelated incident signals at the receiver (such as MUSIC) are less suitable to the high latitude propagation environment. Whereas Loaded Capon and INS are not limited in this way.

The directional behaviour of the incoming signal varied dramatically from one day to another and also within the same day. Large and rapid bearing swings often occurred, when the direction of arrival (DOA) of the signal was well displaced from the Great Circle Path (GCP) direction (188.5°). When the ionosphere did not support the GCP mode, the data were sometimes consistent

with the signal being reflected by structures of over-dense plasma (i.e. patches and arcs), which were located far from the GCP. The DOA of the signal varied according to the motions of these structures. Bearing swings occurred especially when the relevant ionograms showed features detached from the main trace, which could be evidence of patches or arcs of enhanced ionisation. The observations were consistent with those previously obtained with a goniometric system located at Alert (Warrington *et al*, 1997a and b). The drift of a mono-dimensional reflector across the GCP has been simulated and the bearing swings reproduced using realistic values of model parameters such as velocity of drifting features. The results show that the point reflector followed the direction of the measured convection flow (NASA Coordinated Data Analysis Web Site) during the time intervals relevant to the simulations. Simulations were also undertaken to reproduce the elevation swings observed simultaneously with fairly constant bearing. These were consistent with the signal being reflected by a patch drifting along the GCP from Alert to Iqaluit. These observations showed that HF direction finding systems can be highly affected and severely degraded by the propagation environment at high latitudes, whereas instrumental errors ($\pm 0.1^\circ$) are negligible in comparison.

Observations undertaken with Loaded Capon, INS, MUSIC, Beam and, limited to some examples, SML, and with two antenna arrays of 12 and 8 elements have been compared. Whereas the gross structure of the wanted signal was similar in all cases, a more detailed investigation of the traces showed important differences. The results obtained with the 8-antenna sub-array generally exhibited traces less spread in azimuth as a result of less secondary energy detected. Multiple traces were often seen, with different angular separations from the wanted signal. These could be evidence of a multisignal environment or a multimoded propagation, or could result as artefacts of the algorithms. Comparing the results obtained with two different array geometries was usually able to resolve this issue. In particular, closely separated multiple traces, apparent with Loaded Capon and INS, were often likely to be caused by the diffuse nature of the signal, in which case Loaded Capon and INS detected two (or more) peaks of energy, which were essentially portions of the same spread signal. Much of this behaviour was reproduced by modelling studies (see below).

The effect of the integration time and the time over which bearing estimates are averaged on the standard deviation of the bearing estimates (i.e. bearing spread) has been investigated for time intervals with both low and high azimuth spreads. In all cases, for a fixed integration time value, an increase in the time/average resulted in a decrease in the bearing standard deviation.

Loaded Capon generally performed best, and especially in the higher spread intervals with the full array. When the time/average was fixed, the bearing standard deviation values were generally independent of the integration time value up to 1-2 sec. For a better understanding of the results for higher integration times, a longer data set than the one available for this research would be needed. It is important to take into account that relatively high integration time values would be unsuitable in the case of rapid bearing deviations (a frequent occurrence at high latitudes), which could result in poor accuracy of the bearing estimates. However, low integration time values can dramatically increase the computational burden, which might not be acceptable to some DF operators. Furthermore, although a higher time interval over which bearing estimates are averaged implies lower standard deviation values, a too high value can be unsuitable when real-time DF is a crucial factor for DF operators, which is often the case. The information contained in this investigation can be useful for DF operators in order to make the proper parameter choices according to their requirements and to propagation conditions.

Modelling studies have been undertaken, which considered the effect of symmetrical and asymmetrical diffuse reflections from the ionosphere, in azimuth and elevation. This model simulated very disturbed propagation environments, where signals are likely to arrive spread at the receiver, as often occurs at the high latitudes. The aim of these simulations was to determine the types of behaviour that may be expected from the DF algorithms, when the signal is spread. In order to achieve this, the expression of the complex amplitude of the incoming signal needed to be integrated with respect to azimuth and elevation. However, since this expression could not be integrated analytically, a numerical solution was adopted. A multiple discrete source distribution (MDS) was modelled as a grid of point sources in both azimuth and elevation. The simulations produced similar results for separations of 0.2° and below. An angular separation of 0.1° in both azimuth and elevation was used for most simulations.

By comparing the input and output azimuth spreads of the modelled signal, important information was obtained. In particular, it was seen that Loaded Capon and INS produced azimuth spreads less than the input spread of the signal. This was true for both array geometries (8 and 12 antennas), but in particular the lowest spreads were obtained for INS with 8 antennas in use. In contrast, MUSIC produced output values which were spread about the same or slightly more than the input signal spread. These results were similar in many respects to those at times seen in the experimental data.

Several MDSDs have been utilised aiming to reproduce experimental data: raised cosine (for both the amplitude and the power of the signal), gaussian and square. The raised cosine and gaussian distributions were able to reproduce the irregularly scattered experimental data, while the square distribution was not. The raised cosine distributions showed the lowest standard deviations of the bearing estimates.

As well as observing the spread in signal azimuth from the DF results, it has been possible to infer it from the delay spread in the ionogram trace. By running the model with the spread derived from the ionogram and comparing this with the experimental data results in a time interval around the 5-minute sounder slot corresponding to the ionogram, a good agreement was obtained when the ionogram trace delay spread was greater than about 0.4 ms. A poor agreement was found when the ionogram showed either multiple traces or other features at around the operational frequency, or when the trace was narrower than 0.4 ms. However, more examples are needed for a statistically valid analysis.

The visibility of two spread signals arriving at the receiver was also studied. A multisignal propagation environment is expected to be a common occurrence, and at high latitudes the signals will often arrive at the receiver considerably spread. A large number of simulations have been undertaken with all algorithms and both the 8 and 12-element array, aiming to take the SRDF algorithms to their limits in terms of the ability of detecting two spread signals arriving at the receiver for several values of azimuth spread and with the lowest possible angular separation and with the highest possible amplitude ratio. In the case of two signals with the same amplitude and same azimuth spread, Loaded Capon performed the best up to about 15-20° of azimuth spread and INS for higher spreads. In the case of two signals with different amplitudes (2:1, 5:1 and 10:1 were used), but equal azimuth spread, Loaded Capon generally performed slightly better than INS. The higher the amplitude ratio, the higher the angular separation needed by the algorithms for detecting the weaker signal, for a given azimuth spread value. In some simulations, Loaded Capon remarkably succeeded in detecting the two traces even in the case in which the two signal distributions overlapped. MUSIC performed well for unspread signals only, independently of the value of the amplitude ratio. For the vast majority of the cases, and especially for azimuth spreads up to 20°, the angular separation of the mean azimuth values of the two traces was less than the input of the two modelled signals, by 1-2°, whereas it was generally the opposite for azimuth spreads of 30°. Therefore, in case of multisignal environment this information can be useful for a better estimate of the real

separation between two signals, when two clear, independent traces are detected by a DF algorithm.

In these modelling studies (both with one and two spread signals) closely separated traces occurred both with Loaded Capon and INS for values of azimuth spread of 15° or more, whereas they did not occur when the azimuth spread of the distribution was lower. These types of traces, which appear to arise as artefacts of the algorithm, are often found in the experimental data, especially in the case of disturbed propagation environment, which causes the signals to arrive spread at the receiver. This result showed that, when closely separated multiple traces are present in the real data analysis, the relevant signals might be spread in azimuth by 15° or more. Furthermore, the estimated direction of arrival of the signal was more accurate when the bearings belonging to both traces were included to obtain the mean value, both for Loaded Capon and INS. In order to investigate how two traces might arise from one spread signal, a theory about the two-signal representation of one signal spread in azimuth has been derived. This theory (developed from Zatman's, 1998) was obtained by relating the second eigenvalue of the covariance matrix to the azimuth spread of a simulated signal, based on the principle that the higher the spread, the higher the chances for the second eigenvalue to pop up from the noise floor. This occurrence would cause two eigenvalues to represent one spread signal. Theoretical values relating the separation of the two traces to the azimuth spread of the incoming signal were obtained. This theory closely refers to eigenanalysis methods such MUSIC, whereas, with modelling studies, a two-trace representation of one spread signal occurred only with Loaded capon and INS. Nevertheless, a comparison between theoretical and simulation azimuth separation values for a given azimuth spread showed interesting results. The modelling studies have shown two traces arising from one spread signal for azimuth spread values comprised between about 15° and $35-40^\circ$. For a 15° azimuth spread value, the theoretical azimuth separation values were slightly lower than the simulation results, in those circumstances when two traces from one spread signal occurred. For increasing azimuth spread values of the incoming signal, the azimuth separation values of the two traces also increased for both the model data and the theoretical calculations, and were in fairly good agreement, in particular for about 30° azimuth spread. However, a high number of factors influenced the model data results, other than the azimuth spread. These are the algorithm and array geometry in use, the type of signal distribution, and to a lesser extent the elevation angle of arrival, elevation spread, azimuth and elevation Doppler spread. Furthermore, the model data results showed that the two-trace representation of one spread signal occurred (or not) randomly when algorithm (Loaded Capon or INS) and array geometry

(8 or 12 antennas) were varied, keeping constant all other parameters. This implies a high unpredictability of when the two-trace representation is likely to occur, and does not suggest a particular algorithm or array geometry to utilise in order to avoid this undesirable effect.

The overall results about the performance of the various SRDF algorithms utilised for this research suggest the following important conclusions. As expected, MUSIC was unable to detect correlated signals. Furthermore it consistently exhibited highly spread traces across the whole data set, which is not a desirable feature, since it does not help obtaining a precise DOA estimate. Modelling studies showed good MUSIC performances only for a single unspread incoming signal. However, at high latitudes highly correlated and/or spread signals often arrive at the receiver, therefore the use of the MUSIC algorithm for HF direction finding at high latitudes is not recommended.

The Loaded Capon algorithm exhibited good performances at high latitudes, as well as the INS algorithm, which was previously successfully tested at mid-latitudes (Warrington, 1995). These two algorithms have shown the lowest bearing estimates standard deviation values with both experimental and model data, in addition to an excellent agreement between them. Therefore the author recommends utilising Loaded Capon and INS for HF direction finding at high latitudes. The choice of which algorithm would be better to use between them will depend on the propagation characteristics and the DF operator's needs, bearing in mind that INS has the flexibility of the threshold setting. However, the investigation about the effect of integration time, and the time interval over which the bearing estimates are averaged, suggests a criterion according to the value of spread of the bearing traces. The recommended criterion is to use the INS algorithm with a relatively low threshold for azimuth spreads up to 15-20°, whereas the Loaded Capon algorithm is recommended for higher azimuth spread values, in order to obtain the lowest standard deviation of the bearing estimates.

Bearing measurements have also been obtained by means of the SML algorithm (Read, 1999). As stated by Read, SML is unreliable for incoming signals more spread than 15° (not unusual at high latitudes). Nevertheless, SML attempts to model the propagation environment more realistically than Loaded Capon and INS, particularly at high latitudes, as it assumes a spread source instead of a point source generating the received signal. As such, SML has been utilised in order to test its capabilities. The results show the unreliability of SML for signals more spread than 15°, whereas for lower spread values SML often detected a number of secondary traces displaced from the strongest trace, which was inconsistent with the results obtained with Loaded Capon and INS.

Therefore the author does not recommend the utilisation of this version of SML. However, the reader should be aware of further developments of SML and another algorithm developed by Read (1999), i.e. the Spatial Incoherent Region Estimator (SPIRE), which was not available for this research.

The comparison undertaken between the SRDF algorithms utilised for this research, with two array geometries, with both experimental and simulated data and in different propagation environments, can be helpful to DF operators, in order to make the best possible choice in any possible situation that may be encountered. Furthermore, the modelling studies have reproduced much of the behaviour of the SRDF algorithms with experimental data. The results obtained, together with proper further work, could be used to obtain correction factors for experimental results in various propagation environments, and especially when spread signals or spread modes of the same signal arrive at the receiver. This occurrence is particularly frequent in disturbed regions such as the high latitude ionosphere.

7.3 Future work.

The observations presented in this thesis have shown a dependence of the results on the algorithm and array geometry utilised. It would therefore be useful to undertake further investigations aimed at making the best possible choice of which SRDF algorithm to adopt in different propagation environments. In particular, the reader is referred to further developments of the SRDF algorithms created by Read, SML and SPIRE, which at present suffer from a lack of efficiency (see Section 7.2). Furthermore, it would be appropriate to carry out further investigations, possibly with large availability of data, on the effect of the array aperture and geometry on the results, in order to obtain an optimal array size and placement of antennas for given propagation characteristics (such as signal spread and signal separation in case of multisignal or multimoded propagation). This could be obtained by further comparison between modelling studies and experimental results for a high number of array geometries.

An investigation has been undertaken about the effect of the integration time and the time over which the estimates are averaged on the azimuth spread of the results. This was carried out for experimental data and for greater or lesser azimuth spreads, and has given useful information which could be of interest for DF operators. Modelling studies would be useful to investigate the effect of the signal spread on the results, for different integration times and times over which the results are averaged.

The relationship between delay spread in the ionogram traces and azimuth signal spread has been investigated, showing interesting results. However, a longer period of data collection, with the corresponding ionograms, is necessary for a more complete study.

The simulation of a mono-dimensional drifting reflector presented in this thesis showed excellent agreement with the occurrence of bearing swings and the direction of the convection flow. However, the model could be developed into a shaped reflector, which would more realistically simulate a large-scale electron-density structure.

A theory about the two-signal representation of one signal spread in azimuth has been developed, which relates to eigenanalysis methods such as MUSIC. However, a good agreement between the theoretical values and simulation results was found for non-eigenanalysis methods, such as Loaded Capon and INS. It would then be interesting to investigate how the theory presented in this thesis would affect the operation of these algorithms.

Modelling studies of one and two spread signals have been undertaken by assuming several types of multiple discrete source distributions for the incoming signals. The results proved to be dependent on the precise distribution (the raised cosine, the gaussian and the square were used). Simulations with distributions other than those utilised for this work could provide useful results.

The modelling studies of two spread signals have shown that the separation between the traces is usually slightly different than the separation between the peaks of the input distributions. A number of simulations with the incoming signals at various DOAs would be needed for an understanding of how the irregularly shaped beam pattern would influence this result.

REFERENCES

- Adcock F. (1919): Improvements in means for determining the direction of a distant source of electro-magnetic radiation. British patent 130490.
- Adcock F. (1959): Radio Direction Finding in three dimensions. *Proceedings of the Institute of Radio Engineering, Australia*, 20, pp 7-11.
- Agy V. (1970): Ionospheric forecasting. AGARD Conf. Proc. No 49, Advisory Group for Aerospace Research and Development, North Atlantic Treaty Organisation.
- Baden Fuller A.J. (1993): *Engineering Electromagnetism*. John Wiley & Sons Ltd.
- Bain W.C. (1956): The theoretical design of direction-finding systems for high frequencies. *IEE Proceedings*, 103B, pp 113-119.
- Balance W.P. and Miller T.W. (1991): Impact of mainlobe interference angular extent on adaptive beamforming. *25th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, 4-6 November, pp 989-993.
- Banks P.M. and Kockarts G. (1973): *Aeronomy, A and B*. Academic Press, New York, 2 vols.
- Bienvenu G. and Kopp L. (1980): Adaptivity to background noise spatial coherence for high resolution passive methods. *ICASSP 80*, Denver, 9-11 April (IEEE, New York), pp 307-310.
- Bienvenu G. and Kopp L. (1981): Source power estimation method associated with high resolution bearing estimator. *ICASSP 81*, Atlanta, 30 March-1 April, (IEEE, New York), pp 153-156.
- Blagoveshchenskij D.V., Egrova L.V. and Lukashkin V.M (1992): High latitude ionospheric phenomena diagnostics by high frequency radiowave propagation observations. *Radio Science*, 27, pp 267-274.
- Borgiotti G.V. and Kaplan L.J. (1979): Superresolution of uncorrelated interference sources by using adaptive array techniques. *IEEE Transactions on Antennas and propagation*, 27, pp 842-845.
- Brandwood D.H. (1989): Multiple signal DF with coherent signals using a general array. *Passive DF Colloquium*, IEE Digest No 1989/16, pp 5/1-4.
- British Standard (1971): Radio location and navigation terminology. BS 4727, Part 3, Group 06.
- Buchau J., Reinische B.W., Weber E.J. and Moore J.G. (1983): Structure and dynamics of the winter polar cap F region. *Radio Science*, 18, (6), pp 995-1010.

- Burg J.P. (1972): The relationship between maximum entropy spectra and maximum likelihood spectra. *Geophys.*, **37**, pp 375-376.
- Burg J.P. (1975): Ph.D. Thesis: Maximum entropy spectral analysis. Stanford University, California, U.S.A.
- Cannon P.S. (1989): Morphology of the high latitude ionosphere and its implications for HF communications systems. *IEE Proceedings, Part I, Communications, speech and vision* **136**, No.1, pp 1-10.
- Capon J. (1969): High-Resolution Frequency-Wavenumber spectrum analysis. *IEEE Proceedings*, **57**, pp 1408-1419.
- Capon J. (1970): Applications of detection and estimation theory to large array seismology. *IEEE Proceedings*, **58**, pp 760-770.
- Carlson H.C., Wickwar V.B., Weber E.J., Buchau J., Moore J.G. and Whiting W. (1984): Plasma characteristics of polar cap F layer arcs. *Geophysical Research Letters*, **11**, pp 895-898.
- Cawsey D.C. (1974): Linear antenna arrays with broadened nulls. *Electronics Letters (GB)*, **10**, pp 291-292.
- Childers D.G. (Ed) (1978): Modern spectral analysis. IEEE, New York.
- Clarke I.J. (1991): The IMP algorithm. *Advances in Spectrum Analysis and Array Processing*, Vol II, edited by S. Haykin (Englewood Cliffs, New Jersey, U.S.A.: Prentice Hall), Chapter 2.
- Compton R.T. (1988): *Adaptive antennas*. Prentice Hall.
- Cooper D.C. (1973): Zero-steering antenna system for receiving a signal close to the direction of strong interference. *Electronics Letters (GB)*, **9**, pp 140-141.
- Crampton C. (1947): Naval radio direction-finding. *IEE Journal*, **94**, part IIIA, pp 132-153.
- Davies K. (1990): *Ionospheric Radio*. Peter Peregrinus.
- Davies K. and Rush C.M. (1985): High-frequency ray paths in ionospheric layers with horizontal gradients. *Radio Science*, **20**, pp 95-103.
- Dumas D.J. (1997): High-latitude HF direction finding. *Proceedings of the 2nd Symposium on Radiolocation and Direction Finding*. Southwest Research Institute, San Antonio, Texas, 4-6 November 1997 (29 pages).
- Earp C.W. and Godfrey R.M. (1947): Radio direction-finding by the cyclical differential measurement of phase. *J. IEE*, **94**, part IIIA, pp 705-721.
- Ellis J.H. (1980): Fundamental limitations of wave-front analysis by least squares. *IEE Proceedings*, Part H, **127**, pp 219-227.

Featherstone W., Strangeways H.J. and Mewes H. (1997): Improved Capon estimator for HF DF using limited sample data sets. *IEE Colloquium Digest No 1997/390*, pp 2/1-2/5.

Feldstein Y.I. and Starkov G.V. (1967): Dynamics of auroral belt and polar geomagnetic disturbances. *Planetary Space Science*, 15, pp 209-229.

Fistas N. and Manikas A. (1994): A new general global array calibration method. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol IV, pp 73-76.

Gabriel W.F. (1976): Adaptive arrays – an introduction. *IEEE Proceedings*, 64, pp 239-272.

Gabriel W.F. (1980): Spectral analysis and adaptive array superresolution techniques. *IEEE Proceedings*, 68, pp 654-666.

Gassman G. J. (1963): The effect of disturbance of solar origin on telecommunications. *AGARDograph 59*, Macmillan Co., New York.

Gething P.J.D. (1973): Concept of phase centre of an array, applied to elevation-angle measurements. *IEE Proceedings*, 120, pp 541-543.

Gething P.J.D. (1978): *Radio direction finding and the resolution of multicomponent wave-fields*. Peter Peregrinus, Chapter 11.

Gething P.J.D. (1991): *Radio Direction Finding and Superresolution*. Peter Peregrinus.

Gething P.J.D. and Haseler J.B. (1974): Linear antenna arrays with broadened nulls. *IEE Proceedings*, 121, pp 165-168.

Gething P.J.D., Morris J.G., Shepherd E.G. and Tibble D.V. (1969): Measurements of elevation angles of HF waves. *IEE Proceedings*, 116, pp 185-193.

Goldberg J. and Messer H. (1998): Inherent limitations in the localization of a coherently scattered source. *IEEE Transactions on Signal Processing*, 46 (12), pp 3441-3444.

Goodman N.R. (1963): Statistical analysis based on a certain multivariate complex Gaussian distribution. *Ann. Math. Stat.*, 34, pp 152-177.

Goodman J.M. (1992): *HF Communications Science and Technology*. Van Nostrand Reinhold, New York.

Griffiths J.W.R. (1983): Adaptive array processing – a tutorial. *IEE Proceedings*, Pt. F 130, pp 3-16.

Griffiths L.J. (1977): A comparison of quadrature and single-channel receiver processing in adaptive beamforming. *IEEE Transactions on Antennas and Propagation*, 25, pp 209-218.

Hapgood M.A., Bowe G.A., Lockwood M., Tulunay Y. and Willis D.M. (1991): Variability of the interplanetary magnetic field over 24 years: 1963-1986. *Planetary and Space Science*, 39, pp 411-423.

Hargreaves, J.K. (1992): *The solar-terrestrial environment: an introduction to geospace- the science of the terrestrial upper atmosphere, ionosphere, and magnetosphere*. Cambridge University Press.

Haykin S. (1986): *Adaptive filter theory*. Prentice Hall, New York.

Hayward S.D. (1997): Effects of motion on adaptive arrays. *IEE Proceedings on Radar, Sonar and Navigation*, 144, No 1, pp 15-20.

Hill D.S. (December 1990): Multiple signal DF using superresolution: a practical assessment. *Electronics and Communication Engineering Journal*, pp 221-232.

Hockley J. (1973): A goniometer for use with high frequency circularly disposed aerial arrays. *Radio & Electron. Eng.*, 43, pp 475-485.

Hudson J.E. (1981): *Adaptive array principles*. Peter Peregrinus.

Hunsucker R.D. (1992): Mini review - Auroral and polar-cap ionospheric effects on radio propagation. *IEEE Transactions on Antennas and Propagation*, 40, No 7, pp 818-828.

Hunsucker R.D. and Bates H. F. (1969): Survey of polar and auroral effects on HF propagation. *Radio Science*, 4, p 347.

Jenkins R.W. (1994): A simulation study of HF Direction Finding in the presence of F-region scattering and Sporadic-E. CRC Report 93-004.

Jenkins R.W. (1997): The effects of array geometry and pattern uncertainties on high-latitude direction finding. *Proceedings of the 2nd Symposium on Radiolocation and Direction Finding*. Southwest Research Institute, San Antonio, Texas, 4-6 November 1997 (17 pages).

Johnson R.L. and Sherrill W.M. (1982): Modern signal processing applied to radio direction finding. Final Report for SWRI Project 16-9284.

Johnson R.L. and Miner G.E. (1986): Comparison of superresolution algorithms for radio direction finding. *IEEE Transactions on Aerospace and Electronic Systems*, 22, No 4, pp 432-442.

- Jones E.C., Schlicht H.C. and Ernst E.W. (1966): Elevation angle of arrival determination with circular antenna arrays. *Proc. Nat. Elec. Conf. Chicago*, **22**, pp 7-12.
- Kay S.M. and Demeure C. (1984): The high resolution estimator – a subjective entity. *IEEE Proceedings*, **72**, pp 1815-1816.
- Krolik J.L. and Swingler D.N. (1994): On the mean-square error performance of adaptive minimum variance beamformers based on the sample covariance matrix. *IEEE Transactions on Signal Processing*, **42**, (2), pp 445-448.
- Lagunas-Hernandez M.A. and Gasull-Llampallas A. (1984): An improved maximum likelihood method for power spectral density estimation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **32**, pp 170-173.
- Landmark B. (1968): The Ionosphere. *AGARD Lecture series 29 – Radiowave Propagation*.
- Lang S.W. (1981): Ph.D. Thesis: Spectral estimation for sensor arrays. M.I.T., Cambridge, MA.
- Lang S.W. and McClellan J.H. (1982): Multidimensional MEM spectral estimation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **30**, pp 880-887.
- Lee H.B. (1992): Eigenvalues and eigenvectors of covariance matrices for signals closely spaced in frequency. *IEEE Transactions on Signal Processing*, **40**, pp 2518-2535.
- Lee S.R., Song I., Lee Y.U., Chang T. and Kim H.M. (1996): Estimation of two-dimensional DOA under a distributed source model and some simulation results. *IEICE Transactions Fundamentals*, **E79** (9), pp 1475-1485.
- Lee Y.U., Choi J., Song I. and Lee S.R. (1997): Distributed source modeling and direction-of-arrival estimation techniques. *IEEE Transactions on Signal Processing*, **45** (4), pp 960-969.
- Lim J.C. (1977): Introduction to a sharp null response in an otherwise omnidirectional pattern using a circular array. *Radio & Electronic Engineering*, **47**, pp 30-32.
- Lockwood M. (1991a): Incoherent scatter radar measurements of the cusps. Barron, C.I. (Ed.): 'CLUSTER – dayside polar cusp'. ESA SP-330 (European Space Agency Publications, Noordwijk, The Netherlands), pp 57-66.
- Lockwood M. (1991b): On flow reversal boundaries and cross-cap potential in average models of high latitude convection. *Planetary and space science*. **39**, pp 397-409.

- Lockwood M. (1993): Modelling high-latitude ionosphere for time-varying plasma convection. *IEE Proceedings*, H, 140, No 2, pp 91-100.
- Lockwood M. and Carlson H.J. Jr. (1992): Production of polar cap electron density patches by transient magnetopause reconnection. *Geophysical Research Letters*, 19, pp 1731-1734.
- McClellan J.H. (1982): Multidimensional spectral estimation. *IEEE Proceedings*, 70, pp 1029-1039.
- Manikas A. and Turner L.F. (1991): Adaptive signal parameter estimation and classification technique. *IEEE Proceedings*, Part F, 138, pp 267-277.
- Manikas A., Ratnarajah T. and Jinsock Lee (1997): Evaluation of superresolution array techniques as applied to coherent sources. *Int. J. Electronics*, 82, No 1, pp 77-105.
- Mather J. (1991): The incremental multi-parameter algorithm. *Asilomar Conference on Circuits, Systems and Computers, Monterey*. Maple Press, San Jose, California, USA, pp 368-372.
- Mermoz H.F. (1981): Spatial processing beyond adaptive beamforming. *Journal of the Acoustics Society of America*, 70, pp 74-79.
- Milan S. (1994): Ph.D. Thesis: Observations of HF radiowaves propagated over high latitude paths. University of Leicester, U.K.
- Mitra A. P. (1974): Ionospheric effects of solar flares. Reidel Publishing Company, Dordrecht, Holland.
- Monzingo R.A. and Miller T.W. (1980): *Introduction to adaptive arrays*. Wiley, New York, Ch. 3.
- Moyle D.E. and Warrington E.M. (1997): Some superresolution DF measurements within the HF band. *IEE Conference Publication*, 2, No 436, pp 71-74.
- NASA Coordinated Data Analysis Web Site: <http://cdaweb.gsfc.nasa.gov>.
- Nickel U. (1987): Angular superresolution with phased array radar: a review of algorithms and operational constraints. *IEE Processing*, Part F, 134, pp 53-59.
- Oh S.K. and Un C.K. (1989): Computationally efficient maximum likelihood algorithm for localising multiple sources. *Electronics Letters*, 25, pp 1540-1542.
- Parkinson W.D. (1983): *Introduction to geomagnetism*. Scottish Academic Press, Edinburgh.

- Paulraj A., Roy R. and Kailath T (1986): Estimation of signal parameters via rotational invariance techniques - ESPRIT. *19th Asilomar Conference on Circuits, Systems and Computers*, Pacific Grove, 6 November, pp 83-89 (IEE, New York).
- Pike C. P. (1971): A latitudinal survey of the daytime polar F-layers. *Journal of Geophysical Research*, **76**. P 7745.
- Pillai S. U. (1989): *Array Signal Processing*. Springer-Verlag.
- Pisarenko V.F. (1972): On the estimation of spectra by means of non-linear functions of the covariance matrix. *Geophys. J.R. Astr. Soc.*, **28**, pp 511-531.
- Pisarenko V.F. (1973): The retrieval of harmonics from a covariance matrix. *Geophys. J.R. Astr. Soc.*, **33**, pp 347-366.
- Raich R., Goldberg J. and Messer H. (2000): Bearing estimation for a distributed source: modeling, inherent accuracy limitations and algorithms. *IEEE Transactions on Signal Processing*, **48** (2), pp 429-441.
- Ratcliffe J.A. (1959): *The magneto-ionic theory*. Cambridge University Press, Cambridge.
- Read W.J. (1999): Advanced DF algorithms for high latitude operation. *Proceedings of the 3rd Symposium on Radiolocation and Direction Finding*. Southwest Research Institute, San Antonio, Texas, 2-4 November 1999 (33 pages).
- Reddi S.S (1979): Multiple source location – a digital approach. *IEEE Transactions on Aerospace and Electronic Systems*, **15**, pp 95-105.
- Reed I.S., Mallet J.D. and Brennan L.E. (1974): Rapid convergence rate in adaptive arrays. *IEEE Transactions on Aerospace and Electronic Systems*, **10**, (6), pp 853-863.
- Rice D.W. (1971): The CRC high frequency direction finding research facility. *Conference Digest, 1971 International Electrical and Electronics Conference, Toronto*. (IEEE, New York).
- Richmond A.D., Kamide Y., Ahn B.-H., Akasofu S.-I., Alcayde D., Blanc M., De La Beaujardiere O., Evans D.S., Foster J.C., Friis-Christensen E., Fuller-Rowell T.J., Holt J.M., Knipp D., Kroehl H.W., Lepping R.P., Pellinen R.J., Senior C. and Zaitzev A.N. (1988): Mapping electrodynamic features of the high-latitude ionosphere from localised observations: combined ioncoherent-scatter radar and magnetometer measurements from 18-19 January 1984. *Journal of Geophysical Research*, **93**, pp 5760-5776.
- Rindfleisch H. (1956): The Wullenweber wide aperture direction finder. *Nachrichtenech. Z.*, **9**, pp 119-123.

Roesler D.P. and Bliss D.H. (1988): Analysis of Data from a Transauroral HF Experiment. *Working Paper – Document No. WP88-2010*, Rockwell International.

Rogers N.C., Warrington E.M. and Jones T.B. (1996): Comparison of predicted and measured times of propagation over high latitude paths for which direction of arrival measurements are available. *Electronics Letters*, **32**, No 5, pp 490-491.

Rogers N.C., Warrington E.M. and Jones T.B. (2001): Oblique ionogram features associated with off-great-circle HF propagation at high latitudes. *11th International Conference on Antennas and Propagation - IEE Conference Publications*, Volumes 1 and 2, **480**, pp 398-402.

Ross W., Bramley E.N., and Ashwell G.E. (1951): A phase-comparison method of measuring the direction of arrival of ionospheric radio waves. *IEE Proceedings*, **98**, pp 294-302.

Roy R. and Kailath T. (1989): ESPRIT – estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **37**, pp 984-995.

Schmidt R.O. (1979): Multiple Emitter location and signal parameter estimation. *Proc. RADC Spectrum Estimation Workshop*, Griffiths AFB, Rome, NY, pp 243-258; republished in *IEEE Transactions on Antennas and Propagation*, **34**, pp 276-280.

Schmidt R.O. (1981): Ph.D. Thesis: A signal subspace approach to multiple emitter location and spectral estimation. Stanford University, California, U.S.A.

Schmidt R.O. (1986): Multiple Emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, **34**, pp 276-280; republication of Schmidt (1979).

SESC (1988): SESC Glossary of solar-terrestrial terms. Space Environment Services Center, NOAA, Boulder, CO 80303.

Sharp R.D. and Johnson R.G. (1968): Satellite measurements of auroral particle precipitation. *Earth's particles and fields*. Reinhold Book, New York, pp 113-116.

Sherrill W.M. (1971): A survey of HF interferometry for ionospheric propagation research. *Radio Science*, **6**, pp 549-566.

Silberstein R. (1959): The origin of the current nomenclature of the ionospheric layers. *Journal of Atmospheric and Terrestrial Physics*, **13**, p 382.

Smith N. (1939): The relation of radio sky-wave transmission to ionosphere measurements. *IRE Proceedings*, **27**, p 332.

Smith-Rose R.L. and Ross W. (1947): The use of earth-mats to reduce the polarisation error of U-type Adcock direction finders. *IEE Journal*, **94**, part III, pp 91-98.

Stoica P. and Sharman K.C. (1990): Novel eigenanalysis method for direction estimation. *IEE Proceedings, Part F*, **137**, pp 19-26.

Sweeney L.F. Jr. (1970): Spatial properties of ionospheric radio propagation as determined with half degree azimuthal resolution. Thesis, Stanford University.

Tedd B.L., Strangeways H.J. and Jones T.B. (1985): Systematic ionospheric electron density tilts at mid-latitudes and their associated HF bearing errors. *Journal of Atmospheric and Terrestrial Physics*, **47**, pp 1085-1092.

Trump T. and Ottersen B. (1996): Estimation of nominal direction of arrival and angular spread using an array of sensors. *Signal Processing*, **50** (1-2), pp 57-69.

US Department of Commerce, National Geophysical Data Center FTP: <ftp://ftp.ngdc.noaa.gov>.

US Department of Commerce, National Geophysical Data Center Web Site: <http://www.ngdc.noaa.gov>.

Valaee S., Champagne B. and Kabal P. (1995): Parametric localization of distributed sources. *IEEE Transactions on Signal Processing*, **43**, pp 2144-2153.

Van Blaricum M.L. and Mitra R. (1978): Problems and solutions associated with Prony's method for processing transient data. *IEEE Transactions on Antennas and Propagation*, **26**, pp 174-182.

Van Trees H.L. (1968): *Detection, estimation and modulation theory*. Wiley, New York, Ch. 6.

Van Veen B.D. (1991): Adaptive convergence of linearly constrained beamformers based on the sample covariance matrix. *IEEE Transactions on Signal Processing*, **39**, (6), pp 1463-1470.

Vincent W. R., Daly R.F. and Sifford B. M. (1968): Modelling communication systems. *Ionospheric Radio Communications*, ed. K. Folkestad, Plenum Press, New York, pp 321-340.

Warrington E.M. (1986): Ph.D. Thesis: Ionospheric radiowave propagation effects observed with a large aperture antenna array. University of Leicester, U.K.

Warrington E.M. (1995): Measurements of the direction of arrival of HF sky wave signals by means of a wide aperture antenna array and two superresolution direction finding algorithms. *IEE Proceedings on Microwave, Antennas and Propagation*, **142**, pp 136-144.

Warrington E.M. (1998): Observations of the directional characteristics of ionospherically propagated HF radio channel sounding signals over two high

latitude paths. *IEE Proceedings on Microwaves, Antennas and Propagation*, **145**, No.5, pp 379-385.

Warrington E.M. and Jones T.B. (1991): Measurements of the direction of arrival of HF sky wave signals and single site location using a seven element wide aperture (294 m) interferometer array. *IEE Proceeding on Microwaves, Antennas and Propagation*, Part H, **138**, pp 121-130.

Warrington E.M., Jones T.B., Rogers N.C. and Rizzo C. (1997a): Directional characteristics of ionospherically propagated HF radio signals and their effect on HF DF systems. *IEE Colloquium on Propagation Characteristics and Related System Techniques for Beyond Line-of-Sight Radio*, London, U.K.

Warrington E.M., Rogers N.C. and Jones T.B. (1997b): Large HF bearings errors for propagation paths contained within the polar cap. *IEE Proceedings on Microwaves, Antennas and Propagation*, **144**, (4), pp 241-249.

Warrington E.M., Stocker A.J., Rizzo C. and Jackson C.A. (1997c): The directional structure of HF signals received over high latitude paths and the effect on superresolution direction finding. *Proceedings of the Third Symposium on Radiolocation and Direction Finding*, Southwest Research Institute, San Antonio, Texas.

Warrington E.M., Stocker A.J. and Rizzo C. (1999): The effect of the high latitude ionosphere on measurements of the direction of arrival of HF signals using superresolution direction finding. *IEE Conference on Antennas and Propagation*, pp 349-352.

Warrington E.M., Stocker A.J. and Rizzo C. (2000): The behaviour of superresolution direction finding algorithms for HF signals propagating through the high latitude ionosphere. *IEE Proceedings on Microwaves, Antennas and Propagation*.

Weber E.J., Buchau J., Moore J.M., Sharber J.R., Livingston J.D., Winningham J.D. and Reinische B.W. (1984): F layer ionization patches in the polar cap. *Journal of Geophysical Research*, **89**, pp 1683-1694.

Whale H.A. (1954): A rotating interferometer for the measurement of the direction of arrival of short radio waves. *Proc. Phys. Soc.*, **67B**, pp 553-562.

Zatman M. (1995): Production of adaptive array troughs by dispersion synthesis. *Electronics Letters*, **31**, No 25, pp 2141-2142.

Zatman M. (1998): How narrow is the narrowband? *IEE Proceedings on Radar, Sonar and Navigation*, **145**, No 2, pp 85-91.

Zatman M., Strangeways H.J. and Warrington E.M. (1993): Resolution of Multimoded HF transmissions using the DOSE superresolution direction finding algorithm. *Proceedings of the IEE International Conference on Antennas and Propagation, IEE Conference Publication 370*, pp 415-418.

Ziskind I. and Wax M. (1988): Maximum likelihood localization of multiple sources by alternating projection. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **36**, pp 1553-1560.