Modelling of VLF Radio Propagation in the Earth-Ionosphere Waveguide

by

Stephen G. Meyer

Submitted in fulfilment of the requirements for the degree of Master of Science in the School of Physics, University of KwaZulu–Natal.

As the candidate's Supervisor I have/have not approved this thesis for submission

Signed: ..... Name: ..... Date: .....

Abstract

Understanding the propagation of Very Low Frequency radio waves in the waveguide formed between the surface of the Earth and the lower edge of the ionosphere is of great interest and practical use. Very Low Frequency radio waves serve as a useful tool to investigate the ionosphere and different factors influencing it, most notably solar conditions and events. In this thesis, two methods of simulating the propagation of narrowband Very Low Frequency radio waves are presented, namely Waveguide Mode Theory and the Finite-Difference Time-Domain method. The principles behind these methods are introduced and the advantages and drawbacks of each are discussed. Wait's two ionospheric parameters, H' and β, are used in both cases to simulate the electron density and conductivity of the lower ionosphere; probably the most important factors in this field of study. The two models are first used to investigate some theoretical aspects and concepts. Applications of the models are then illustrated through the use of data from UltraMSK and the DEMETER satellite.

PREFACE

The experimental work described in this dissertation was carried out at the Hermanus Magnetic Observatory, Hermanus, and in the School of Physics, University of KwaZulu–Natal, Durban, from February 2010 to December 2011 under the supervision of Dr Andrew B. Collier.

These studies represent original work by the author and have not otherwise been submitted in any form for any degree or diploma to any tertiary institution. Where use has been made of the work of others it is duly acknowledged in the text.

DECLARATION 1 - PLAGIARISM

I, ..., declare that

1. The research reported in this thesis, except where otherwise indicated, is my original research.
2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
   (a) Their words have been re-written but the general information attributed to them has been referenced
   (b) Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.
5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Signed: .....

DECLARATION 2 - PUBLICATIONS

DETAILS OF CONTRIBUTION TO PUBLICATIONS that form part and/or include research presented in this thesis (include publications in preparation, submitted, in press and published and give details of the contributions of each author to the experimental work and writing of each publication)

- Dahlgren, H., T. Sundberg, A. B. Collier, E. J. Koen, and S. G. Meyer (2011) Solar flares detected by the new narrowband VLF receiver at SANAE IV South African Journal of Science, 107(9/10).

Signed: .....

Acknowledgements

I would like to thank the National Research Foundation (NRF) for the funding I received that made my studies possible. I would also like to thank everyone at SANSA Space Science for making my time in Hermanus so pleasant, especially those that I have shared an office or flat with. Thank you to Dr. Craig Rodger from the Univeristy of Otago for the suggestions related to the DEMETER data. Thanks to the others in our research group, especially Brett and Etienne for their help and suggestions with many things over the last two years. A massive thanks to my supervisor, Dr. Andrew B. Collier for all the support and guidance he provided throughout my studies and for having confidence and faith in me and my work. And finally, an incredibly large thank you to my parents who allowed me to get to this point and for their endless support and interest invested in me.

Contents

1 Introduction **1**

1.1 Outline . . . 1

1.2 History of Ionospheric Modelling Techniques . . . 2

1.2.1 Waveguide Mode Theory . . . 2

1.2.2 The FDTD Method . . . 3

1.3 The Ionosphere . . . 3

1.4 Modelling: Wait's Parameters and Chapman Layers . . . 6

1.5 LWPC . . . 7

2 Theory **9**

2.1 Waveguide Theory . . . 9

2.2 Mode Theory Formulation . . . 12

2.3 The FDTD Method . . . 17

2.3.1 Finite Differencing and Grid Allocation . . . 17

2.3.2 Nearly Perfectly Matched Layer . . . 25

2.3.3 Surface Impedance Boundary Condition . . . 26

3 Data Collection **29**

3.1 VLF Antenna and UltraMSK . . . 29

3.2 DEMETER and GOES . . . 33

4 Mode Theory **34**

4.1 Program Results . . . 35

4.1.1 Height Gain Function . . . 35

4.1.2 Attenuation . . . 35

4.1.3 Dependence on Height H' and Sharpness Parameter β . . . 36

4.2 Temporal Variation . . . 38

4.3 DEMETER Comparison . . . 39

5 The FDTD Method **51**
5.1 Free Space 53
5.2 Basic Waveguide 53
5.3 Nearly Perfectly Matched Layer 55
5.4 Surface Impedance Boundary Condition 59
5.5 Background Plasma, Magnetic Field, Electron-Neutral Collisions 61
5.6 Realistic Example 62
5.6.1 Comparison with DEMETER observation 66
5.7 Numerical Stability of the FDTD method 70

6 Discussion and Conclusions **75**

Bibliography **78**

List of Figures

1.1 Ionospheric layers for day and night 5
1.2 Example of a single Chapman Layer 5
1.3 Electron density profiles with varying height parameter 7
1.4 Electron density profiles with varying sharpness parameter 8

2.1 Horizontal parallel plate waveguide 10
2.2 TM and TE modes in a waveguide. 12
2.3 Layout of problem 14
2.4 Refractive index profile and reflection condition 15
2.5 1D FDTD Grid. 20
2.6 3D Yee Cube. 21
2.7 E-J Collocated grid. 22

3.1 VLF loop antenna. 30
3.2 UltraMSK schematic. 30
3.3 UltraMSK phase and amplitude: NWC. 32
3.4 UltraMSK phase and amplitude: NAA. 32

4.1 Height gain function, real component 36
4.2 Height gain function, imaginary component 37
4.3 Attenuation rates 43
4.4 Influence of height parameter on modal interference pattern 44
4.5 Influence of sharpness parameter on modal interference pattern 44
4.6 Modelled temporal change of height parameter 45
4.7 X-ray flux and corresponding narrowband VLF signal 46
4.8 Simulated amplitude response to a small flare 47
4.9 Simulated amplitude response to a larger flare 47
4.10 DEMETER observation over NWC 48
4.11 Location of NWC transmitter 48
4.12 DEMETER profiles of eastward and westward propagation 49

4.13 Comparison of model with data 50

5.1 Field patterns for a vertical and horizontal electric dipole source in free space . . . 54
5.2 Field patterns for a vertical and horizontal electric dipole source in a small waveguide 55
5.3 Field patterns for a vertical and horizontal electric dipole source in a larger waveguide 56
5.4 Field patterns for a vertical source located at different heights 57
5.5 Conductivity of the NPML 58
5.6 Real and imaginary components of the NPML 59
5.7 Example of SIBC implementation for sea and land 60
5.8 Influence of a background plasma, with and without electron-neutral collisions . . 63
5.9 Current density with and without electron-neutral collisions 63
5.10 FDTD results of field patterns with different values of H' 64
5.11 FDTD results of field patterns with different values of β 64
5.12 FDTD results for propagation during daytime and nighttime 67
5.13 FDTD results for propagation over land and sea 68
5.14 FDTD results at different altitudes 69
5.15 FDTD results compared with data 71
5.16 Example of numerical dispersion 74

List of Tables

5.1 Effect of frequency and grid sampling density on discretisation. 52
5.2 SIBC Conditions 61
5.3 Numerical dispersion analysis results. 74

Acronyms

ABC Absorbing Boundary Condition

EIWG Earth-Ionosphere Waveguide

FDTD Finite-Difference Time-Domain

LWPC Long Distance Propagation Capability

NPML Nearly Perfectly Matched Layer

PEC Perfect Electric Conductor

PML Perfectly Matched Layer

PPW Points Per Wavelength

TE Transverse Electric

TM Transverse Magnetic

TEM Transverse Electromagnetic

SIBC Surface Impedence Boundary Condition

VLF Very Low Frequency

Chapter 1

Introduction

Very Low Frequency (VLF) radio waves fall within the frequency range of 3 to 30 kHz, with corresponding wavelengths of 100 to 10 km. VLF wave propagation has been a field of interest since Marconi performed the first Trans-Atlantic radio broadcast in 1901. The concept of the Earth-Ionosphere Waveguide (EIWG) arose after it was realised that a normal ground wave could not account for the observed field strength [Green, 1974]. Government, and especially the military, became aware of the importance of understanding radio wave propagation for its use as a communications tool. It was later recognised that this is especially relevant at VLF, since these frequencies could be used to communicate with submerged submarines. This fact alone was a blessing for scientists as the maritime super powers built massive VLF transmitters around the world to communicate with their submarines. Being able to communicate with submarines and help them navigate was a major driving force behind understanding radio wave propagation in the EIWG. During the time since the Cold War, several methods have been developed to understand how these radio waves propagate and to predict their behaviour. Since the original, hostile motivation for the study of VLF propagation, the focus has shifted to more peaceful and scientific goals. It was realised that these waves could be employed as an incredibly useful tool to study the ionosphere and the great many factors that influence it. Simulated and experimental results have been used to investigate many of these factors, ranging from events on the Sun such as solar flares, coronal mass ejections and solar proton events to terrestrial phenomena like earthquakes.

1.1 Outline

This thesis looks at modelling narrowband VLF radio wave propagation in the EIWG. The focus will be on two methods, namely Mode Theory [Wait, 1962] and the Finite-Difference Time-Domain (FDTD) method [Taflove and Hagness, 2005], which are discussed in Chapter 2. Firstly, a Mode Theory model is used in Chapter 4 to gain an understanding of the different factors that determine properties of the waveguide and how these affect propagation. It also allows for the theoretical

investigation of other waveguide properties such as attenuation rates and refractive index profiles. The effects of a solar flare are examined and explained in terms of temporal changes in the modal interference pattern due to perturbations in the electron density profile from the enhanced solar radiation. To validate the Mode Theory simulation, results are compared to data from the Detection of Electro-Magnetic Emissions Transmitted from Earthquake Regions (DEMETER) satellite. The FDTD method is introduced and examples of propagation in free space are shown in Chapter 5. The FDTD model is then used to illustrate basic waveguide propagation as well as propagation in a plasma. Changes in the interference pattern due to waveguide parameters and observation height are examined. Results from this model are then compared to the DEMETER data.

1.2 History of Ionospheric Modelling Techniques

Over the last 50 to 60 years, there has been a great amount of effort put into modelling the ionosphere and radio wave propagation in the EIWG. The initial attempts at explaining wave propagation used Mode Theory. This approach delivered satisfactory results and has become more advanced over the years, taking more complex effects into account. Mode Theory was the favoured choice of modelling and explaining observations until the advance of computers and the advent of the FDTD method allowed for more intensive numerical simulations. Both the FDTD method and Mode Theory are still widely used today, with each having its own advantages and drawbacks.

1.2.1 Waveguide Mode Theory

Without a doubt, James R. Wait was the most prolific researcher in the field of VLF propagation, publishing numerous papers over more than 50 years [Wait, 1960, 1964; Wait and Spies, 1965]. He understood that certain situations required a numerical approach while he analytically solved many fundamental propagation problems. Many of today's methods are based on or derived from work presented in the classic book of Wait [1962], including the Mode Theory technique employed in this thesis. Wait, as well as K. G. Budden both pursued the Mode Theory approach independently but ended up with essentially equivalent formulations [Wait, 1962; Budden, 1962]. Budden's work formed the basis for a numerical model developed by the U.S. Navy. Mode Theory has two obvious advantages: firstly, for narrowband propagation it is essentially an exact description. Secondly, once the full set of modes has been calculated, the field can be easily calculated at any point in the waveguide. One of the drawbacks of Mode Theory is that it is difficult to deal with a change in the waveguide boundaries. At such an interface, mode conversions need to occur. This is the main factor determining the Mode Theory simulation time: it scales linearly with the number of waveguide slabs and hence mode conversions.

1.2.2 The FDTD Method

Although Mode Theory is fast, allows field calculation at any point, and can provide very good results, FDTD is generally considered the future of modelling propagation in the EIWG. This is due to its ability to generalise and customise the problem without major changes to the code. The fact that computational power is rapidly increasing is another reason for FDTD becoming the popular choice. The FDTD method was first proposed by Yee [1966], although the acronym FDTD was not used then. Yee [1966] introduced the original concept of discretising the fields through finite differences and placing the electric and magnetic field components at half integer indices on the grid, although only perfectly conducting surfaces could be implemented. The acronym FDTD was first coined by Taflove [1980] who validated FDTD models of sinusoidal steady-state electromagnetic wave penetration into a metal cavity. Mur [1981] released the first numerically stable absorbing boundary condition for Yee's grid. This boundary condition was improved when Liao et al. [1984] released their version which is based on the space-time extrapolation of the field adjacent to the boundary. Berenger [1984] produced the Perfectly Matched Layer (PML), considered to be the most advanced and successful method of implementing boundary conditions. The Nearly Perfectly Matched Layer (NPML) which was proposed by Cummer [2003], is used in this work and discussed in further detail. One of the advantages of the FDTD method which makes it very popular is that it is a technique that has an incredibly large range of applications. Some of the interesting fields of study where the FDTD method has been applied include wireless personal communication devices, detection of early stage breast cancer using ultrawideband microwaves, biophotonics and propagation in the EIWG. Due to the FDTD method being a time domain method, it allows a wide range of frequencies to be modelled. This is of great interest in this discipline as it allows for the study of the propagation of broadband electromagnetic energy released during a lightning stroke. At the same time, it is very well suited to investigate the propagation of narrowband signals emitted from VLF transmitters around the world. One of the motivations for using a FDTD method to study propagation in the EIWG is that it inherently calculates any evanescent waves which Mode Theory does not account for.

1.3 The Ionosphere

The ionosphere is the region in the atmosphere extending from around 60 km to more than 1000 km altitude where there are a significant number of free electrons. These free electrons arise from radiation, mostly solar in origin, ionising atmospheric neutral molecules. Due to the complicated distribution of species in the atmosphere as well as the broad spectrum of ionising radiation, the free electron density profile can be categorised into three layers, known as the D, E and F layers.

Figure 1.1 shows the electron density profile in these three regions, with the D-region the lowest in altitude, starting from around 60 km in altitude. This extends up to an altitude of about 90

to 100 km where the E-region starts, here there is a dramatic jump in the difference between day time and night time. From an altitude of around 200 km, the F-region starts, splitting up into the F1- and F2-layers during the day. Figure 1.1 also shows the variation in electron density due to the solar cycle, which has a stronger influence at higher altitudes.

D-region:

The D-region of the ionosphere is mainly responsible for the reflection of VLF radio waves [Nicolet and Aikin, 1960]. These (partial) reflections take place in the D-region since the electron density, and therefore refractive index, changes dramatically within the space of one wavelength. It is the lowest layer in the ionosphere, ranging from around 60 to 90km and has the lowest free electron density of 10^2 to 10^4 electrons cm^{-3}. The main source of ionisation is Lyman-α hydrogen radiation (Vacuum UV/ Far UV, $\lambda = 121.6$nm) from the Sun ionising atmospheric neutral NO. Since the main source is solar in origin, the D-region effectively disappears at night, except at very high latitudes. This also means that the D-region height has a strong dependence on season and solar cycle. Galactic cosmic rays and Lyman-α radiation scattered off the geocorona still provide some ionisation at D-region altitude when there is no sunlight [Thomson et al., 2007]. It is the D-region that creates the extra attenuation experienced by VLF waves during the day compared to at night. The fact that VLF waves are reflected from the D-region makes them a very useful tool for studying what is otherwise a very difficult region to investigate. It is too low for satellites since the air is too dense leading to drag on the satellite, and is too high for balloons [Cheng et al., 2006; Han and Cummer, 2010]. Rocket borne measurements are possible and provide accurate results. The problem with rockets is that they are expensive, and provide once off localised data, meaning they are inappropriate for observing diurnal and seasonal variations in the ionosphere.

E-region:

The E-region is the middle layer of the ionosphere and extends from around 90 to 140 km. It is mostly produced from soft X-rays (110nm) and far UV radiation ionising molecular oxygen (O_2). The free electron density here is about 10 to 10^3 cm^{-3} . It is responsible for reflection of radio waves up to a few MHz and also weakens at night due to decreased radiation.

F-region:

This layer is the highest in the ionosphere and can be split into the F1 and F2 layers in the day. The F1 layer ranges from around 140 km to 200 km. The F2 layer is above F1 and stretches from 200 km upwards. Figure 1.1 shows that the F1 layer disappears during the night, leaving a single F layer. The highest electron densities are found in this region and is therefore responsible for the reflection of the highest frequency waves. The highest frequency wave will be reflected from the region in the layer where the electron density is at a maximum. This frequency is known as the Maximum Usable Frequency (MUF). Any frequencies above the MUF will propagate through the ionosphere into space while all below it will be reflected.

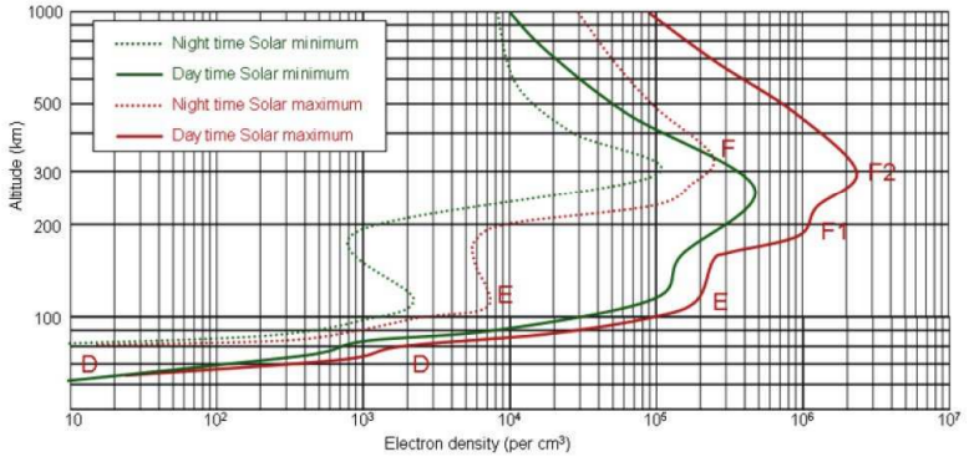


Figure 1.1: Ionospheric layers for day and night. Stanford AWESOME Tutorial

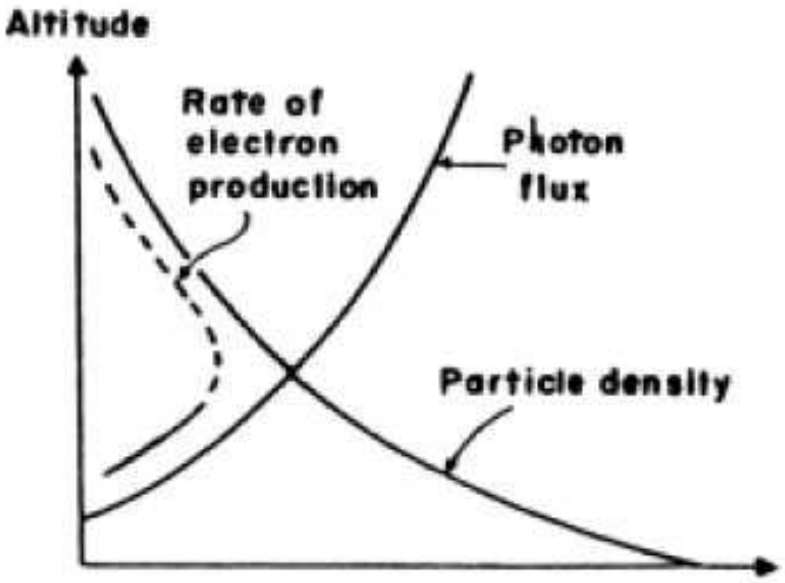


Figure 1.2: Example of a single Chapman layer

1.4 Modelling: Wait's Parameters and Chapman Layers

The ionosphere consists of many layers, leading to it being quite difficult to model mathematically. The general idea behind modelling the ionosphere is to think of it in terms of Chapman layers [Chapman, 1931]. Chapman considered that the density of a gas in the atmosphere decreases with increasing height while the total solar radiation increases with altitude as it is more highly absorbed at lower altitudes. The idea that at high altitudes, there are many photons but few particles, while on the other hand, there are many particles but few ionising photons at lower heights is illustrated in Figure 1.2. A single Chapman layer is applicable for a single species of gas being ionised by a monochromatic source. Since the atmosphere is composed of many species and solar radiation covers a very wide frequency range, the actual ionosphere can be imagined in terms of a superposition of Chapman layers, for each gas species and for each wavelength. Summing various Chapman profiles will lead to a realistic electron density profile as seen in Figure 1.1.

Since the work here focuses on VLF waves in the D-region, a full description based on Chapman layers will not be needed. For VLF waves, only the lower edge of the ionosphere is of interest. It is in this region that the electron density, N_e in m^{-3}, increases roughly exponentially with altitude. This means that when looking at the D-region alone, modelling the electron density with an exponential function is appropriate. Wait and Spies [1964] proposed doing this with an exponential profile characterised by two parameters, H' and β. H' is an ionospheric reference height, measured in km and gives an indication of the height of the D-region. β is known as the steepness, or sharpness, and is measured in km^{-1}. The value of β determines the gradient of the electron density profile. These parameters appear as follows in what is generally known as Wait's exponential electron density profile [Wait and Spies, 1964]:

$$N_e(z) = 1.43 \times 10^{13} e^{-0.15H' + (\beta - 0.15)(z - H')} m^{-3}. \tag{1.1}$$

where z is the altitude in km. The values of H' and β can vary a lot, depending on latitude, time of day and current stage in the solar cycle. H' generally varies between 65 and 90 km, with lower heights in the day time and during periods of high solar activity where the increased solar radiation lowers the height of the D-region. H' will also be slightly lower at the poles due to the increased ionisation from the solar wind that is guided down the magnetic field lines. β tends to vary between 0.3 and 0.6 km^{-1} with higher values occurring when there is less radiation incident on the ionosphere. Examples of the electron density profile under different conditions are shown are Figures 1.3 and 1.4. Figure 1.3 shows that for a constant β, changing H' shifts the the profile up or down. Figure 1.4 shows the changes in the electron density profile when H' is kept constant and β changes. We see that lower values of β lead to higher electron densities below H' and lower densities above H'. Another important factor that influences radio wave propagation in the Earth-Ionosphere Waveguide is the effect of free electrons colliding with neutral molecules [Cummer et al., 1998]. These electron-neutral collisions act as an energy sink by transferring

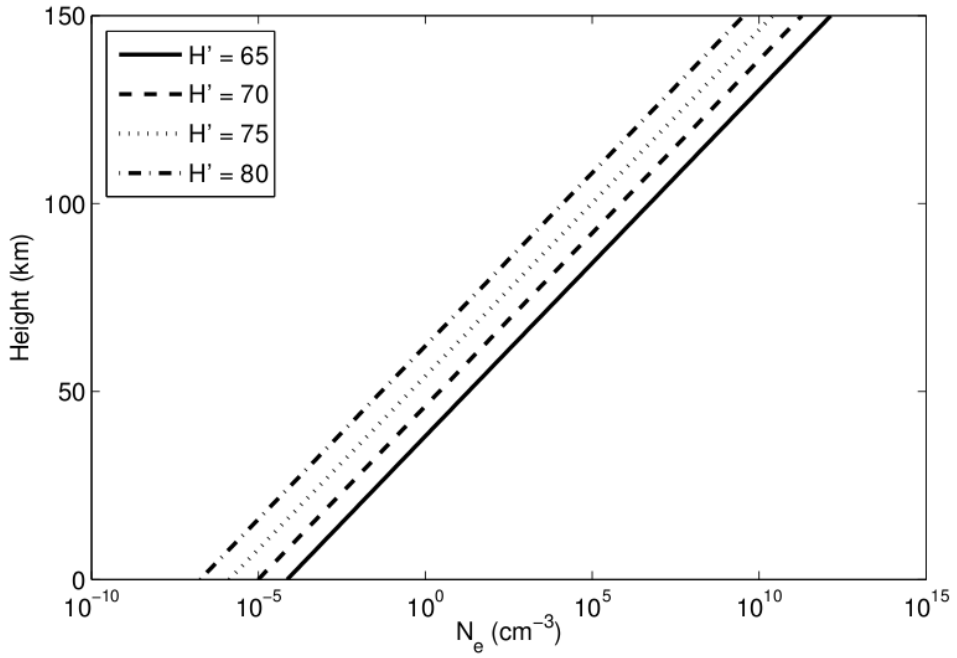


Figure 1.3: Electron density profiles using (1.1) with varying height parameter, $\beta = 0.4km^{-1}$

energy to the neutral atmosphere from free electrons originally excited by the radio wave's electric field. These collisions are characterised by the electron-neutral collision frequency, ν_e. The vertical profile used for this is an exponential one that has become widespread and common in the field since it was first proposed by Wait and Spies [1964] and takes the form:

$$\nu_e(z) = 1.816 \times 1.011e^{-0.15z}s^{-1} \tag{1.2}$$

where z is the altitude in km.

1.5 LWPC

The U.S. Navy's Long Distance Propagation Capability (LWPC) was developed by the Space and Naval Warfare Systems Center in San Diego [Ferguson and Snyder, 1980; Ferguson, 1992; Morfitt, 1977]. The most current documentation of LWPC can be found in Ferguson [1998]. LWPC works with the EIWG where the ionosphere is characterised by a user-defined conductivity profile, an exponential profile being the default. This exponential profile is defined by Wait's parameters, H' and β. The propagation path is broken up into segments or slabs, determined by changes in the ionosphere, ground conductivity and geomagnetic field. LWPC then uses a mode conversion model from Ferguson and Snyder [1980] to connect the horizontal segments along the propagation path. The most accurate method of implementing the mode conversion model is through the

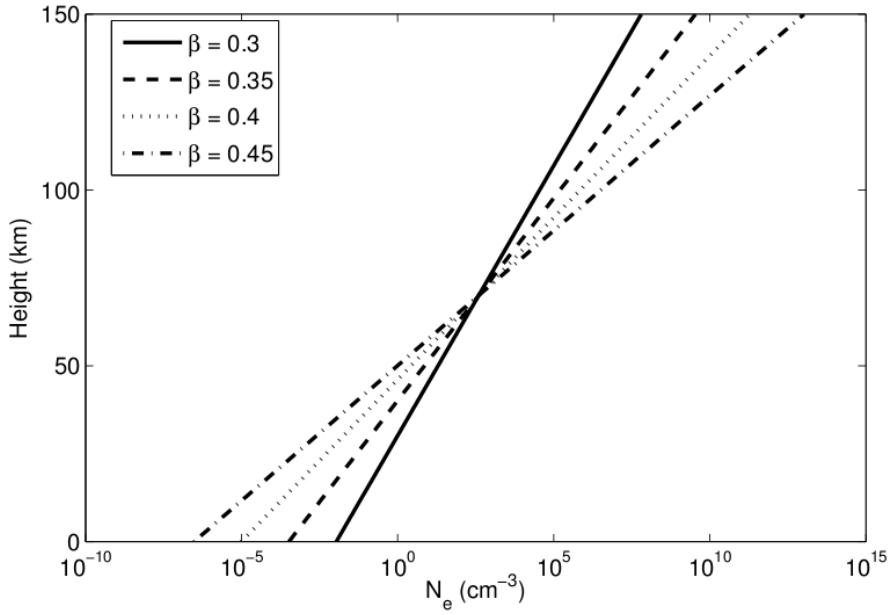


Figure 1.4: Electron density profiles using (1.1) with varying steepness parameter, $H' = 70$ km

vertical integration of the radio fields at the interface between two slabs. Although LWPC has been successfully used by many over the years, it is complicated to use and is a bit of a "black box". This is one of the motivations for creating one's own model, the inner workings are known and can be adjusted if need be.

Chapter 2

Theory

2.1 Waveguide Theory

A waveguide is any structure that controls and directs the propagation of a wave. It has the defining property that it has walls that reflect the waves, generally resulting in very little loss. Consider a two dimensional parallel plate waveguide in free space with a wave propagating in the $\hat{\mathbf{x}}-$direction with perfectly reflecting boundaries at $z = 0$ and $z = a$, as seen in Figure 2.1. The wave vector for such a wave, incident on the boundaries at an angle θ, is given as

$$\mathbf{k} = -(\cos\theta\hat{\mathbf{z}} + \sin\theta\hat{\mathbf{x}}). \tag{2.1}$$

This waveguide theory is purely spatial and time independent. For simplicity, no $\exp(-i\omega t)$ shall be included for now by effectively setting $t = 0$ since the following derivation is true for all times. The electric field incident on the lower boundary at $z = 0$ is then given by

$$\mathbf{E}_i = E_i e^{ik(-z\cos\theta + x\sin\theta)}\hat{\mathbf{y}} \tag{2.2}$$

and the reflected field

$$\mathbf{E}_r = E_r e^{ik(zcos\theta + x\sin\theta)}\hat{\mathbf{y}}. \tag{2.3}$$

The perfectly reflecting boundary condition requires that the total field on the boundary be zero since no energy is transmitted into the walls, $\mathbf{E} = \mathbf{E}_i - \mathbf{E}_r = 0$. Since the boundaries are perfectly reflecting, the magnitude of the reflected field has to equal the magnitude of the incident field, $|\mathbf{E}_r| = |\mathbf{E}_i|$.

The total field $\mathbf{E}$ is thus

$$\mathbf{E} = Ee^{ikx\sin\theta}(e^{-ikz\cos\theta} - e^{ikz\cos\theta})\hat{\mathbf{y}}. \tag{2.4}$$

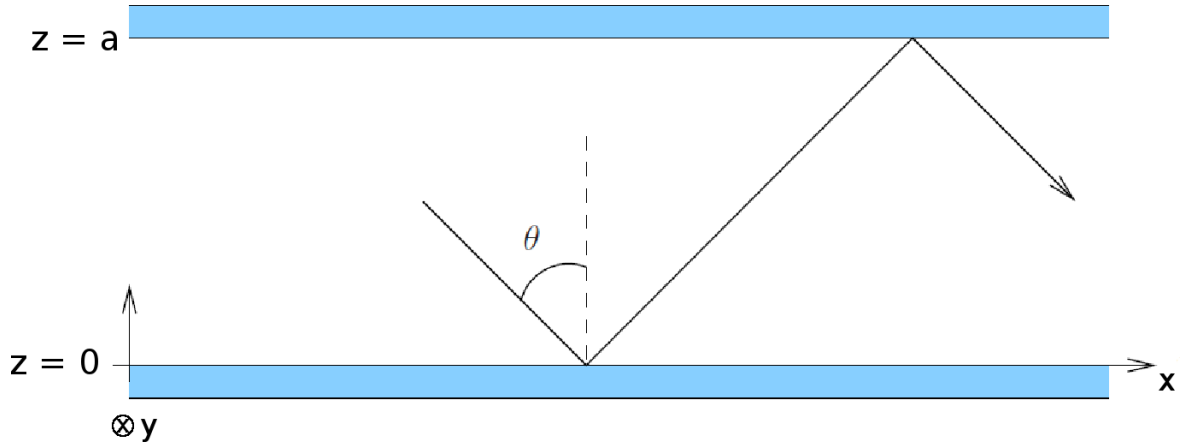


Figure 2.1: Horizontal parallel plate waveguide with incident at angle θ

or

$$\mathbf{E} = 2iE\sin(kz\cos\theta)e^{ikx\sin\theta}\hat{\mathbf{y}} \tag{2.5}$$

where $E = E_i = E_r$ since $|\mathbf{E}_r| = |\mathbf{E}_i|$.

The exponential part at the end of (2.5) refers to the spatial portion of the wave traveling in the z direction with wave number

$$k_x = k\sin\theta = \frac{2\pi}{\lambda_x}. \tag{2.6}$$

This gives us the guide wavelength, related to the free space wavelength, $\lambda_0 = c/f$ where c is the speed of light in vacuum and f is the frequency, by

$$\lambda_x = \frac{\lambda_0}{\sin\theta} \tag{2.7}$$

so that $\lambda_x > \lambda_0$.

The $\sin(kz\cos\theta)$ part of (2.5) indicates this is not a traditional plane wave and has varying amplitude depending on position in the waveguide. Implementing the boundary condition at the upper boundary, $z = a$, leads to the following condition:

$$\sin(ka\cos\theta) = 0 \tag{2.8}$$

which is satisfied if

$$k\cos\theta = \frac{m\pi}{a} \text{ for } m = 1, 2, 3, ... \tag{2.9}$$

From 2.6 and 2.9,

$$k_x^2 = k^2 - \pi\frac{m^2}{a^2} \tag{2.10}$$

For a wave-like motion to occur in the x direction as the wave travels in the waveguide, $k_x \in \Re$,

$$k_x^2 = k^2 - \pi\frac{m^2}{a^2} > 0 \text{ so } k > \frac{m\pi}{a} \tag{2.11}$$

where m is the mode number. For free space propagation, $\omega = |k|c$, which gives the cutoff frequency for propagation for the mth mode as

$$\omega_c = \frac{m\pi}{c}a \tag{2.12}$$

which is the minimum frequency for which k_x is real. Below the cutoff frequency k_x is imaginary, and the wave experiences exponential attenuation. Above this cutoff frequency k_x is real, and the waves will propagate. Exactly at the cutoff frequency $k_x = 0$ so that $\sin\theta = 0$, so that $\theta = 0$ and

$$\lambda = \frac{2a}{m} \tag{2.13}$$

which represents a standing wave with m half wavelengths between the two plates.

The phase velocity along the x direction in the waveguide is given by

$$v_\phi = \frac{\omega}{k_x} = \frac{c}{\sin\theta} > c \tag{2.14}$$

while the group velocity is given by

$$v_g = c\sin\theta \tag{2.15}$$

since the product of the group and phase velocity is equal to the speed of light in the medium,

$$v_g v_\phi = c^2. \tag{2.16}$$

Different types of modes can propagate within a waveguide. These are known as Transverse Electric (TE), Transverse Magnetic (TM) and Transverse Electromagnetic (TEM) modes. The modes that propagate in a waveguide depend on the orientation of the source relative to the direction of propagation and the boundary conditions. When there is no wave magnetic field in the direction of propagation, it is a TM mode, while it is TE if there is no wave electric field in the direction of propagation. An example of the these two field orientations is given in Figure 2.2. It can be seen that if the waves originated from a vertical electric source, a TM mode would propagate in the waveguide while a horizontal electric source would results in TE mode propagation. A TEM mode exists when neither the wave electric nor wave magnetic fields have components in the direction of propagation.

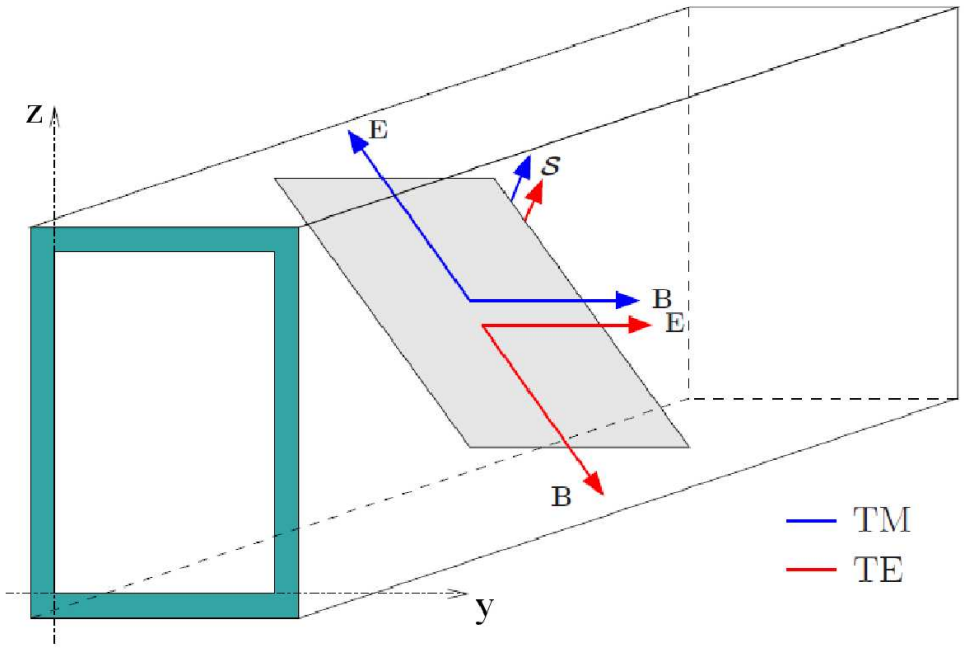


Figure 2.2: TM and TE modes propagating down a waveguide. Illustration courtesy of A. B. Collier.

2.2 Mode Theory Formulation

The mode theory approach used for this thesis is based on the description given by Wait [1962]. A 2D cylindrical coordinate system is chosen in the $z-\rho$ plane with the surface of the Earth at $z = 0$ and the ionosphere at $z = h$ being the reflective boundaries of a free space waveguide. A cylindrical coordinate system is now used rather than a cartesian one due to the circular symmetry around the the transmitter in the $\rho - \phi$ plane. Another reason is the solution equation is appropriate for cylindrical coordinates, as shall be seen later. The transmitter is modelled as a vertical electric dipole placed at the origin as shown in Figure 2.3. This model works on the assumption of a flat Earth and sharply defined ionospheric reflection layer, with free space in between the two boundaries. These assumptions are required to define the waveguide boundaries in the model. Although the problem is truly spherical in nature, the approximation of a flat Earth is valid for propagation paths shorter than the radius of the Earth [Cummer, 2000]. These are the major approximations in the model.

For the radio waves to propagate over hundreds and even thousands of kilometres, the reflections from the waveguide boundaries have to be in phase with each other. This general reflection condition occurs at angles where the ray crosses the waveguide twice, is reflected at each boundary and suffers a net phase shift of $2\pi m$ [Wait, 1962]:

$$R_g(C_m)R_i(C_m)e^{-i2khC_m} = e^{-i2\pi m} \tag{2.17}$$

where R_g and R_i are the reflection coefficients of the ground and ionosphere respectively. $C_m = \cos\theta_m$ and $S_m = \sin\theta_m$ will be used for simplicity while θ_m represents the angle of incidence for mode m. If no loss is assumed, the reflection coefficients of the boundaries are real, and therefore θ_m will be as well. In reality, the waves propagating in the EIWG do experience loss in the ionosphere, which has a complex refractive index as is shown in the following paragraph. This complex refractive index results in a complex reflection coefficient, attenuating the waves and leading to a complex angle of incidence.

The reflected fields can be imagined as fields originating from images of the original dipole located at $z = \pm 2h$, $\pm 4h$, ... [Wait, 1962]. The reflection coefficients are seen to depend on the angle of incidence θ_m, which in turn, depends on the reflection coefficients. This means that without some approximation, it is impossible to analytically solve (2.17) for C_m. We now first assume that the reflection coefficients for both the land and the ionosphere are independent of the angle of incidence. Equation 2.17 reduces to one of the two following equations, depending on the assumed values of R_g and R_i , for

$$\begin{aligned} R_g = +1, R_i = +1 : \\ 2hC_m = m\lambda \\ m = 0, 1, 2, ... \end{aligned} \tag{2.18}$$

and for

$$\begin{aligned} R_g = +1, R_i = -1 : \\ 2hC_m = (m - \frac{1}{2})\lambda \\ m = 1, 2, 3, ... \end{aligned} \tag{2.19}$$

In reality the ionosphere is far from a perfect conductor, but treating it as a perfect magnetic conductor and assuming it has a reflection coefficient of -1 is not a bad approximation [Wait, 1962]. The Earth on the other hand, is a good conductor, and $R_g = +1$ is valid initial assumption. The height used in the reflection condition is found by looking at the complex refractive index, n, for a wave propagating in a cold, magnetised plasma which satisfies:

$$n^2 = 1 - \frac{X}{1 - iZ - \frac{Y_s^2}{2(1-X-iZ)} \pm \frac{1}{1-X-iZ}(\frac{1}{4}Y_s^4 + Y_c^2(1 - X - iZ)^2)^{1/2}} \tag{2.20}$$

This is known as the Appleton-Hartree equation [Budden, 1985]. The three parameters used to define the plasma are:

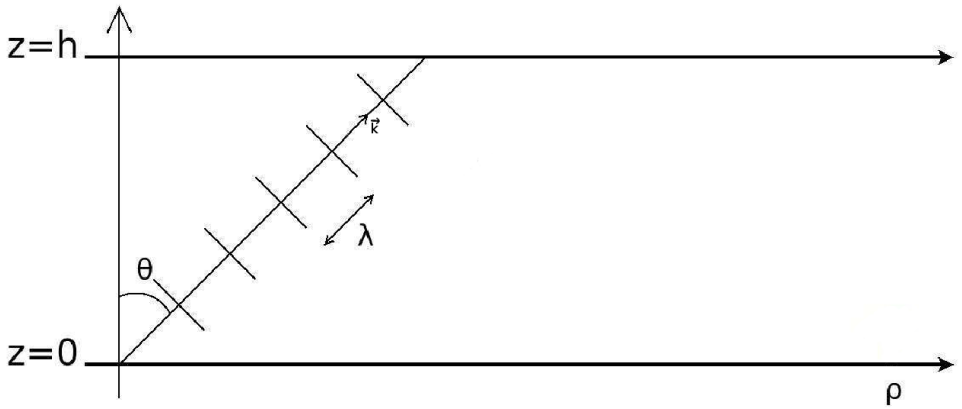


Figure 2.3: Layout of problem where the ionosphere is located at height h and the wave incident at angle θ .

$$X = \frac{\omega_p^2}{\omega^2};\ Y = \frac{|\omega_c|}{\omega};\ Y_s = Y\sin\theta_B;\ Y_c = Y cos\theta_B;\ Z = \frac{\nu_e}{\omega} \tag{2.21}$$

where ν_e is the electron-neutral collision frequency; $\omega = 2\pi f$ is the angular wave frequency; $\omega_p = \sqrt{N_e e^2/m_e\epsilon_0}$ is the plasma frequency; $\boldsymbol{\omega}_c = \mathbf{B}_0|e|/m_e$ is the electron gyro frequency; θ_B is the angle between ambient magnetic field and wave vector. X and Z represent ω_p^2 and ν_e from (1.1) and (1.2), simply scaled by ω^2 and ω respectively. Y is based on the Earth's magnetic field which we shall assume to be constant with height, which is valid at the altitudes encountered in VLF radio propagation in the EIWG. The + sign refers to the ordinary mode wave while the − sign refers to the extraordinary mode. The ordinary mode wave is the component of the wave that is polarised with its electric field component parallel to the ambient magnetic field, $\mathbf{B}_0$, while the extraordinary mode wave is the component polarised with electric field perpendicular to $\mathbf{B}_0$.

The upper boundary of the waveguide is taken as the level where the VLF waves are (partially) reflected. This happens when the refractive index changes dramatically in the space of a wavelength. To investigate this, we look at the behaviour of (2.20). For these purposes, we shall only consider the effect of X and Z. This is appropriate since X is the major defining parameter of the plasma, and in the lower ionosphere, the electron-neutral collisions (Z) have more of an effect than the geomagnetic field (Y). Firstly, let us consider the refractive index when there are no collisions and the magnetic field is also ignored. The Appleton-Hartree equation then reduces to

$$n^2 = 1 - X. \tag{2.22}$$

When X exceeds unity, the refractive index becomes imaginary and the wave will no longer propagate. Physically, this means that at this point, the electron density becomes high enough that the plasma frequency matches the frequency of the wave. This means that the Debye shielding

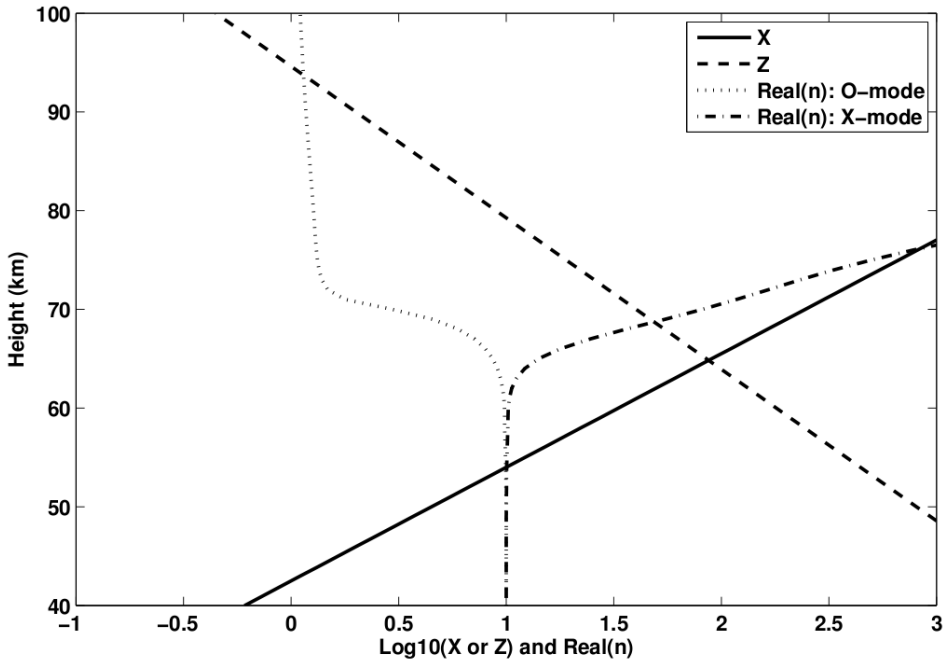


Figure 2.4: Using (1.1), (1.2), (2.20) and (2.21), it is seen that the refractive index changes dramatically in the region where $X = Z$.

from the plasma is fast enough to block the applied electric field from the wave, stopping it from propagating further. If Z is ignored, the altitude at which this would happen is seen in Figure 2.4 to be just above 40km ($\text{Log10}(X) = 0, X = 1$).

If the effect of electron-neutral collisions is to be included, Figure 2.4 shows that Z is far larger than X at low altitudes (< 50km). This means that the electron-neutral collisions suppress the Debye shielding, allowing the wave to propagate when $X > 1$. When $Z = X$, the electron-neutral collision effect weakens, and no longer suppresses the Debye shielding, and the wave is reflected. Looking at (2.20), when $Z \gg X$, the second term on the right hand side has a negligible effect, and the refractive index stays close to one. Near $Z = X$, the second term grows rapidly, and therefore so does the deviation of the refractive index from one as can be seen in Figure 2.4. The height at which $X = Z$ is therefore used as the height of the upper boundary of the waveguide since this is the height at which the refractive index changes dramatically within the space of a wavelength. Y does not have an effect since it has been assumed that the magnetic field does not vary significantly with height at these altitudes.

This height is used in the reflection condition described in (2.19). This is then used to find the modes, or eigenangles, but doesn't take the complex refractive index of the ionosphere or the surface of the Earth into account, although the effect of the surface of the Earth is far smaller. To do this, a perturbative method is used to recalculate the eigenangles that incorporates the

complex valued refractive indices of Earth and Ionosphere [Wait, 1962]:

$$C_m \cong \frac{\pi(m-\frac{1}{2}) + i/(N_g \bar{C_m})}{kh - in^2[n^2 - 1 + (\bar{C_m})^2]^{-1/2}} \tag{2.23}$$

Where $\bar{C_m}$ is the cosine of the angle determined from the original reflection condition in (2.19) and N_g is the refractive index of the surface of the Earth given by [Wait, 1962]:

$$N_g = \sqrt{\frac{\sigma_g + i\epsilon_g\epsilon_0\omega}{i\epsilon\omega}} \tag{2.24}$$

Where σ_g and ϵ_g is the conductivity and dielectric constant of the surface of the Earth.

These eigenangles essentially contain all the information required for the problem, such as the electric field at any point along the propagation path. The field for any mode for a vertical electric dipole is known and can be described in terms of a 0-order Hankel function of the 2nd kind, $H_0^{(2)}$. The Hankel function is used since it expresses the solution of inward and outward propagating waves in cylindrical coordinates. The total field at any height and distance is then calculated by summing all the contributions of the different modes [Wait, 1962]:

$$E_z(\rho, z) = \frac{\mu_0\omega I ds}{2h} \sum_{m=1}^{\infty} S_m^2 H_0^{(2)}(kS_m\rho) f_m(z) \tag{2.25}$$

Where I is the current and ds the length of the radiating antenna. For the sake of this study, this can be taken as an arbitrary constant and will not be mentioned again. S_m is calculated from $S_m = \sqrt{1 - C_m^2}$ and ω is the angular frequency of the wave in question. (2.25) says that the total field at any point in the waveguide is dependent the contributions from different modes, and that each mode's contribution is dependent on the frequency, distance from source, height in waveguide and incidence angle. The number of modes m determines the level of accuracy, with 6-8 normally being more than sufficient for practical uses in this field where the maximum value of ρ will be on the order of a few thousand km. This was confirmed by comparing the results using 8 modes and 20 modes under a variety of waveguide parameters, and finding no significant difference. From this it was concluded that 8 modes will be set as the practical limit used.

Here f_m is an additional height gain function that allows for the field to be calculated at any height in the waveguide and takes the form [Wait, 1962]:

$$f_m(z) = \frac{e^{ikC_m z} + R_g(C_m)e^{-ikC_m z}}{2[R_g(C_m)]^{1/2}} \tag{2.26}$$

This form of f_m assumes that the source is at the surface of the Earth and also takes the effect of the conductivity of the surface of the Earth into account, such as with (2.23).

S_m, $H_0^{(2)}$ and f_m are all complex values, therefore each having influencing the amplitude and phase of each mode propagating in the waveguide. The phenomenon of modal interference is due

to the interaction of the phases of the different modes.

2.3 The FDTD Method

The FDTD method is one of the primary computational electrodynamic modelling techniques and belongs to the general class of grid-based differential time-domain numerical modelling methods. Time-dependent versions of Maxwell's equations in partial differential form are discretised using a central differencing scheme. After electric and magnetic field equations are discretised, they are placed on a lattice grid, with one of the fields shifted so that there is a half a grid step in between an electric and magnetic field component. This is done for both the spatial and temporal index of the electric and magnetic fields and allows for the use of a Leap-Frog time marching scheme as illustrated in Figure 2.5. Additionally, to successfully model electromagnetic propagation, other aspects have to be included or considered. These include the implementation of materials with non-zero conductivity and the effects of a magnetised plasma. This can either be done by inherently incorporating the effects onto the grid by taking them into account when discretising Maxwell's equations or by implementing regions where the fields are calculated using different equations. These aspects of the FDTD method as well as the use of an absorbing region at the edges of the computational domain will now be discussed in further detail.

2.3.1 Finite Differencing and Grid Allocation

Finite differencing is the method applied to transforming the derivatives of the fields in Maxwell's equations from the continuous to the discrete. It is based on the definition of the derivative but is a numerical approximation, which is important to remember for stability and accuracy. The derivative of a function, $f(x)$, around a point, x_0 can be written as

$$\left.\frac{df(x)}{dx}\right|_{x=x_0} = \lim_{\delta\to 0} \frac{f(x_0+\frac{\delta}{2}) - f(x_0-\frac{\delta}{2})}{\delta} + O(\delta^2) \tag{2.27}$$

Taking this into discrete form where $f_m = f(m\Delta x)$ assuming $\delta \equiv \Delta x$ is small enough to ignore higher order terms, this now becomes:

$$\left(\frac{\Delta f}{\Delta x}\right)_m = \frac{f_{m+1/2} - f_{m-1/2}}{\Delta x} \tag{2.28}$$

This is the basis of the finite differencing method and can be applied to any set of differential equations such as those governing electromagnetic wave propagation as shall now be illustrated following the initial procedures of Yee [1966] and a more recent, complete description of Taflove and Hagness [2005].

Consider two of Maxwell's equations, namely the Maxwell-Faraday and Ampere's circuital

equations:

$$\nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t} \tag{2.29}$$

$$\nabla \times \mathbf{H} = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \tag{2.30}$$

or

$$\mu_0 \frac{\partial H_x}{\partial t} = \left(\frac{\partial E_y}{\partial z} - \frac{\partial E_z}{\partial y} \right) \tag{2.31}$$

$$\mu_0 \frac{\partial H_y}{\partial t} = \left(\frac{\partial E_z}{\partial x} - \frac{\partial E_x}{\partial z} \right) \tag{2.32}$$

$$\mu_0 \frac{\partial H_z}{\partial t} = \left(\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} \right) \tag{2.33}$$

$$\epsilon_0 \frac{\partial E_x}{\partial t} = \left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} \right) \tag{2.34}$$

$$\epsilon_0 \frac{\partial E_y}{\partial t} = \left(\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} \right) \tag{2.35}$$

$$\epsilon_0 \frac{\partial E_z}{\partial t} = \left(\frac{\partial H_y}{\partial x} - \frac{\partial E_x}{\partial y} \right) \tag{2.36}$$

For simplicity, let us only consider a grid with one spatial dimension. Maxwell's equations (2.32, 2.36) simplify to:

$$\mu_0 \frac{\partial H_y}{\partial t} = \frac{\partial E_z}{\partial x} \tag{2.37}$$

$$\epsilon_0 \frac{\partial E_x}{\partial t} = \frac{\partial H_y}{\partial x} \tag{2.38}$$

By introducing the following convention

$$(H_y)^q_m = H_y(q\Delta t, m\Delta x) \ , \ (E_z)^q_m = E_z(q\Delta t, m\Delta x), \tag{2.39}$$

And applying the finite differencing scheme, the partial derivatives in (2.37) - (2.38) can be rewritten in discrete form as:

$$\frac{\partial H_y}{\partial x}(q\Delta t, m\Delta x) \equiv \left(\frac{\Delta H_y}{\Delta x} \right)^q_m = \frac{(H_y)^q_{m+1/2} - (H_y)^q_{m-1/2}}{\Delta x} \tag{2.40}$$

$$\frac{\partial H_y}{\partial t}(q\Delta t, m\Delta x) \equiv \left(\frac{\Delta H_y}{\Delta t} \right)^q_m = \frac{(H_y)^{q+1/2}_m - (H_y)^{q-1/2}_m}{\Delta t} \tag{2.41}$$

and

$$\frac{\partial E_z}{\partial x}(q\Delta t, m\Delta x) \equiv \left(\frac{\Delta E_z}{\Delta x}\right)_m^q = \frac{(E_z)_{m+1/2}^q - (E_z)_{m-1/2}^q}{\Delta x} \tag{2.42}$$

$$\frac{\partial E_z}{\partial t}(q\Delta t, m\Delta x) \equiv \left(\frac{\Delta E_z}{\Delta t}\right)_m^q = \frac{(E_z)_m^{q+1/2} - (E_z)_m^{q-1/2}}{\Delta t} \tag{2.43}$$

Once discretised, the fields can be applied to the grid in one of two possible ways: the fields can either be placed on the boundaries of the cells or in the centres. In this case, let the electric field values be placed on the boundaries (integer values) and the magnetic field in the middles of the cells (half-integer values). Apply this for both spatial and temporal indices. Let us now combine the spatial and temporal indices by assigning the spatial index as the main argument with the temporal index in superscript. Equations (2.37) and (2.38) then become

$$\mu_0 \frac{(H_y)_{m+1/2}^{q+1/2} - (H_y)_{m+1/2}^{q-1/2}}{\Delta t} = \frac{(E_z)_{m+1}^q - (E_z)_m^q}{\Delta x} \tag{2.44}$$

$$\epsilon_0 \frac{(E_z)_m^{q+1} - (E_z)_m^q}{\Delta t} = \frac{(H_y)_{m+1/2}^{q+1/2} - (H_y)_{m-1/2}^{q+1/2}}{\Delta x} \tag{2.45}$$

After re-arranging, these yield

$$(H_y)_{m+1/2}^{q+1/2} = (H_y)_{m+1/2}^{q-1/2} + \frac{\Delta t}{\mu_0 \Delta x}\left((E_z)_{m+1}^q - (E_z)_m^q\right) \tag{2.46}$$

$$(E_z)_m^{q+1} = (E_z)_m^q + \frac{\Delta t}{\epsilon_0 \Delta x}\left((H_y)_{m+1/2}^{q+1/2} - (H_y)_{m-1/2}^{q+1/2}\right) \tag{2.47}$$

As can be seen, the temporal evolution of E (or H) at a certain point depends on the adjacent H (or E) values from half a time step before as well as the E (or H) value at the same spatial point, one full time step ago, this is known as the Leap-Frog scheme. The update equations and the Leap-Frog scheme are illustrated in Figure 2.5. The idea of separating electric and magnetic field components by half integer values and using a Leap-Frog time stepping scheme was proposed by Yee [1966] when he first published the FDTD method. Each block is known as a Yee Cell or Cube, depending on the dimensions. A typical 3D Yee Cube is shown in Figure 2.6. The method used here is slightly different to the standard method as it incorporates the current density $\mathbf{J}$ as well, with $\mathbf{E}$ and $\mathbf{J}$ collocated in space [Hu and Cummer, 2006; Simpson, 2010]. Subscripts i and j will refer to x and y positions on the grid respectively. A 2D $xy-$grid is used where the x components of $\mathbf{E}$ and $\mathbf{J}$ are found at $(i + 1/2, j)$, the y components of $\mathbf{E}$ and $\mathbf{J}$ are found at $(i, j + 1/2)$ and H_z is found at $(i + 1/2, j + 1/2)$ as illustrated in Figure 2.7. The time index will be denoted by n, $\mathbf{E}$ and $\mathbf{J}$ will be at integer values of n and $\mathbf{H}$ at half-integer values. The update equations are no longer based purely on Maxwell's equation, but are now also coupled

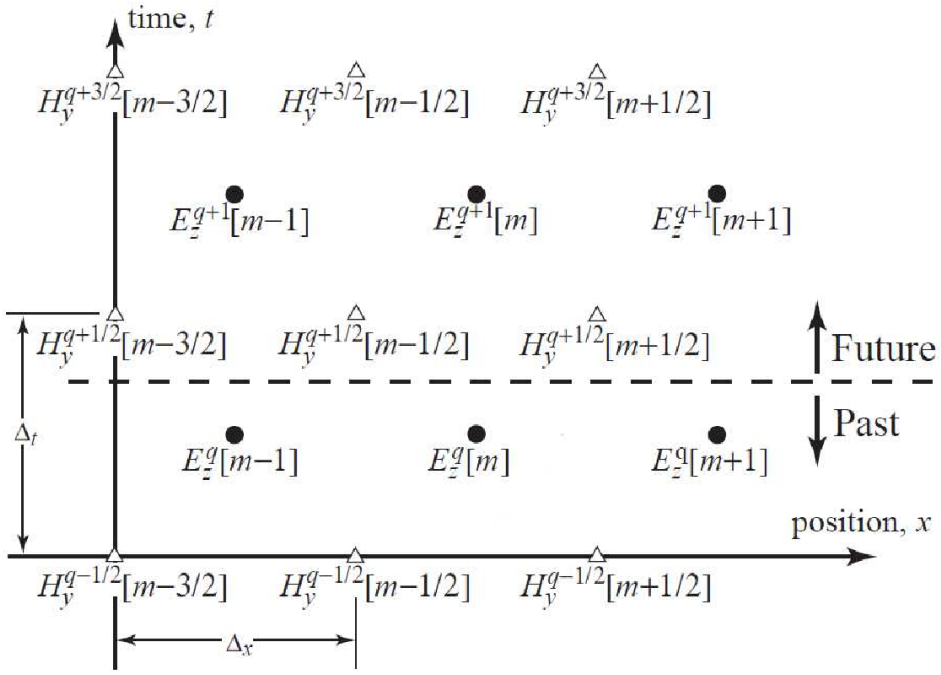


Figure 2.5: E_z and H_y grid layout with position on the horizontal axis and time on the vertical.

to the Lorentz equations of motion due to the incorporation of a background plasma [Simpson, 2010].

$$\nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t} \tag{2.48}$$

$$\nabla \times \mathbf{H} = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \mathbf{J} \tag{2.49}$$

$$\frac{d\mathbf{J}}{\partial t} + \nu_e \mathbf{J} = \epsilon_0 \omega_p^2 \mathbf{E} + \boldsymbol{\omega}_c \times \mathbf{J} \tag{2.50}$$

where ν_e , ω_p and $\boldsymbol{\omega}_c$ are the electron-neutral collision frequency, plasma frequency and electron cyclotron frequency, respectively. The current density refers to the contribution from the electron plasma. The full 3 dimensional vector equations are not needed since the aim is to construct a 2 dimensional model, similar to that of the mode theory model, illustrated in Figure 2.3. After separating equations (2.48 - 2.50) into their individual spatial components and reducing the system to 2 dimensions, the following equations emerge:

$$\frac{\partial H_z}{\partial t} = \frac{1}{\mu_0}\left(\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x}\right) \tag{2.51}$$

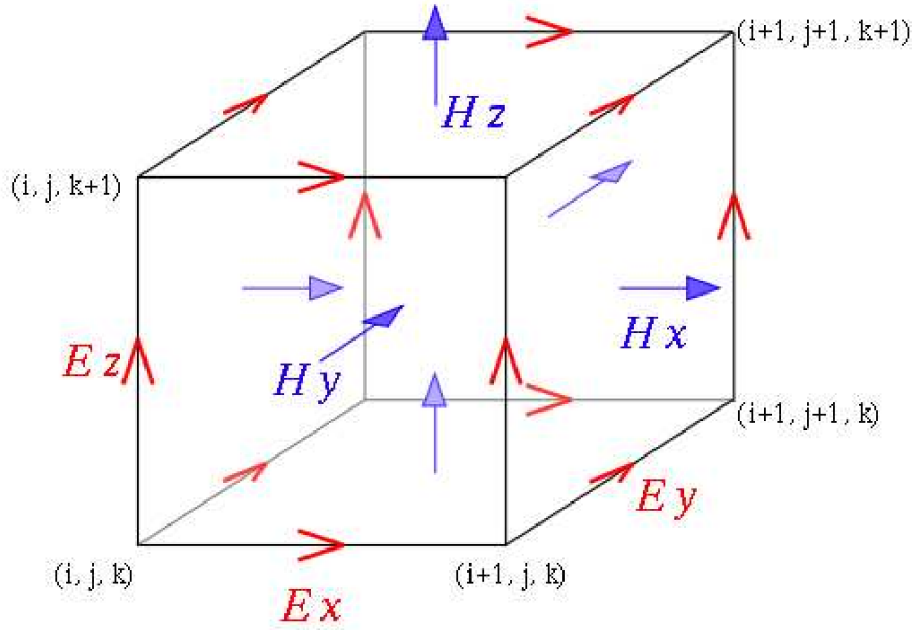


Figure 2.6: 3D Yee Cube. Stanford AWESOME Tutorial,

$$\frac{\partial E_x}{\partial t} = \frac{1}{\epsilon_0}\left(\frac{\partial H_z}{\partial y} - J_x\right) \tag{2.52}$$

$$\frac{\partial E_y}{\partial t} = \frac{1}{\epsilon_0}\left(-\frac{\partial H_z}{\partial x} - J_y\right) \tag{2.53}$$

$$\frac{dJ_x}{dt} + \nu_e J_x = \epsilon_0 \omega_p^2 E_x - \omega_c J_y \tag{2.54}$$

$$\frac{dJ_y}{dt} + \nu_e J_y = \epsilon_0 \omega_p^2 E_y + \omega_c J_x \tag{2.55}$$

The system is now 2 dimensional and only concerned with the z-component of the Earth's magnetic field because there is no J_z for the x- and y-components of the electron cyclotron frequency to interact with. This means $\mathbf{B}_0 = B_0 \mathbf{e}_z$, and therefore $\boldsymbol{\omega}_c$ can be written as,

$$\boldsymbol{\omega}_c = \frac{|e|B_0}{m_e}\mathbf{e}_z \text{ and hence } \boldsymbol{\omega}_c = \omega_c \mathbf{e}_z. \tag{2.56}$$

The finite differencing scheme is then applied to the derivatives and the grid as seen in Figure 2.7 to yield the following equations

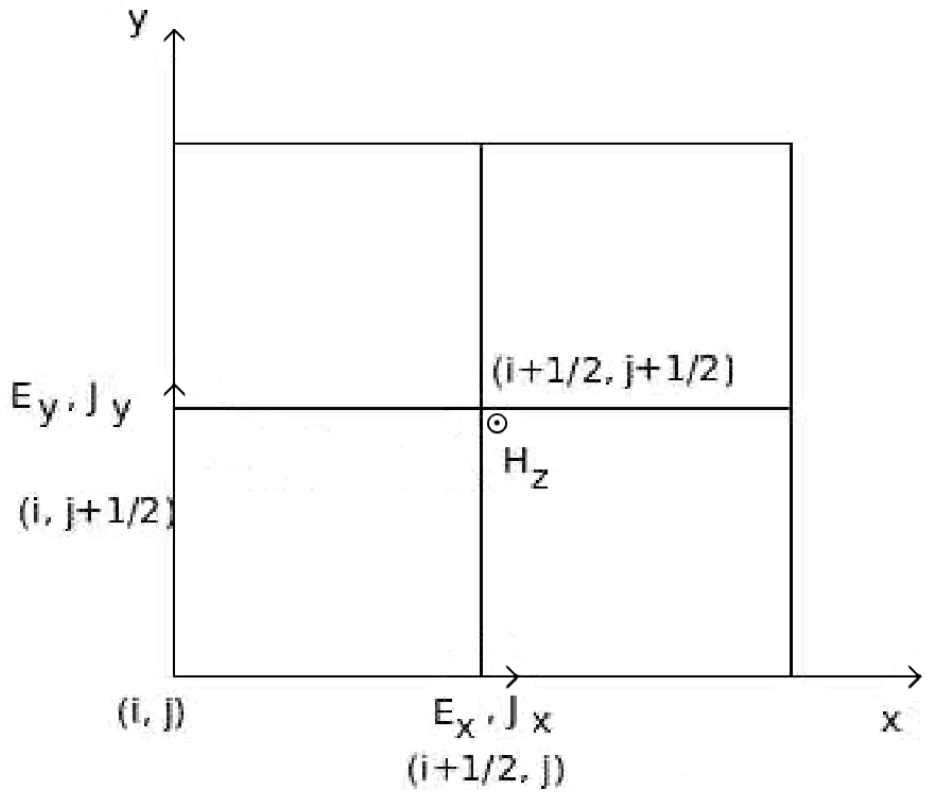


Figure 2.7: Layout of E-J collocated grid with the H_z with each electric field or current density component half a grid step away from a magnetic field component and *viceversa*.

$$(H_z)^{n+1/2}_{i+1/2,j+1/2} = (H_z)^{n-1/2}_{i+1/2,j+1/2} + \frac{\Delta t}{\mu_0}\left[\frac{(E_x)^n_{i+1/2,j} - (E_x)^n_{i+1/2,j+1}}{\Delta y} - \frac{(E_x)^n_{i+1/2,j} - (E_x)^n_{i+1/2,j+1}}{\Delta x}\right]$$

$$(E_x)^{n+1}_{i+1/2,j} + \frac{1}{2}(J_x)^{n+1}_{i+1/2,j} = (E_x)^n_{i+1/2,j} - \frac{1}{2}(J_x)^n_{i+1/2,j} + \frac{dt^2}{\epsilon_0\mu_0\Delta y^2}\left[(H_z)^{n+1/2}_{i+1/2,j+1/2} - (H_z)^{n+1/2}_{i+1/2,j-1/2}\right]$$

$$(E_y)^{n+1}_{i,j+1/2} + \frac{1}{2}(J_y)^{n+1}_{i,j+1/2} = (E_y)^n_{i,j+1/2} - \frac{1}{2}(J_y)^n_{i,j+1/2} + \frac{dt^2}{\epsilon_0\mu_0\Delta x^2}\left[(H_z)^{n+1/2}_{i+1/2,j+1/2} - (H_z)^{n+1/2}_{i-1/2,j+1/2}\right]$$

$$\frac{(\omega_p dt)^2}{2}(E_x)^{n+1}_{i+1/2,j} + \left(1 + \frac{\nu_e dt}{2}\right)(J_x)^{n+1}_{i+1/2,j} + \frac{\omega_c dt}{2}(J_y)^{n+1}_{i,j+1/2} = \frac{(\omega_p dt)^2}{2}(E_x)^n_{i+1/2,j} + \left(1 - \frac{\nu_e dt}{2}\right)(J_x)^n_{i+1/2,j} - \frac{\omega_c dt}{2}(J_y)^n_{i,j+1/2}$$

$$-\frac{(\omega_p dt)^2}{2}(E_y)^{n+1}_{i,j+1/2} - \frac{\omega_c dt}{2}(J_x)^{n+1}_{i+1/2,j} + \left(1 + \frac{\nu_e dt}{2}\right)(J_y)^{n+1}_{i,j+1/2} = \frac{(\omega_p dt)^2}{2}(E_y)^n_{i,j+1/2} + \frac{\omega_c dt}{2}(J_x)^n_{i+1/2,j} - \left(1 - \frac{\nu_e dt}{2}\right)(J_y)^n_{i,j+1/2}$$

The equations for calculating the newest values of the four **E** and **J** components can be re-arranged and written in matrix form. This is useful as it allows for easy calculation of the coefficients used to update the fields. The update equation for H_z remains the same and is performed before the **E** and **J** matrix equation updates. The equations in matrix form are

$$A \begin{bmatrix} E_x^{n+1} \\ E_y^{n+1} \\ \bar{J}_x^{n+1} \\ \bar{J}_y^{n+1} \end{bmatrix} = B \begin{bmatrix} E_x^{n} \\ E_y^{n} \\ \bar{J}_x^{n} \\ \bar{J}_y^{n} \end{bmatrix} + C \begin{bmatrix} \frac{\Delta \bar{H}_z}{\Delta y} \\ \frac{\Delta \bar{H}_z}{\Delta x} \\ 0 \\ 0 \end{bmatrix} \tag{2.57}$$

or

$$\begin{bmatrix} E_x^{n+1} \\ E_y^{n+1} \\ \bar{J}_x^{n+1} \\ \bar{J}_y^{n+1} \end{bmatrix} = A^{-1}B \begin{bmatrix} E_x^{n} \\ E_y^{n} \\ \bar{J}_x^{n} \\ \bar{J}_y^{n} \end{bmatrix} + A^{-1}C \begin{bmatrix} \frac{\Delta \bar{H}_z}{\Delta y} \\ \frac{\Delta \bar{H}_z}{\Delta x} \\ 0 \\ 0 \end{bmatrix} \tag{2.58}$$

where

$$A = \begin{bmatrix} 1 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & \frac{1}{2} \\ -\frac{(\omega_p dt)^2}{2} & 0 & (1 + \frac{\nu_e dt}{2}) & \frac{\omega_c dt}{2} \\ 0 & -\frac{(\omega_p dt)^2}{2} & -\frac{\omega_c dt}{2} & (1 + \frac{\nu_e dt}{2}) \end{bmatrix} \tag{2.59}$$

$$B = \begin{bmatrix} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & -\frac{1}{2} \\ \frac{(\omega_p dt)^2}{2} & 0 & (1 - \frac{\nu_e dt}{2}) & -\frac{\omega_c dt}{2} \\ 0 & \frac{(\omega_p dt)^2}{2} & \frac{\omega_c dt}{2} & (1 - \frac{\nu_e dt}{2}) \end{bmatrix} \tag{2.60}$$

$$C = \begin{bmatrix} \frac{dt^2}{\epsilon_0 \mu_0 \Delta y} & 0 & -1 & 0 \\ 0 & \frac{dt^2}{\epsilon_0 \mu_0 \Delta x} & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{2.61}$$

By multiplying with the inverse of A, the update equations emerge in a way that is easy to implement in code. If the ionospheric properties remain constant and do not have a time dependence, these matrices only need to be calculated once. In the final update equations that are implemented in the code (2.58), $\bar{\mathbf{J}}$ and $\bar{H}$ are scaled versions of **J** and H, while **E** is not scaled. The scaling scheme is as follows:

$$\bar{H}_z = \frac{\mu_0 \Delta x}{\Delta t} H_z \tag{2.62}$$

$$\bar{\mathbf{J}} = \frac{\Delta t}{\epsilon_0} \mathbf{J}. \tag{2.63}$$

The fields are scaled to avoid extremely large numbers in the coefficient matrices. This avoids stability and accuracy problems that can arise due to the large numbers. It should be noted that there is a coupling between the x and y components of the electric field through the interaction of the x and y current densities with the ambient magnetic field. This means that spatial interpolation of the field values needs to be done to account for this since there is no specific value for a x field at a y position and *vice versa*. As an example, the E_y field around $E_{x\ i+1/2,j}$ is the average of the four surrounding E_y values.

$$(E_y)^n_{i+1/2,j} = \frac{1}{4}((E_y)^n_{i,j+1/2} + (E_y)^n_{i,j-1/2} + (E_y)^n_{i+1,j+1/2} + (E_y)^n_{i+1,j-1/2} \tag{2.64}$$

2.3.2 Nearly Perfectly Matched Layer

There are various methods of implementing boundary conditions in FDTD. The boundaries need to be able to deal with an incoming wave without creating non-physical reflections, or at least minimising them. There are two different boundary conditions which have been popular in the past, namely the Absorbing Boundary Condition (ABC) and the Perfectly Matched Layer (PML), with different variations of each. The ABC generally works by estimating the field at the edge of the computational domain by looking at the incident fields adjacent to the boundary at previous time steps. These kind of ABCs can work well but normally only if specifically designed for a certain situation. The PML was first proposed by Berenger [1984] and has now become the preferred method for most users. There are many variations of the PML but all versions effectively act as a lossy layer that absorbs the fields traveling out, away from the simulation space. The PML is implemented in such a way that there is no loss in the direction tangential to the interface while there is loss in the direction normal to the interface. In a continuous space, the PML absorber is perfectly matched to the host medium and would totally absorb the incident waves. However, the FDTD method works in a discrete space where the electric and magnetic fields are separated in space. This leads to discretisation errors that will result in the PML not acting ideally. In this thesis, the version implemented is the Nearly Perfectly Matched Layer (NPML), introduced by Hu and Cummer [2006]. The NPML is very similar to the regular PML except that it has the additional goal of maintaining the form of the original update equations. Consider Maxwell's equations (2.33, 2.34, 2.35) for a system with 2 spatial dimensions, looking only at field variations in the x and y directions. If the fields are harmonic with an $exp(i\omega t)$ variation, they can be written as:

$$i\omega H_z = \frac{1}{\mu_0}\left(\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x}\right) \tag{2.65}$$

$$i\omega E_x = \frac{1}{\epsilon_0}\left(\frac{\partial H_z}{\partial y}\right) \tag{2.66}$$

$$i\omega E_y = -\frac{1}{\epsilon_0}\left(\frac{\partial H_z}{\partial x}\right) \tag{2.67}$$

For example, to apply the NPML that absorbs waves going in the x-direction, a complex coordinate stretching is done [Cummer, 2003], in which

$$\partial x \Rightarrow \partial \tilde{x} = (1 + \frac{\sigma_x(x)}{i\omega})\partial x. \tag{2.68}$$

This coordinate conversion maintains the real part of the fields, while attenuating through the $\frac{\sigma_x(x)}{i\omega}$ term. Through the implementation of an effective NPML, the waves propagating away from the computational space of the main problem will be attenuated enough as to not cause artificial reflections when they reach the boundaries of the global computational space. The other components of the stretched field are handled in an analogous fashion.

For the NPML region, the update equations remain the same, but the stretched fields are used instead. The conductivity of the NPML, which is somewhat arbitrary, increases as you go deeper in the layer. The exact values for $\sigma_x(x)$ that will work effectively depend on the frequency of the wave and thickness of the NPML as well as grid discretisation size. Equations (2.65) to (2.67) now take the form

$$i\omega H_z = \frac{1}{\mu_0}\left(\frac{\partial E_x}{\partial y} - \frac{\partial \tilde{E}_y}{\partial x}\right) \tag{2.69}$$

$$i\omega E_x = \frac{1}{\epsilon_0}\left(\frac{\partial H_z}{\partial y}\right) \tag{2.70}$$

$$i\omega E_y = -\frac{1}{\epsilon_0}\left(\frac{\partial \tilde{H}_z}{\partial x}\right) \tag{2.71}$$

and

$$\tilde{F} = (1 + \frac{\sigma_x}{i\omega})^{-1} F \text{ where } F \in \{E_y, H_z\} \tag{2.72}$$

The update equations used in the NPML have the same form as the original equations (2.58), making it computationally easy to implement. The only difference being that, in the NPML region the fields are damped by the introduction of the conductivity associated with the NPML.

2.3.3 Surface Impedance Boundary Condition

Simulating a material of finite conductivity can be challenging in FDTD if the wavelength in the material is significantly smaller than that in free space. This would require that the grid cells be made even smaller, which then makes the code more computationally expensive. An effective method of avoiding this is to employ a Surface Impedence Boundary Condition (SIBC) at the

interface. The SIBC replaces the lossy material by calculating the impedance at the boundary and finding the relationship between the fields tangential to the interface. It is based on the impedance equation that relates the tangential components of the electric and magnetic field at the interface [Maloney and Smith, 1992].

$$\mathbf{E}_{\parallel}(\omega) = Z_s(\omega)[\hat{\mathbf{n}} \times \mathbf{H}_{\parallel(\omega)}] \tag{2.73}$$

where $\mathbf{E}_{\parallel}$ and $\mathbf{H}_{\parallel}$ are the tangential electric and magnetic fields and $\hat{\mathbf{n}}$ is the unit normal vector pointing out at the interface.

The surface impedance of a material with conductivity σ, permeability μ and permittivity ϵ adjacent to free space is given by:

$$Z_s(\omega) = \sqrt{\frac{i\omega\mu}{\sigma + i\omega\epsilon}} \tag{2.74}$$

for an electromagnetic wave of radial frequency ω. Assuming that $\sigma \gg \omega\epsilon$, which is known as the Leontovich impedance boundary condition [Maloney and Smith, 1992; Soloviev, 2003], (2.74) can be approximated as

$$Z_s(\omega) \cong \sqrt{\frac{i\omega\mu}{\sigma}} = (1+i)\sqrt{\frac{\omega\mu}{2\sigma}} \tag{2.75}$$

The implications of this are discussed later on.

From now on we will assume that the interface is in the $x-z$ plane, so the tangential fields are now E_x and H_z. Following Beggs et al. [1992], (2.73) can now be written as

$$E_x(\omega) = (R_s(\omega) + i\omega L_s(\omega))H_z(\omega) \tag{2.76}$$

where R_s and L_s are the surface resistance and inductance defined by

$$\begin{aligned} R_s(\omega) &= \sqrt{\frac{\omega\mu}{2\sigma}} \\ L_s(\omega) &= \sqrt{\frac{\mu}{2\sigma\omega}} \end{aligned} \tag{2.77}$$

To be able to implement this in FDTD code, the SIBC (2.73) has to be taken to the time domain. This is simple for the monochromatic case since ω is constant:

$$E_x(t) = R_s H_z(t) + L_s \frac{\partial}{\partial t} H_z(t) \tag{2.78}$$

After discretising, performing finite differencing, averaging of the fields and rearranging, (2.78)

can be written as

$$\begin{aligned}(H_z)^{n+1/2}_{i+1/2,j+1/2} &= \left[\frac{\mu_0\Delta y+L_s-R_s\Delta y/2}{\mu_0\Delta y+L_s+R_s\Delta y/2}\right] \\ &+ \frac{\Delta t}{\mu_0\Delta y+L_s+R_s\Delta t/2}[(E_y)^n_{i,k+1/2} \\ &- (E_y)^n_{i+1/2,k+1/2}+(E_x)^n_{i+1/2,k+1}].\end{aligned} \tag{2.79}$$

The appearance of an E_y term is the result of the temporal derivative of H_z in (2.78) being written in terms of spatial derivatives of E_x and E_y, such as in (2.33).

Chapter 3

Data Collection

3.1 VLF Antenna and UltraMSK

Narrowband data from VLF transmitters are collected using a pair of orthogonal induction loop antennas or a whip antenna. The orthogonal loop setup has the two loops in an E-W and N-S orientation, as seen in Figure 3.1. This has the capability to perform direction finding on a signal by looking at the ratio of power flux and phase through the two loops. A single loop is used when gathering narrowband data, with the choice of the loop depending on which loop is receiving higher power for the frequency in question. UltraMSK is a software package developed by James Brundell of the University of Otago, Dunedin, New Zealand. It is used in scientific research applications that involve measuring the phase and amplitude of a Minimum Key Shifted (MSK) modulated narrowband VLF radio signal and has the ability to receive both 100 and 200 baud MSK signals. UltraMSK is able to monitor multiple stations simultaneously. It is relatively simple to set up and has minimal hardware requirements. All that is needed is a VLF antenna, a computer with a sound card and a GPS receiver to accurately record the time. The GPS Pulse Per Second (PPS) is also used to calibrate the computer's clock for accurate sampling of the signal. The advantage of using the UltraMSK system is that it uses a normal audio sound card in a computer and not specialised, expensive scientific equipment. This is due to the useful coincidence that VLF frequency range falls within the human audio range and sound cards are developed to operate at these frequencies. The only consideration with the sound card is the number of available inputs. One input is required for each antenna channel and an additional one for the GPS PPS signal.

Examples of the UltraMSK data are given in Figures 3.3 and 3.4 which show the amplitude and phase of the transmitters with call signs NWC (North West Cape, Australia, 21.816° S, 114.166° E, 19.8 kHz) and NAA (Cutler, Maine, USA, 44.646° N, 67.281° W, 24 kHz) detected at Marion Island. The NWC path is slightly more than 7200 km long while the NAA path is double that at 14400 km. Both graphs are similar in that the amplitude and phase signal can be divided into

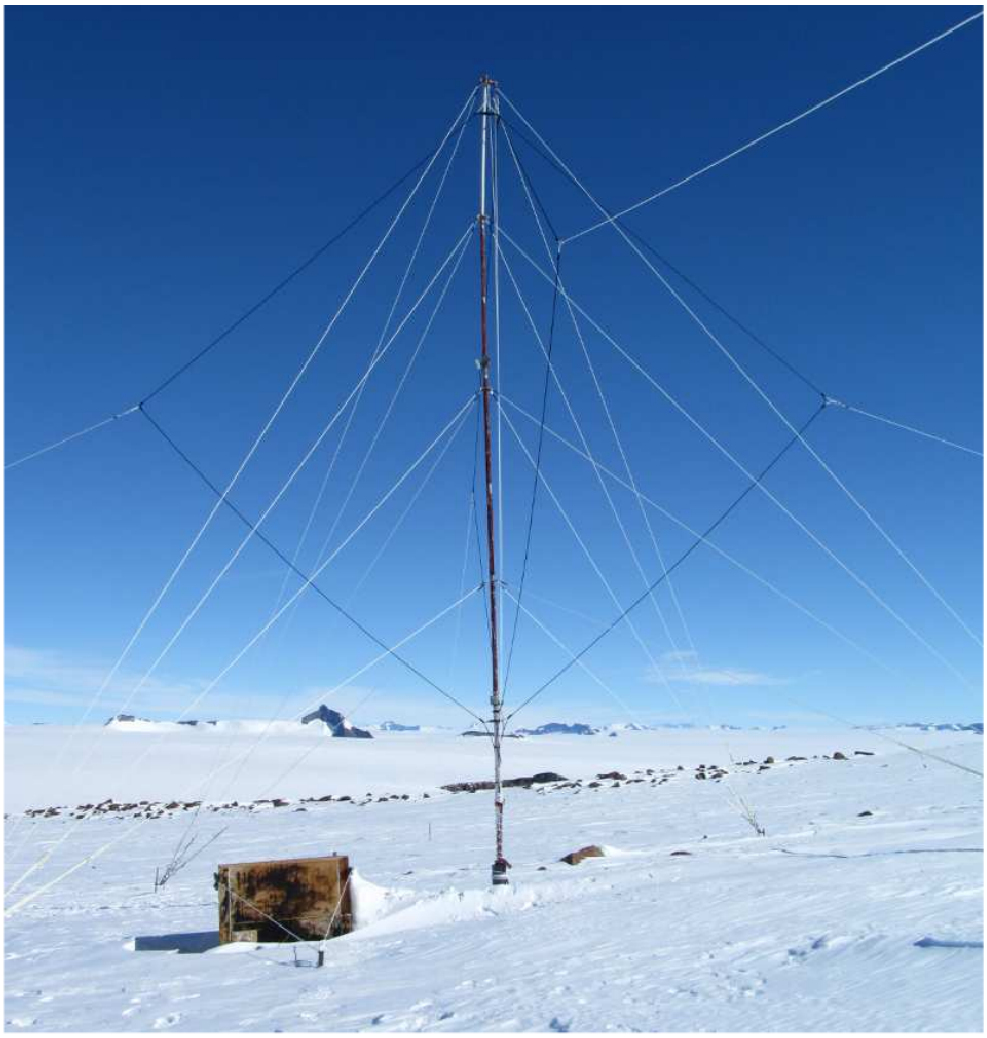


Figure 3.1: The VLF loop antenna at SANAE IV, Antarctica.

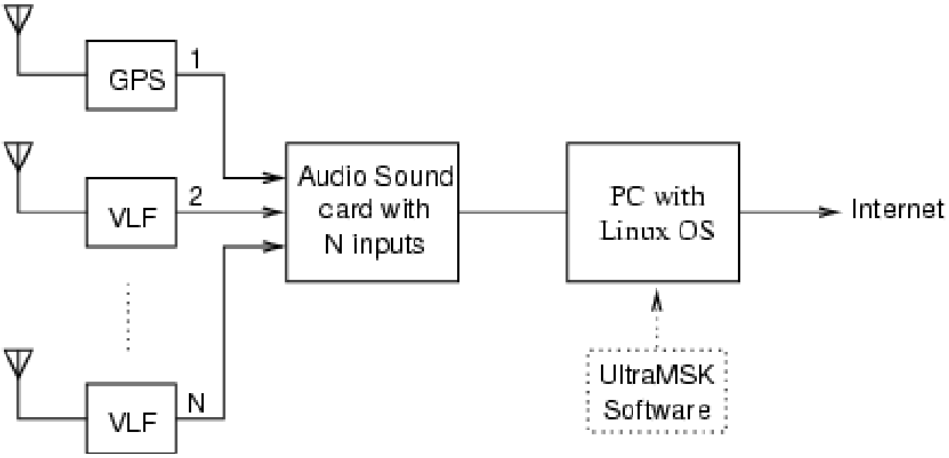


Figure 3.2: Block diagram schematic of hardware layout for operation of UltraMSK.

three regions; a full day path, a full night path, and a combination day-night path. The full day profile is visible from around 01 to 12 UTC in Figure 3.3 and 10 to 17 UTC in Figure 3.4. During this time, the entire path from transmitter to receiver was being exposed to solar radiation. The constant stream of solar radiation results in waveguide conditions remaining unchanged, leading to a steady, stable amplitude as well as phase. The sudden changes in amplitude are due to the passage of the day-night terminator over the transmitter or receiver. Another noticeable feature here is that at the edges of the all-day region, the phase changes dramatically with the passage of the terminator, being constant until this point. The amplitude on the other hand shows a slight decrease before the sudden effect of the terminator passing over. These differences can be attributed to the fact that the amplitude of a signal has been found to be more sensitive to the electron density gradient, while the phase is more sensitive to the height [Thomson, 1993]. This is due to the amplitude of the partial reflections and the absorption in the ionosphere which depends on the electron densities below the reflective layer. The phase on the other hand is dependent on the total distance traveled by the wave which is affected by the height of the ionosphere. At either side of the all day region, the profile for a part day and part night path can be seen. This region can easily be noticed with a steady increase or decrease in phase. This can be attributed to the terminator traveling over the propagation path, changing the total propagation distance at a constant rate as the day/night ratio of the path changes. The part day - part night region in the morning in the amplitude plot of Figure 3.4 exhibits some interesting behaviour. The amplitude is seen to periodically rise and drop. It should be noted that the maximum amplitude at the peak of these periodic structures is equal to the constant amplitude experienced during an all day path. This can be attributed to a combination of two things, the long propagation distance and mode conversion taking place at the day/night boundary. For a long path like this, the first mode will be the most dominating factor to the observed signal, be it for day or night. When mode conversion takes place, some or even all of the energy in a propagating mode will remain in that same mode on the other side of the terminator. This depends on the position of the terminator and the behaviour of reflected waves in that region. If the terminator is in such a position so that there is little or no energy lost in the first mode when crossing the day-night interface, the amplitude will remain roughly the same. This is what is happening at the maxima of the periodic structure. In between these points, the mode conversion takes place in such a way so that the amplitude of the first mode after conversion is less than before. The night time region is seen to be more stable in both phase and amplitude for the longer NAA path than the NWC. A night time ionosphere allows for more modes to propagate due to the lower attenuation but is also more variable. For the short NWC path, there would be multiple modes contributing to the observed signal. On a very long path, only the first one or two modes will have any significant contribution. The fact that a single mode is dominant will result in the phase being relatively constant while the absorption will still vary at night without the constant solar wind.

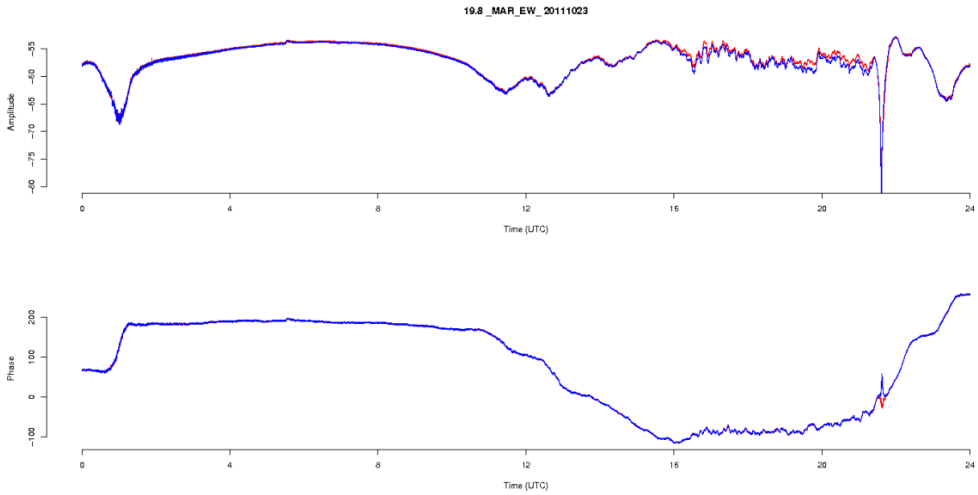


Figure 3.3: Amplitude and Phase of the NWC transmitter detected at Marion Island.

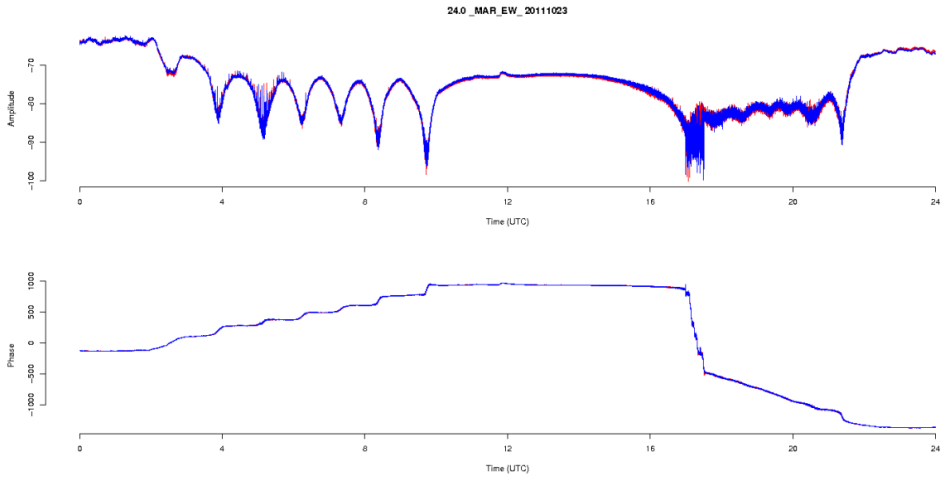


Figure 3.4: Amplitude and Phase of the NAA transmitter detected at Marion Island.

3.2 DEMETER and GOES

DEMETER (Detection of Electro-Magnetic Emissions Transmitted from Earthquake Regions) is a French satellite that was launched in 2004 with the main objective of studying ionospheric disturbances due to seismic and volcanic activity [http://smsc.cnes.fr/DEMETER/]. It is in a quasi Sun-synchronous orbit at 10:30 and 22:30, shortly before local noon and midnight with an altitude of around 710km. It completes 14 orbits a day. It was launched on 29 June2004 and stayed in operation until December 2010. In December 2005, the orbit was lowered to 660 km. There are a few scientific payloads on board the satellite, but of interest to this thesis are the electric sensors that can measure fields from DC up to 3.5 MHz.

GOES (Geostationary Operational Environmental Satellite) is a system of satellites operated by the U.S. National Oceanic and Atmospheric Administration (NOAA). Satellites in geostationary orbit are located directly above the equator (0°) and orbit at the same angular velocity as the Earth, meaning they remain directly above the same location on the Earth. Several GOES satellites have been put into orbit, with GOES-14 being the source of some data in this thesis. The main mission of the GOES satellites is to observe the Earth's infrared radiant energy as well as visible reflected solar energy. The GOES payload also includes several pieces of equipment designed to monitor the near-Earth space environment. These include proton and alpha particle detectors, a magnetometer, and an X-ray sensor which provides the X-ray flux used in this thesis.

Chapter 4

Mode Theory

The implementation of the 2D waveguide Mode Theory model will be discussed here with some illustrative examples. A number of input variables are decided on by the user, these include the ionospheric parameters, H' and β, and ambient magnetic field strength and direction, (dip angle and azimuth) while the transmitter is assumed to be located at the origin ($\rho = 0, z = 0$). The other parameters needed are the signal frequency, the conductivity and dielectric constant of the surface of the Earth along the propagation path as well as the bearing of the propagation path to enable calculation of the angle between the wave vector and ambient magnetic field. The user also specifies the height at which the results will be calculated. This is normally ground level or the reflection height of the ionosphere. This height dependence arises from the two height dependent used to quantify the conductivity, and hence refractive index, of the ionosphere. These height dependent parameters being the electron density N_e, and electron-neutral collision frequency ν_e, coming from (1.1) and (1.2) respectively. The program now has all the necessary information required to calculate the electron density profile, and therefore the three ionospheric parameters X, Y and Z (2.21). Using the height where $X = Z$ as the upper boundary of the EIWG, the mode angles are calculated with the original assumption that $R_i = -1$ and $R_g = +1$, as in (2.19). These initial angles allow for the calculation of the angle between each mode of propagation and the ambient magnetic field. All required information is then available to calculate the complex refractive index at the height of reflection through the Appleton-Hartree equation in (2.20). The eigenangles for the different modes are then recalculated using (2.23), but with the complex refractive index of the ionosphere taken into account. These angles, along with the height gain function (2.26) are then used to calculate the electric field at any point in the waveguide in terms of a 0-order Hankel function of the second kind (2.25).

4.1 Program Results

One of the advantages of using one's own program is that the results from individual components can be investigated. Examining the different properties for different modes is one of the advantages of this program which LWPC, for example, does not have.

4.1.1 Height Gain Function

The height gain function, $f_m(z)$, allows for the calculation of the field at any height in the waveguide. This will allow for model results to be compared to DEMETER satellite data. Figures 4.1 and 4.2 show the real and imaginary parts of (2.26) for the first three odd modes for a certain k. (2.26) has an some inherit dependencies on factors such as H', β and ω through the calculation of C_m, but Figures 4.1 and 4.2 serve to illustrate the variation in height of (2.26) due to mode number. The imaginary part arises due to the complex refractive index leading to complex modal angles θ_m, from (2.23). This then in turn leads to the height gain function having a real and imaginary part. It should be noted that the real and imaginary parts start at 1 and 0, respectively, for all modes. This indicates that when the transmitter and receiver are at the same level, all modes are weighted equally. Figure 4.1 indicates how different modes have different contributions to the total field at different heights in the waveguide. For $m = 1$, we see that that the real part of the height gain function remains between 0.5 and 1 all the way up to 100km, indicating that the first mode will always account for a large part of the total field. Higher order modes have more variable contributions across the waveguide as is expected due to the higher number of nodes associated with higher modes. Figure 4.2 arises due to the complex modal angles that lead to attenuation in the waveguide. This figure is also therefore an indication of the variation of attenuation experienced by different modes in the waveguide. Another height gain function can be included to account for the variation of the height of the transmitter. This was not done since the work here focuses on the modeling of man made VLF transmitters which are all found at ground level. The height gain function for the transmitter has the exact same form as (2.26).

4.1.2 Attenuation

With this model, the theoretical attenuation experienced by each mode can also be investigated. In waveguide theory, once the modes or eigenangles are known, essentially all the information can be acquired, including the attenuation [Wait, 1962]. The waves are attenuated since the refractive index has an imaginary component [Wait, 1962]. This complex refractive index means that the modal angles will also be complex. The attenuation rate is then equal to $k\mathrm{Im}(S_m)$, [Wait, 1962], where the group velocity, $v_g = S_m c$ (2.15). The effect of different parameters on the attenuation rate can investigated by changing these parameters that change the modal angles. Changes due to mode number and ionospheric reflection height are given as an example here, different modes

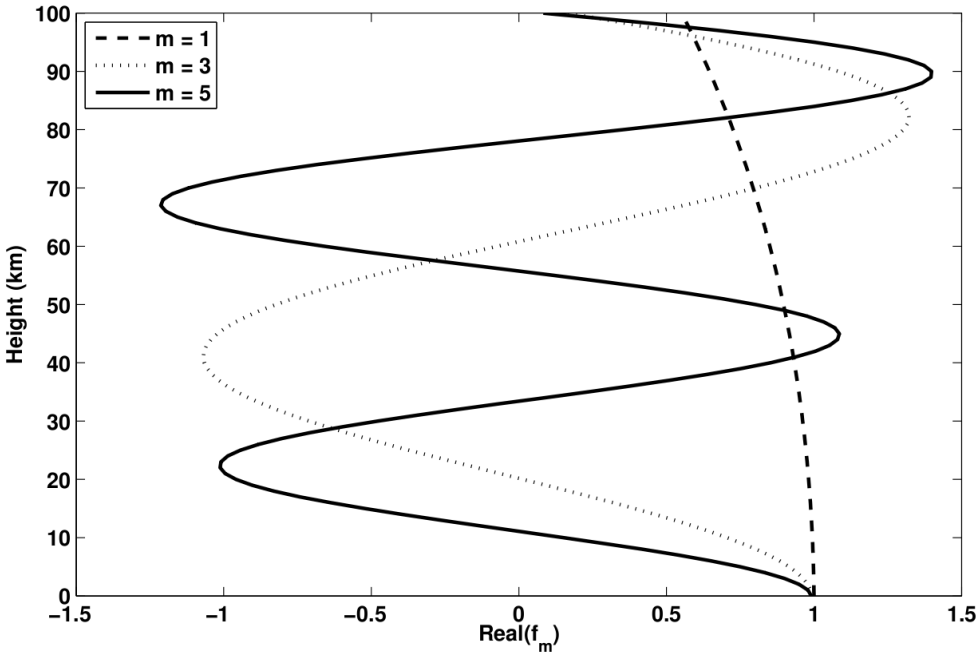


Figure 4.1: Real component of $f_m(z)$ for $m = 1, 3, 5$ modes.

each have different modal angles, and changing the reflection height changes the modal angles, and hence θ_m as well. This is done for the VLF range of frequencies, since the modal angles are wavelength dependent as well. Figure 4.3(a) shows the attenuation rate for the first four modes at different frequencies for an ionospheric height of 70km. This shows that, especially at lower frequencies, higher modes suffer far greater attenuation than lower modes. It is clear from this graph why the lower modes are the main contributors and why it is said that at large distances, only the first and second modes will make any significant contribution. Not only can one say that the higher modes experience high attenuation, but that the attenuation rate for the primary mode is very low, explaining why VLF signals can propagate such enormous distances around the globe. Figure 4.3(b) shows the variation in attenuation rate due to changes in the height of the waveguide for $m = 1$. Here we see that the lower the ionosphere, the higher the attenuation rate. This result agrees with the analysis of the refractive index, that the waves suffer less attenuation when propagating in a more night time ionosphere.

4.1.3 Dependence on Height H' and Sharpness Parameter β

The observed field strength at any point in the waveguide is the sum of the fields of different modes, equation. Due to the spatial dependence of the modes and modal angles, there will be regions where the contributions from the modes interfere more constructively with each other than in other regions. This leads to what is known as a modal interference pattern, which can have multiple local minima and maxima. The parameters that have the most influence over the

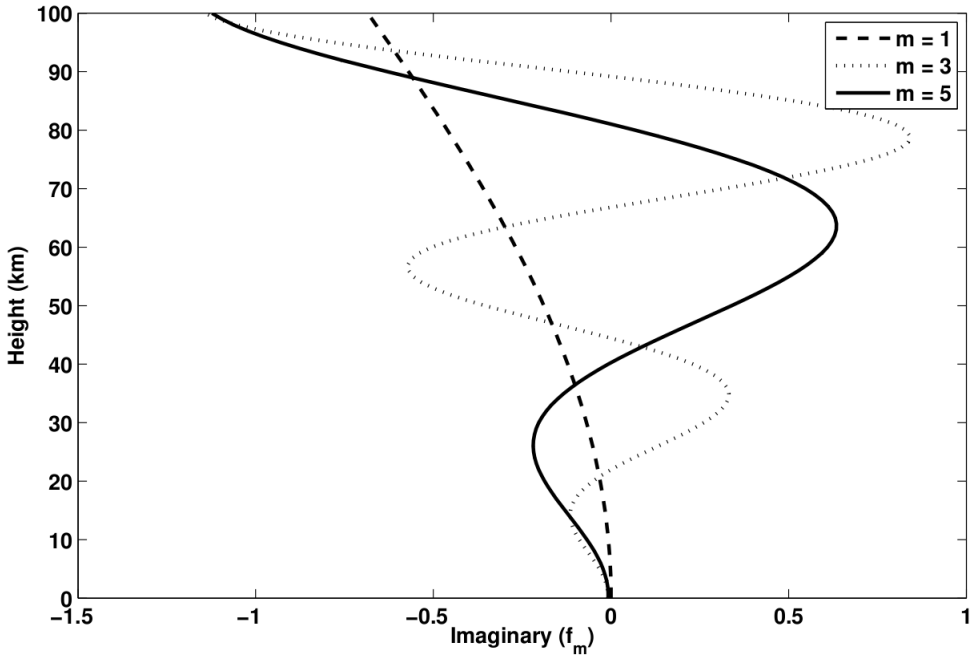


Figure 4.2: Imaginary component of $f_m(z)$ for $m = 1, 3, 5$ modes.

final modal interference pattern are H' , β and frequency. Changes to each of these parameters and the effect that they have on the interference pattern will be discussed here.

The nature of the D-region ionosphere is defined through H' and β. It should therefore be no surprise that the model is sensitive to these values. These values are dependent on many things, mostly time of day and year, latitude and solar activity. Summer and daytime, higher solar activity and being located at a high latitude lead to lower H'. This is expected as all of these are associated with a higher flux of solar radiation. This increased forcing on the ionosphere leads to equivalent levels of ionisation at lower altitudes than during winter or solar minimum. Figure 4.4 shows the changes in the modal interference pattern when the upper boundary of the waveguide is lowered. It can be seen that when the reflective layer in the ionosphere is lower, the modal interference pattern is slightly compressed towards the transmitter with the modal minima and maxima appearing closer to the transmitter. Figure 4.5 shows the effect of changing β while keeping the height fixed. The field strength in both Figures 4.4 and 4.5 is plotted as $log10(|E_z|^2)$ where E_z is calculated from (2.25). This was done since we are more interested in relative field strengths due to different sets of parameters than absolute values. Figures 4.4 and 4.5 aim to serve as examples that illustrate the effect of different parameters, for this reason, the value of other parameters such as frequency can be somewhat arbitrary. In this case, 15 kHz was chosen since it falls in the middle of the VLF range.

We see that increasing β can also lead to the modal interference pattern being compressed towards the transmitter, similar to Figure 4.4, but with more of a change to the overall form and

level of the interference pattern. The difference between $\beta = 0.35km^{-1}$ and $\beta = 0.45km^{-1}$ shows that a lower value of β results in the overall field strength being slightly less due to a higher electron density and therefore high absorption below the reflection height. A notable feature on both Figures 4.4 and 4.5 is the large modal minima around 500 km.. In Figure 4.4, decreasing the reflection height from 80 km to 75 km to 70 km moves the minima closer to the transmitter as well as gradually increasing the field strength at the minima. In Figure 4.5 on the other hand, increasing β from 0.35 km^{-1} to 0.40 km^{-1} to 0.45 km^{-1} also moves the minima closer to the transmitter, but first decreases the field strength with β going from 0.35 km^{-1} to 0.40 km^{-1}, then increases it again from 0.40 km^{-1} to 0.45 km^{-1}. The difference in this behaviour illustrates the more complex influence β has on the propagation of VLF radio waves in the EIWG. Whereas changing the reflection height simply lowers the ionosphere and therefore changes the modal angles, changing β has a greater influence on the reflective properties of the ionosphere. These differences can also be seen in other modal minima and maxima, although it is more clear in the minima around 500 km.

4.2 Temporal Variation

Until now, all ionospheric parameters had been kept constant. By adding a time dependence to H' and β, the response to temporal changes at a fixed location can be investigated. This allows for the approximate simulation of a fixed VLF receiver monitoring the strength of a VLF signal from some transmitter where a transient event modifies propagation conditions along the path. The approximation in the simulation is due to the fact that in reality, the propagation path has varying waveguide conditions due to changing latitude and longitude and possibly different surfaces. This model cannot implement changes in the waveguide like this, therefore its results are more useful in gaining conceptual insight and theoretical understanding, rather than attempting to replicate experimental results. The height and sharpness are modelled by having a sudden rise or drop in value with a steady return to quiet time value as shown in Figure 4.6. This is to replicate the rapid onset and gradual recovery of additional X-ray flux that is typically experienced during a solar flare. Figure 4.7 (a) shows the narrowband signal strength of the 18.3 kHz HWV transmitter in France, detected at SANAE-IV, the South African base in Antarctica. On 11 February 2010 there were no events that caused perturbations in the signal and what is observed is a typical quiet day curve. The following day however, there were a few solar flares of varying magnitude as seen in Figure 4.7 (b) which shows the 0.10.8 nm X-ray flux on the GOES-14 satellite. Three flares in particular stand out, one a bit before 08:00 UTC, another around 12:00 UTC and the final at around 18:00 UTC. The smaller two were registered as C-class flares while the larger, middle one being M-class. Firstly we notice that the time of these three flares coincides perfectly with the perturbations of the narrowband signal. Looking at the exact form of these responses, we see that for the two smaller flares, the VLF perturbation has a fast rise and slow exponential

decay back to the quiet day curve. The large, midday flare on the other hand, also has a fast rise, but then decays to below the quiet day level and then has an increasing return to quiet day value.

These observed changes can be pictured in terms of changes in the modal interference pattern, and how this changes in time, keeping the temporal form of the change in mind. When the X-rays from a solar flare hit the Earth, the enhanced ionisation reduces the height of the waveguide and modifies the modal interference pattern. The response in the signal strength at a receiver depends on the position of the receiver relative to the modal interference pattern. Since solar flares lower the ionosphere, we expect the modal interference pattern of the disturbed ionosphere to be similar to the quiet pattern, except compressed towards the transmitter as seen in Figure 4.4. This will cause the transmitter to ride up or down the slope of the modal interference pattern, depending on the original position. If the transmitter is after a modal maximum and before a modal minimum, it is on a downward slope, and the response to the flare will be negative. Once it reaches the value that corresponds to the maximum perturbation, it will start to move back towards its original value, but at a slower rate, just like the X-ray flux. On the other hand, if the transmitter is after a minimum and before a maximum, it is on an upward slope. The perturbation caused by the flare and the resulting squeezing of the interference pattern will then lead to an increase in the signal, before slowly returning down to the original level. An example of this is given in Figure 4.8. For a larger flare, such as the midday one, the response in the VLF signal can be a bit more complicated as is seen in Figure 4.7. In this case, the perturbation is large enough so that it moves over a modal minimum or maximum, but due to the very fast rise time, the transmitter does not resolve the modal extremum. Once the perturbation has reached its peak, it starts to slowly decay. Due to the slow return to quiet day levels, the transmitter now spends more time in the modal extremum, which it skipped during the onset of the perturbation. Figure 4.9 gives an example of this happening where the temporal response is of a similar form to the response to the large flare in the data. Note that the smaller flare is modelled by changing the ionospheric height by 2 km and β by 0.01 km^{-1} and the larger flare by 4 km and 0.03 km^{-1} respectively.

4.3 DEMETER Comparison

The model's use has been demonstrated in terms of gaining theoretical understanding of the propagation of VLF waves in the EIWG. However, the results still need to be compared to some experimental data to check their validity. To accurately do this is not as simple as one might imagine, keeping model constraints in mind. The best method would be to continuously measure the field strength from a VLF transmitter, while moving away from the transmitter in a straight line. Moving in a straight line and measuring the field strength, while avoiding man made noise is practically impossible. Measuring the field in an aircraft could theoretically serve this purpose, but is obviously very impractical. Another option is to use satellite data. This seems counter-intuitive since satellites orbit at hundreds of kilometers above the surface of the Earth, well above

the EIWG.

The idea for this is that even though the majority of the VLF energy remains trapped in the waveguide, some of it leaks out of the waveguide at the reflection layer and up into space. This leaves a fingerprint of the modal interference pattern above the transmitter that satellites with a radio receiver on board are able to observe. Figure 4.10 shows the observed electric field strength for day time (10:30 LT) and night time (22:30 LT) passes over the NWC transmitter. The colouring scale is logarithmic with arbitrary reference not provided with the data. NWC is located at 21.816° S, 114.166° E and operates at 19.8 kHz at 200 baud with a power level of 1 MW, making it one of the most powerful VLF transmitters. To examine the data more closely, horizontal rows (i.e. constant latitude) are extracted from the data. This is done for both day and night, and for eastward and westward propagation, as indicated by the red blocks in Figure 4.10. The main reason for choosing horizontal rows is for convenience, but another advantage is the NWC transmitter is on the North-West coast of Australia as seen in Figure 4.11. This means that westward propagation is directly over sea, while eastward propagation is over land, at least until the other side of Australia. This provides a useful comparison for the model which cannot account for discontinuities in the boundaries of the waveguide.

Figures 4.12(a) and 4.12(b) show the two horizontal sections for day and night time, with the same logarithmic units as in Figure 4.10. It is evident that westward propagation over the ocean is slightly stronger than eastward propagation over land, for both day and night. Looking at the day time plot in Figure 4.12(a), modal minima and maxima from the interference of different modes are clearly visible. Firstly, one must remember that this data is averaged over an entire year so that the observed pattern is not that from a unique ionospheric configuration, but rather the average of multiple ionospheres. Looking at the theoretical examples in Figures 4.4 and 4.5, there is definite similarity. Both Figures 4.4, 4.5 and 4.12(a) first show a sharp decrease in field strength close to the transmitter. In the theoretical plots, this is in the range of around 0 to 500 km, and 0 to 1000 km for the DEMETER data, with this difference being due to different transmitter frequencies, as well as different ionospheric parameters. Both the theoretical plots and the physical data show a strong modal minima after the initial sharp drop, with the field strength dropping at a lot lower rate at distances further than the first large minimum. There is another minimum before this, seen in the theoretical plots at around 300 km, although this minimum is not nearly as pronounced as the one at 1000km. This initial, small minimum, is not totally clear on the DEMETER data, although is probably the cause for the slight plateau at around 700 km in Figure 4.12(a). A second clear minimum is seen in the DEMETER data at around 2600 km, and 1500 to 2000 km in the theoretical plots.

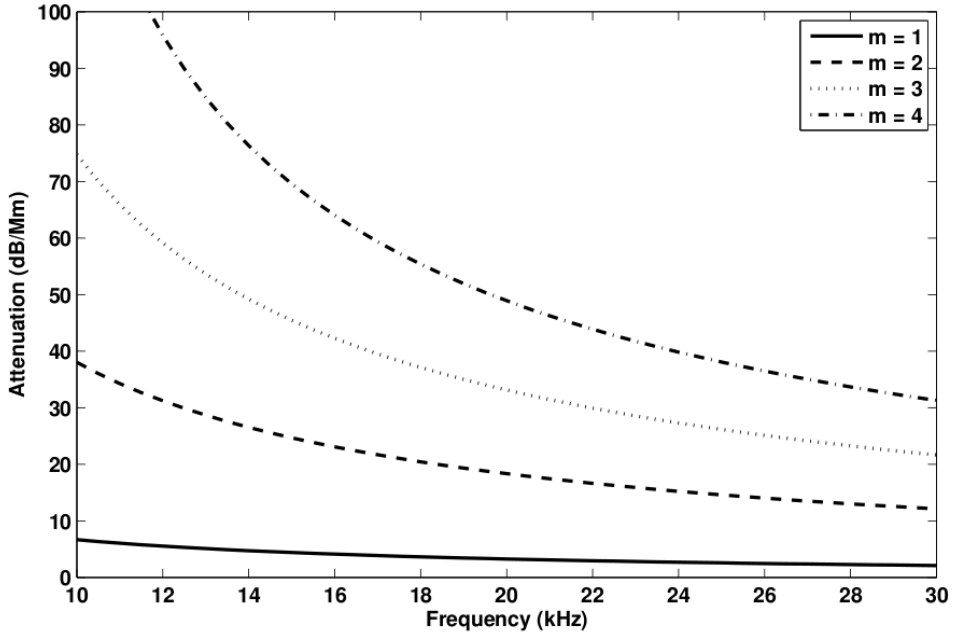
For night time observation in Figure 4.12(b), the modal minima and maxima are not so clearly defined. During night time, the reflective layer of the ionosphere is far higher than during the day. The higher the upper boundary of the waveguide, the more spread out the modal minima and maxima of the interference pattern. This idea is illustrated in Figure 4.4. For the first main

modal minima around 500 km, the distance between the minima of reflection heights of 70 and 80 km is around 100 km. At the second modal minima between 1500 and 2000 km distance, the difference between the minima for reflection heights 70 and 80 km is now almost 500 km, showing that the higher heights lead to increasing distances between modal minima. Another explanation can be given in terms of the gradient of the electron density, β. As discussed in Chapter 1, β has higher values when there is less radiation, as is the case at night. Figure 4.5 shows that for a low β such as 0.35 km^{-1}, the modal minima at 600, 1800 and 3100 km are all well defined. For a higher beta, such as 0.45 km^{-1}, the modal minima at 500 km is well defined, however the minima after that, at around 1500 and 2500 km are not as well defined as those for β =0.35 km^{-1}. When considering these observations that when the height of the ionosphere as well as the gradient of the electron density are both higher, such as the case during night time, then the modal features of the interference pattern are less clear. Another factor to consider is the attenuation, since the attenuation is lower at night there will also be contributions from higher order modes than during the day. In Figure 4.12(b), it can be seen that these factors combined together with the averaging effect will lead to the modal structure of the interference pattern being less well defined. For these reasons, only the day time data will be examined further.

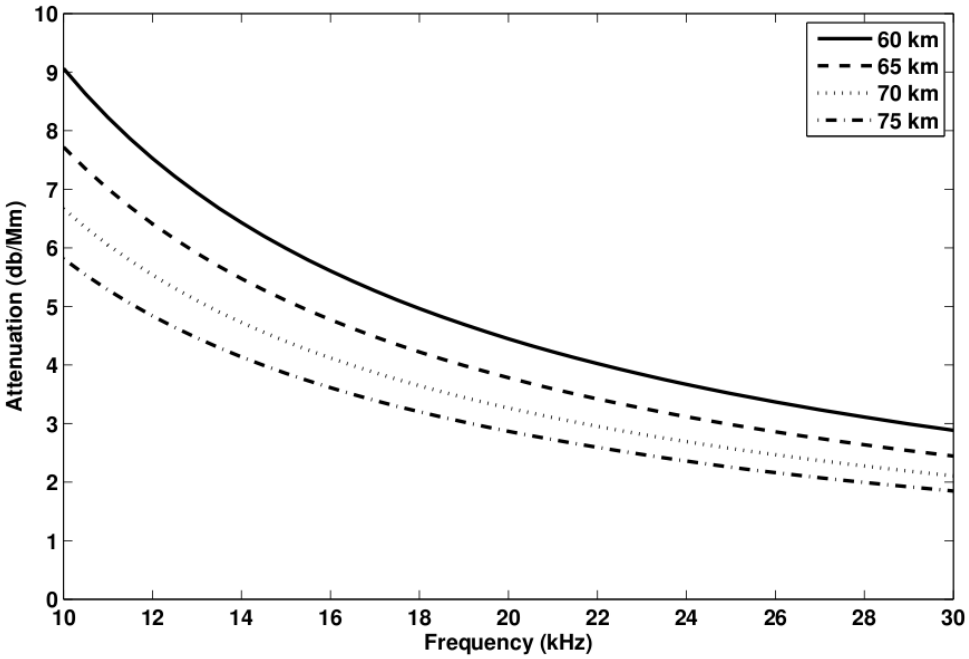
To compare results from the Mode Theory model described in Chapter 2.2 with this data, the optimisation toolbox in MATLAB was utilised. This was done by minimising the error between the model and the data with respect to changes in ionospheric height, keeping model constraints in mind. The two main constraints to consider are that at short distances, the model does not account for any evanescent modes that would add to the field close to the transmitter. At larger distances, the model starts to do less well since it assumes a flat Earth. When the propagation distance becomes comparable to the radius of the Earth, this approximation becomes invalid. For these two reasons, the minimisation was only performed over intermediate distance between 1000 km and 4000 km. When performing the minimisation, the only parameter in the model that changed between east- and westward propagation was the conductivity of the surface of the Earth. Commonly accepted conductivities were chosen for the sea (4 S/m) and land (4.5×10^3 S/m). Coincidentally, the magnetic declination at NWC is roughly 0 , meaning that for east- and westward propagation, there will be no difference between the bearing angle and the Earth's magnetic field due to symmetry. After much trial and error, it was found that when looking at propagation in just one direction, the model could be quite well matched to the observed field with the averaging of only 4 or 5 ionospheric configurations with different values of H'. Minimisation with respect to β was also investigated, but the results were inconsistent, and it was found that just modifying H' while keeping β constant delivered acceptable results. To successfully compare the performance of the model over land and sea to the data with one set of values, it was found that more heights were needed, 8 or 9 to deliver satisfactory results. Figure 4.13 shows the results of the model comparison with the data after the minimisation was performed. Thomson [2010] investigated the height of the ionosphere around NWC and reported heights of 70 and 72km

whereas this model found an appropriate match to the data with heights ranging from 65 to 69km. This is due to very different methods being used, a different interpretation of H' and the fact that the results are taken from different times in the solar cycle meaning that the D-region would be at a different altitude.

Even though there is some disagreement between these results and those of Thomson [2010], the results here do agree fairly well with the data. Looking in the white region of interest in Figure 4.13, the model is successful in calculating the location of the modal minima, as well as the difference in field strengths due to Earth's surface for the region of interest. In the near- and far-field there is less agreement, but this is to be expected due to the model constraints and is the reason for the minimisation of the error being in the white region of interest in Figure 4.13.



(a) Attenuation rate for different modes



(b) Attenuation rate at different heights

Figure 4.3: Effect of varying mode number and reflection height for a range of VLF frequencies on the attenuation rate, $-k\mathrm{Im}(S_m)$.

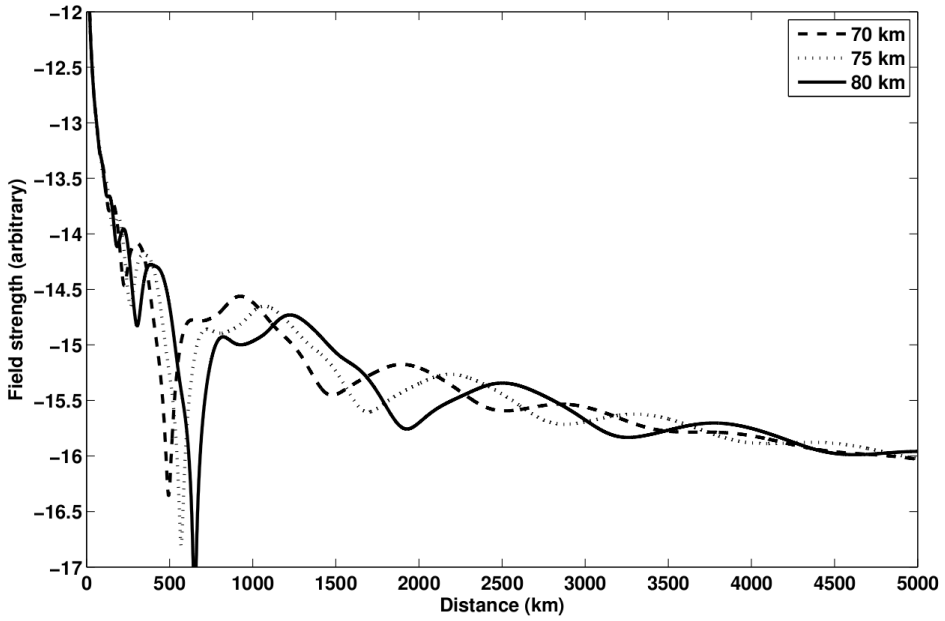


Figure 4.4: Changes in the modal interference pattern at $z = 0$ due to changes in the reflection height for a 15 kHz wave over sea water against distance from transmitter, $\beta = 0.35km^{-1}$. The reflection heights here are those where $X = Z$ from (2.21).

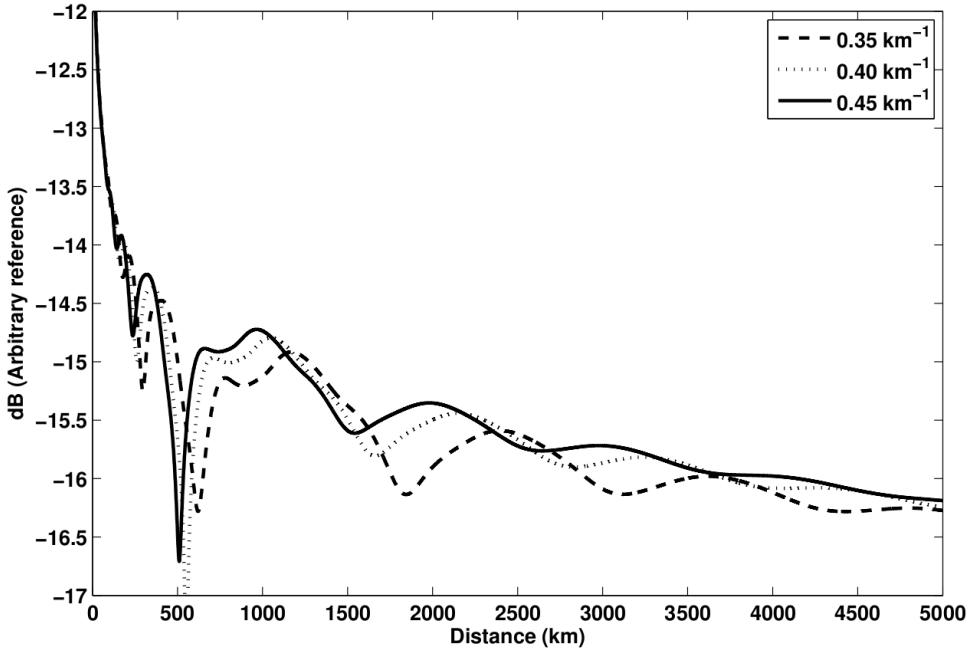


Figure 4.5: Changes in the modal interference pattern at $z = 0$ due to changes in β for a 15 kHz wave over sea water against distance from transmitter. $H' = 49$ km with respective reflection heights of 78, 74 and 71km.

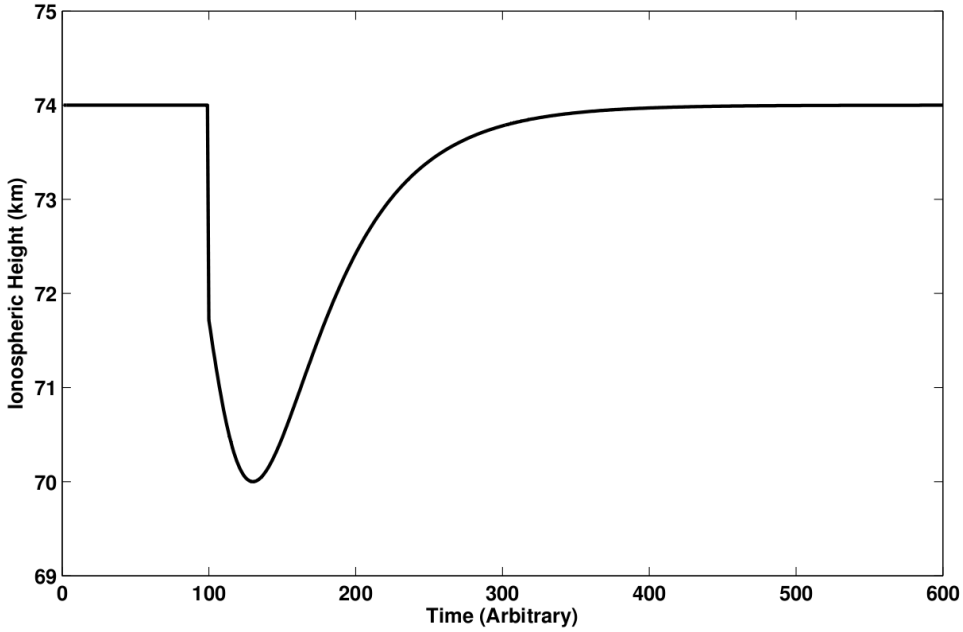


Figure 4.6: Temporal change in height used in simulation of transient event, β changes in a similar way.

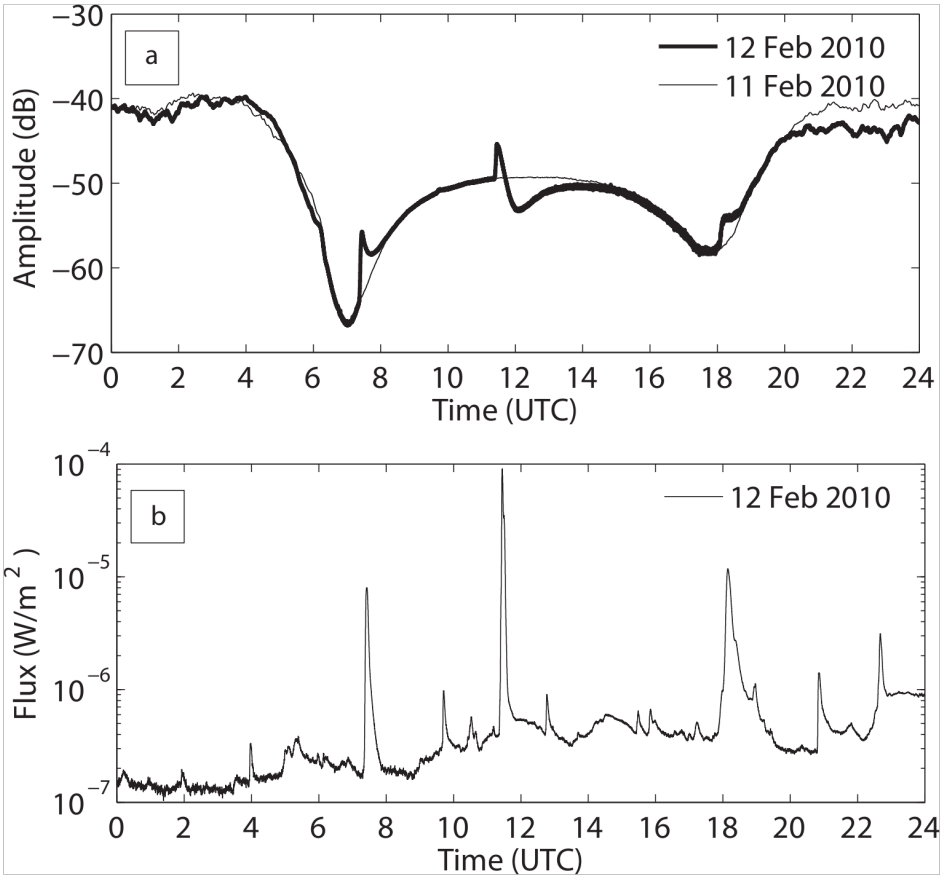


Figure 4.7: 18.3 kHz signal from HWV transmitter in France detected at SANAE IV, Antarctica for a quiet day and a day with solar activity. (b): X-ray flux measured on GOES 14 satellite for the day with solar activity.

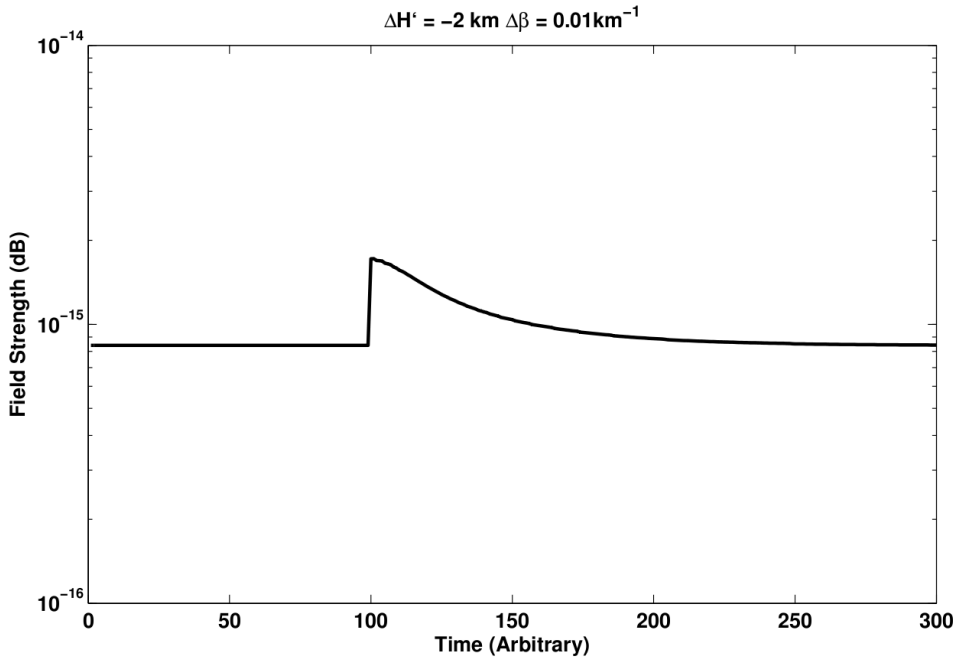


Figure 4.8: Simulated positive signal response to a small solar flare.

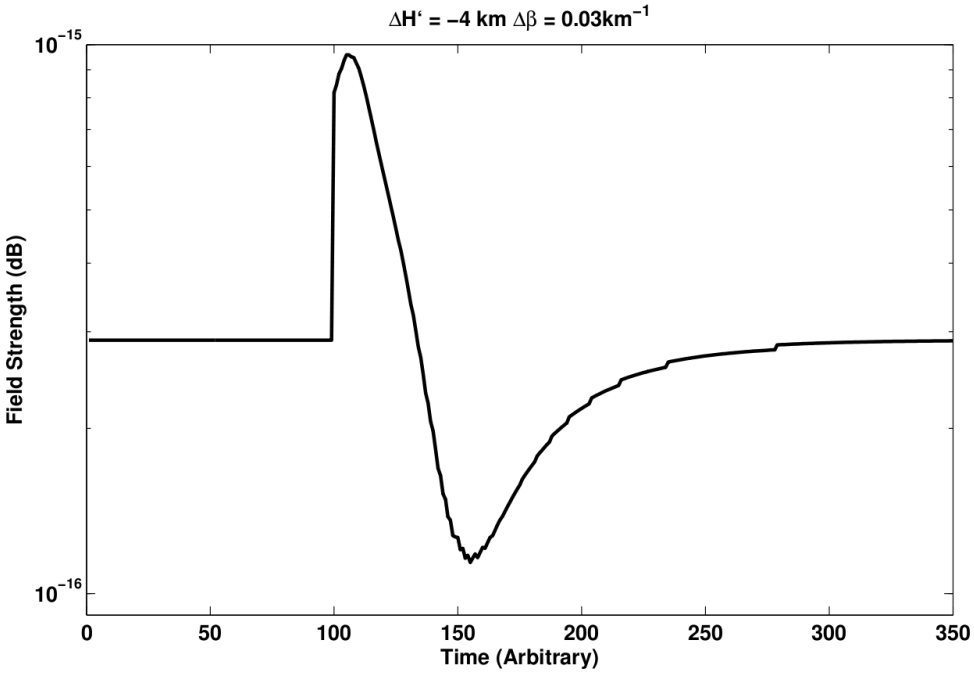


Figure 4.9: Simulated positive followed by negative response to a larger flare.

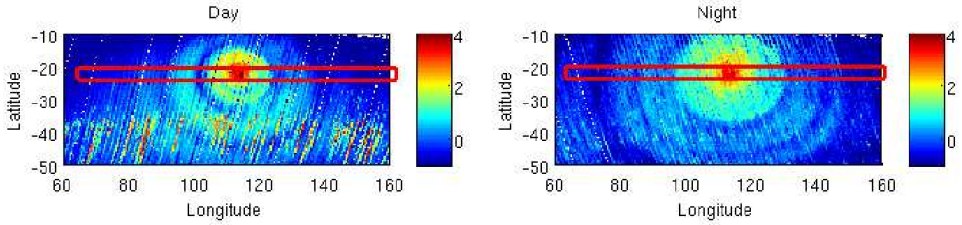


Figure 4.10: Power over the NWC transmitter averaged over a year measured on DEMETER. Red blocks show the areas extracted for further analysis.

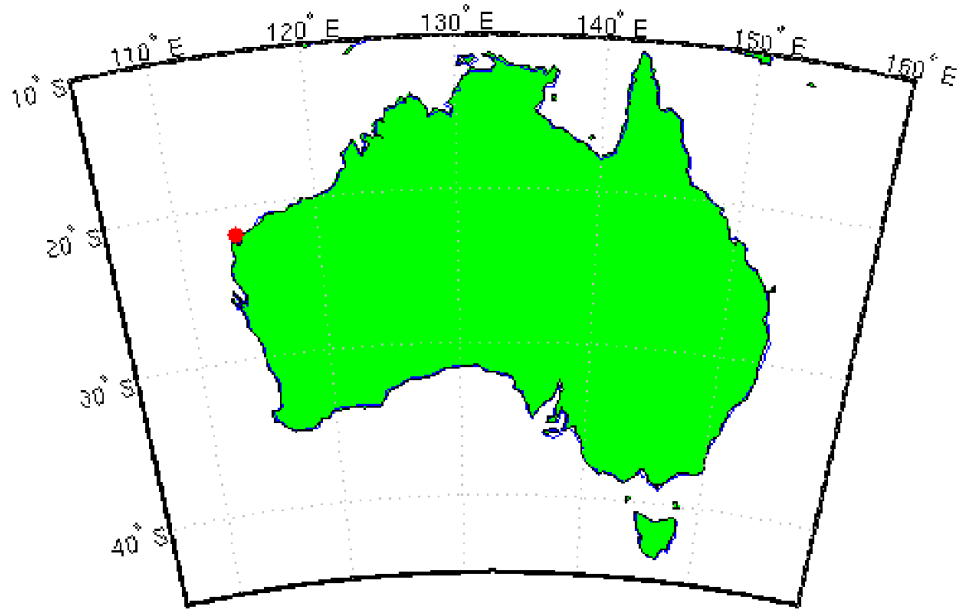
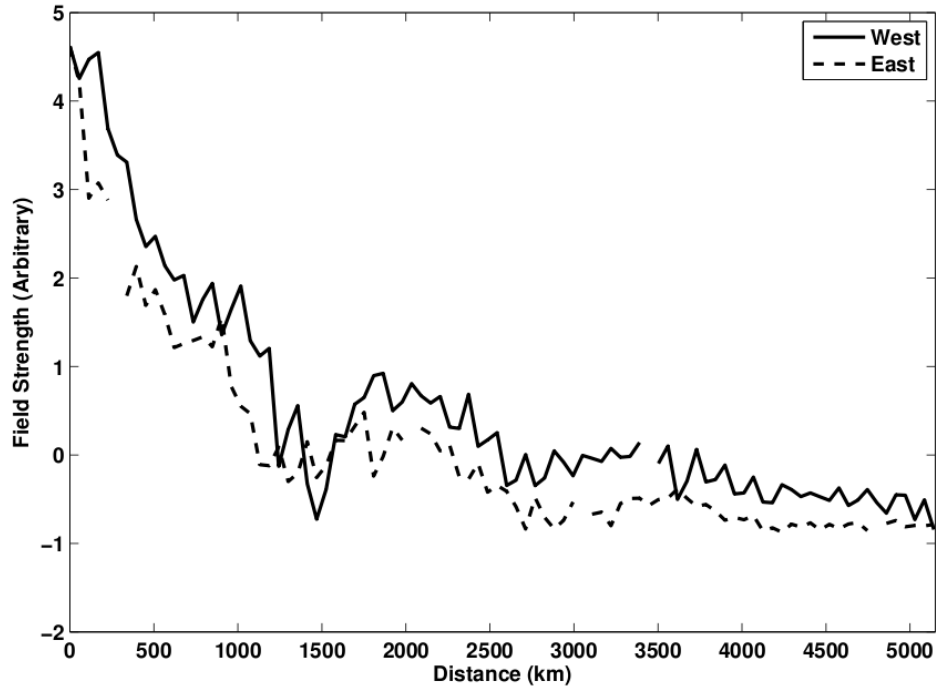
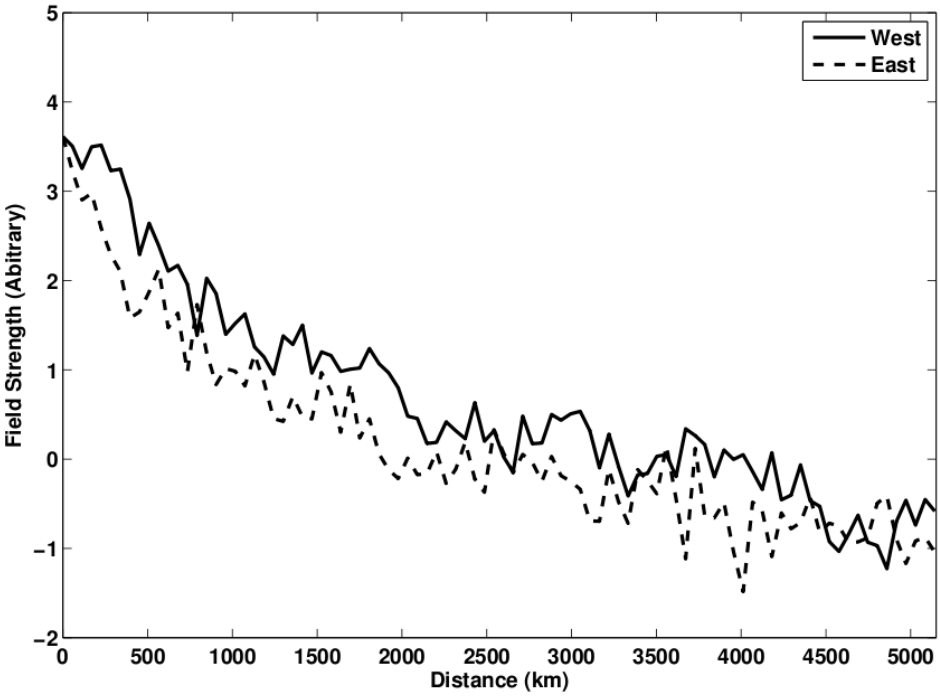


Figure 4.11: Position of NWC transmitter in Australia.



(a) Day time modal interference pattern.



(b) Night time modal interference pattern.

Figure 4.12: Field strength plots extracted from red horizontal sections from Figure 4.10 of DEMETER data for (a) day time and (b) night time observations over NWC for east- and westward propagation.

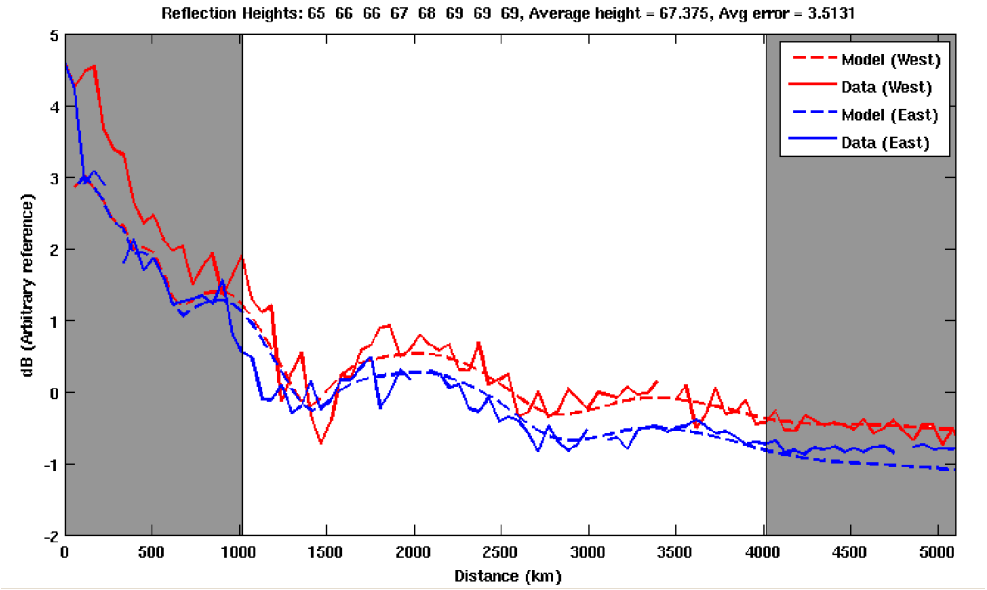


Figure 4.13: Vertical electric field intensity $\log 10|E_z|^2$, comparing interference patterns from model results with the data after minimisation of the error between model and data in the white region. Model results adjusted with a constant value to be directly comparable to DEMETER data. Both model results and DEMETER data share location of modal minima as well as relative difference between East- and Westward propagation.

Chapter 5

The FDTD Method

For basic free space propagation, the user enters the frequency of the signal, the maximum number of spatial and temporal steps and the value for another variable, Points per Wavelength (PPW), which determines the grid sampling density. The frequency, f and PPW value, N_λ are used to determine the length of a spatial step,

$$\Delta x = \frac{c}{N_\lambda f} \tag{5.1}$$

for a wave traveling at the speed of light, c.

The appropriate time step is then calculated using the Courant number. The Courant number is used to impose a necessary condition on numerical time-marching schemes, such as FDTD, to maintain convergence. It enforces that the time between two consecutive discrete time steps can be no more than the time taken for the wave to travel between two adjacent grid points. In d-dimensions, the largest distance between spatial grid points with indices that differ by no more than one is $\sqrt{d}\Delta x$, assuming isotropic grid spacing. During one time step, a numerical electromagnetic wave will fully propagate to its nearest neighbours (eg: in 2D, a numerical wave at point (x_i, y_i) can propagate to $(x_{i\pm1}, y_i)$ and $(x_i, y_{i\pm1})$ in one time step, but not $(x_{i\pm1}, y_{i\pm1})$), it takes $d\Delta t$ to propagate the maximum $\sqrt{d}\Delta x$ mentioned previously. Formally,

$$\sqrt{\sum_{i=1}^{d}(\Delta x_i)^2} \geq dc\Delta t. \tag{5.2}$$

where Δx_i is the grid spacing for dimension i.

Assuming isotropic grid spacing, $\Delta x_i = \Delta x$, the Courant number C_s is defined from (5.2) as

$$C_s = \frac{c\Delta t}{\Delta x} \leq \frac{1}{\sqrt{d}}, \tag{5.3}$$

with maximum value

$$C_s = \frac{1}{\sqrt{d}}. \tag{5.4}$$

The time step Δt is then

$$\Delta t = \frac{C_s \Delta x}{c}. \tag{5.5}$$

By implementing the relation between spatial and temporal step through the Courant number, scaling to other frequencies and determination of the required spatial and temporal step size is done automatically. The larger N_λ is, the more accurate the simulation since the field values are closer together, temporally and spatially, leading to smaller numerical errors. Increasing N_λ however also results in an increase in computation time. When N_λ is below 20, notable errors occur quite quickly. Above 30, the increase in computation time becomes too large to validate the accuracy gain, meaning this is a suitable value. Table 5.1 gives some indication as to how increasing N_λ can greatly increase computation time due to an increase in the number of spatial and temporal steps required.

Table 5.1: Effect of frequency and grid sampling density on discretisation.

(a) Δx values in m for different frequencies and N_λ

Frequency (Hz) \ N_λ	10	20	30	40
10000	2980	1490	993.3	745
15000	1986.6	993.3	662.2	496.6
20000	1490	745	496.6	327.5
25000	1192	596	397.3	298

(b) Δt values in ms for different frequencies and N_λ

Frequency (Hz) \ N_λ	10	20	30	40
10000	7.07	3.54	2.36	1.77
15000	4.71	2.36	1.57	1.18
20000	3.54	1.77	1.18	0.88
25000	2.83	1.41	0.94	0.71

(c) Number of cells required to simulate 500 km

Frequency (Hz) \ N_λ	10	20	30	40
10000	168	336	503	671
15000	252	503	755	1007
20000	336	671	1007	1342
25000	419	839	1258	1678

Basic examples of wave propagation in free space will first be shown. Then in subsequent steps, further levels of complexity will be added to the model until it is able to simulate VLF propagation in the EIWG. After the free space example, simple reflecting boundaries will be added, these will illustrate the formation of modes within a waveguide and the effect of transmitter position, orientation and waveguide height. The Nearly Perfectly Matched Layer will then be introduced, serving as a more realistic boundary, able to model different surfaces, such as sea water or land. The effect of a background plasma and electron-neutral collisions similar to those in the EIWG will then be added, but without a reflecting surface. Finally, the different aspects will be combined to form a model of the Earth-Ionosphere waveguide, including surfaces of different conductivities and different background ionospheres. This model will then be used to produce results that can also be compared to the DEMETER results, like those from the Mode Theory model in the previous chapter. After the DEMETER comparison, a small section is added that discusses the phenomenon of numerical dispersion, and the effect of changing the Courant number.

5.1 Free Space

Two examples are given for a 19.8 kHz signal propagating in free space for both a vertical and horizontal source and shown in Figure 5.1, with the source located in the centre of the computational region. The choice of frequency is arbitrary, since these are illustrative examples, but 19.8 kHz was used since this is the frequency corresponding to a physical situation, that of the NWC transmitter. The highest values (light blue - yellow) of the respective electric field components are found broadside of the source while the field is a minimum (dark blue) directly in line with the source (a, e). For a vertical E_y source, there is minimal horizontal electric field both broadside and in line with the source (b) and *vice versa* (d). The strongest electric fields orthogonal to the source (E_x/E_y field from a E_y/E_x source) are found along the diagonal directions from the source location (b, d). The magnetic field has a similar but wider distribution to the electric field parallel to the source (c, f). These results agree with the theory of dipole radiation that the strength of an electric field from a dipole source is proportional to $\sin\theta$ where θ is the angle between the orientation of the dipole and direction in question [Griffiths, 1996], so that a source orientated along the $y-$axis will have a maximum field along the x direction, where $\theta = \pi/2$. This indicate the successful basic operation of the model.

5.2 Basic Waveguide

The most basic waveguide that can be implemented is one where the walls are perfect electric conductors. Figure 5.2 shows the field patterns that emerge for two different configurations of the source. Note how the vertical source (a) is better at injecting energy into stable waveguide modes that allow for improved propagation. The horizontal source (e) is sending most of its energy at

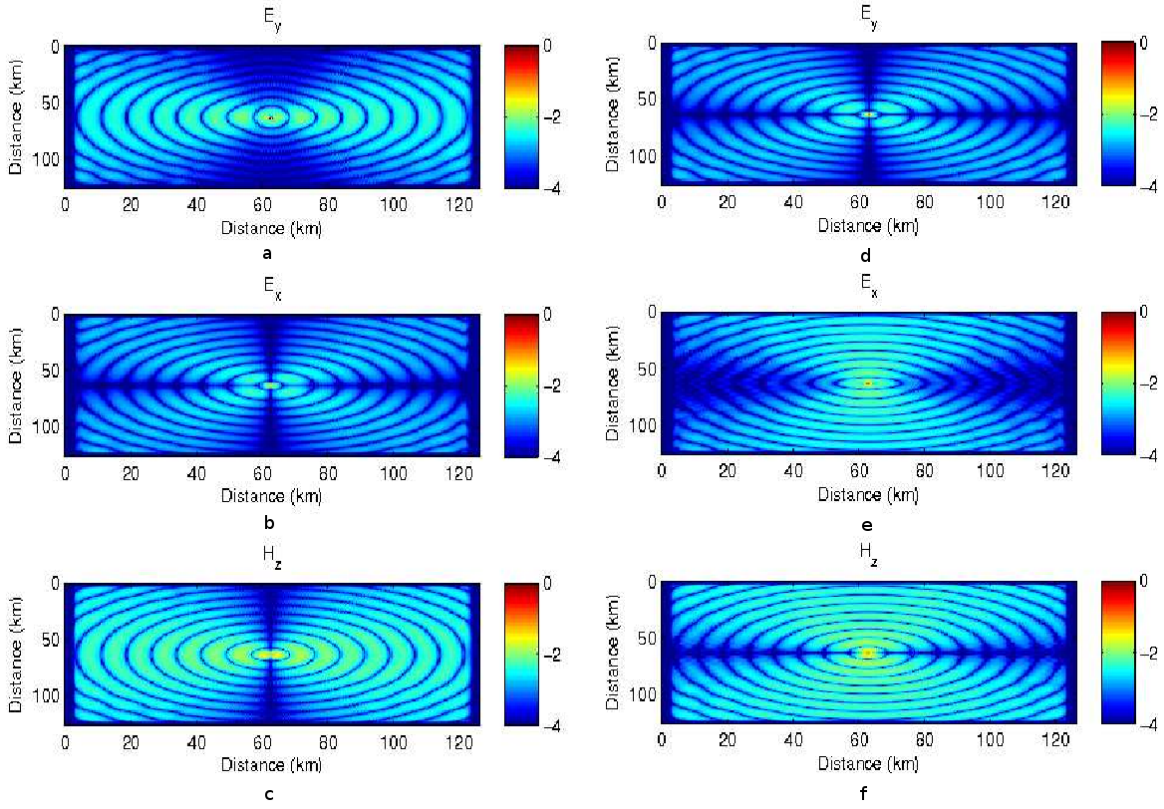


Figure 5.1: Electric and scaled magnetic fields for a vertical E_y (a - c) and horizontal E_x (d - f) electric dipole source located in the centre. The colour scale is logarithmic, normalised to the maximum value at the source of $\log 10 E_z = 0$. The magnetic field plots are those of the scaled magnetic field, according to (2.62).

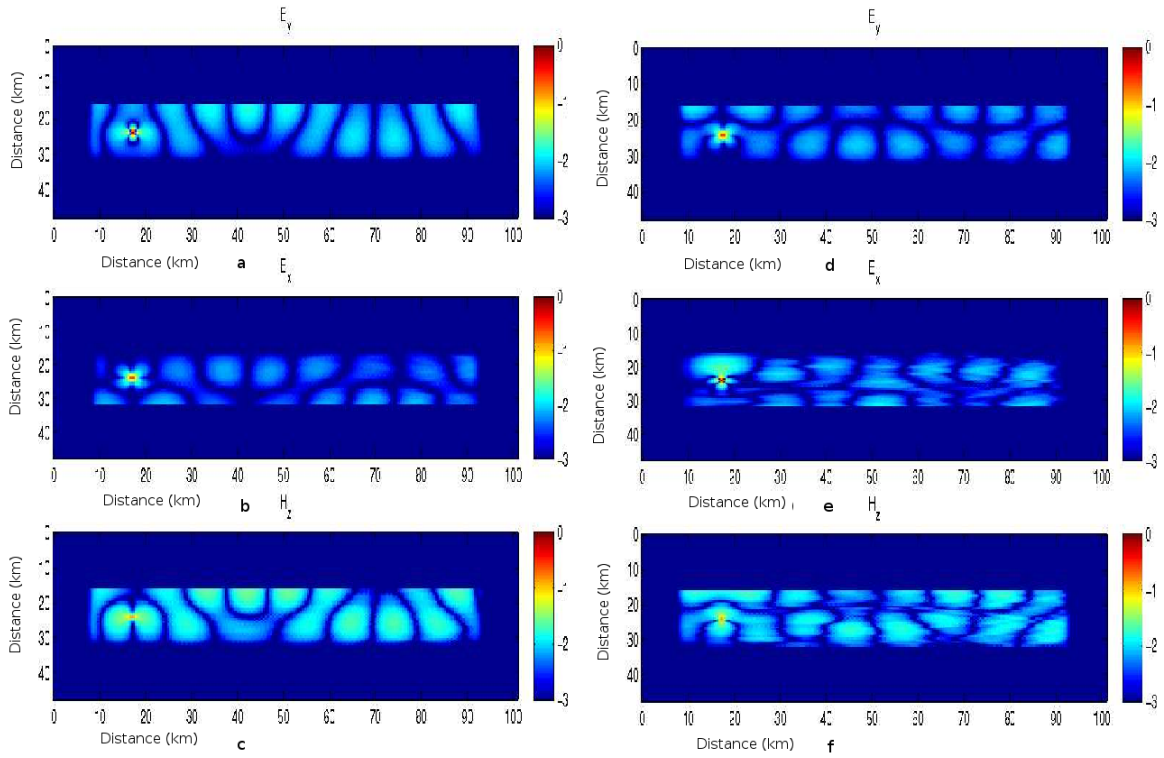


Figure 5.2: Waveguide field patterns for a vertical (a - c) and horizontal (d - f) electric dipole source located near the left end of the waveguide, in the middle between the two boundaries. Colouring consistent with Figure 5.1.

normal incidence to the boundary, leading to poor reflections down the waveguide. Figure 5.3 shows how increasing the size of the waveguide leads to the formation of higher order modes, with 3 to 4 peaks of maximum amplitude across the height of the waveguide, compared to 1 in 5.2. Figure 5.4 demonstrates how the position of the source can influence the modes that form in the waveguide. When the source is located in the middle of the waveguide, the energy directed broadside is the dominant factor in the observed field (d). Conversely, when the source is at a boundary the observed fields have more influence from reflections along diagonal paths (a).

5.3 Nearly Perfectly Matched Layer

The NPML is implemented around the boundaries in such a way as to absorb incoming waves with minimal reflection. The NPML operates in a region surrounding the computational domain of the main problem by introducing an artificial conductivity, $\sigma_x(x)$, as described in Chapter 2. This conductivity is a function of the position in this region between the problem domain and computational boundary. Taking this into discrete space, $x = j\Delta x$, $\sigma_x(j) \equiv \sigma_x(j\Delta x) = \sigma_x(x)$, this function can be written as

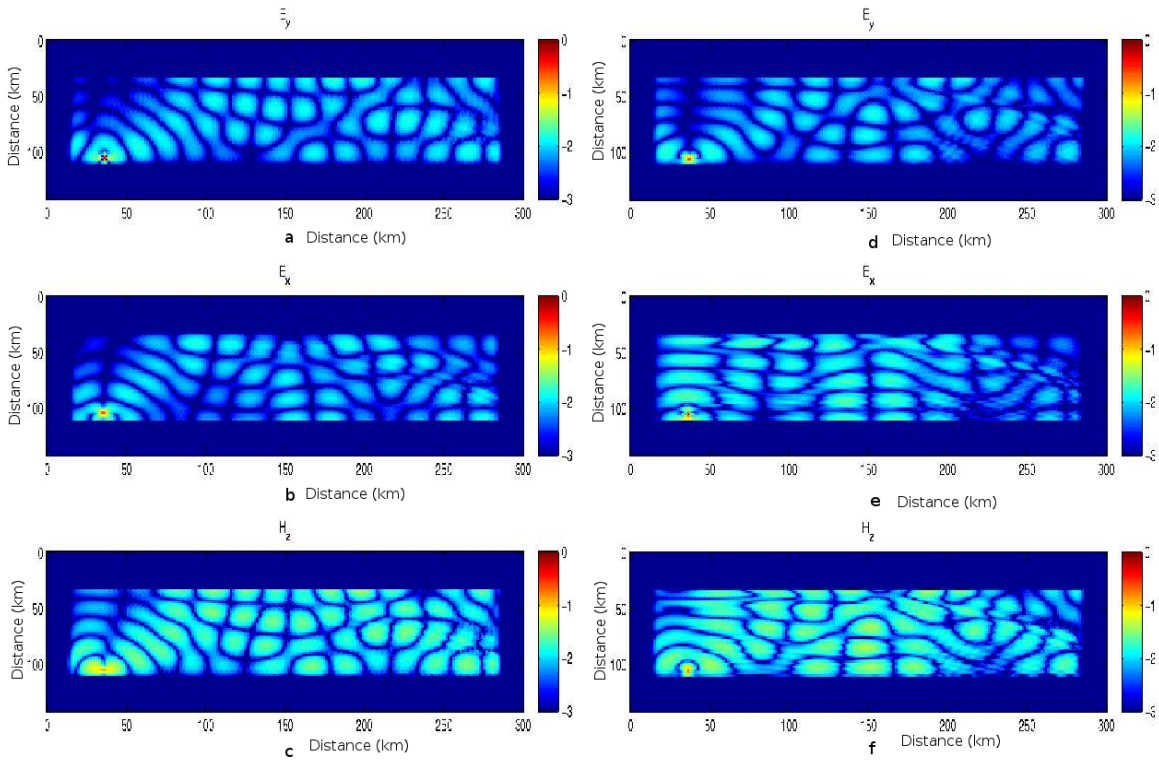


Figure 5.3: Waveguide field patterns for a vertical (a - c) and horizontal (d - f) electric dipole source in a larger waveguide with source located on the lower boundary, on the left side. Colouring consistent with Figure 5.1.

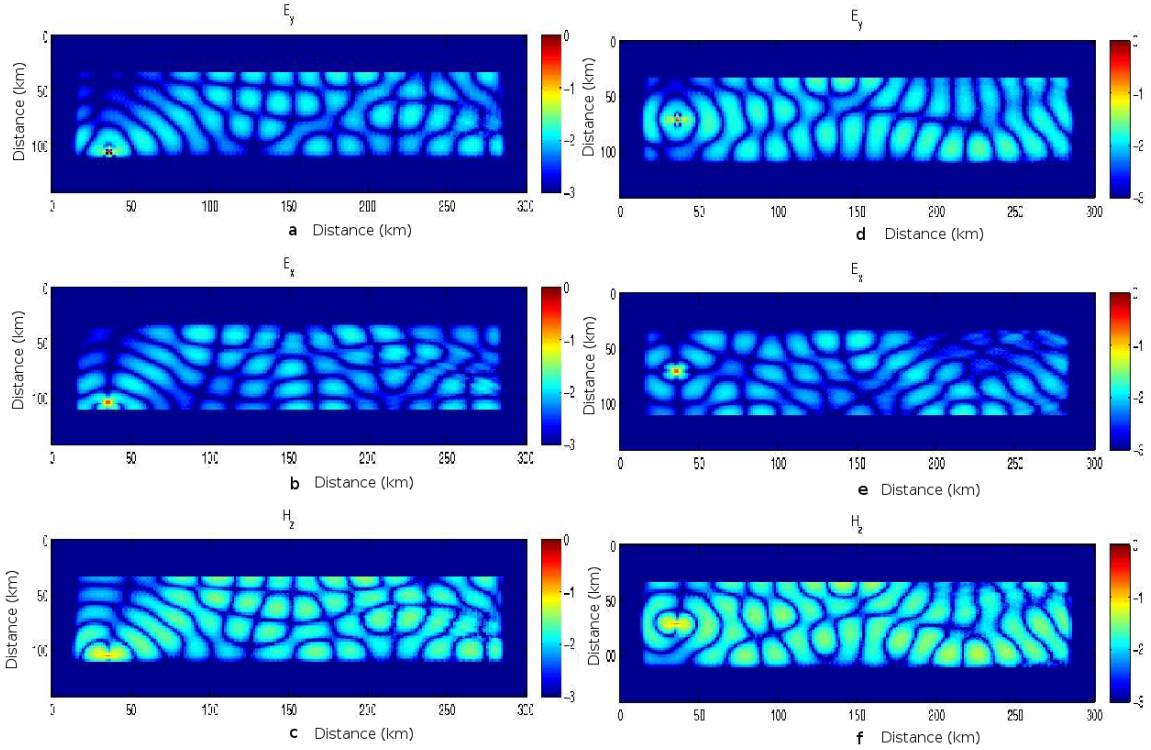


Figure 5.4: Waveguide field patterns for a vertical source but located at different heights in the waveguide, at the lower waveguide boundary (a - c) and in the middle of the waveguide (d - f). Colouring consistent with Figure 5.1.

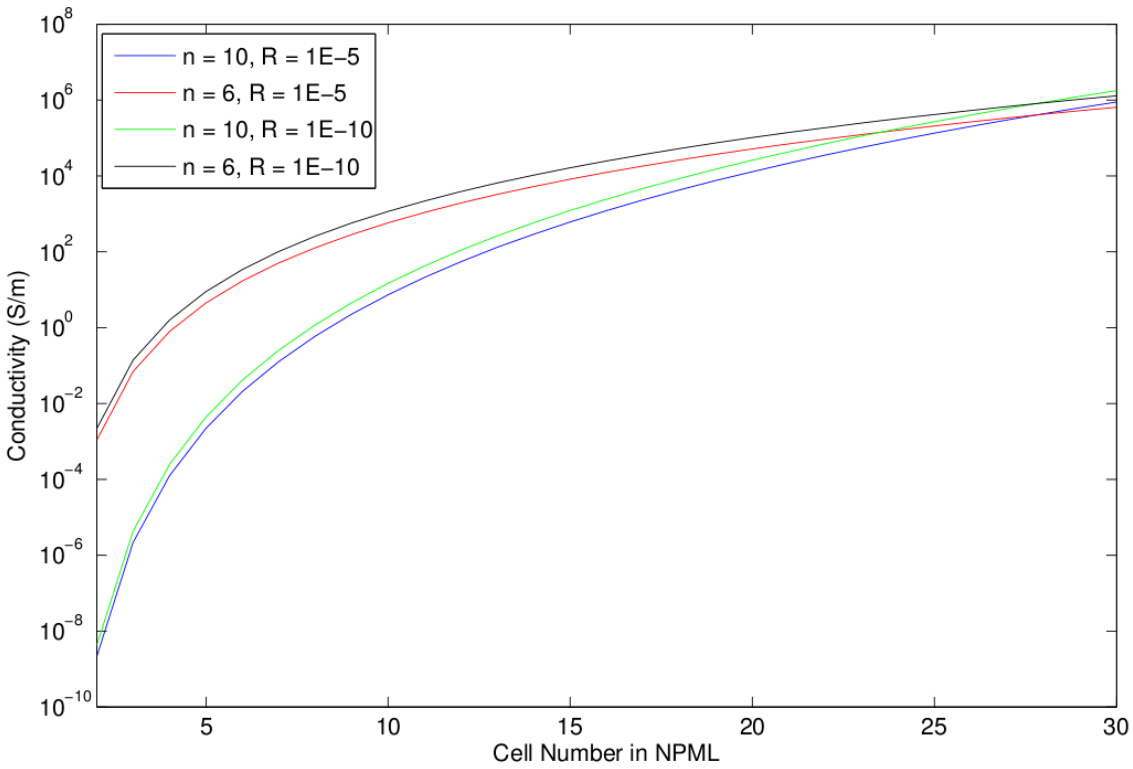


Figure 5.5: Conductivity of the NPML at different points in the layer for different parameters. Cell number refers to the position in the NPML, where 0 is adjacent to the computational region of the problem and 30 is adjacent to edge of the entire computational domain.

$$\sigma_x(j) = \sigma_M \left(\frac{j-1}{N^j_{pml}} \right)^n \tag{5.6}$$

where j is the spatial index in the x direction, N^j_{pml} is the thickness of the NPML in the x direction [Simpson, 2010]. σ_M is the maximum conductivity in the absorbing region and is given by

$$\sigma_M = \frac{c(n+1)\log(R)}{2N^j_{pml}\Delta x} \tag{5.7}$$

where R is the minimum desired reflection coefficient and n the fractional variation. Figure 5.5 shows how this conductivity varies with different n and R values. This conductivity is then applied to the NPML grid which stretches the fields as described in (2.60), or

$$C_{NPML}(j) = \left(1 + \frac{\sigma_x(j)}{i\omega} \right) \tag{5.8}$$

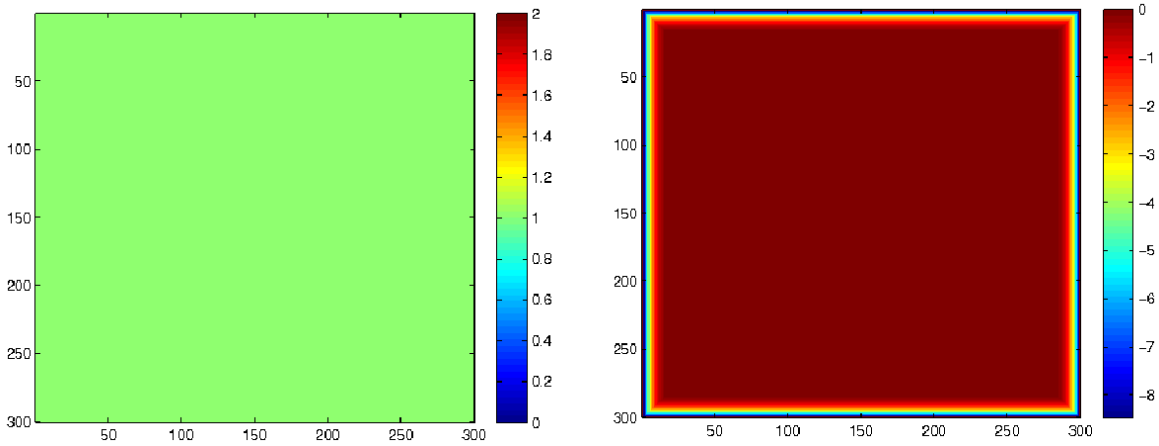


Figure 5.6: Real (left) and imaginary (right) components of C_{NPML}, the field scaling coefficient. The colour scale represents the value of the C_{NPML} component in question.

where $C_{NPML}(j)$ is the complex valued field scaling coefficient at point j in the NPML. These real and complex components are illustrated in Figure 5.6. It can be seen that the real component of C_{NPML} is always one while the imaginary component is zero in the main computational problem region and decreases steadily towards the edge of the computational space. It is this imaginary component that attenuates the incoming wave to such a small level that by the time it reaches the boundary of the computational space it is too small to create artificial reflections. In a realistic model of the EIWG, the NPML would only be required on the left and right boundaries due to the reflective waveguide boundaries at the top and the bottom. Due to the model being built up from a free space system, the NPML is implemented along all the edges of the computational domain despite this being unnecessary.

5.4 Surface Impedance Boundary Condition

Implementation of different surfaces and boundaries can be challenging when using the FDTD method under certain circumstances. As discussed in Chapter 2, simple implementation of non-zero conductivity is not straight forward, especially if the wavelength in the medium is a lot smaller than for free space. The SIBC implemented here is based on the assumption that $\sigma \gg \omega\epsilon$, where σ and ϵ are the conductivity and dielectric constant of the material in question, ω is the angular frequency of the wave. The validity of this assumption is examined in Table 5.2, where the "1" and "2" on land and ice represent the maximum and minimum values of the conductivity ranges for these surfaces. The approximation assumes that the value in the second last two columns is far greater than one. It is seen that this is not the case with ice, can be valid for land, and is always valid for sea water. This is not too much of an issue since we are mostly interested in propagation over sea and land. Attention must be paid when modelling land when using the

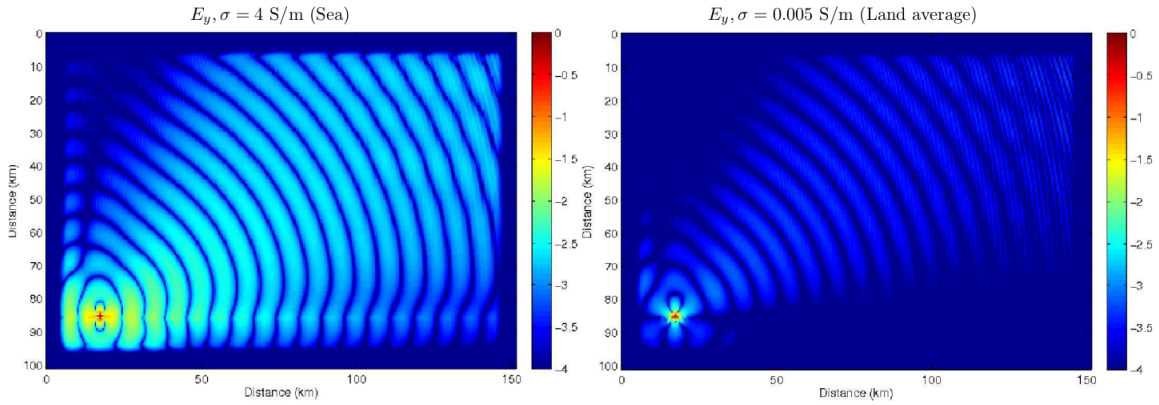


Figure 5.7: Emerging field patterns for sea (left) and land (right) SIBC for a vertical electric dipole just above the surface, in the lower left of the plot, at the small red area. Colouring consistent with Figure 5.1.

SIBC that the combination of conductivity and frequency meet the requirements. Conductivity and dielectric constant values we have used are those that are commonly accepted to fall within this range [Thomson, 2010; Han et al., 2011; Cummer, 2000].

Figure 5.7 shows the effect of different surface conductivities on the reflection of the radio waves originating from a E_y source near the lower left corner, near (15, 85). The outer 15 kms are occupied by the NPML, accounting for different field pattern observed below the source. The source was placed slightly above the lower boundary, as this is the case for real-life situations, which will be looked at shortly. Placing it at the left side of the region was purely for convenience of maximising the horizontal distance that the wave travels from the source. All FDTD plots of 2D electromagnetic fields from this point will have the source located here, as the aim is to build up to a realistic situation.

Due to the source being located directly above the surface, the surface has a very strong influence on the waves propagating upwards away from the source. Waves reflecting off the ocean will have a lot higher amplitude than off land due to the difference in conductivity, seen by the light blue colour dominating propagation over sea (Figure 5.7 left) compared to the dark blue for propagation over land (Figure 5.7 right). For waves propagating directly above the surface, the effect of the different conductivities is even greater. When propagating over ocean, the electric field is still quite strong, directly above the surface. While over land on the other hand, the field strength directly above the surface decreases far more rapidly as the wave propagates away from the source. These results indicate how the implementation of the Surface Impedance Boundary Condition (SIBC) successfully incorporates surfaces of different conductivity.

Table 5.2: Testing the SIBC approximation. Frequencies ($\omega_i/2\pi; i = 1, 2$) correspond to 10 and 30 kHz respectively.

Surface	**Conductivity σ (S/m)**	**Dielectric Constant $\frac{\epsilon}{\epsilon_0}$**	$\frac{\sigma}{\omega_1\epsilon}$	$\frac{\sigma}{\omega_2\epsilon}$
Sea	4	81	88787.9	29596.0
Land1	10^{-2}	15	1198.6	399.5
Land2	10^{-3}	15	119.9	40.0
Ice1	10^{-4}	5	36.0	12.0
Ice2	10^{-5}	5	3.6	1.2

5.5 Background Plasma, Magnetic Field, Electron-Neutral Collisions

The final step to successfully modelling the ionosphere is the implementation of the background magnetised plasma. This is done by using the full set of update equations and matrices in (2.58). The plasma and electron neutral collision frequencies come from Wait's exponential electron density and electron-neutral collision frequency profiles given in (1.1) and (1.2). Figure 5.8 shows the effect a magnetised plasma has on the propagation of radio waves. The plot on the left ignores collisions between electrons and atmospheric neutrals. At low altitudes ($\sim$<30km), the electron density in the ionosphere is low (Figures 1.3 and 1.4), and has little to no effect on the propagation of the radio waves. As altitude increases, the electron density increases, so does its effect on the refractive index of the ionosphere. Comparing to the Mode Theory model, this is characterised by parameter X in (2.20).

With the addition of a background ionosphere, the waves reach a point where they can no longer propagate, and are reflected back down. On the left we see that when there are no collisions, the waves reflect sharply off the ionosphere, at a height of around 70 km. Looking at Mode Theory, this would happen where the refractive index from the simplified Appleton-Hartree equation (2.22) changes rapidly within the space of a wavelength. In the right hand plot, which includes collisions, the effect of the reflected waves is barely noticeable. This is due to the collisions between the excited electrons and neutral molecules absorbing energy at higher altitudes, and not re-radiating strong reflections. The lack of these strong reflections leads to the electromagnetic field being dominated by the incident upward propagating waves. This can be seen by looking at the similarity of right side of Figure 5.8 and the left side of Figure 5.7 which consists only of upward propagating waves since there is no ionosphere to cause reflections. There is no surface implemented in these two plots as they aim to examine the effect of reflections off the ionosphere only and not a waveguide bounded on both sides.

The propagation of radio waves in a magnetised plasma leads to the formation of currents in the ionosphere. With the full description of the background plasma, these currents can also be observed. Figure 5.9 looks at the effect of electron-neutral collisions on these currents that are

formed. On the left, the collisions are ignored and the currents can be seen to form at very low altitudes. While on the right, we see that if collisions are included, the currents only form at higher altitudes. If collisions are ignored, the free electron density is the dominating factor at low altitudes, even if it is low. Since the magnetic field is assumed constant at all altitudes, currents will form as long as there are some free electrons, as is the case on the left. If electron-neutral collisions are included, their effect is far stronger than the free electron density at low altitudes, this blocks the formation of currents. With their inclusion, the formation of currents is only allowed at higher altitudes, where these effects start to become comparable. Comparing to the mode theory model, this is seen in Figure 2.4, through the parameters X and Z from (2.20) This highlights the importance of the inclusion of electron neutral collisions for a realistic ionosphere.

Now that the effects of a background plasma have been incorporated, variations to the ionospheric parameters can also be investigated. As discussed earlier, Wait's ionospheric parameters characterise the height of the ionosphere and the gradient of the electron density profile (1.1). Figures 5.10 and 5.11 give an indication as to how H' and β influence the propagation of the radio waves. Figure 5.10 looks at variation in the emerging wavefield from changing the height of the ionosphere and keeping β constant. As one would expect, but changing the height of the ionosphere, the waves travel to a higher altitude before being blocked and reflected. Another observation, although not as clear, is that when the ionosphere is higher, the wavefield appears to be slightly more distorted. This is due the reflected waves being slightly stronger. Figure 5.8 showed that when there are no collisions, the reflections from the ionosphere are a lot stronger. By increasing the height of the ionosphere, the reflection of the waves is now occurring where the electron-neutral collisions are weaker, seen by parameters Z in 2.4.

For β, however, the effect is not as obvious. Figure 5.11 shows that for a higher value of β, the waves travel to a slightly higher altitude before reflection. Looking back at Figure 1.4, we see that a lowering β corresponds to a increasing the electron density below H' and a decreasing the electron density above H' . This higher electron density at a lower altitude could then lead to the waves also being reflected at a lower altitude, as is the case in Figure 5.11, although this also depends on H', as well as the frequency of the propagating wave. This means that the influence of β on the propagation of the radio waves is not as simple as that of H'; the combined effect of H' and β has to be considered.

5.6 Realistic Example

The FDTD model is now able to effectively treat the fields at the boundaries of the computational domain, include surfaces of varying conductivity and the effects of a magnetised plasma with electron-neutral collisions. By incorporating these aspects together, the code is able to model a realistic situation of a narrowband VLF transmitter radiating radio waves traveling over land or sea with a user defined background ionosphere. The 2D snapshots that have been included

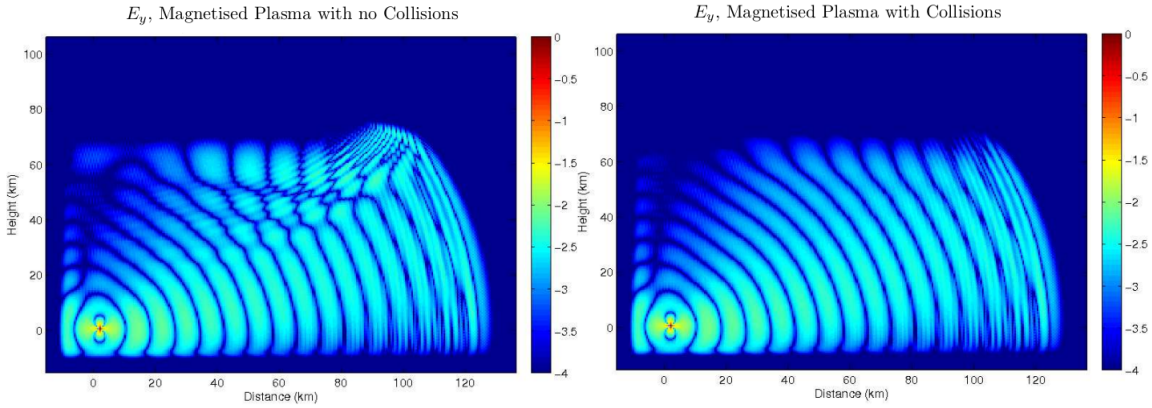


Figure 5.8: Field patterns from a vertical electric dipole with a background plasma and no ground. Left: Magnetised plasma with no collisions, Right: Magnetised plasma with collisions. Colouring consistent with Figure 5.1.

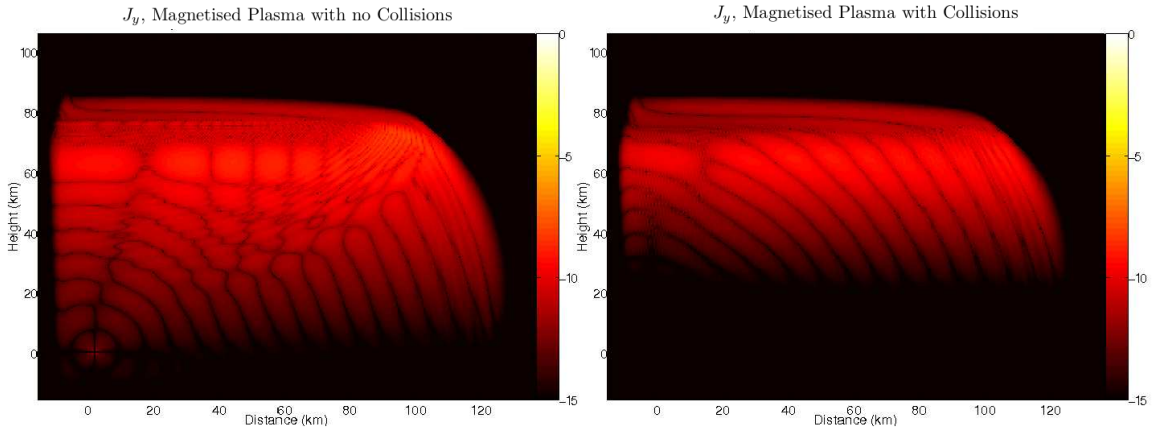


Figure 5.9: Current density patterns from a vertical electric dipole with a background plasma and no ground. Left: Magnetised plasma with no collisions, Right: Magnetised plasma with collisions. Colouring is of $\log 10 J_y$ that is that corresponds to the colouring of the electric fields in other plots.

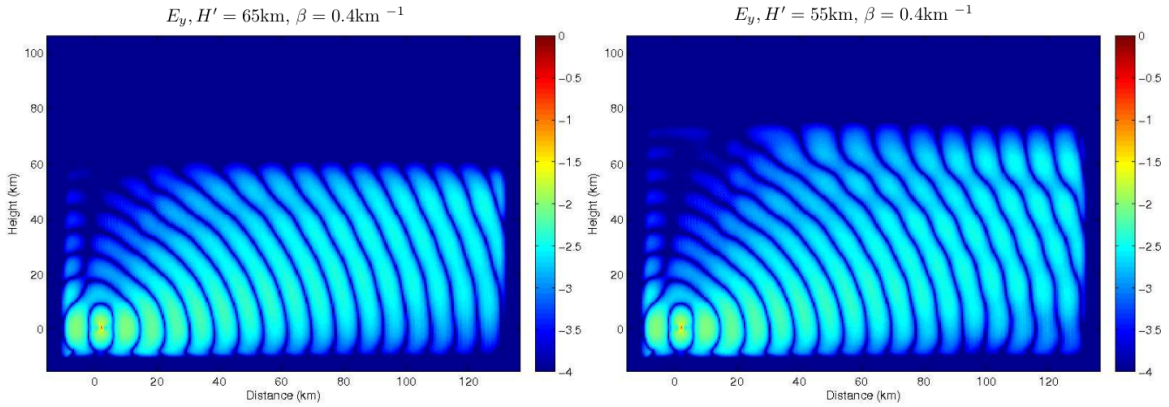


Figure 5.10: Field patterns that emerge with different ionospheric heights, (Left) $H' = 65$km, (Right) $H' = 75$km with $\beta = 0.4$ km $^{-1}$. A higher value of H' results in the wave propagating to higher altitudes. Colouring consistent with Figure 5.1.

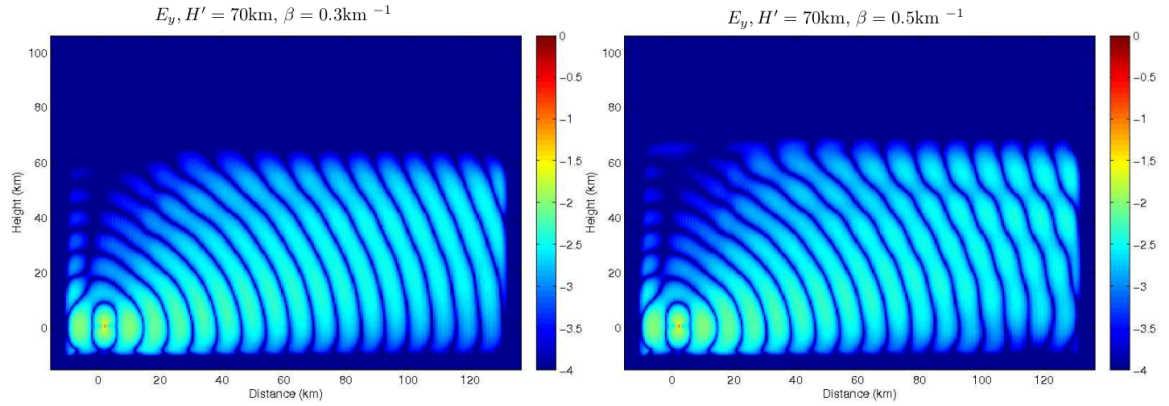


Figure 5.11: Field patterns that emerge with different electron density gradients, (Left) $\beta = 0.3$ km $^{-1}$, (Right) $\beta = 0.5km^{-1}$ with $H' = 70km$. In this case, a higher value of β leads to the wave traveling to a slightly higher altitude before reflection. Colouring consistent with Figure 5.1.

up until now help understand the general behaviour of the waves in the EIWG. They do not however give much indication of the horizontal modal interference pattern profile and the overall steady state strength of the field at different positions in the waveguide. To produce 1D modal interference pattern plots such as those in Chapter 4, from the 2D FDTD simulations, the field values at each grid point along the length of the waveguide at the altitude in question are recorded and added together over a number time steps to observe the steady state field. By performing multiple simulations with different parameters, the results such as those in Figures 5.10 and 5.11 can be displayed like those from the Mode Theory model in Chapter 4, as in Figures 4.4 and 4.4. The addition of the field values over multiple time steps has to be done to achieve a smooth curve that represents the steady state field values. Care must be taken to ensure that the code has been able to run through enough time steps for the fields to reach steady state values, especially at larger distances.

Figure 5.12 was generated in this way, by performing two simulations, one characterised by electron density from (1.1) with typical day-time ionospheric parameters ($H' = 72km$, $\beta = 0.33km^{-1}$), and one with typical night-time ionospheric parameters ($H' = 85km$, $\beta = 0.55km^{-1}$). The simulations were then run by inserting the appropriate values into (2.58). For each simulation, the vertical electric field values just above the surface of the Earth and along length of the waveguide (200 km) were recorded for a number of time steps after the system had reached steady state. It can be seen that for the first 60 km, the field is dominated by the ground wave that falls off with a typical r^{-1} trend for a two dimensional system. This shows that the properties of the background ionosphere play no role in the ground wave. It is also seen that for nighttime propagation, the signal is generally stronger, as expected from previous theory and experimental observation.

Figure 5.13 was created in the same manner, except now looking at differences due to different surfaces while keeping the ionosphere constant. The effect of sea water having a far higher conductivity ($4\ S/m$ compared to $5 \times 10^{-3} S/m$) is clearly visible. This is especially evident for the ground wave where the high conductivity of the sea water allows the wave to travel with little attenuation apart from the expected decay due to distance. For land, the lower conductivity results in high attenuation of the ground wave. This is the reason for the modal interference effects being noticeable at a lot shorter distance than over ocean where he ground wave is still powerful enough to mask the initial modal effects. It can be seen that once the interference pattern dominates over the decay of the ground wave, propagation over both sea and land show very little attenuation, highlighting the ability of the EIWG to let VLF waves travel enormous distances.

The results in Figure 5.14 keep the properties of the waveguide constant (sea, day-time), but look at differences in the steady state field at different altitudes in the waveguide. The clearest variation with height is closest to the transmitter. For the interference pattern along the ground, there is a strong field in the vicinity close to the transmitter (< 30 km). This is due to the ground

wave being emitted directly broadside from the vertical dipole, where the maximum energy is emitted. For the interference pattern at higher altitudes, it is seen that at this close range, the higher the altitude, the lower the field strength. This is exactly what is expected due to the radiation pattern from a vertical electric dipole, as can be seen in the Figure 5.1 (a). At a distance of around 50 to 60 km away from the source, the field strength is similar at all altitudes. From this point on, waves that have reflected off the ionosphere start to influence the electromagnetic field as well. It is seen that the change in the modal interference pattern due to changes in observation height are not the same at all altitudes. For an observation altitude of 0 km, there are modal minima just before a propagation distance of 150 km, and a bit after 200km. When increasing the altitude to 10 km, we clearly see that these modal minima have shifted towards the transmitter. However, going up another 10 km to an altitude of 20 km, this shift of modal minima is not clear. At a propagation distance of 150 km, the highest field strength can be found at an altitude of 30 km, while at a distance of 250 km, the field strength is strongest at 10 km. These observations show that the variation of field strength with position in the waveguide is not simple, but slightly complex, based on the combined effect of the contribution of different modes at different distances and heights. In the Mode Theory model, this is incorporated by (2.25) and (2.26), with Figures 4.1 and 4.2 showing how different modes can have varying influence at different heights.

5.6.1 Comparison with DEMETER observation

Results from the FDTD model will now also be compared to the NWC data from the DEMETER satellite. As with the Mode Theory model, to compare the results to the data, averaging needs to be done over multiple heights to successfully emulate the averaged year long data from the satellite. The field here is also taken at the reflection height in the ionosphere. Figure 5.15 shows the comparison of simulated FDTD results to satellite data. The simulation results are for propagation over sea, and averaged over 10 ionospheric profiles. The initial values chosen for H' and β were for a typical daytime ionosphere and were then experimented with in different combinations to reach this selection. (H' and β) values in (km, km^{-1}) used to generate the model results in Figure 5.15 are (72, 0.38), (72, 0.40), (72, 0.42), (73, 0.38), (73, 0.40), (73, 0.42), (73, 0.44), (74, 0.40), (74, 0.42) and (74, 0.44).

There is some agreement between the model and data. It can be seen that both reveal a local minimum at around 1300km, for propagation over land and sea. For propagation over sea, the results agree better with the data than for land, especially between 2000 and 2500km where the model creates a local minimum not seen in the data while for sea there is far less of a stark difference. This is most likely due to the implementation of the Surface Impedance Boundary Condition. The use of the SIBC is based on the assumption that $\sigma \gg \omega\epsilon$ (2.75), Table 5.2 shows that this assumption is far more valid for sea than for land due to sea water being more conductive than land. Any small differences or errors that arise from the SIBC approximation

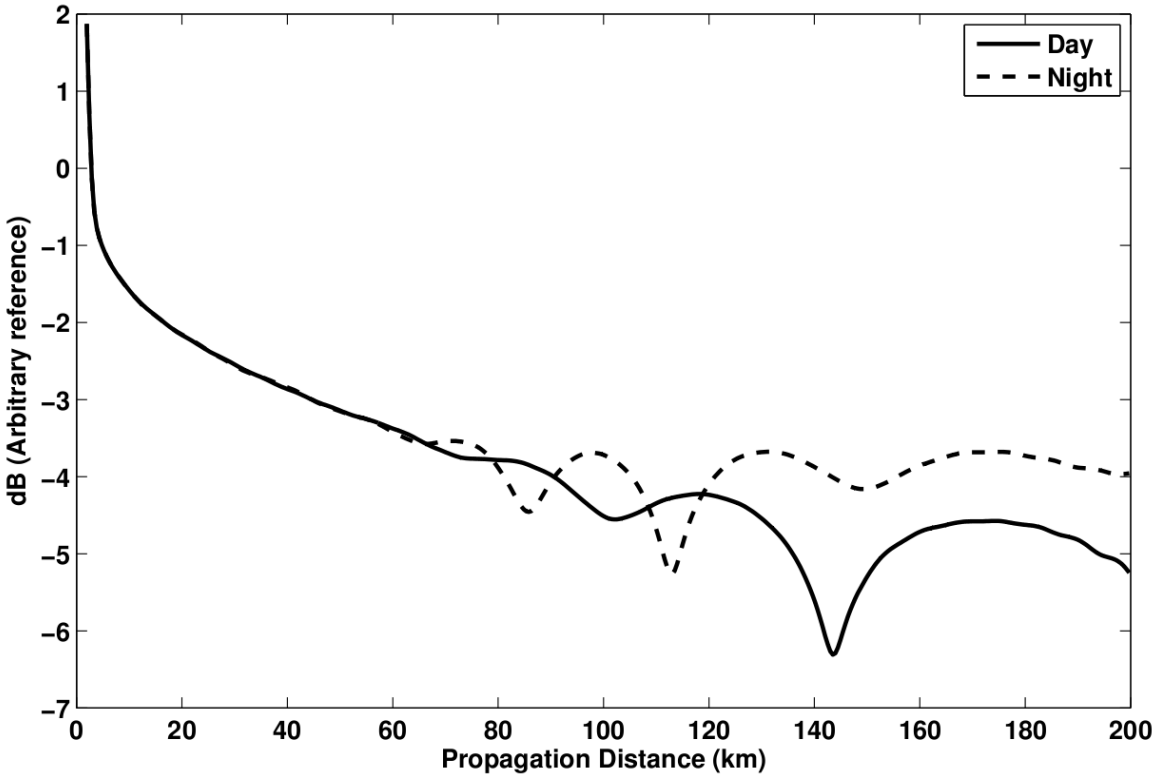


Figure 5.12: Interference pattern for daytime ($H' = 72km$, $\beta = 0.33km^{-1}$) and night time profiles ($H' = 85km$, $\beta = 0.55km^{-1}$) for a 21 kHz wave over sea. The units on the $y-$axis are the same as in previous plots like in Figure 5.1 except now added over multiple steps to achieve steady state. The relative values resulting from different simulations with different sets of parameters are of more interest than absolute values.

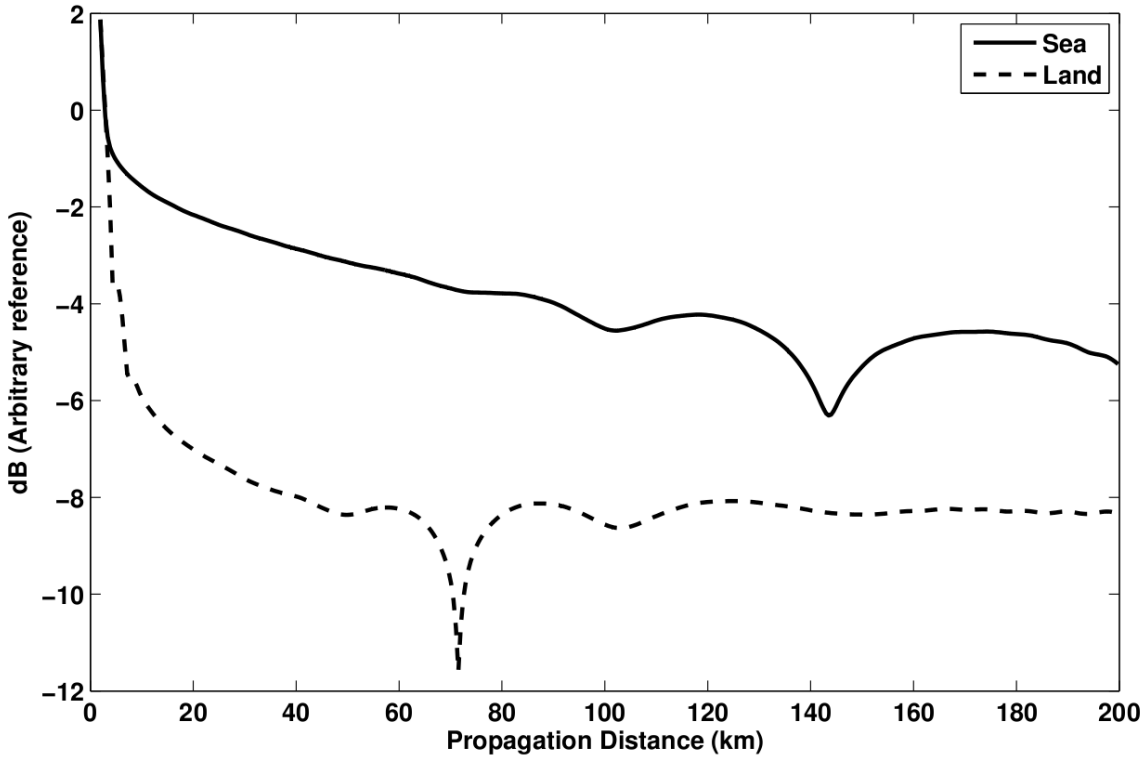


Figure 5.13: Interference pattern for daytime ($H' = 72km, \beta = 0.33km^{-1}$) propagation of a 21 kHz signal over sea and land. Units of the $y-$axis are consistent with Figure 5.12.

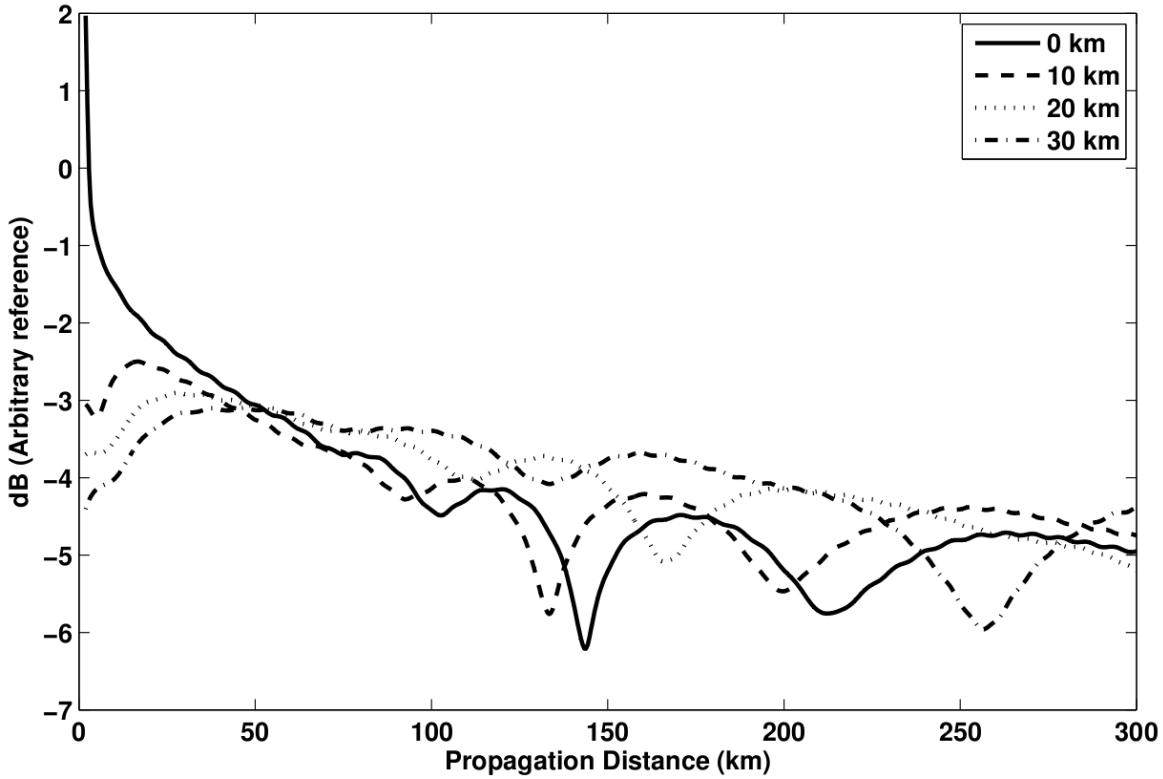


Figure 5.14: Interference pattern for daytime ($H' = 72km, \beta = 0.33km^{-1}$) propagation of a 21 kHz signal over sea at different altitudes. Units of the $y-$axis are consistent with Figure 5.12

will accumulate and become more noticeable at larger propagation distances. The SIBC is seen to correctly account for the fact that propagation over land yields a lower field strength than over sea, but creates undesirable differences in the interference of different propagation modes.

The simulation results over land show an irregularity near the far end of the region. For a smooth curve to be obtained, the fields have to be averaged over a sufficient number of periods while the system has reached a steady state. By allowing for the simulations to carry on for longer and by averaging over more periods, this part could have also been smoother. This was not deemed worthwhile due to time constraints as it takes many hours for a single run to be performed, and many runs need to be performed and then summed and averaged. Although these irregularities are undesirable, is it mostly an aesthetic issue with little consequence to the actual result. Overall, the model performs reasonably well at matching the data, especially when keeping certain aspects of the data in mind. This dataset was created from averaging the results from a year's worth of orbits over the NWC region and stored in 0.5° by 0.5° cells. At the latitude of NWC (21° 47' S), 0.5° of longitude corresponds to roughly 51.4km, that is equal to 2.5 wavelengths. The combination of the averaging of separate orbits and large spatial resolution of the data mean that these results should not be expected to be exact, especially considering this is the image of the modal interference pattern. The FDTD simulation however has an very small spatial resolution of 504m compared to the satellite data. This is clear to see in the fine structure of the modal interference pattern.

5.7 Numerical Stability of the FDTD method

Due to the FDTD algorithm projecting the propagation of electromagnetic waves in the continuous world onto a numerical, discretised grid, nonphysical dispersion takes place. This implies that the phase velocity of the simulated waves, even in the absence of a plasma medium, can differ from c, with the variation dependent on the wavelength, direction of propagation and discretisation. One way to think about this view is that the algorithm embeds the electromagnetic waves onto a "numerical ether" which has properties very similar to vacuum, but not exactly the same [Taflove and Hagness, 2005]. This leads to the numerical waves accumulating delays or phase errors which can lead to nonphysical results.

To illustrate this numerical dispersion, consider a 2D TM polarised free space system, governed by (2.30a), (2.30b) and (2.31c). The FDTD method is applied to yield the following set of equations

$$\frac{H_x{}^{n+1/2}_{i,j+1/2} - H_x{}^{n-1/2}_{i,j+1/2}}{\Delta t} = -\frac{1}{\mu_0}\left(\frac{E_z{}^{n}_{i,j+1} - E_z{}^{n}_{i,j}}{\Delta y}\right) \tag{5.9}$$

$$\frac{H_y{}^{n+1/2}_{i,j+1/2} - H_y{}^{n-1/2}_{i,j+1/2}}{\Delta t} = \frac{1}{\mu_0}\left(\frac{E_z{}^{n}_{i+1,j} - E_z{}^{n}_{i,j}}{\Delta x}\right) \tag{5.10}$$

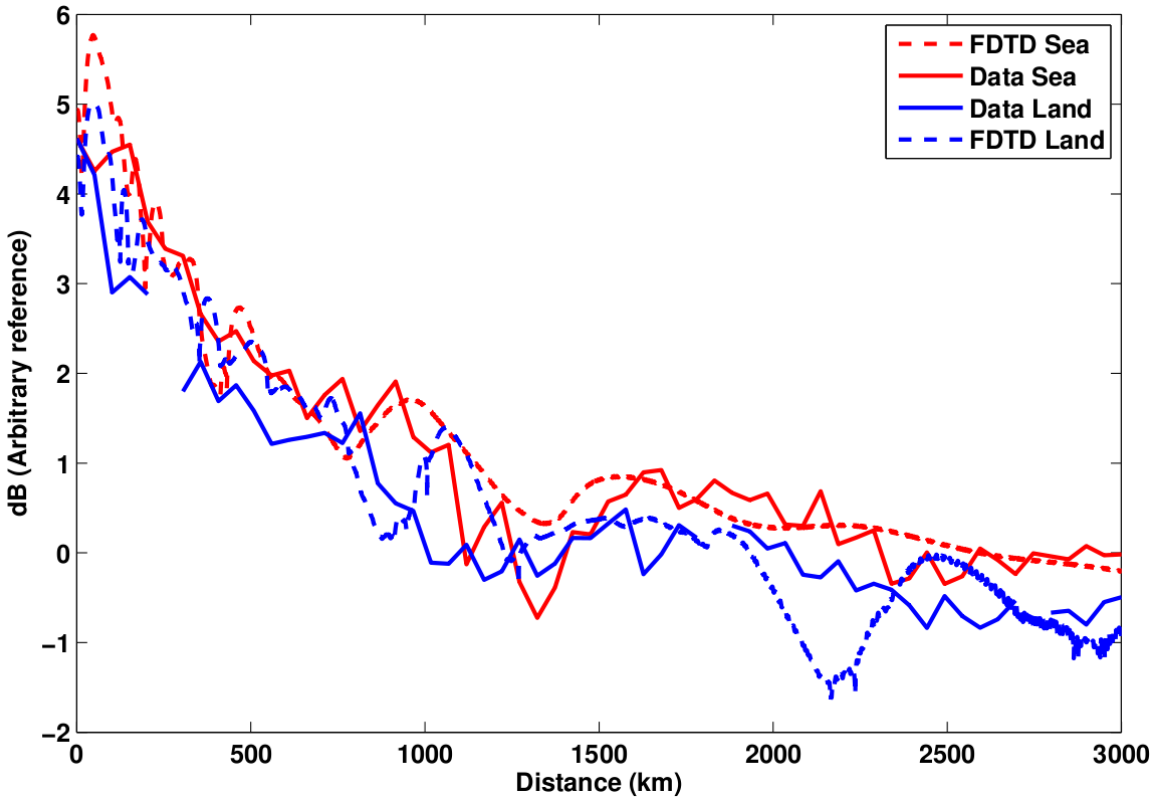


Figure 5.15: Simulated FDTD results of the average of multiple ionospheres compared with DEMETER data.

$$\frac{E_z{}_{i,j}^{n+1} - E_z{}_{i,j}^{n}}{\Delta t} = \frac{1}{\epsilon_0}\left(\frac{H_y{}_{i+1/2,j}^{n+1/2} - H_y{}_{i-1/2,j}^{n+1/2}}{\Delta x} - \frac{H_x{}_{i,j+1/2}^{n+1/2} - H_x{}_{i,j-1/2}^{n+1/2}}{\Delta y}\right). \tag{5.11}$$

To investigate the numerical dispersion, a monochromatic traveling-wave trial solution is substituted into the equations. This trial wave solution takes the form:

$$E_z{}_{i,j}^{n} = E_{z_0} e^{i(\omega n\Delta t - \tilde{k}_x i\Delta x - \tilde{(k)}_y j\Delta y)} \tag{5.12}$$

$$H_x{}_{i,j}^{n} = H_{x_0} e^{i(\omega n\Delta t - \tilde{k}_x i\Delta x - \tilde{(k)}_y j\Delta y)} \tag{5.13}$$

$$H_y{}_{i,j}^{n} = H_{y_0} e^{i(\omega n\Delta t - \tilde{k}_x i\Delta x - \tilde{(k)}_y j\Delta y)} \tag{5.14}$$

where $\tilde{k}_x$ and $\tilde{k}_y$ are the $x-$ and $y-$ components of the numerical wave vector and ω is the angular frequency. By substituting (5.12) to (5.14) into (5.9) to (5.11) and simplifying, the following relations emerge:

$$H_{x_0} = \frac{\Delta t E_{z_0}}{\mu_0 \Delta y}\frac{\sin(\tilde{k}_y \Delta y/2)}{\sin(\omega\Delta t/2)} \tag{5.15}$$

$$H_{y_0} = -\frac{\Delta t E_{z_0}}{\mu_0 \Delta x}\frac{\sin(\tilde{k}_y \Delta y/2)}{\sin(\omega\Delta t/2)} \tag{5.16}$$

$$E_{z_0}\sin\left(\frac{\omega\Delta t}{2}\right) = \frac{\Delta t}{\epsilon_0}\left[\frac{H_{x_0}}{\Delta y}\sin\left(\frac{\tilde{k}_y\Delta y}{2}\right) - \frac{H_{y_0}}{\Delta x}\sin\left(\frac{\tilde{k}_x\Delta x}{2}\right)\right]. \tag{5.17}$$

Substitution of (5.15)and (5.16) into (5.17) yields

$$\left[\frac{1}{c\Delta t}\sin\left(\frac{\omega\Delta t}{2}\right)\right]^2 = \left[\frac{1}{\Delta x}\sin\left(\frac{\tilde{k}_x\Delta x}{2}\right)\right]^2 + \left[\frac{1}{\Delta y}\sin\left(\frac{\tilde{k}_y\Delta y}{2}\right)\right]^2 \tag{5.18}$$

Eq. 5.18 is now rewritten for the square cell grid with $\Delta x = \Delta y = \Delta$ and with the use of the Courant number, C_s , and the grid sampling density, or PPW, $N_\lambda = \lambda_0/\Delta$,

$$\frac{1}{C_s^2}\sin^2\left(\frac{\pi C_s}{\lambda}\right) = \sin^2\left(\frac{\Delta\tilde{k}\cos\phi}{2}\right) + \sin^2\left(\frac{\Delta\tilde{k}\sin\phi}{2}\right) \tag{5.19}$$

where ϕ is the angle between the propagation direction and the $x-$axis. For propagation along a major vertical or horizontal axis ($\phi = 0°, 90°, 180°$ or $270°$), (5.19) reduces to

$$\frac{1}{C_s}\sin\left(\frac{\pi C_s}{\lambda}\right) = \sin\left(\frac{\tilde{k}\Delta}{2}\right) \tag{5.20}$$

or equivalently

$$\tilde{k} = \frac{2}{\Delta}\sin^{-1}\left[\frac{1}{C_s}\sin\left(\frac{\pi C_s}{N_\lambda}\right)\right] \tag{5.21}$$

with corresponding numerical phase velocity

$$\tilde{v}_p = \frac{\omega}{\tilde{k}} = \frac{\pi}{N_\lambda \sin^{-1}\left[\frac{1}{C_s}\sin\left(\frac{\pi C_s}{N_\lambda}\right)\right]}c. \tag{5.22}$$

By solving the numerical dispersion relation, (5.19) for propagation along a diagonal axis ($\phi = 45, 135^\circ, 225^\circ$ or 315°), we get

$$\frac{1}{C_s}\sin\left(\frac{\pi C_s}{N_\lambda}\right) = \sqrt{2}\sin\left(\frac{\tilde{k}\Delta}{2\sqrt{2}}\right) \tag{5.23}$$

or equivalently

$$\tilde{k} = \frac{2\sqrt{2}}{\Delta}\sin^{-1}\left[\frac{1}{C_s\sqrt{2}}\sin\left(\frac{\pi C_s}{N_\lambda}\right)\right] \tag{5.24}$$

with corresponding numerical phase velocity

$$\tilde{v}_p = \frac{\omega}{\tilde{k}} = \frac{\pi}{N_\lambda\sqrt{2}\sin^{-1}\left[\frac{1}{C_s}\sqrt{2}\sin\left(\frac{\pi C_s}{N_\lambda}\right)\right]}c. \tag{5.25}$$

To illustrate a basic example, if we assume the grid to have $C_s = 0.5$ and $N_\lambda = 30$, the dispersion relations yield numerical phase velocities of $0.998625c$ and $0.999542c$ for major axis and diagonal propagation respectively. Both of these velocities are less than c, indicating a phase lag relative to a physical wave. If we then compare these two results with each other, we notice they are not equal. We find that a wave traveling obliquely has a speed $0.999542/0.998625 = 1.00092$ times the speed of a wave propagating along a major axis. This represents a velocity anisotropy of around 0.092% between oblique and axis-long propagation. The wavefront distortion due to this anisotropy would be about 0.9 cells for every 1000 cells the wave has propagated. For a 20 kHz wave with $N_\lambda = 30$, this means that the diagonal wave would have traveled around 1 km further than the horizontal wave over a distance of roughly 500 km. This seems like a small amount, but could have an effect for example on the interference between reflected and direct waves in a waveguide. The effect becomes more pronounced at large distances as it scales linearly with propagation path length. A wave traveling over a distance of $10\lambda_0$ with a velocity anisotropy of 0.092% could suffer a phase lag of 3.3° . If $N_\lambda = 20$ instead, then the velocity anisotropy would be 0.208% which would result in a phase lag of 7.5° over the space of $10\lambda_0$. This highlights the importance of setting N_λ high enough as to minimise numerical errors. This effect of numerical dispersion is an inherent property of grid based numerical simulations and cannot be totally avoided, only minimised. More comprehensive results are given in Table 5.3 which shows the effects of different parameters on numerical dispersion. Note here that the diagonal phase speed

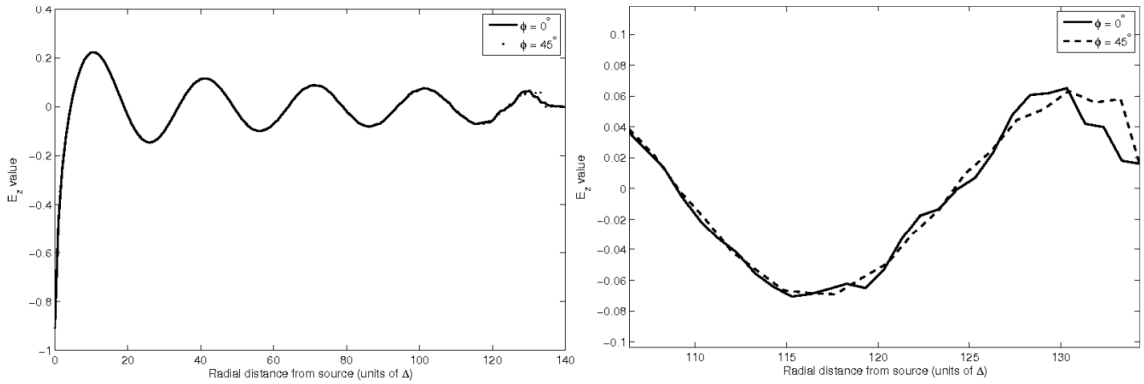


Figure 5.16: Left: E_z field distribution along diagonal and vertical axis. Right: Expanded view between 105 and 135 Δ.

always exceeds that of propagation along a major axis. We see here that when C_s is set to its maximum value (5.4), the diagonal phase velocity matches the speed of light perfectly. If the Courant number is set above its limit, $1/\sqrt{d}$, where d is the number of dimensions, then the phase velocity for diagonal propagation would exceed c, resulting in superluminal propagation. Figure 5.16 was generated by simulating the electric and magnetic fields from (5.9) to (5.11) and recording the normalised E_z values along two paths, one along the $x-$direction from the source, and the other along a diagonal, xy-direction. The distance of propagation along each path was calculated and normalised to the size of Δx $(= \Delta y)$. At short distances, the field strength appears to be the same for the two directions, as it should be. Further away from the source, the consequences of the slightly different phase velocities become apparent with the values along the horizontal and diagonal paths starting to diverge. The effect is small, on the order of 1 cell for every 1000 that the wave has propagated, but it is clear that the waves are starting to diverge.

Table 5.3: Numerical dispersion analysis results.

	$\tilde{v}_p/c, C_s = 0.5$			$\tilde{v}_p/c, C_s = 1/\sqrt{2}$		
N_λ	$\phi = 0°$	$\phi = 45°$	$\phi = \frac{\phi=0°}{\phi=45°}$	$\phi = 0°$	$\phi = 45°$	$\phi = \frac{\phi=0°}{\phi=45°}$
10	0.987264	0.995817	1.008663	0.991485	1.0	1.008588
20	0.996892	0.998968	1.002082	0.997926	1.0	1.002078
30	0.998625	0.999542	1.000919	0.999083	1.0	1.000918
40	0.999227	0.999743	1.000516	0.999485	1.0	1.000515

Chapter 6

Discussion and Conclusions

The propagation of VLF waves in the EIWG can be a complicated topic with many factors influencing it. Understanding these factors is of scientific interest, as well as for other reasons such as communication. This thesis set out to produce two models that simulate the propagation of Very Low Frequency (VLF) radio waves in the EIWG using Mode Theory and the Finite-Difference Time-Domain (FDTD) technique. The main aim being to use these models to investigate theoretical as well as experimental aspects of VLF wave propagation. The theoretical background of the two methods is discussed as well as the concepts of the EIWG and the important factors influencing it such as electron density. Characterisation of the electron density profile is presented in terms of Wait's parameters, H' and β. The Mode Theory model is first used to illustrate the formation of a modal interference pattern and how changes to Wait's parameters modify it. We see that reducing the height of the ionosphere causes the modal interference pattern to be slightly compressed and shifted towards the transmitter while the overall field strength remains the same. Increasing β also changes the shape of the modal interference pattern in a way similar to reducing the height except changes in β also result in a noticeable change in the field strength, with smaller values of β leading to lower field strengths. This is due to the fact that the shape of the pattern is mostly determined from the resulting interference between reflected waves. This interference depends on the phase of the fields and this is mostly dependent on the height of the ionosphere. The strength of the field on the other hand is more dependent on how much absorption of the waves occurs, this absorption is more closely related to the gradient of the electron density profile. The individual components of the Mode Theory model are then used to examine other theoretical aspects of waveguide propagation. The refractive index profiles for the X- and O-mode waves are examined for night and day ionospheres. Results from the model agree with established theory such as higher attenuation rates during the day and for higher order modes. The effect of mode number and ionospheric height on the attenuation rate for a range of frequencies is also presented.

A version of the model that includes temporally varying ionospheric parameters was then introduced and used to investigate the effect of a solar flare on a narrowband VLF signal. The

form of the amplitude response of the 18.3 kHz HWV transmitter in France detected at SANAE-IV, Antarctica to 2 different sized solar flares was examined. Temporal changes to H' and β are included in the same manner as the X-ray flux from a solar flare, fast onset and slow recovery. A small and larger flare were examined with respective changes to H' and β of $\Delta H' = 2km$, $\Delta\beta = 0.01km^1$ and $\Delta H' = 4km$, $\Delta\beta = 0.03km^1$. The changes to the amplitude are explained in terms of changes in the modal interference pattern and the relative position of the receiving station as well as the magnitude of the flare. The model was able to successfully replicate the form of the unusual double-peak response to a flare that occurred on 12 February 2010.

Comparison to DEMETER data was performed to validate the model. The model was in reasonable agreement with the data. The direct comparison with the data was good but there is some doubt with respect to the parameters used to achieve these results. The model found agreement using ionospheric reflection heights ranging from 65 to 68 km, lower than results from a previous study which found the height to be between 70 and 72 km [Thomson, 2010]. This discrepancy is due to the very different methods used to estimate the height as well as the fact that the measurements are from different times in the solar cycle. The FDTD model is initially used to illustrate some basic concepts of electromagnetic propagation such as the distribution of an electric field around a dipole source. Waveguide boundaries are then introduced to show the formation of waveguide modes. The effects waveguide height as well as orientation and position of source are examined. Different aspects of propagation in the EIWG such as a magnetised plasma and different surfaces are first examined individually. These different aspects are then combined so that the model can simulate a realistic example of VLF propagation in the EIWG.

First the differences between day and night time propagation as well as propagation of over sea and land are illustrated. These results from the FDTD model agree with the Mode Theory model as well established theory; waves travel better at night and over sea. The FDTD model is also used to look at the interference pattern at different heights in the waveguide. It clearly illustrates that at short distances, the pattern is determined by the direct rays and their dipole distribution. Past a certain point, the structure is determined from the modal interference pattern.

The FDTD model is then compared to the same DEMETER data as the Mode Theory model with the same process of averaging the results from multiple unique ionospheres. The FDTD results of the observed interference pattern do not agree with the data as well as the Mode Theory model. The agreement for propagation over sea water is better than for that over land. This could be due to the implementation of the SIBC. The approximation used to validate the use of the SIBC is far stronger for sea than land due to the higher conductivity of sea water. The spatial resolution of the data is just over 100 times lower than that of the model, 51.4km compared to 504m. The FDTD model did however agree more with the previous study of Thomson [2010] with respect to the height of the ionosphere than the Mode Theory model. Compared to the 7072km range Thomson [2010] reported, the FDTD method is only 2km higher with heights ranging from 7274km. Although both models proved to be useful, there is definitely room for improvement. The

waveguide model could be modified as to include a correction for the curvature of the Earth. It would also be very desirable for the Mode Theory model to be able to perform mode conversions so that slabs of varying waveguide parameters could be included, although this is quite challenging. There are two large and clear possible ways to improve the FDTD model. Firstly, the model can be extended to three dimensions. This in itself would not be more challenging as the two dimensional case since extending to higher dimensions is trivial, yet tedious. This would however dramatically increase the computational requirements of the program. Having a 3 dimensional model would allow for the investigation of many interesting phenomena such as indirect, scattered signals between transmitter and receiver. Secondly, the addition of temporally varying ionospheric parameters would lead to the possible simulation of many time dependent phenomena such as the passage of the day-night terminator and other traveling ionospheric disturbances. This would however also cause a massive increase in computational time since the coefficients determining the propagation of the waves will have to be constantly recalculated as opposed to the static version in this thesis. Both of these models presented in this thesis are shown to be useful in both theoretical and experimental situations. When looking at the results, it must be kept in mind that each of the models suffer from some sort of weakness. This generally arises from simplifications or approximations made. The results from the two models show that the FDTD model is more suited to investigating propagation over short and medium distances where evanescent modes can still have significant contributions. For the FDTD model to simulate propagation over large distances, the computational requirements would dramatically increase, which are already large for a standard computer. The effect of numerical dispersion will also be far more pronounced due to phase errors and therefore incorrect interference of the reflected waves. Another positive for the FDTD model is that the grid can be customised with little trouble. The Mode Theory model on the other hand is seen to be more useful when investigating long distance propagation where the modal aspect of waveguide propagation is dominant. It is for this reason that the Mode Theory model is better at describing the overall, long distance behaviour as opposed to the small, finer details.

Bibliography

Beggs, J. H., Luebbers, R. J., Yee, K. S., and Kunz, K. S. (1992). Finite difference time-domain implementation of surface impedance boundary conditions. *IEEE Transactions on Antennas and Propagation*, 40(1):49–56.

Berenger, J. P. (1984). A Perfectly Matched Layer for the Absorption of Electromagnetic Waves. *Journal of Computational Physics*, 114:185–200.

Budden, K. G. (1962). The influence of the earth's magnetic field on radio propagation by waveguide modes. *Proceedings of the Royal Society of London, Series A*, 265(1323):538–553.

Budden, K. G. (1985). *The propagation of radio waves.* Cambridge University Press.

Chapman, S. (1931). The absorption and dissociative or ionizing effect of monochromatic radiation in an atmosphere on a rotating earth. *Proceedings of the Physical Society*, 43:26–45.

Cheng, Z., Cummer, S. A., Baker, D. N., and Kanekal, S. G. (2006). Nighttime D region electron density profiles and variabilities inferred from broadband measurements using VLF radio emissions from lightning. *Journal of Geophysical Research*, 111(A5):1–9.

Cummer, S. (2000). Modeling electromagnetic propagation in the Earth-ionosphere waveguide. *IEEE Transactions on Antennas and Propagation*, 48(9):1420–1429.

Cummer, S. (2003). A simple, nearly perfectly matched layer for general electromagnetic media. *IEEE Microwave and Wireless Components Letters*, 13(3):128–130.

Cummer, S. A., Inan, U. S., and Bell, T. F. (1998). Ionospheric D region remote sensing using VLF radio atmospherics. *Radio Science*, 33(6):1781–1792.

Ferguson, J. A. (1992). A Review of the Ionospheric Model for the Long Wave Prediction Capability A Review of the Ionospheric Model for the Long Wave Prediction Capability. *Control*.

Ferguson, J. A. (1998). Computer Programs for Assessment of Long- Wavelength Radio Communications , User's Guide and Source Files. *Assessment*, (May).

Ferguson, J. A. and Snyder, F. (1980). *Approximate VLF/LF Waveguide Mode Conversaion Model. Computer Applications: FASTMC and BUMP*. Defense Technical Information Center.

Green, A. L. (1974). Early history of the ionosphere. *Journal of Atmospheric and Terrestrial Physics*, 36(12).

Griffiths, D. J. (1996). *Introduction to electrodynamics*. Prentice-Hall.

Han, F. and Cummer, S. A. (2010). Midlatitude daytime D region ionosphere variations measured from radio atmospherics. 115(May):1–13.

Han, F., Cummer, S. A., Li, J., and Lu, G. (2011). Daytime ionospheric D region sharpness derived from VLF radio atmospherics. *Journal of Geophysical Research*, 116(A5):1–11.

Hu, W. and Cummer, S. A. (2006). An FDTD Model for Low and High Altitude Lightning-generated EM Fields. *IEEE Transactions on Antennas and Propagation*, 54(5):1513–1522.

Liao, Z., Wong, H., Yang, B., and Yuan, Y. (1984). A transmitting boundary for transient wave analyses. *Scientia Sinica (series A)*.

Maloney, J. G. and Smith, G. S. (1992). The use of surface impedance concepts in the finite-difference time-domain method. *IEEE Transactions on Antennas and Propagation*, 40(1).

Morfitt, D. (1977). *Effective Electron Density Distributions Describing VLF/LF Propagation Data*. Defense Technical Information Center.

Mur, G. (1981). Absorbing Boundary Conditions for the Finite-Difference Approximation of the Time-Domain Electromagnetic-Field Equations. *IEEE Transactions on Electromagnetic Compatibility*, EMC-23(4):377–382.

Nicolet, M. and Aikin, A. C. (1960). The Formation of the D Region of the Ionosphere. *Journal of Geophysical Research*, 65(5).

Simpson, J. (2010). An E-J Collocated 3-D FDTD Model of Electromagnetic Wave Propagation in Magnetized Cold Plasma. *IEEE Transactions on Antennas and Propagation*, 58(2):469–478.

Soloviev, O. (2003). 3D modelling of VLF radio wave propagation in terrestrial waveguide allowing for localized large-scale ionosphere perturbation. *Journal of Atmospheric and Solar-Terrestrial Physics*, 65(5):627–634.

Taflove, A. (1980). Application of the Finite-Difference Time-Domain Method to Sinusoidal Steady-State Electromagnetic-Penetration Problems. *IEEE Transactions on Electromagnetic Compatibility*, EMC-22(3):191–202.

Taflove, A. and Hagness, S. (2005). *Computational Electrodynamics, the Finite-Difference Time-Domain Method*. Artech House, 3rd edition.

Thomson, N. (1993). Experimental daytime VLF ionospheric parameters. *Journal of Atmospheric and Terrestrial Physics*, 55(2):173–184.

Thomson, N. R. (2010). Daytime tropical D region parameters from short path VLF phase and amplitude. *Journal of Geophysical Research*, 115(A9):1–11.

Thomson, N. R., Clilverd, M. A., and Mcrae, W. M. (2007). Nighttime ionospheric D region parameters from VLF phase and amplitude. *Journal of Geophysical Research*, 112:1–14.

Wait, J. R. (1960). Terrestrial propagation of very-low frequency radio waves, a theoretical investigation. *Journal of Research of the National Bureau of Standards*, 64D:153.

Wait, J. R. (1962). *Electromagnetic Waves in Stratified Media*, volume 3 of *International Series of Monographs on Electromagnetic Waves*. Pergamon Press, New York.

Wait, J. R. (1964). On phase changes in very-low freqeucnt propagation induced by an ionospheric depression of finite extent. *Journal of Geophysical Research*, 64(3):441–445.

Wait, J. R. and Spies, K. P. (1964). *Characteristics of the Earth-Ionosphere waveguide for VLF radio waves.* National Bureau of Standards, technical note 300. edition.

Wait, J. R. and Spies, K. P. (1965). Influence of finite ground conductivity on the propagation of vlf radio waves. *Journal of Research of the National Bureau of Standards*, 64D:1359.

Yee, K. S. (1966). Numerical solution of initial boundary value problems involving Maxwells equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, 14(3):302–307.