Simulation of Ionospheric Irregularities on Transionospheric Radio Systems for the Enhancement of Safety in the Aviation Industry

By

AKALA, Andrew Oke-'Ovie

Matric No. 029076033

A Thesis Submitted to the School of Postgraduate Studies, University of Lagos, Akoka, Lagos, Nigeria in Partial Fulfillment of the Requirements for the Award of Doctor of Philosophy (Ph.D) Degree in Applied Physics (Ionospheric and Radio Propagation)

Department of Physics

University of Lagos

Akoka, Lagos

December, 2009

SCHOOL OF POSTGRADUATE STUDIES
UNIVERSITY OF LAGOS

CERTIFICATION

This is to certify that the Thesis:

"SIMULATION OF IONOSPHERIC IRREGULARITIES ON TRANSIONOSPHERIC RADIO SYSTEMS FOR THE ENHANCEMENT OF SAFETY IN THE AVIATION INDUSTRY"

Submitted to the
School of Postgraduate Studies
University of Lagos

For the award of the degree of

DOCTOR OF PHILOSOPHY (Ph. D)

is a record of original research carried out

By

AKALA, ANDREW OKE-'OVIE

In the Department of Physics

AKALA, ANDREW OKE-'OVIE		2/12/09
AUTHOR'S NAME	SIGNATURE	DATE
1ST SUPERVISOR'S NAME	SIGNATURE	DATE
Dr. A.B. Adelaye		02/12/09
2ND SUPERVISOR'S NAME	SIGNATURE	DATE
Dr E.O. Somoye		02/12/09
3RD SUPERVISOR'S NAME	SIGNATURE	DATE
Prof J.A. Adebojo		2/12/2009
1ST INTERNAL EXAMINER	SIGNATURE	DATE
Dr. E.A. Ayoola		2/12/09
2ND INTERNAL EXAMINER	SIGNATURE	DATE
Dr. A.B. Rabiu		2/12/09
EXTERNAL EXAMINER	SIGNATURE	DATE
Dr S.O. Adeosun		02/12/2009
SPGS REPRESENTATIVE	SIGNATURE	DATE

APPROVAL

This is to certify that **AKALA Andrew Oke-'Ovie** carried out the entire work described in this thesis under our supervision in the Department of Physics (Ionospheric and Radio Propagation Option), University of Lagos, Akoka, Lagos, Nigeria.

Dr J. Akinrimisi (late)

Major Supervisor

Dr A. B. Adeloye

Major Supervisor

Dr E. O. Somoye

Co-supervisor

Dr (Mrs) D. O. Olorode

Ag. Head of Department

DEDICATION

This thesis is dedicated to **ALL** lives that have been lost to aviation disasters in Nigeria.

ACKNOWLEDGEMENTS

I thank God for His mercy and protection throughout the period of this research. My sincere thanks go to my supervisor, Dr J. A. Akinrimisi (late) for his guidance and painstaking effort in reading my reports. It is painful that death has decided to snatch you away from me when I needed you most! Indeed, epitome of humility and embodiment of knowledge is gone. Rest in the bosom of our Lord Jesus Christ.

I express special thanks to the Postgraduate Committee of the Department of Physics, University of Lagos for coming to my aid while the need to fill the vacuum that was created by the demise of my supervisor arose. This Committee appointed Dr A. B. Adeloye and Dr E. O. Somoye as my supervisors, and the duo demonstrated high level of dexterity in reviewing my thesis at a very short time. This effort and guidance is acknowledged with unreserved gratitude.

At times, we experienced huddles before seeing light at the end of the dark tunnel. At these points, many colleagues (home and abroad) gave us encouragements. The magnanimity of Dr L. L. N. Amaeshi in exposing me to useful literature must be mentioned. I am very much indebted to Dr Patricia Doherty (Director) of the Institute for Scientific Research, Boston College, Boston, Massachusetts, USA, and other two of her colleagues at the Institute, Dr Charles Carrano (formerly with Atmospheric and Environmental Research Inc, Lexington, Massachusetts, USA) and Dr Cesar Valladares for their useful guidance and contributions to the aviation application segment of this research. Dr Carrano reviewed this segment over two times, and also

provided the TEC records from Nairobi, Kenya and scintillation records from Ascension Island. He created contact on our behalf with Dr Ronald Woodman of Jacamarca Radio Observatory, Peru for the provision of the spread F records. Special thanks are also due to Dr Todd Walter and Dr Jiwon Seo of the Department of Aeronautics and Astronautics, Stanford University, Stanford, California, USA for granting the permission to use scintillation records that they acquired sometimes ago by Garmin 480 WAAS receiver from Sao Jose dos Campos, Brazil. I acknowledge with gratefulness, all the useful discussions on numerical methods with Prof. J. A. Adepoju of the Department of Mathematics, University of Lagos. I also appreciate the guidance rendered by Dr L. O. Chukwu of the Department of Marine and Fisheries, University of Lagos during the preparation and compilation of this thesis. The useful comments and corrections on this thesis by Prof. J. O. Adeniyi of the Department of Physics, University of Ilorin is acknowledged with gratefulness.

These acknowledgements will not be adequate if I refuse to thank the former Head of Department of Physics, Dr E. A. Ayolabi for his encouragement towards the successful completion of my work. Special thanks are also due to the present Head of Department, Dr (Mrs) D. O. Olorode for her support towards my final *viva*. I thank all my senior colleagues in the Department, Drs. M. A. C. Chendo, E. O Oyeyemi, B. S. Badmus, H. O. Boyo, K. F. Oyedele, L. Adeoti and all other staff of the Department that supported us in one way or the other.

In course of this research, some agencies supported us with travel grants to attend conferences. I am indebted most of all to the management of the University of Lagos for the Graduate Fellowship Award (two times) and for the travel grant to attend the Institute of Navigation Conference (January, 2009) in California. Prof. (Mrs.) M. O. Ogunlesi, Deputy Vice Chancellor (Academic & Research), University of Lagos was instrumental to the speedy approval of this grant. I thank the sponsors of the International Heliophysical Year (IHY) Workshop (November, 2008) held in Enugu, Nigeria for supporting us with grant to attend the meeting.

At this Juncture, I must express my profound gratitude to my parents for their support from my first day in school till date. Finally, my deepest appreciation goes to my wife, Salome Akala for her understanding and perseverance during this long process. I must apologize to all that helped us in one way or the other but whose names have been omitted in these acknowledgements. Please, it is an oversight!

ABSTRACT

This thesis highlights the role that transionospheric radio propagation (with particular interest on global positioning system (GPS)) plays in enhancing aviation safety. The ionosphere over equatorial latitudes is characterized with high dynamics (irregularities) that may challenge the performance of transionospheric radio system, especially on phenomenal days. The objectives of this work are to: investigate how ionospheric irregularities grow and evolve, and thereafter assess their potential consequences on transionospheric radio system; develop mitigation strategies to reduce effects of ionospheric irregularities on systems' performances; elucidate operational principles of GPS using numerical approach, with a view to assisting future modeling efforts; and finally design a robust safety technique for aviation industry using GPS-based initiatives.

The methodology follows two approaches: (i) mathematical modeling, and (ii) analyses of experimental observations. The growth and evolution of ionospheric irregularities is investigated from nonlinear dynamics perspective, and revealed to have bifurcation patterns. This was thereafter explained to have interpretable correlations with scintillations. The model result shows good agreement with experimental observations obtained by previous workers. The phase screen model part of this study concludes that sustainability of transmission coherency depends on (i) centre frequency of propagation, and (ii) extent of fluctuations (scintillations) in the ionosphere.

Measured ionospheric data at different stations within the equatorial anomaly region during solar maximum and solar minimum are analyzed and interpreted. The analyses lead to conclusions that the impacts of scintillation on GPS-based navigation system can only be hazardous when too many satellites' channels (leaving less than four) are lost to deep fading (greater than 22dB-Hz) at exact the same time.

However, for the overall reliability and robustness of navigation systems in event of anomalous scintillation conditions which may give rise to simultaneous deep fading on too many satellites' channels, the study suggests that intended GPS avionics for transequatorial flights should be enhanced to cope with fading of over 25dB-Hz. The study also indicates that short re-acquisition time of 10 seconds or less, which corresponds to the International Civil Aviation Organization's (ICAO) requirements for all precision approach categories is necessary for optimal receiver's tracking capability.

The underlying principle of operations of GPS is presented. Under a clear view, uniquely positioned GPS receivers have direct line of sight (LOS) with four to twelve satellites. A software code in FORTRAN language that reduces superfluous variables to a system of four variables by multidimensional Newton-Raphson method and finally solved by Gaussian elimination approach to provide the navigation solution was developed.

Lastly, the study is applied to enhance aviation safety by designing a GPS-driven airspace availability optimization technique whereby airplanes can be flown at: minimum multilateral separations, reduced emissions and efficient scheduling. Also, a robust aeronautical search and rescue mechanism is developed with background of the Nigerian airspace in grid calibration mode for easy referencing during emergencies.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

1.0 INTRODUCTION

1.1 Background to Study

The application of transionospheric radio propagation in enhancing safety in the aviation sector has continued to be a subject of discussion within the aviation community in recent times. The Federal Aviation Administration (FAA) of the United States and European Space Agency (ESA) of Europe have supported series of research efforts in order to improve the performance of this technology in the past few years (Van Dyke, 2001; Doherty et al., 2002; and 2004).

One fundamental problem that has been noted to be a challenge to this technology is the scintillation effect that such radio wave suffers as it traverses the ionosphere (McClure, 1964; Elkins and Papagiannis, 1969; Rufenach, 1972; Whitney and Basu, 1977; Basu and Kelly, 1979; Serafim, 1980; Aarons, 1982; Yeh and Liu, 1982; Groves et al., 1997; Wernik et al., 2003; de-Paula et al., 2003; Dandekar and Groves, 2004; Kintner and Ledvina, 2005; and Ray et al.,2006). Both phase and amplitude of radio wave fluctuate as it passes through ionospheric irregularities. These irregularities introduce stochastic features on the electrons, causing them to behave as scatterers to transient waves (Kelly, 1989; and Rama Rao et al., 2006), even at GHz frequency, before they are finally diffracted to free space below the ionosphere. Levels of scintillations suffered by transionospheric radio waves are usually further complicated by the geographical location of the link, with the worst effect around the equatorial anomaly region-around the geomagnetic latitude where Nigeria is located (Aarons, 1982). Other potentially active regions are the aurora and polar cap latitudes

(Yeh et al., 1975; and Brown et al., 1998). Scintillations also undergo diurnal, seasonal, solar cycle (Koster, 1963; and Rama Rao et al., 2006) and geomagnetic storm time variations (Singleton, 1968; 1977; and Gherm et al., 2005). Although, scintillation is not too frequent in the mid-latitude region, it portends hazard to application of transionospheric radio systems in terms of availability and continuity of service in equatorial region (Seo et al., 2007).

Previous studies on scintillation theories (Yeh and Liu, 1982; Yeh et al., 1975; Groves et al., 1997; and Wernik et al., 2003), have related observed signal statistics to the statistics of ionospheric density irregularities. Consequently, adequate understanding of the formation and evolution of these irregularities is a sine-qua-non for proper solution of the problem of scintillation (Calvert and Cohen, 1961; and Woodman and La Hoz, 1976).

One of the emerging applications of transionospheric radio propagation is the Global Positioning System (GPS). Historically, GPS technology started as far back as 1973 but declared operational by the US Department of Defence (DoD) in 1994, after successful inauguration of reliable constellation (Tsui, 2000; Kintner et al., 2001; and Kintner and Ledvina, 2005). It was originally intended for Defence purposes: to provide range for the military at any point on the earth surface at all times (Davies and Smith, 2000). Globally, territorial conflicts are being drastically reduced by peace campaigns, and this has led manufacturers of GPS to extend the scope of their production by including civil society on their production lists. Today, GPS system has found ubiquitous civilian/scientific applications, ranging from navigation solution, (including intended application for precision navigation), probing the earth's

ionosphere, and in mapping tectonic movements on a finer scale than hitherto conceivable (Kintner, 1999). GPS receivers appear as hand-held units, built into cars, mobile phones, airplanes, ship etc.

GPS technology is made up of three segments: (i) the space segment comprises of the satellite, signals and the medium of propagation. The hardware unit of the space segment consists of a constellation of twenty-four satellites that are positioned on six orbits so that each orbit accommodates four satellites. Each orbit makes a 55° angle with the equatorial plane. The radius of the satellite orbit is 26,560 km and it rotates around the earth twice in a side-real day (Hofmann-Wellenhof et al., 1997; and Tsui, 2000). GPS satellites broadcast on two dedicated carrier frequencies; L1: 1.57542 GHz and L2: 1.2276 GHz (Leitinger, 1989; and Brown et al., 1998), which are spectrum modulated (Code Division Multiple Access: CDMA) so that satellites are distinguished by their pseudorandom noise (PRN) numbers (Carrano and Groves, 2006; and Carrano, 2007), (ii) the control segment controls the satellite, including their orbits and operations, and (iii) the ground segment, basically made up of the receiver, otherwise the user.

As a radio navigation system, a GPS receiver should have a direct line of sight (LOS) and be able to receive signals from four to twelve satellites (Carrano and Groves, 2006; and Carrano, 2007) under clear view, at any time and any location on the earth. The underlying challenges on operations of GPS are: (i) How the receiver knows the location of the satellites, and (ii) how pulses are created and sent to the receiver for the onward determination of the user's position. These challenges are addressed in this study. It is important to mention that deep signal fading, caused by

scintillation can lead to loss of lock of the carrier tracking loop in the GPS receiver. GPS avionics use phase data to estimate position. If a significant number of carrier tracking signals lose lock simultaneously, navigation service delivery will be hindered until the system re-acquires a sufficient number of signals. Figure 1.1 shows how scintillation events can impose hazards to air navigation.

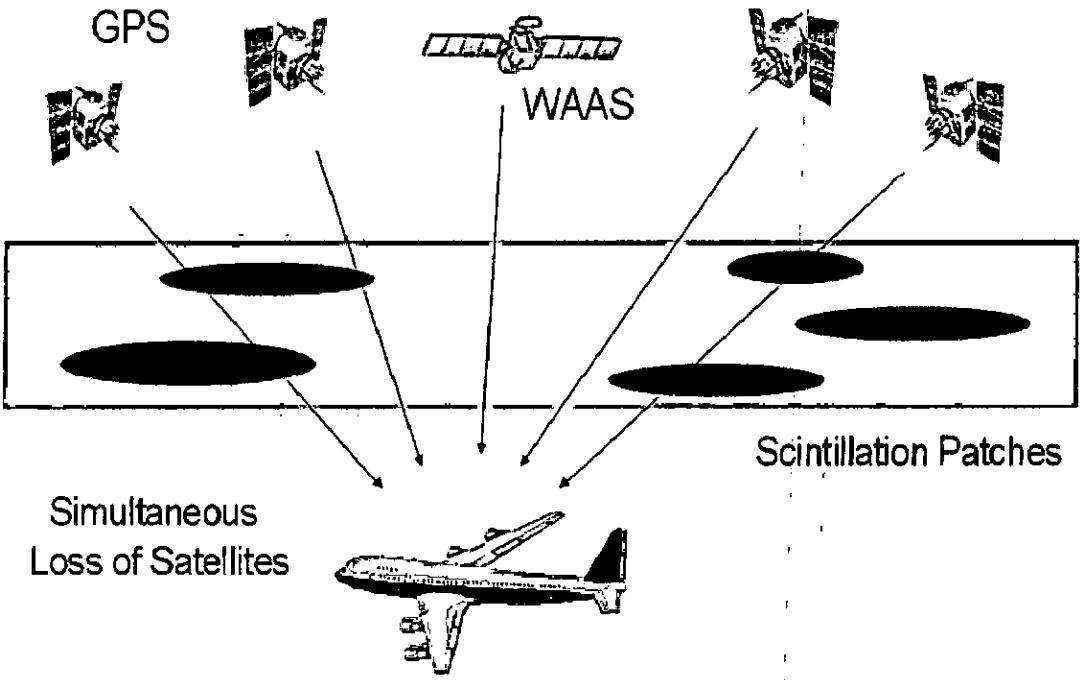


Figure 1.1: Potential Hazard to Air Navigation due to Ionospheric Scintillation

This thesis also simplifies the underlying principles of GPS operations, especially as to how it determines the position of its user. To further elucidate the computational framework, the study simulated with arbitrarily chosen input parameters, a case of eleven satellites that are being tracked by a receiver and the position of the receiver were consequently determined.

Under experimental observations, the study analyzed some ionospheric observables measured at some stations within the equatorial (Appleton) anomaly and concluded that scintillation is a post sunset event, and that it could be hazardous to GPS-based air navigation service delivery when intense and impacted simultaneously on lock system of signals from too many satellites. It must be mentioned that these stations have close ionospheric characterization with the Nigeria environment, being co-located within the Appleton's anomaly.

One of the challenges inherent in the use of GPS for aviation application is the need to ensure user integrity in the face of unbounded array of possible satellite failures (Viarrsson et al., 2001). This failure, which is often caused by scintillation results to adverse consequences on service integrity. The International Civil Aviation Organization (ICAO) Standards and Recommended Practices for Global Navigation Satellite System (GNSS) spells out that availability must be within 99.999% to 100% for all categories, and time for continuity is 10 seconds for approach with vertical guidance-I (APV-I) and 6 seconds time to alert for approach with vertical guidance-II (APV-II) (ICAO, 2000; and Van Dyke, 2001), same for CAT I precision approach (instrument guidance down to decision height of 200 ft (60m)), and just 2 seconds for CAT II/III precision approaches (instrument guidance down to decision height of 100 ft (30m)) (Viarrsson et al., 2001). In other words, GPS avionics must be able to alert pilots of impending failure within this time limit so as to execute an alternative approach, or else, the system must be ready to acquire the satellite back (continuity) before the elapse of this time.

At the end, the study applied GPS-based initiatives to develop two mechanisms of enhancing aviation safety in Nigeria. (i) a GPS-driven technique that will mitigate over-bearing effects of management of large traffic on air traffic systems and, (ii) a robust aviation search and rescue (SAR) mechanism for each state of the world, which may also be operable by two neighbouring states on bilateral understanding.

1.2 Statement of the Problem

Ionospheric irregularities have been found to impact severely on transionospheric radio signals (communication and navigation system) (Mannucci et al., 1998; and Aquino et al., 2002). Therefore, understanding the physics of ionospheric irregularities is sine-qua-non to the problem of transionospheric radio systems. Also, it has been known that a GPS receiver must view at least four satellites before correct navigation solution can be obtained. In situations where the receiver is in view of many satellites, the receiver must reduce the superfluous variables to four variables for onward determination of the user's position. In recent times, all over the world, volume of air traffic in the sky has increased tremendously, leading to over-bearing impact on existing infrastructure. Many times, controllers are left with the only option of managing traffic on first come, first served basis: stack and hold airplanes at low altitudes on arrival, or at times diverted, only to be given slots to land in turns. The consequential effects of this pattern are: fuel wastage, environmental pollution, and ineffective flight scheduling. Other worrisome incidents in the past are the situations where airplanes are declared missing in an airspace, only to be located after many days when all occupants have perished, or in some worse situations, not seen at all; even when

seen on time, mobilization and coordination of emergency resources is often very poor.

1.3 Aim and Objectives of Study

Overall Aim of Study:

The aim of this work is to employ transionospheric radio propagation in enhancing aviation safety by first of all exploring challenges that ionospheric dynamics may impose on such system, and thereafter develop mitigation strategies for optimal performance.

Objectives of Study:

1. To investigate the formation and evolution of ionospheric irregularities using non-linear dynamics model.
2. To assess the potential consequences of these irregularities on transionospheric radio signals, with particular emphasis on equatorial latitudes, and thereafter develops mitigation strategies to reduce the effects of the irregularities on systems' performances.
3. To elucidate theoretical operations of Global Positioning System (GPS) using numerical approach.
4. To design a robust safety mechanism for aviation industry in Nigeria using GPS-based technology.

1.4 Research Questions

1. How are ionospheric irregularities formed and evolved?
2. What are the consequences of these irregularities on communication and navigation (GPS) systems. Are there possibilities of mitigating the effects of

these irregularities on systems' performances?

3. What are the basic operational principles of GPS? How does GPS receiver determine the position of its user? How are pulses created and sent to the receiver for onward determination of the user's position?
4. How can an efficient and robust aviation safety (traffic management system; search and rescue (SAR) mechanism) be designed using GPS-based technology.

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 The Earth's Atmosphere

The atmosphere of a planet is the gaseous envelop that surrounds it and revolves with it. For the planet earth, it has a mass of about 5.15×10^{15} tons held together by the earth's gravitational attraction (Budden, 1985). The proportion of gases excluding water vapour is nearly uniform up to approximately 80km above the earth's surface. The major components of this region by volume are oxygen (21%), nitrogen (78%) and argon (0.93%). Small traces of other gases are also present. The atmosphere is divided into several concentric spherical strata separated by narrow transition boundaries (Kelly, 1989). The uppermost boundary at which gases disperse into space lies at an altitude of approximately 1000km above sea level.

The closest layer to the earth surface is the troposphere which also has the largest percentage of the atmospheric mass. Followed by it is the stratosphere which has an altitude between 10-50km. Next to the stratosphere is the mesosphere with an altitude that extends about 80km. After the mesosphere is the thermosphere with an altitude of 100-200km. At this extreme altitude, gas molecules are widely separated, creating lesser collision regimes. The exosphere is the outermost region of the earth's atmosphere (Davies, 1990). Within the exosphere, atoms follow ballistic trajectories and rarely undergo collisions because the density of atoms is very low. The exosphere begins at approximately 500km altitude and extends outwards until it has a transition with interplanetary space at an altitude of roughly 10000km (Budden, 1985).

The part of the earth's atmosphere that is fully ionized above about 60km where free electrons exist in numbers sufficient to influence the travel of radio waves is known as the ionosphere (Rishbeth and Garriot, 1969; Ratcliffe, 1972; and Beer, 1976). Near the ground, the air is almost unionized and its electrical conductivity is negligibly small, because the ionizing radiations have all been used up at greater heights. When any part of it is in equilibrium, its state is controlled by the earth's gravitational field so that it is a horizontally stratified system. Although, it is never in complete equilibrium, gravity has a powerful controlling effect up to about 1000km from the ground (Budden, 1985; and Thomas et al., 2001). The molecules of neutral atmosphere have an electric polarizability which means that the refractive index for radio waves is very slightly greater than unity (1.0026) near the ground. Thus, the neutral air refracts radio waves slightly.

Incoming ultraviolet radiation and x-rays from the sun ionize the atmosphere. At greater heights of the order of a thousand kilometers or more, the medium is almost fully ionized, but so tenuous that the ion and electrons concentrations are small. Jordan and Balmain (1978) described the existence of the ionosphere in form of a layer on the following basis; at great heights, the ionizing radiations are very intense, but the atmosphere is rare and there are few molecules present to be ionized. Therefore in this region, the ionization density is very low. As the height decreases, the atmospheric pressure and ionization density increase until a height is reached where the ionization density is maximum. Below this height the atmospheric pressure continues to increase, but ionization density decreases because the ionizing radiation has been absorbed or used up in the process of

ionization. The existence of layers within a region is accounted for by the fact that the atmosphere is a mixture of several gases that differ in their susceptibility to the ionizing radiations, and so produces maximum ionization at different altitudes (Beer, 1976; and Budden, 1985).

The height where maximum ionization occurs depends on the absorption coefficient of the air for the ionizing constituents of the sun's radiation and on the mechanism by which ions and electrons are removed. Different ionizing constituents give maxima at different heights. Thus, there are several different ionospheric layers of ionization. The principal ones are the D layer at altitude less than 90km, E layer at about 100km and F layer at about 300km (Rishbeth and Garriot, 1969; and Ratcliffe, 1972). Although the number and height of the layer vary in time, the E and F layers have permanent existence which makes possible radio communication by principle of reflection. The D layer is present only during the day, although, it does not normally reflect high frequency waves, its presence decreases the intensity of signals reflected from the higher layers. Other layers within the E region that do not have permanent existence are called the sporadic E layer (E_s).

2.2 Layer Formation Theory and Loss Processes

For better appreciation of the reflective properties of the ionosphere on radio waves, it is necessary to know how the rate of electron production depends on height above the ground. To study this, some assumptions must be made about how ionospheric layers are formed. In doing this, Chapman (1931) assumed that the ionosphere is isothermal, and horizontally stratified. Also the radiation falling on

the ionosphere from the sun is assumed to be monochromatic. When this radiation is absorbed in the atmosphere, it heats it, dissociates its molecules and librates electrons. The rate of production at any level is proportional to the product of the gas concentration and the intensity of the radiation. The approximate peak of production can be deduced from simple arguments as follows: suppose that the atmosphere consists of molecules with the same absorbing cross-section σ and that radiation falls on it from outside, possibly in an oblique direction. Then, in a column of cross-section drawn in the direction of incidence, and sufficient to contain N molecules, the total projected absorbing cross-section is $N\sigma$, and when this equals to unity, the radiation will in an oversimplified picture be completely absorbed (Ratcliffe, 1972).

The radiation thus descends to a level where the total number of molecules in a possible oblique column of unit area above is $N = \sigma^{-1}$: the peak of production is near that level. Chapman (1931) used the above argument to derive a profile for the rate of production which is today referred to as Chapman's law;

$$q(x,z) = q_o \exp\{1 - z - \sec\chi \exp(-z)\} \qquad (2.1)$$

The peak rate of production q_o is expressed as;

$$q_o = \frac{\eta S_\infty}{2.72H} \qquad (2.2)$$

where η is the ionization efficiency, S_∞ is the original radiation intensity from the sun, H is the scale height of the gas under consideration, and z is the normalized height expressed as;

$$z = \frac{h - h_o}{H} \qquad (2.3)$$

where h_o is the reference height, h is the height of interest and χ is the solar zenith angle.

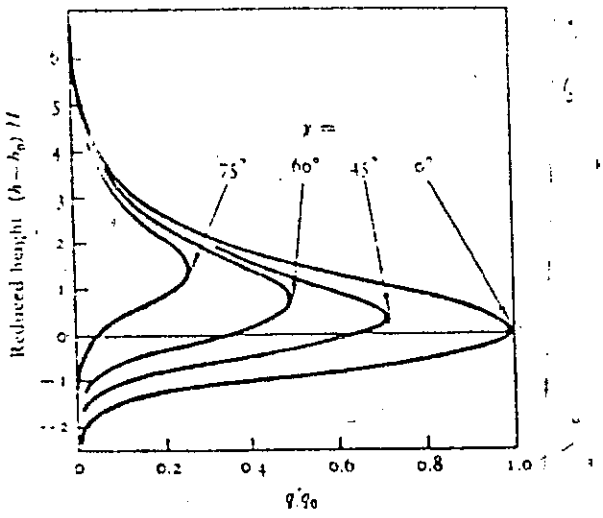


Figure 2.1: Chapman Production Curves at Different Solar Zenith Angles

Figure 2.1 shows the Chapman production curves for various zenith angles. The rate of production has reached maximum when the sun is overhead. In addition to the Chapman layer model, other models like exponential, sech-squared, parabolic and quasi-parabolic layer models have also gained prominence in radio propagation studies (Davies, 1990).

Ratcliffe (1972) explained loss processes in two broad categories; firstly, destruction within the column and then movement away from the column.

(1) Destruction within the column

(a) Radiative Recombination: Recombination of free electrons with an atomic positive ion can result to photon emission.

$$e^- + X^+ \rightarrow X + h\nu \qquad (2.4)$$

(b) Dissociative Recombination: Recombination accompanied by dissociation of a molecular positive ion.

$$e^- + XY^+ \rightarrow X + Y \qquad (2.5)$$

(c) Attachment: This is the process where an electron fuses into a neutral particle to form the negative ion of the particle.

$$e^- + Z \rightarrow Z^- \qquad (2.6)$$

At greater heights, the process of attachment to form negative ions are rapidly detached by other reactions, hence, loss by attachment can be neglected at greater heights. In lowest part of the ionosphere, loss of electron by attachment is however very important.

When electrons are lost by radiative and dessociative recombination, the rate of loss from a unit volume is given by Ratcliffe (1972) as;

$$L = -d[e]/dt = -\alpha[e][p^+] \qquad (2.7)$$

α is the recombination coefficient and p^+ is the concentration of the atomic or molecular positive ion. For a neutral ionospheric plasma, i.e. when $[e] = [p^+]$;

$$L = -d[e]/dt = -\alpha[e]^2 \qquad (2.8)$$

This is prevalent at nights whereby the predominant loss process is by dissociative recombination. The electrons produced in a unit volume by the radiation may leave the volume either by being destroyed inside the column or by moving out of it. The number destroyed in a unit volume in a unit time is normally referred to as the rate of loss (Rishbeth and Garriot, 1969). It depends on the concentration, n of the electrons and can be written as $L(n)$. If vertical velocity is

denoted by W, the rate of removal of electrons from the unit volume can be expressed as;

$$\frac{-dn}{dt} = \frac{d(nW)}{dt} = n\frac{dW}{dt} + W\frac{d(n)}{dt} \qquad (2.9)$$

The movements may be by diffusion, electric and magnetic field or by the effect of wind on surrounding gases (Rishbeth et al., 1963). The electrons that are produced continuously in the unit volume at a rate q are also removed by the loss processes and movements with the result that the concentration increases at a rate given by the continuity equation (Ratcliffe, 1972).

$$\frac{dn}{dt} = q - L(n) - \frac{d(nW)}{dt} \qquad (2.10)$$

It is obvious that inclusion of radiative recombination reduces the difference between the rates of production and loss (Rishbeth et al., 1963). If no new electrons are being produced (as may be at nights), q=0, (with the assumption that solar radiation is the only source of production). Also, if there are no movements away from the column, W=0. Then the electron concentration decays as:

$$\frac{dn}{dt} = -L(n) \qquad (2.11)$$

During the day, as the sun rises and sets, the rate of production by the solar radiation at any level increases to a maximum and then decreases: at some stages the resulting electron concentration also reaches maximum when $dn/dt = 0$.

(2) Movement away from the column

(a) Diffusion: As long as the atmosphere is homogenous, no diffusion will occur. However, if there is greater quantity of a particular gas at one location, then the pressure there will be greater and the gas will diffuse

away from the region of high concentration in an attempt to equalize the pressure (Rishbeth et al., 1963; and Shimazaki, 1964).

(b) $\mathbf{E} \times \mathbf{B}$ Drifts: An electric field generated in the dynamo region travels up the high conductivity field lines and appears undiminished in the region. Because of this, the region is permeated by horizontal electric fields generated in the lower region and these electric fields produce vertical $\mathbf{E} \times \mathbf{B}$ drift of ionization (Fesen et al., 2000; and Ray et al., 2006).

2.3 The Magnetosphere

The sun emits gases known as the solar wind which stream out towards the earth. It tends to exert pressure on the earth's magnetic field and induces current on it (Beer, 1976; Gonzalez et al., 1999; and Ostgaard, 2008). The force exerted by the earth's magnetic field on the induced current causes the solar wind to change its direction so that it avoids a particular region that surrounds the earth.

The magnetosphere is part of the ionosphere where the earth's magnetic field has a dominant control over the motions of the charged particles. It is like a blob of conducting fluid surrounding the earth, with the earth's magnetic field frozen into it. The shape of the magnetosphere is greatly influenced by the solar wind (see Figure 2.2).

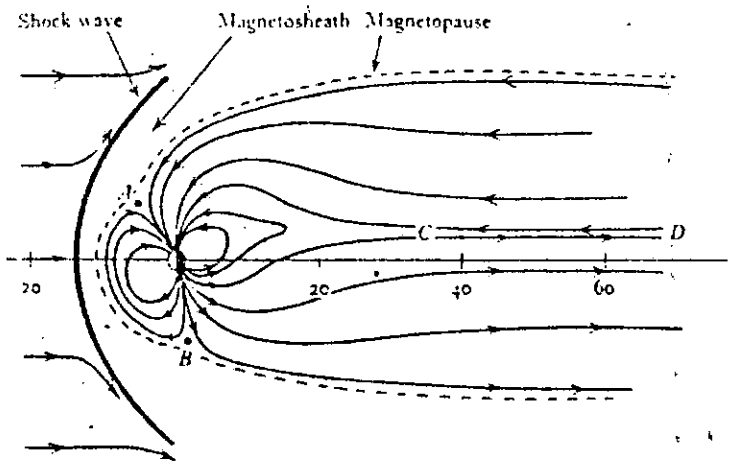


Figure 2.2: An Illustration of the main Regimes of the Magnetosphere. The Geomagnetic Field Lines are distorted by Solar Wind Effects.

The solar wind itself is fully ionized plasma, and therefore a good conductor. Thus the earth's magnetic field cannot quickly penetrate into it because it is 'frozen out'. The magnetosphere is usually distorted by the solar wind as it stretches out into a long tail known as the magnetotail in the anti-sunward direction. This extends out to a distance of about 100 earth radii ($100R_e$) (Kelly, 1989). On the sunward direction, there is a surface at about 10 to $12R_e$ from the centre of the earth where the earth's magnetic field ends abruptly but still defines an external course at which the solar wind glides. This boundary that surrounds the earth is known as the magnetopause. Beyond this region is another region of turbulence known as the magnetosheat and outside this on the sunward direction is a shock front on which the solar wind impinges. The region inside the magnetopause is the magnetosphere, which contains the compressed geomagnetic field from which the solar wind is

excluded. The plasma within this region moves with the earth and is regarded as part of the earth's upper atmosphere.

The exosphere is the uppermost region of the atmosphere where collisions between molecules are so infrequent that neutral particles move in trajectories subject only to gravity, whereas ionized particles are constrained by magnetic field.

Beer (1976) explained that any particle entering the exosphere from below must satisfy some conditions of upward velocity that is related to the gas temperature. As the upward movement progresses, a time is reached when it will either escape from the earth or re-enters the atmosphere from above. Theoretically, if a particle is to escape from the earth's gravitational field, its kinetic energy must be greater than the gravitational potential energy at the radial distance, r of the boundary under consideration.

Below the exosphere is the thermosphere. In the thermosphere, velocities of gases are distributed by Maxwellian distribution. As they approach the exosphere, those with greatest velocities will escape into the exosphere (Ratcliffe, 1972). For a given temperature, the highest atoms will have the highest velocities. It therefore implies that the velocity of hydrogen atoms is sufficiently close to the escape velocity for an important number to leave into the top of the atmosphere. The number becomes less for helium atoms but may be improved by a significant rise in temperature. The number remains negligible for oxygen and heavier gases which hardly escape.

The scale height for hydrogen atoms is greatest because of its lightness. At greatest heights, atomic hydrogen is the main constituent of the atmosphere. Even

though hydrogen and helium atoms may escape with the above principle, natural processes still help in balancing the loss to maintain equilibrium (Ratcliffe, 1972). At nights, hydrogen and helium will escape in the manner described above. The situation becomes entirely different and complicated during the day because most hydrogen and helium atoms now maintain ionic status, and as such difficult to escape even at velocities above the escape velocity. As soon as this occurs, the H^+ and He^+ ions can no longer move freely but are forced to follow the magnetic field lines which later dump them back at the exosphere in the opposite hemisphere.

In the magnetosphere, some field lines are closed while some are opened. The closed field lines have their ends terminating on the earth's surface and the opened field lines have one end at the earth and other end in the interplanetary space (Rishbeth and Garriot, 1969; and Budden, 1985). Opened field lines are found in the high latitudes on the earth's surface. At heights of $4R_e$ from the centre of the earth at the equator at night, the earth's magnetic field lines are opened.

In the high latitude polar regions, the H^+ and He^+ can still escape at night, since the anti-sunwards direction's high latitude magnetic field lines are opened. At the threshold of the escape velocity, the hydrogen or helium particle becomes ionized and follows the opened magnetic field lines to the depths of space behind the earth. The magnetic field lines behave as the carrier path for the ionic particles (Gonzalez, 1999). The resultant plasma drifts away from the high latitude region in a manner always referred to as the polar wind.

The loss of charged particles caused by the polar wind seriously depletes the high latitude exosphere of ions and electrons (Carpenter, 1966; and Ratcliffe, 1972).

This is explained in Figure 2.3. Since the ionospheric plasma is in diffusive equilibrium along geomagnetic field lines, there is a sharp gradient in the electron concentration in other parts of the ionosphere that lie on the field line that coincides with the knee (Beer, 1976). Thus, an inner region of greater electron concentration separated from an outer region of lesser concentration. The inner region is the plasmasphere and the boundary that separates it from the outer magnetosphere is known as the plasmapause.

From Figure 2.3, there appear a sharp kink in electron density profile over the equator at about $4R_e$ from the centre of the earth. This kink is known as the Carpenter's knee named after the originator of the concept (Carpenter, 1966), and it marks the location of the plasmapause. Above the plasmapause, the electron density is depleted to about $10^7 m^{-3}$ as a result of polar wind effect. Below the plasmapause is the plasmasphere, here the electron density profile is in the neighbourhood of $10^{10} m^{-3}$. The plasmapause is demarcated by the magnetic field lines that extend four earth radii above the earth's centre at the equator. This can be represented in a short notation as L=4 field lines (Elkins and Papagiannis, 1969; and Basu et al., 2001).

If L=4 field line is examined, then the plasmpause is being examined. Maps of the earth's magnetic field reveal that the L=4 field line is at altitude of 2000km at a geographic latitude of 55°N, at 1000km altitude at 57°N and at 100km altitude at 60°N (Beer, 1976). This implies that at 45° geographic latitude, the plasmapause is in the F region of the ionosphere and it is almost vertical.

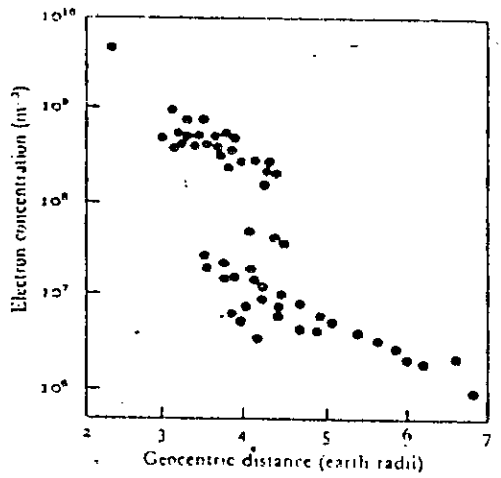


Figure 2.3: The Carpenter Knee. Electron Concentration as a Function of Geocentric Distance above the Equator.

2.4 Ionospheric Events

Solar disturbances sometimes occur accompanied by enhancement of the solar photon radiation in parts of the spectrum or by an increase in the velocity and concentration of the solar wind to such an extent that the modified portion behaves like a cloud of denser plasma (Gonzalez et al., 1999), or by the ejection of more energetic protons and electrons with such small concentrations that they behave not as plasma but as independent charge particles. These disturbances are of three types: sudden ionospheric disturbances (SIDs), ionospheric storms and the magnetic storms. They sometimes occur simultaneously and sometimes separately (Okeke, 2005).

Observations of the interplanetary magnetic field emanating from the sun have made it evident that the field either points towards the sun or away from the sun (Gonzalez et al., 1999). To an observer at the equator at noon, the field is either straight up or straight down. There are also small east-west and north-south components of the field. Normally, this field points northwards and when

occasionally points southwards, there are changes in the earth's magnetic field called magnetic storms. There are also changes in the ionosphere called ionospheric storms and these lead to increase in the number and intensity of aurorae.

Frequently, this magnetic field reversal is also accompanied by an enhancement of the solar wind in both speed (Rama Rao, 2006) and concentration (Ratcliffe, 1972) thereby making the storm effect to be more intense.

When the interplanetary magnetic field has a southward component, it manages to connect with the earth's magnetic field. It is important to mention that the solar wind plasma glides along the interplanetary magnetic field lines as it moves through space. There are now two ways in which the solar plasma can penetrate into the magnetosphere. It can slide straight down a field line or it can come in through the magnetotail. The particles coming in from the neutral point in the magnetotail are accelerated towards the earth with energies up to a few Mev. Some of these accelerated particles become fully trapped in the Val Allen radiation belts. The rest remain in a quasi-trapped state, wandering around the magnetotail until they finally escape again or dumped at the poles to produce aurora. Since these particles are more energetic, the resulting aurora appears more energetic with violent oscillations (Beer, 1976).

The energetic nature of the solar wind plasma causes it to carry the solar magnetic field along with it, and this is the reason why the interplanetary magnetic field lines which are said to be frozen to the plasma move from the left to the right. Because the solar magnetic field lines are attached to the geomagnetic field lines, they drag the geomagnetic field lines along with them producing magnetospheric

convection. Ions and electrons dwell along the geomagnetic field lines. They are therefore also made to convect, adding slightly to the polar wind. New geomagnetic field lines are then formed on the day side and are in turn convected to the nightside of the earth. At their base, these field lines are attached to the highly conducting dynamo region at 100km. As we know, a good conductor will not want to retain fields and therefore releases field orthogonally. Thus the particles moving at the base of the geomagnetic field lines set up a current in the high latitude dynamo region, that is, the polar current in upper atmosphere where collisions are unimportant.

In as much as the collision frequency is less than the gyro-frequency, the charged particles follow the moving lines of force. As a line penetrates further into the ionosphere, this condition loses its validity at a height of about 140km for ions and about 80km for electrons. There is thus a region between heights 140km and 80km where electrons are dragged round by the lines of force whereas ions are not. The currents flowing in a direction opposite to the electrons motions constitute the polar current system.

2.4.1 Magnetic Storms

In situations whereby the interplanetary magnetic field points southwards, the effect always results to changes in the earth's magnetic field known as the magnetic storms. The changes are very little in order of 2% of the original field.

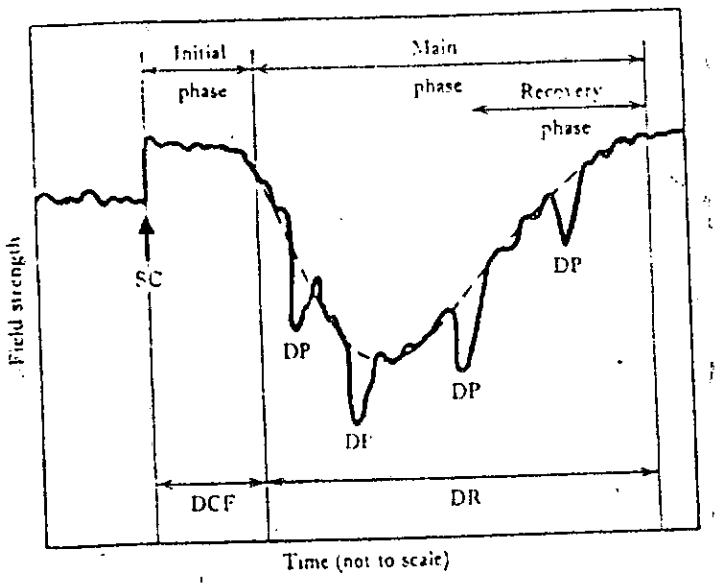


Figure 2.4: Geomagnetic Field Storm Variations (D_{st} Measurement)

The first feature of a storm on the magnetic records is a sudden rise in the magnetic field strength by about 20nT (Okeke, 2005). The abrupt rise in the strength of the magnetic field at initial time zero is known as the sudden commencement of the storm. It is the consequence of the impact of the arrival of an enhanced solar wind at the magnetopause (Ostgaard, 2008). This enhances the compression of the magnetic field lines already caused by the normal solar wind. The compression exerts magnetic pressure that increases the magnetic field strength (Matsushita, 1963). Figure 2.4 shows a schematic diagram of geomagnetic storm variations (not drawn to scale).

The sudden arrival of the enhanced solar wind has the form of a shock wave; this is followed by a more-or-less steady enhanced solar wind that results in a continuing compression of the magnetic field lines (Dandekar and Groves, 2004).

The corresponding increase of the field at the surface of the earth is called the initial phase of the storm.

The onset of a storm may not be necessarily be marked by an enhanced solar wind, though in most cases it is. Occasionally there may not be a sudden commencement, and then the magnetic field variation follows the dotted part of Figure 2.4, which is due to the enhanced ring current. Six to ten hours after the beginning of a magnetic storm, there is an increase in the number of energetic particles in the Van Allen radiation belts delivered therein by the solar wind particles through the magnetotail. As the number of these energetic particles increase in their trapped motions around the earth, electrons move towards the east, and ions towards the west. The consequence is an increase in the ring current (Gonzalez, 1999). The westward ring current produces a southward magnetic field. It therefore implies that an increased ring current produces a decrease in the magnetic field at the ground. This decrease in the magnetic field strength is called the main phase of the magnetic storm (Beer, 1976; and Dandekar and Groves, 2004). After new particles cease to be injected into the Val Allen radiation belts, the ring current gradually decays away for one or two days, and the magnetic field gradually returns to its normal value. This is the recovery phase of the storm. It will be important to note that if there is an enhanced solar wind when the interplanetary magnetic field is not having a southward component, the main phase will be missing. There will only be an initial phase for the storm and nothing more. In other words, if the interplanetary magnetic field has a northward component, calm ionosphere is experienced.

The main phase of a magnetic storm is often interspersed with magnetic bays which are rather abrupt disturbances lasting for an hour or two. Magnetic bays are related to the magnetospheric field lines connection with the interplanetary magnetic lines and the subsequent magnetospheric convections. The movement of electrons in the high latitude dynamo region following these field lines produces the polar current system (Mazaudier, 2008). The number of electrons in the polar dynamo region fluctuates depending on the number of particles from the solar wind sliding down the magnetic field lines. These fluctuations show themselves as bays in Figure 2.4.

2.4.2 Sudden Ionospheric Disturbances (SID)

Ionospheric effects resulting from the solar activity show up as soon as the enhanced radiation from the sun reaches the earth. The onset of a sudden ionospheric disturbance (SID) is the first warning that a severe magnetic storm is brewing. The increase in the ultraviolet radiation during a storm increases the ionospheric temperature which controls most of the chemistry of the ionosphere.

Sudden ionospheric disturbances are usually caused by a sudden increase in the number of solar flares. This abrupt increase in solar activity increases the amount of energetic radiation reaching the ionosphere. This in turn increases the D region ionization and consequently the radio wave absorption. A burst of radiation of this kind would ionize only on the day-side of the ionosphere.

During sudden ionospheric disturbances, absorption of medium- and high-frequency radio waves is often great enough to result in complete removal of a

signal reflected from the E or F regions, and it is then said that there is a short wave fadeout (Beer, 1976).

Corona mass ejection (CME) is accompanied with enhanced visible H-alpha lines over a small region usually near a complex group of sunspots known as solar flares (Gonzalez, 1999). When this occurs the solar X-radiation with wavelength less than about 1nm increases in strength so as to produce increased ionization in the D region at heights around 80km. The increase in electron concentration leads to several phenomena grouped together under the name sudden ionospheric disturbances (SIDs). Their effects are most intense when the sun is overhead. It is interesting to note that, although the H-alpha line and short wavelength X-radiation are considerably enhanced during a solar flare. The strength of the Lyman-alpha line in the ultraviolet region usually changes very little. In some occasions, they lead to increase in electron concentrations of the E and F regions.

The electron content of the D-region increases significantly within a few minutes and afterwards slowly recovers to its normal value within a time of order of three-quarter to one and a half hours. Ratcliffe (1972) explained that the intensities of low and very low frequency waves change during sudden ionospheric disturbances in a way that depends on the frequency and on the angle of reflection. For instance, waves of frequency near 27kHz reflected obliquely from the ionosphere are increased in strength. The increase is clearly noticeable when atmospherics originates at a distant flash light. For waves of very low frequency reflected from the ionosphere, the most clearly marked effect of sudden ionospheric

disturbances is a sudden change of phase usually referred to as a sudden phase anomaly (SPA).

2.5 Atmospheric Dynamo and Motor

The extent of generated conductivity of the ionospheric plasma as caused by drift of charges depends fundamentally on whether the collision frequencies (ν_i) of the plasma components (electrons and ions) are greater or less than their angular gyro-frequencies (Ω). As height decreases through the middle ionosphere, from 300km to 70km, the collision frequencies of electrons increase from about $10^3 s^{-1}$ to $10^7 s^{-1}$ and those of ions from about $0.5 s^{-1}$ to $10^5 s^{-1}$, whereas the angular gyro-frequencies remain approximately constant, about $8 \times 10^6 s^{-1}$ for electrons and $160 s^{-1}$ for ions (Ratcliffe, 1972). Within this height range, the behaviour of each type of particle changes radically, with the result that a layer of relatively good direct current conductivity exists at heights around 110km (the E-layer). This conductivity layer is sufficient to behave as the armature of an atmospheric dynamo to produce observable angular changes in geomagnetism that are associated with the solar and lunar day.

The sun and the moon are responsible for dictating the manner in which the conducting layer of the ionosphere moves across the geomagnetic field. The sun contributes to this by heating the ozone layer, while the moon acts through its gravitational attraction. The heat input from the sun has a period of one solar day, and a strong harmonic component with period equal to half a day. The atmosphere is distributed in such a way that over most of the earth, it responds to this component to produce horizontal air movements with a period of twelve hours; they

are called solar semi-diurnal tide. A tide is also driven by the gravitational attraction of the moon; like tides in the oceans, it has a period of half a lunar day.

Like lunar tide, the atmospheric solar tide were the result of gravitational attraction, it might be expected to be smaller, because of the smaller gravitational force; the situation would then be as it is in the oceans where the lunar tide is more important. In the atmosphere, the thermal origin of the solar tide, coupled with a quasi-resonance in the atmosphere cause the tide to be greater than the lunar one.

The drift of the conducting layer across the permanent magnetic field of the earth induces electromotive forces whose components, in the plane of the layer, have directions perpendicular to the horizontal component of the tidal velocity and magnitudes proportional to the product of that component and the vertical component of the geomagnetic field.

Ratcliffe (1972) gives conceptualized explanations of electrodynamics processes that are involved in creating atmospheric dynamo, using a simple model of "current flow mechanism".

Ionospheric dynamo is usually discussed by assuming a simple model of a conductivity layer that is in form of an isolated slab that is bounded above and below by horizontal planes, and that only the horizontal components of the fields were responsible for driving the horizontal currents. In a model of this kind, vertical components of induced fields (arising from horizontal movements across the horizontal component of the earth's field) produce space charges on the upper and lower boundaries (Ratcliffe, 1972).

In discussing the effect of these space charges, it will be necessary to consider the way in which they modify the tensor conductivity of the ionospheric slab. The effective conductivity in the horizontal plane becomes comparatively large near the geomagnetic equator, where the magnetic field is horizontal (Kelly, 1989). The current in the atmospheric dynamo becomes correspondingly great around this region (Mazaudier, 2008). This large current is referred to as the ***Equatorial Electrojet (EEJ).*** The conductivity at the equator and elsewhere has a peak magnitude at a height near 110km. Experiments with rockets have confirmed that the currents of the dynamo flow near this height.

Having known that the atmospheric dynamo is supposed to be located at a height near 110km by the simplified model above, it implies that the dynamo must be bounded above and below, which leads to conceptualization of the atmospheric motor at greater heights and in the F-region of the ionosphere so that it can be driven by the dynamo below (Ratcliffe, 1972). It is supposed that the conductivity of that part of the ionosphere above the dynamo is much greater along the direction of the magnetic field than across it, so that electrostatic fields that are developed in the dynamo region are transferred along slopping lines of force to the F-region at heights around 250km (i.e. electrodynamics coupling process). These fields, acting in conjunction with the geomagnetic field, can produce movements of the F-region. Since the collision frequencies of ions and electrons are much less than the corresponding gyro-frequencies at these heights, the electric field '**E**' moves the ions and electrons together, in a direction perpendicular to the magnetic field '**B**'.

2.6 Currents in the Ionosphere

The earth's magnetic field approximates to that of a dipole at its centre (i.e. the centre of the earth) with strength and direction that change slowly. These slow changes are accompanied by more rapidly changing currents in the ionosphere at small or great (magnetospheric) heights, or by hydromagnetic waves. They are in turn produced by mechanical forces or electric fields, and their magnitude depends upon conductivities of the ionized medium in which they flow. These currents are:

(i) Currents near a height of 110km and are driven by movement of the neutral air (the atmospheric dynamo).

(ii) Currents near a height of 110km and in high latitude regions, driven by forces originating high in the ionosphere (the polar current system).

(iii) Currents in the ionosphere (magnetosphere) at a geocentric distance of about 4 earth radii (the ring current).

(iv) Currents in the magnetopause; resulting in a compression of the geomagnetic field

The magnitude of the ring current, or of the compression of the field changes only if the energy of the trapped particles, or the strength of the solar wind changes. The associated magnetic fields do not change much except during storms. Even in quite non-storm times, however, these constant fields may be asymmetrically distributed so that to an observer rotating with the earth, they (fields) appear to vary during 24 hours (Ratcliffe, 1972).

2.6.1 Ionospheric conductivities

In principle, specific conductivities are obtained by multiplying the conductivities per ion pair by the electron concentration which increases with height, and by making proper allowance for height-variations of temperature and the corresponding variations of collision frequency, ν.

When an electric field is applied normal to a magnetic field, σ_1 corresponds to the current that flows in the direction of the electric field; it is called the *Pedersen conductivity*. σ_2 corresponds to the current that flows in a direction normal to both the electric and the magnetic fields; it is called the *Hall conductivity*. In the ionosphere this conductivity is due to drift motion of the electron $(\mathbf{E}\times\mathbf{B})$ and maximum in the E layer where only electron practically drifts to the direction of $\mathbf{E}\times\mathbf{B}$. When an electric field is applied along the direction of a magnetic field, σ_0 corresponds to the conductivity that is produced in the absence of a magnetic field, i.e. the current that flows along the direction of the magnetic field, and it is known as parallel conductivity.

If an electric field with components E_x, E_y, E_z is applied and the magnetic field is along z direction. From the three conductivities, the tensor conductivity can be expressed as:

$$\begin{vmatrix} J_x \\ J_y \\ J_z \end{vmatrix} = \begin{vmatrix} E_x \\ E_y \\ E_z \end{vmatrix} \times \begin{vmatrix} \sigma_1 & \sigma_2 & 0 \\ \sigma_2 & \sigma_1 & 0 \\ 0 & 0 & \sigma_0 \end{vmatrix} \qquad (2.12)$$

The theory above actually discusses what occurs in infinitely unbounded plasma. The ionosphere is, however bounded above and below; therefore, the situation becomes different. To examine it, it is assumed that the plasma previously

discussed is bounded by two parallel planes perpendicular to a known point, and that a field E_x is applied. A current 'J_y' then flows and builds up charges on the boundary planes, until they produce an additional field E_y, just sufficient to prevent any further generation of current J_y. From eq. (2.12), the situation can be represented as:

$$J_x = \sigma_1 E_x + \sigma_2 E_y$$

$$J_y = -\sigma_2 E_x + \sigma_1 E_y = 0 \qquad (2.13)$$

When the field E_y is produced, generation of current j_y stops, at this point, j_y vanishes. From equation (2.13),

$$J_x = \left(\sigma_1 + {}^{\sigma_2^2}\!/_{\sigma_1}\right) E_x = \sigma_3 E_x \qquad (2.14)$$

The conductivity σ_3 describes the current that flows along the direction of the electric field when current in the perpendicular direction is inhibited. This conductivity is sometimes called the *Cowling conductivity*.

If only one kind of charged particle were free to move, the magnitude of σ_1 and σ_2 can be used to show that $\sigma_3 = \sigma_0$. It therefore follows that in a bounded medium containing only one kind of moving charge, currents flow just as they would if there were no magnetic field. This is what happens when current flows in a bounded piece of metal. Although the superposition of a magnetic field does not alter the conductivity, instead it produces an electrostatic field (E_y) perpendicular to the boundaries. This is the *Hall Effect*.

The situation is however, different in the ionosphere, where both ions and electrons can move; σ_1 and σ_2 are then, each the sum of ion and electron conductivities and σ_3 is no longer equal to σ_0.

An enhanced conductivity of this kind near the geomagnetic equator in a slab-like ionosphere is supposed to be responsible for a relatively large current in the atmospheric dynamo, known as the Equatorial Electrojet (Mazaudier, 2008).

2.7 Ionospheric Irregularities

One important research objective of the Space Weather Program is to understand the structure and formation of ionospheric plasma density irregularities that cause scintillation to transionospheric radio links. Since scintillations originate from random electron density irregularities acting as wave scatterers, understanding the formation and evolution of irregularities means understanding scintillations (Wernik et al., 2003).

Over the past four decades, research interest on the study of ionospheric irregularities and how they impact on radio waves has been on the increase (Calvert and Cohen, 1961). Many reviews on the study have been published (Woodman and La Hoz, 1976; Aarons, 1982; Yeh and Liu, 1982; Basu et al., 2001; and Thomas et al., 2001). Also, comprehensive reviews on the physics of ionospheric irregularities have been provided by Keskimen and Ossakow (1983), Kelly (1989), and Wernik et al. (2003). Previous studies on scintillation theories (Yeh et al., 1975; Groves et al., 1997; and Wernik et al., 2003) have related observed signal statistics to the statistics of ionospheric density irregularities.

The general condition for instability is that the $\mathbf{E} \times \mathbf{B}$ direction must be parallel to the plasma density gradient. At the equator, the zonal component of the electric field often increases to a large eastward value just after sunset, lifting the F-layer to very high altitudes via Rayleigh Taylor instability mechanism (Kelly, 1989; and Rama Rao et al., 2006). During early hours, the E region ionization is usually quickly eroded due to the recombination of electrons and ions in the absence of production by solar radiations. Therefore, soon after the local sunset, the F-region ambient ionization becomes relatively high with sharp gradients that are anti-parallel to gravity at the bottom side of the F layer (Woodman and La Hoz, 1976; and Kelly, 1989).

As the plasma-depleted bubbles rise from the bottom side of the F layer, they get elongated along the magnetic field lines to off equatorial latitudes (Aarons et al., 1980), which thereafter causes enhancement in the anomaly gradient. This process re-energizes equatorial fountain by depleting the ionization near the equator and simultaneously increasing the ionization at off equatorial regions (Rama Rao et al., 2006). In polar latitudes, aurora activities coupled with the precipitations of field aligned particles are the sources of irregularities generations while in mid-latitudes, overflow of events from equatorial and polar latitudes are responsible (Brown et al., 1998).

Another important factor for the generation of irregularities is the occurrence of geomagnetic storms. During these events, high ionospheric irregularities gradients cause fast variations in total electron content (TEC) of the ionosphere (Brown et al.,

1998; and Aquino et al., 2003), which impact severely on satellite communication and navigation systems.

2.7.1 Equatorial (Appleton) Anomaly

From the view point of electron density variations, the equatorial region around the dip magnetic equator displays a complex pattern. During the day, an increase in maximum electron density occurs away from the equator.

In principle, one would have expected greatest electron concentration and hence, highest critical frequency at the equator, being the place where the sun's intensity is highest but contrary is the case. This departure from what could have been obtained by principle is referred to as anomaly.

The existence of the equatorial anomaly depends on the fact that magnetic field lines are good conductors. This has two consequences: firstly, electrons can easily diffuse along the magnetic field lines. This implies that the electrons loss processes will also be very high along the magnetic field lines. Since the greatest production of electrons occurs at the equator, the greatest diffusion also occurs away from the equator (Beer, 1976).

Secondly, an electric field generated in the dynamo region travels up the highly conducting magnetic field lines and appears undiminished in the F-region. As a result of this, the F-region is permeated by horizontal electric fields generated in the E-region and these electric fields produce vertical $\mathbf{E}\times\mathbf{B}$ drifts of ionization. Since E-region electric fields are westward at low latitudes during the day, the resulting drift will be upward (Rama Rao et al., 2006).

These two effects; diffusion and $\mathbf{E}\times\mathbf{B}$ drifts produce an overall motion of the ionization that is upwards and away from the magnetic equator. This mechanism is analogous to the working principles of a fountain, so that this phenomenon of equatorial anomaly is referred to as "fountain effect" (see Figure 2.5). The presence of a meridional (north-south) wind produces an asymmetrical distribution for the anomaly.

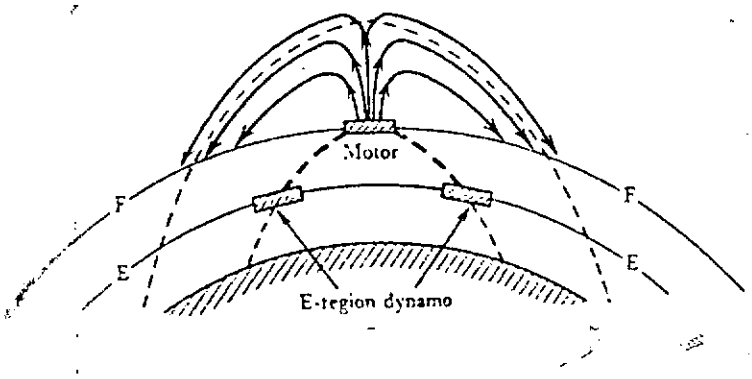


Figure 2.5: Fountain Effect. Electric Fields Produced in the Dynamo region travel along the Earth's Magnetic Field Lines, which are equipotentials to the F Region. Then a Combination of $\mathbf{E}\times\mathbf{B}$ Drifts and Diffusion Enhances Ionization Concentration on Each Side of the Equator but a Decrease at the Equator Itself (Adapted from Ratcliffe (1972)).

The electric fields produced in the dynamo region travel along the earth's magnetic field lines, which are equipotentials, to the F-region. Then a combination of $\mathbf{E}\times\mathbf{B}$ drifts and diffusion produces a decreased concentration of ionization at the equator and enhanced concentrations also on each side of the equator through electrodynamics processes. In general, atmospheric wave motions couple into plasma dynamics through electrodynamics processes. This generates large scale electric fields that cause plasma drifts forming the equatorial (Appleton) anomaly ($\approx \pm 15^{\circ}$ away from the geomagnetic equator) (Aarons, 1982). The magnitude of the

generated current depends upon the direction of the geomagnetic flux **B** at the equator, the geomagnetic flux is horizontal to the magnetic equator and the current is greatest in this region. This current is known as the equatorial electrojet. In the ionosphere, near the magnetic equator, the geomagnetic field is of the order of 0.3x 10^{-4} Wb/m^2 with gyrofrequency of almost 0.84MHz (Oyinloye, 1987).

2.7.2 Equatorial Sporadic E (E_s)

The most important type of ionospheric anomaly at E-region heights is sporadic-E, conventionally written in short form as E_s. It is a dense cloud of drifting ionization which randomly and unpredictably appears within the E-region of the ionosphere. As discussed earlier, the conductivity of the ionosphere reaches its peak value within the E-region in the equatorial regions. Here the strength of the equatorial electrojet is greatest (Ray et al., 2006). It implies that the strength of the equatorial electrojet at any particular time has a very strong control on the irregularities density within the E-region of the ionosphere than normal, and it is therefore sufficient to return electromagnetic waves that would normally have passed straight through the ionosphere to the earth. Occurrence of sporadic-E plays severe havoc to radio communications.

Sporadic-E is believed to be caused by combinations of several ionospheric mechanisms amongst which is the shear in the horizontal wind. If there is an eastward wind at one height and a westward wind at another height, then the charged particle will converge into a thin layer. The theory behinds this, is that, the neutral particles of the wind collide with the charged particles and set them moving. However, the charged particles start circling around the almost horizontal magnetic

field line. In the E-region, large number of neutral-charged particle collisions takes place, such that charged particles are not even able to complete one half revolutions before they are pushed on by the neutral wind (Beer, 1976). One set of particles move downwards at a slant and another set moves upwards at a slant. This forms the wind shear mechanism of sporadic-E. Good knowledge and predictions of wind shears will serve as efficient instrument for practical predictions of sporadic-E formation for enhancement of radio systems performances.

In practice, a really intense day time sporadic-E patch can have a critical frequency, f_xE_s which is greater than the maximum F2-region critical frequency, f_xF_2. But the sporadic-E layer usually occur at a height range of about 100-120km, and like the normal E-region, it produces a straight line echo on ionogram traces (MaClure, 1964).

For intense sporadic-E during periods of high solar activity, the maximum usable frequency for a 200km transmission path can exceed 100MHz. Radio taxis use frequencies in this range, so that at high solar activity when sporadic-E occurs, microwaves and television receptions are efficiently transmitted without disruptions (Beer, 1976). Close to the equator, sporadic-E is essentially a day time phenomenon with little seasonal variation, while in the aurora zone; it is most prevalent during the night hours, though, still dependent on little seasonal variation (Jordan and Balmain, 1978). In Li et al. (2006) investigation, sporadic E has been suggested to be responsible for ionospheric Pedersen conductivity enhancement.

2.7.3 Equatorial Spread F (ESF)

The term equatorial spread F (ESF) encompasses several turbulent processes and different types of plasma instabilities that grow and propagate in the night-time equatorial ionosphere (Valladares et al., 2004). Jacamarca radar has been used to describe the altitude distribution, time evolution, scale size, vertical and zonal drifts (Valladares et al., 2001; and 2004), and the magnetic aspect angle dependence of 3m irregularities associated with ESF (Calvert and Cohen, 1961; and Woodman and La Hoz, 1976). These authors suggested that the time evolution of the radar plumes followed the evolution of the Rayleigh-Taylor instability (RTI). The study undertaken by MaClure et al. (1977) confirmed that the deep plasma depletions that are seen at F region altitudes are actually formed at the bottom side, with the tendency to propagate towards higher altitudes.

The equatorial and low latitude ionosphere is the region where the most significant post sunset ionospheric dynamics play a great role in the generation of ESF irregularities causing scintillations on transionospheric radio signals (Valladares and Doherty, 2009). The distortions caused on such signals can degrade system performance. In navigation system, it may lead to errors in received messages (Seo et al., 2007).

GPS-based ionospheric measurements (scintillations, total electron content (TEC), range delays) are the most simple and efficient way of characterizing ESF irregularities (Rama Rao et al., 2006). Authors like Rishbeth et al. (1963), Kelly (1989) and Valladares et al. (2001) explained that some conditions are responsible for the generation of ESF irregularities. These are: (i) a sharp gradient at the bottom

side of the F layer, anti-parallel to gravity, (ii) the pre-reversal enhancement (PRE) in the upwards $\mathbf{E} \times \mathbf{B}$ drift and associated uplifting of the F layer, (iii) simultaneous decay of the E region conductivity at both ends of the field line, (iv) transequatorial component of the thermospheric winds and the associated symmetry in equatorial ionization anomaly (equatorial plasma electrodynamics) (Valladares et al., 2001; and Rama Rao et al., 2006) and (v) penetration of magnetospheric winds.

Very commonly observed anomaly in the F-region is the ESF. Lawrence (1958) attributed the phenomenon of ESF to the increased drifting velocity of ionospheric wind. It was stated categorically that this velocity pertains to the F-region because scintillations are caused by ESF (Booker, 1958). Though theoretical considerations of instability mechanisms and through radar backscatter and rocket and satellite *in-situ* measurements, it has been established that night time ionospheric equatorial anomaly region emerging after sunset develop from bottom side instabilities, probably of the Rayleigh-Taylor type. The depleted density bubble rises into the region above the peak of the F2 layer (Woodman and LaHoz, 1976; Yeh and Liu, 1982; and Aarons, 1982). Steep gradients on the edges of the hole help to generate the smaller scales irregularities within the patch which produces intense scintillation effect.

Spread F traces are broadened and diffused for heights corresponding to the ionospheric F region on an ionogram. Irregularities responsible for this are commonly referred to as spread F irregularities (Yeh and Liu, 1982). Spread F is usually a night time phenomenon. It is believed that electrodynamic processes of electrojet are coupled into the F region from its dwelling E region to produce spread

F. Thus, equatorial electrojet has influence on the formation of F region irregularities.

In the past, several classifications for types of spread F have been suggested by early workers, such as high-, medium-, and low-latitude spread F but this method is not entirely satisfactory. An alternative division into "range type" and "frequency type" has been suggested (Jordan and Balmain, 1978). Typically, equatorial spread F is marked by a diffusive spreading of the echo traces at lower frequencies called the "range spreading." It only occurs at night, and most spread F occurs just after sunset. At middle and high latitudes, the width of the spread F trace increases with frequency on the ionogram and is a maximum near critical frequency. This non-equatorial type of spread F is called "frequency spreading."

Many experimental techniques have been used in the past to study spread-F irregularities. These techniques have been broadly grouped into two: remote sensing techniques and *in-situ* measurements. Most remote sensing techniques utilize radio wave and can be classified according to whether the radio waves are reflected from, scattered from, or penetrated through the ionosphere. In a low power operation, radio waves are normally reflected from the ionosphere in experiment such as vertical ionosonde, back scatter ionosonde and forward scatter ionosonde. Such experiment is useful in detecting the existence of spread-F irregularities (Valladares and Doherty, 2009).

As radio frequency is increased beyond some value, the radio wave begins to penetrate the ionosphere and almost all its electromagnetic energy escapes into outer space. Nevertheless, there is a very small amount of its energy that is scattered

back. Under quiescent conditions, the back scattering is caused by ionospheric plasma fluctuations under thermal agitations. For sufficiently power radars, the scattered signal may be strong enough to provide us with useful information. Radars operating on this principle are known as incoherent scatter radars. It must be understood that, such radars can only sense the irregularities in a very narrow spectral window. On occasion during the presence of spread-F irregularities, the radar returns have been observed (Yeh and Liu, 1982) to increase in power by 80dB in a matter of few minutes. This means that in few minutes, the irregularity spectral intensity can increase by as much as 10^8 folds. This suggests the high dynamic nature of the phenomenon of spread-F.

2.8 Overview of Global Positioning System (GPS)

One of the emerging applications of transionospheric radio propagation is the Global Positioning System (GPS). Historically, GPS technology started as far back as 1973 when the U.S Department of Defence directed the Joint Program Office (JPO) located at the U.S Air Force System Command's Space Division, Los Angeles to establish, develop, test, acquire and deploy a space borne positioning system, but was declared fully operational by the Federal Aviation Administration (FAA) of USA in 1994, after successful inauguration of reliable constellation (Tsui, 2000; Kintner et al., 2001; and Kintner and Ledvina, 2005). This project culminated in Navigation System with Timing and Ranging (NAVSTAR) Global Positioning System (GPS). It was originally intended for Defence purposes: to provide range for the military at any point on the earth surface at all times (Davies and Smith, 2000).

GPS technology is made up of three segments; (i) the space segment which comprises the satellite, signals and the medium of propagation (ii) the control segment which controls the operation of the satellite, and (iii) the user segment which is the receiver. The receiver appears as hand-held unit, built into cars, mobile phones, airplanes, ship etc (Hofmann-Wellenhof et al., 1997; and Kintner and Ledvina, 2005).

A GPS constellation comprises of twenty-four satellites that are positioned on six orbits so that each orbit accommodates four satellites. Each orbit makes a 55^{o} angle with the equator. The radius of the satellite orbit is 26,560 km and it rotates around the earth twice in a side-real day (Hofmman-Wellenhof et al., 1997; and Tsui, 2000). In other words, each satellite has a revolution time of 11hr58min. The implication of this is that the satellite passes each point on the earth 2 min earlier than normal time, and every day, 4 min earlier for the two daily rovolutions.

GPS operates on the principle of spread spectrum technology using direct sequence code division multiple access (DS-CDMA). GPS satellites broadcast on two dedicated carrier frequencies; L1 (1.57542 GHz) and L2 (1.2276 GHz) (Leitinger, 1989; Brown et al., 1998; and Zernov, 2006). Each carrier frequency is modulated with one or more pseudorandom codes. There are basically two types of codes vis-à-vis; the coarse (clear)/acquisition (C/A) and the precision (P) codes. The actual P code is not directly transmitted by the satellite, but it is modified by a Y code to obtain P(Y) code (Kintner and Ledvina, 2005). The P(Y) code is dedicated for military use only.

The C/A codes are thirty-two in number with each having 1023 bits so that the code repeats every 1ms. The receivers contain replicas of each of the 32 C/A codes. For a receiver to receive signal from a specific satellite, the entire L1 bandwidth of the C/A code is converted to base-band, digitized and then sent to a correlator channel (Van Dierendonck et al., 1993; and Van Dierendonck, 2005). The correlator compares the bit streams with a local C/A code replica of the receiver via cross-correlation. If the C/A code replica synchronizes with the phase and repetition rate of an incoming satellite C/A code, the correlator yields peak value and the receiver is tracked by the satellite.

2.8.1 Control Segment

The operational control system (OCS) consists of a master control station and three ground control stations. The main operational tasks of the OCS are: tracking of the satellites for the orbit and clock determination and prediction, time synchronization of the satellites, and upload of the data message to the satellites. It also gives Selectivity Availability (SA) to the broadcast signals (Tsui, 2000). Aside the operational responsibility, the OCS also performs other responsibilities such as procurement and launch activities.

2.8.1.1 Master Control Station

The location of the master control station was formally at Vandenberg Air Force Base, California, but has been moved to the consolidated space operations centre (CSOC) at Falcon Air Force Base, Colorado Springs, Colorado. CSOC collects the tracking data from the monitor stations and calculates the satellite orbit and clock parameters using Kalman estimation (Hofmman-Wellenhof et al., 1997). These

results are then passed to one of the three ground control stations for eventual upload to the satellites. The satellites control and system operations is also the responsibility of the master control station.

2.8.1.2 Monitor Station

There are five monitor stations located at: Hawaii, Colorado Springs, Ascension Island in the South Atlantic Ocean, Diego Garcia in the Indian Ocean, and Kwajalein in the North Pacific Ocean (Hofmman-Wellenhof et al., 1997). Each of these stations is equipped with a precise cesium time standard and receivers which continuously measure pseudoranges to all satellites in view. Pseudoranges are measured every 1.5 seconds and using the ionospheric and meteorological data. They are smoothed to produce 15 minutes interval data which are transmitted to the master control station. The above tracking network is designed to determine the broadcast ephemerides and to model satellite clocks.

2.8.1.3 Ground Control Station

These stations collocated with the monitor stations at Ascension Island in the South Atlantic Ocean, Diego Garcia, and Kwajalein (Hofmman-Wellenhof et al., 1997) are the communication links for the satellites and mainly consists of the ground antennas. The satellites ephemerides and clock information calculated at the master control station and received via communication links are uploaded to each GPS satellite via S-band radio link on daily basis.

2.8.2 User Segment

The intended purpose of GPS by the U.S Department of Defence is for military activities (Davies and Smith, 2000), by incorporating GPS into virtually every

major defence system. It was envisioned that every military aircraft, ship, land vehicle and even groups of infantry would have an appropriate GPS receiver to coordinate military activities.

The civilian applications came on board several years ahead of schedule in a manner not envisioned by system's planners. The primary focus in the first few years of the system's development was on navigation. Today, GPS receivers have proliferated in civil environments. They are routinely used to conduct all types of land and geodetic control surveys. Other major areas of use are for fleet management and control. In the nearest future, GPS receivers would be mounted on emergency vehicles with a central control system for easy coordination of resources, similar technology is also in the pipeline for Air Traffic Control management whereby aircrafts can be laterally flown very close together (details are provided in chapters three and four). The use of GPS receiver in space research has also grown substantially in recent times (Basu 1978; Basu and Kelly, 1979; Valladares et al., 2001; and 2004).

2.9 Total Electron Content (TEC)

Total electron content is defined as the line integral of electron density along the radio line of sight (LOS) for any satellite-receiver pair (Fremouw et al., 1978). Using dual frequency application, GPS receivers are capable of measuring slant TEC values of the ionospheric path where the signal travels. For practical applications, the slant TEC determined must be approximated to equivalent vertical TEC values (Basu et al., 2001), which forms the centre of our discussion under this section. The dispersive nature of the ionosphere made the estimation of TEC possible via dual frequency

technology (Leitinger, 1989). TEC data are measured by using the carrier phase and group delay measurements of GPS signals at L1 and L2 reported at 30 seconds intervals.

The two frequencies suffer different delays as the signals traverse the ionosphere. These delays show up as a difference in the pseudorange and phase information both of which are determined by receiver with dual frequency technology (Dandeker and Groves, 2004). Carrano and Groves (2006) explained that the ionosphere being weakly ionized plasma imparts group delay and carrier phase advance on traveling radio waves through the ionosphere.

With investigations carried out by several authors (Fremouw et al., 1978; Bürrell et al., 2006; and Carrano and Groves, 2006), the extent of group delay and carrier phase advance depends on the satellite elevation angle at the receiver site. Observations at lower angles pass through longer path in the ionosphere than those at high elevation angles. Normally, the TEC values are approximated into equivalent vertical TEC values by assuming a thin shell ionosphere between height 350-400km (Li et al., 2006; Carrano, 2007). The standard practice is to define a minimum elevation cut-off of 10° during TEC processing, so that multipath induced noisy measurements can be avoided. The group delay is expressed as;

$$\Delta t = 40.3 * \int_R^S Nds / cf^2 \qquad (2.15)$$

where $\int_R^S Nds$ is the total electron content (TEC), f is the frequency of the signal in Hertz and c is the speed of light in m/s (Leitinger, 1989).

2.9.1 Two Coherence Frequency Delay Correction.

Two frequencies that are widely spaced can be used for ionospheric delay corrections for navigation and ranging systems. This method is usually adopted when there is need to reduce effects of ionospheric time delay error to a practicable level (Leitinger, 1989). For the two chosen frequencies f1 and f2, the two line delay errors Δt_1 and Δt_2 can be represented as:

$$\Delta t_1 = 40.3 * TEC/cf_1^2 \; ; \; \Delta t_2 = 40.3 * TEC/cf_2^2 \qquad (2.16)$$

In applying this method, the operating frequency is assumed as f1 so that f2 will be a lower frequency for correction purposes: For GPS applications, these frequencies are conventionally expressed as L1 (f1: 1575.42 MHz) and L2 (f1: 1227.60 MHz).

Thus,

$$\delta(\Delta t) = (K/c) * TEC\left(1/f_1^2 - 1/f_2^2\right) \qquad (2.17)$$

where $\delta(\Delta t)$ is the delay differential correction .

$$\Delta t_1 = \left(f_1^2/f_2^2 - 1\right)^{-1} * \delta(\Delta t)$$

The value of the delay differential correction can be obtained from the difference of the simultaneous measurements of the total range including ionospheric time delay at the two frequencies. The term $\left(f_1^2/f_2^2 - 1\right)^{-1}$ is the ionospheric scaling factor.

2.9.2 Phase Effect on Traversing Radio Wave

Theoretically, frequency can be expressed as the time derivative of phase (Leitinger, 1989). Variations in the total number of electrons encountered along the signal line of sight (LOS) can result to Doppler shift. As radio waves traverse the

ionosphere, they experience change of phase that is advanced relative to its original phase in free space. The phase advance is defined as:

$$\Phi = 40.3(2\pi)/cf * TEC \quad \text{(radians)} \tag{2.18}$$

Thus the Doppler shift can therefore be expressed as:

$$\Delta f = (40.3(2\pi)/cf) * d(TEC)/dt \tag{2.19}$$

For high orbit satellite where decimal changes in TEC are greater, the rate of change of TEC is approximately 0.1 TECU per second (Leitinger, 1989). During severe ionosphere conditions, variations in TEC are rapid enough to cause Doppler shifts that impact severely on radio waves. During these periods, the rate of change of TEC can rise to about 1.2 TECU per second. This can result to a Doppler shift of over 1 Hertz per second even at GHz frequencies (Brown et al., 1998), which may be sufficient to at times rapidly spread out the spectrum of the received signal in a manner that may cause the receiver to lose lock (Brown et. al, 1998; and Dandekar and Groves, 2004). Brown et al. (1998) gave explanations that the remedy to this problem is by designing receivers with correlators to cope with the fast variations of frequencies that do arise from phase scintillations.

2.9.3 Estimation of Ionospheric TEC Using GPS Receiver

Dual frequency GPS receivers are capable of generating two delay observables for each satellite being tracked: pseudorange delay (P) and carrier phase delay (Φ). Each observable is acquired at the two GPS carrier frequencies, L1 (f1: 1575.42 MHz) and L2 (f1: 1227.60 MHz). The frequency difference phase delays provide very precise measurements of TEC changes but contain an overall bias that is associated with integer cycle ambiguities (Mannucci, et al., 1998; and Carrano and

Groves, 2006). The frequency difference pseudorange data provide an absolute measure of total ionospheric delay between satellite and receiver but have significantly more multipath and system noise than the phase based data (Burrell et al., 2006). In addition, the pseudorange delays are biased by interfrequency delays in the satellite and receiver hardware. Manucci et al. (1998) in their investigation inferred that multipath induced error on receiver measurements depends on the total electromagnetic field of the surrounding.

Our explanation in estimating ionospheric TEC follows closely with the approaches of Van Dierendonck (2005) and Carrano (2007). In terms of the pseudorange measurement on L1 and L2 frequencies, TEC along the LOS to each GPS satellite can be estimated as follows;

$$TEC_P = A\{[P_2 - P_1] - [B_R + B_S] + D_P + E_P\} \tag{2.20}$$

where P_1 = Pseudorange on L1 (ns)

P_2 = Pseudorange on L2 (ns)

B_R = Receiver differential code bias (ns)

B_S = Satellite differential code bias (ns)

D_P = Differential pseudorange multipath error (ns)

E_P = Differential pseudorange measurement noise (ns)

The constant A and B that give TEC_P in units of TECU (1 TECU= 10^{16} el/m^2) are: A= 2.854 TECU/ns and B= 1.812 TECU/L1 cycle.

Measuring TEC using pseudoranges involves determination of the hardware differential code biases, B_R and B_S then subtracting them from the differential

pseudorange measurement. Accuracy in the measurement of TEC using the pseudorange alone is usually limited by the multipath and measurement noise terms which are difficult to model and can exceed those of the phase by an order of a magnitude or more (Carrano and Groves, 2006).

An alternative way of estimating TEC along satellite LOS involves the carrier phases on the L1 and L2 frequencies as follows:

$$TEC_L = B\{[L1 - (f_1/f_2)L2] - [N_1 + (f_1/f_2)N_2] + D_L + E_L\} \quad (2.21)$$

where $L1$ = Carrier phase on L1 (cycles)

$L2$ = Carrier phase on L2 (cycles)

N_1 = Integer ambiguity of L1 phase (cycles)

N_2 = Integer ambiguity of L2 phase (cycles)

D_L = Differential phase multipath error (cycles)

E_L = Differential phase measurement noise (cycles)

Measuring TEC using the carrier phase is generally more precise in that multipath error and measurement noise are smaller and may be neglected. The disadvantage is that the integer numbers of accumulated cycles of phase for frequency N_1 and N_2, are unknown and changes dramatically after each cycle slip (Makela et al., 2001).

Standard practice for estimating TEC using dual frequency GPS receivers combine the strengths of both pseudorange and carrier phase approaches (Fremouw et al., 1978). The calculation performed by GPS receiver in real-time is based on the phase formulation with the multipath and noise terms neglected, while the pseudoranges are used to estimate the unknown number of accumulated phase cycles. The method proceeds by defining the calibrated TEC as equal to the relative total

electron content TEC_R minus the satellite and receiver differential code biases. Carrano and Groves (2006) gave the calibrated TEC as,

$$TEC = TEC_R - A(B_R + B_S) \tag{2.22}$$

The relative TEC is determined from the differential carrier phase and a filtered combination of different pseudoranges and carrier phases, such that,

$$TEC_R = DCP + \langle DPR - DCP \rangle_{ARC} \tag{2.23}$$

where the differential carrier phase DCP and differential pseudorange DPR are defined in units of TECU (Carrano, 2007) as,

$$DCP = B[L1 - (f_1/f_2)L2] \tag{2.24}$$

$$DPR = A[P_2 - P_1] \tag{2.25}$$

The notation $\langle\ \rangle_{ARC}$ indicates an average taken over a phase connected arc (between successive cycle slips). Once the relative TEC has been determined, subtraction of the satellite and the receiver differential code biases yields the calibrated TEC (number of electrons encountered along the satellite-receiver LOS) via eq (2.25).

The equivalent vertical TEC is estimated by the principles that slant TEC depends on the elevation angle. If the single layer approximation for the ionosphere is valid, and in absence of spatio-temporal density gradients, then the verticalized TEC from each satellite should be the same (Carrano, 2007). At low to mid-latitudes, this condition is most closely met just before local sunrise, although there are cases where the ionosphere is highly structured even late at nights so that this assumption remains valid (Carrano and Groves, 2006).

The receiver bias is usually estimated once per day using data collected between 03:00-06:00 LT. Given the relative TEC, the satellite bias for each link, the satellite elevation, ε and the ionospheric height, h, then, the verticalized TEC becomes a function solely of the receiver bias.

$$TEC_V(B_R) = [TEC_R - A(B_R + B_S)]/M(\varepsilon, h) \tag{2.26}$$

In the above equation, $M(\varepsilon, h)$ is the single layer mapping function of the ionosphere, defined as,

$$M(\varepsilon, h) = \sec\{\sin^{-1}[(R_e \cos\varepsilon)/R_e + h]\} \tag{2.27}$$

where R_e is the earth radius in km, and ε is the elevation angle in radians. The ionospheric height h is unknown. Burrell et al. (2006) gave a model for determining the height. Alternatively it can be held fixed at a value representative of typical conditions. A height between 300-450km has been proposed by Fremouw et al. (1978), Carrano and Groves (2006), and Burrell et al. (2006).

2.10 Scintillation

The presence of irregularities is the main cause of scintillations on GNSS signals. On phenomenal days, the ionosphere is usually characterized by irregularities on vast spectrum of scales. Ionospheric scintillation is characterized by rapid fluctuations in the amplitude and phase of signals due to variations in the local index of refraction along the propagation path. Intensity scintillation is a purely propagation effect that arises from a combination of focusing/defocusing and diffraction pattern development. Consequently, the spectrum of intensity scintillations is limited at the low-frequency end by Fresnel filtering phenomenon. The extended nature of the phase spectrum and Fresnel cut off of the intensity spectrum require that scintillation

be defined on the basis of amplitude scintillation index, S_4 and phase scintillation index σ_ϕ (Briggs and Parkin, 1963). Equatorial (Africa) intensity scintillation is seasonally dependent and it is limited to local night-time hours (Doherty et al., 2004).

Briggs and Parkin (1963) defined amplitude scintillation index, S_4 as the standard deviation of the received signal intensity normalized to the mean intensity. Similarly, the phase scintillation index "sigma phi", (σ_ϕ), is the standard deviation of the received carrier phase. In practice, the phase and intensity of the received signal are smoothed by means of high/low-pass sixth order (pole) Butterworth filter with a 3dB cut-off at 0.1 Hz (Fremouw et al., 1978; and Van Dierendonck et al., 1993). The basic process of smothering is known as detrending. It implies passing the received signal through a coherent Automatic Gain Control (AGC) circuit which forms an integral part of the receiver.

2.10.1 Phase Scintillation, σ_ϕ

Traditionally, phase scintillation σ_ϕ is monitored by estimating the standard deviation of the phase, in radians, of the signal from GPS satellite at the receiver site. The standard practice requires that σ_ϕs are computed off-line over 1, 3, 10, 30 and 60sec internal in every 60sec. These σ_ϕs data are individually averaged over the 60sec period, so that the five σ_ϕs data are stored and displayed for each satellite on sight by the receiver through their pseudorandom noise code number (PRN), azimuth, elevation angles and the site's earth coordinates, (Van Dierendonck, 2005). For practical analysis, the 60sec σ_ϕ ($\sigma_{\phi 60}$) is often used. In detrending the raw phase measurements, a high pass filter on a modified bandwidth that may be individually varied between 0.01 to 1.0Hz is adopted to remove slowly varying systematic effects

due to satellite motion, clocks, selective availability (SA), and the troposphere (Thomas et al., 2001).

The raw detrended data are normally recorded at a sample rate of 50Hz (implying that 3,000 detrended phase data are computed over 60 sec intervals). In modern day measurements, an Oven Controlled Crystal Oscillator (OCXO) is incorporated into GPS receivers to minimize phase noise effects of low level phase scintillation which the previous Temperature Compensated Crystal Oscillator (TCXO) usually compromises (Van Dierendonck et al., 1993).

2.10.2 Amplitude Scintillation, S_4

Amplitude scintillation, S_4 is monitored by evaluating standard deviation of the received signal intensity normalized to the mean intensity. This is done by measuring the signal intensity of the received signal power in a manner that is devoid of fluctuations with noise power, and it is based on signal-to-noise ratio (Van Dierendonck et al., 1993). The received signal power varies due to errors arising from changing range, antenna patterns and multipath. This is minimized by filtering the intensity measurements in a low pass filter. Because S_4 index is usually normalized, the receiver's absolute gain is unimportant as long as it is relatively constant during the detrending period. It is also important that the intensity measurement is left linear with respect to the signal power over the entire range, including deep scintillation fades.

At GPS frequencies, it becomes important to remove ambient noise effects on the measured S_4 index. The S_4 index values are normally computed over 60sec intervals. In the Ionospheric Scintillation Monitor (ISM), S_4 values are usually stored and

displayed for each satellite being tracked using their respective PRN numbers. The ambient noise effect is removed by estimating the average signal-to-noise ratio over the 60sec interval, and using the estimate to determine the expected S_4 due to ambient noise (Thomas et al., 2001). In the absence of scintillation, S_4 index has values that lie below 0.05, while value of 1.0 indicates 100% Scintillation occurrence (Carrano, 2007). The amplitude scintillation index can be defined as:

$$S_4 = \sqrt{\frac{\langle I^2 \rangle - \langle I \rangle^2}{\langle I \rangle}} \qquad (2.28)$$

where I represents the signal intensity (amplitude squared). The scintillation index may be interpreted as the fractional fluctuation of the signal (e.g. $S_4 = 0.0$ indicates no modulation whereas $S_4 = 1.0$ indicates 100% modulation). Worst case of amplitude scintillation is experienced at GPS frequencies during high solar period, with fading depth of about 20dB in equatorial latitudes, especially during equinox. Obviously, scintillation depth in excess of 20dB has potential capability of undermining GNSS performance.

2.11 Differential GPS

For safety-critical implementation of GPS for aviation applications, there is need for enhancement in accuracy, availability and integrity of signals. This is achieved by augmenting GPS satellites with geostationary satellites. These satellites have ranging capabilities like the conventional GPS satellites, and are equipped with dedicated packages to transmit Wide Area Differential GPS (WADG) corrections and integrity messages. Differential technique rely on one or more reference stations at well known locations for ranging measurements, comparing them with their known

location, computing satellite location and residuals. These residuals represent the ranging errors. Generally, they lump together the satellite clock errors, orbit errors, and propagation errors, including multipath (Kintner and Ledvina, 2005). The ranging errors are then broadcast to mobile receivers for pseudorange corrections purposes. The accuracy of this concept depends on whether the reference receiver and the mobile receiver have common error sources or not. Also, differential systems are usually used to aid single frequency receivers.

Most regions of the world have launched their geosynchronous satellites to relay differential corrections to receivers, and many of them are on the L1 frequency and C/A codes. In the U.S., the Federal Aviation Administration (FAA) has developed Wide Area Augmentation System (WAAS) for differential corrections on GNSS systems with the view of guiding aircraft on approach with vertical guidance in the North American airspace (Doherty et al., 2001, 2004; and Van Dyke, 2001). Development of similar systems, the European Geostationary Navigation Overlay Services (EGNOS) via Galileo satellites are ongoing in Europe, and in the Pacific, and in parts of Asia; Multifunctional Transport Satellite Based Augmentation System (MSAS) is in development.

These are intended to provide continuous global coverage for aviation users. The quality of corrections depends on space environmental factors with tendency to sacrifice validity in equatorial/low-latitude zones during challenging ionospheric conditions. Under these conditions, reference receiver spacing, real-time measurements of ionospheric gradients and scintillations are essential factors to consider for a robust system.

CHAPTER THREE

3.0 MATERIALS AND METHODS

The basic tool for this investigation is numerical modeling and comparison with experimental observations. The work was approached in two parts: The first part explores the physics of the medium of propagation (ionosphere) for transionospheric (satellite) radio signals, and consequently, the impairment that the medium imposes on the traversing signals, especially during disturbed days. Under the numerical modeling, a nonlinear dynamics model is used to investigate the growth and evolution of ionospheric irregularities. Thereafter, phase screen model is used to affirm the dependence of ionospheric irregularities layer's thickness on the performance of traversing radio waves. In a view to establishing adequate theoretical fore-understanding of the equipment (GPS), we elucidated the operational principles of GPS, especially on how it fixes its user's position using numerical approach.

The second part is on experimental observations at some stations within the equatorial (Appleton) anomaly. Nigeria is on close geomagnetic latitude with these stations; hence, data from these stations will characterize excellently the ionosphere over Nigeria. These stations are: Nairobi (1.28°S, 36.83°E, dip lat: 26.20°S) [KENYA], Ascension Island (7.90°S, 14.42°W, dip lat: 16.00°S) [SOUTH ATLANTIC OCEAN], Sao Jose dos Campos (23.20°S, 45.90°W, dip lat: 17.60°S) [BRAZIL]. The basic equipment is the GPS Ionospheric Scintillation and Total Electron Content (TEC) Monitors (GISTMs), which are under the coordination of the United States Air Force Research Laboratory at Hanscom, Massachusetts, and

lastly a station located at Jacamarca (11.95°S, 76.87°W, dip lat: 4.00°S) [PERU] which measured the Spread F records on a phenomenal day. The first three stations observed scintillation effects on traversing GPS signals and simultaneously characterize the ionosphere by measuring the total electron content (TEC) of the ionosphere. The TEC records from the Nairobi station was used to further justify the second peak electron production after sunset which is due to recombination processes. The imposed noise due to scintillation on signals was observed by the Ascension Island, Sao Jose dos Campos, and Jacamarca stations. When the scintillation effect is so deep, it can cause fading effects of the order of 22dB-Hz and above on signals. We investigated the impacts of such deep fades on GPS-based aviation applications in terms of continuity and availability. Finally, we showed how this technology can be used to enhance aviation safety in Nigeria.

3.1 The Nonlinear Dynamics Model for Mapping Ionospheric Irregularities Growth and Evolution

The formation of irregularities in the ionosphere has been studied (Woodman and La Hoz, 1976; Basu and Kelly, 1979; Kelly, 1989; and Ray et al., 2006) to be related to vertical upwelling of $\mathbf{E} \times \mathbf{B}$ drifts of ions and electrons in the region of plasma depletion, resulting in lower perturbations through processes of nonlinear dynamics (Huang 1975; Hudson, 1978).

The approach of Kelly (1989) has been modified to obtain the difference equation. It is believed that energy flows from eddy to eddy across the spectrum of irregularities (Ishimaru, 1977), resulting in nonlinear processes (Hudson, 1978; and Basu and Kelly, 1979). In the study of radio propagation through random media, it

is conventional to define random functions in terms of four-dimensional Fourier transform pair (Yeh and Liu, 1982; and Zernov, 2006).

$$\zeta(\vec{r},t) = \iiint\int_{-\infty}^{+\infty} \Psi(\vec{k},\omega)\exp[+i(\vec{k}\cdot\vec{r}-\omega t)]d\vec{k}d\omega \qquad (3.1a)$$

$$\tilde{\zeta}(\vec{k},\omega) = \frac{1}{(2\pi)^4}\iiint\int_{-\infty}^{+\infty} \Psi(\vec{r},t)\exp[-i(\vec{k}\cdot\vec{r}-\omega t)]d\vec{r}dt \qquad (3.1b)$$

Where $\zeta(\vec{r},t)$ and $\Psi(\vec{k},\omega)$ are Fourier transform pair of the two random functions, k is the wave number and ω is the angular frequency of the wave respectively.

In practice, the integrals in eqs. (3.1a & b) are not always convergent. For instance, if the random function $\zeta(\vec{r},t)$ is statistically homogenous, its spectrum depicts delta function behaviour (Zernov, 2006). In contrast, correlation function spectra are in general convergent in the mean-square sense (Yeh and Liu, 1982; Kelly, 1989; and Davies, 1990).

For mathematical convenience, we attempt to express the correlation function in two-dimensional form, such that the correlation function now becomes entirely spatial. Using eq. (3.1a) and expressing the correlation function in the mean-square sense but in two-dimensional form:

$$\left\langle \Psi_\zeta(\vec{r}_1,\vec{r}_2)\right\rangle = \int\int_{-\infty}^{+\infty}\left\langle \tilde{\tilde{\zeta}}(\vec{k}_1)\tilde{\tilde{\zeta}}^*(\vec{k}_2)\right\rangle \exp[i(\vec{k}_1\cdot\vec{r}_1-\vec{k}_2\cdot\vec{r}_2)]d\vec{k}_1 d\vec{k}_2 \qquad (3.2)$$

It is important to note that the random function and its conjugate inside the integrals have tildes placed above them. This denotes that these random functions are zero-mean functions. The intensity of the spectrum of these functions can therefore be written as:

$$I_k(\vec{k}) = \left|\tilde{\tilde{\zeta}}(\vec{k})\right|^2 \equiv \left\langle \tilde{\tilde{\zeta}}(\vec{k}_1)\tilde{\tilde{\zeta}}^*(\vec{k}_2)\right\rangle \exp[i(\vec{k}_1\cdot\vec{r}_1 - \vec{k}_2\cdot\vec{r}_2)] \qquad (3.3)$$

Therefore, eq. (3.2) can be re-written in a more concise form as:

$$\left\langle \Psi_\zeta(\vec{r}_1,\vec{r}_2)\right\rangle = \int_{-\infty}^{+\infty}\int I_k(\vec{k})d\vec{k} \qquad (3.4)$$

where $\vec{k} = \left|\vec{k}\right| = \left[k_1^2 + k_2^2\right]^{\frac{1}{2}}$.

Assuming the irregularities are isotropic and are in the plane perpendicular to $\vec{B} = B\hat{a}_2$, then, $I_k(\vec{k}) = I_k(k)$. The two-dimensional spatial correlation function $\Psi_\zeta(\vec{r}_1,\vec{r}_2)$ can be used to define the statistical properties of the ionospheric irregularities. In other words, the extent of spatial correlation depends on the extent of irregularities fluctuations in the ionosphere.

We can therefore express eq. (3.4) in terms of the total ionospheric irregularities density fluctuation strength:

$$\left\langle (\Delta N/N)^2\right\rangle = \int_{-\infty}^{+\infty}\int I_k(\vec{k})d\vec{k} \qquad (3.5)$$

Kelly (1989) used eq. (3.5) to derive a differential equation for $I_k(k)$ which describes the flow of energy in the system as a function of k. He explained that energy flows nonlinearly from eddy to eddy through the spectrum. At the end of his derivation (see Kelly (1989) for details), this differential equation was obtained.

$$\frac{d}{dk}\left(k^4 I_k^{3/2}\right) \propto \frac{kI_k}{V_o}\left(A - Bk^2\right) \qquad (3.6)$$

where A is the growth parameter, B is the damping parameter and V_o is the drift velocity of electrons across eddies.

From eq. (3.6), it is obvious that the LHS defines the gradient of the two-dimensional spectrum for the density fluctuation of the irregularities, with a nonlinear factor $\left(A - Bk^2\right)$. This study modifies equation (3.6) to:

$$X_{n+1} = const(1 - pX_n^2) \qquad (3.7)$$

where $X = \left(k^4 I_k^{3/2}\right)$ and $p = B/Ak_p$ $(k_p = k^6 I_k^3)$. P is a parameter that is strength related. The motive here is to obtain numerical approximation for the difference eq. (3.7) through a numerical process. In doing this, we assumed a simple case where the constant in eq. (3.7) is set to unity, such that:

$$X_{n+1} = 1 - pX_n^2 \qquad (3.8)$$

The difference eq. (3.8) is numerically evaluated using iterative technique with an initial guess of 0.01. A FORTRAN code was developed to solve the difference equation. Large numbers of iterations in several thousands are carried out in order to solve the difference equation with the last hundred of iterations retained while others are discarded. The reason for this is that the behaviour of these irregularities on long-term basis is of interest and this behaviour can only be made clear by these long-term iterations.

Viewing eq. (3.8) comprehensively, it becomes obvious that this equation defines the gradient of irregularities drift in the ionosphere (Basu, 1978). In carrying out the numerical computations, it is assumed that this gradient is bounded within plus and minus unity, that is, $X_{n+1} = 1 - pX_n^2$, $X_n \in [1,-1]$. As previously mentioned, p is a parameter that is related to the strength of the irregularities. This parameter was varied from 0.1 through to 2 with incremental step of 0.1. For each

p-parameter, ten thousand iterations were carried out with the last hundred retained while the first nine thousand nine hundred iterations discarded. After the entire iterations for all values of p-parameters, the gradient of the irregularities drift was plotted against p-parameters as shown in Figure 4.1. The figure shows a map of nonlinear regime of the evolution of ionospheric irregularities.

3.2 Mutual Coherence Functions and the Phase Screen Model

In this section, phase screen model was used to relate the statistical properties of the traversing signals through the ionosphere to the statistical characteristics of the irregularities especially those of equatorial origin. Usually, the scattering experienced by a traversing radio wave through the ionosphere is random and in forward direction (Ishimaru, 1977), which also causes the total electron content (TEC) to fluctuate irregularly. To study the effects of random scattering on transient radio wave through the ionosphere, it is relevant to employ some statistical tools (Gherm and Zernov, 1998). To this end, a method is adopted for a situation when radio waves at different frequencies are simultaneously transmitted from a satellite through a slab of irregularity that is bounded by an infinite random screen to a receiver located below the ionosphere, so that the initial stochastic field is given at the boundary.

The correlations between these frequency components after emerging from the irregularity slab to free space environment where the receiver is located, are of interest. For simplicity, two of these frequencies are considered. The usual Helmholtz's wave equation is used to obtain the approximate parabolic equation (Bijukov et al., 2003). In turn, under the Markov random process and parabolic

equation approximations, two frequency and two position mutual coherence functions are derived using the principle of moments. This section is characterized by several mathematical apparatus, and they have been applied in their simplest forms. The physical ideas implied by their solutions are of interest and they are therefore better stressed.

3.2.1 Complex Amplitude of the Traversing Wave

The problem was motivated by the fact that the radio wave that traverses the ionosphere suffers various distortions and consequently degradation of the output signal at the receiver located outside the ionosphere. This effect is not unconnected with the interaction of the traversing wave with the irregularity structures of the ionosphere which introduces stochastic behaviour on the wave at the boundary of the ionosphere before its exit to free space below. The problem is formulated by using the geometry of Figure 3.1.

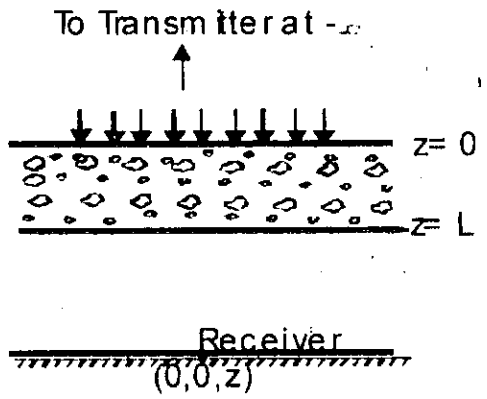


Figure 3.1: Geometry of the Ionospheric Slab

A portion of the ionosphere is assumed to be a slab of random electron density irregularities with thickness L. A time-harmonic wave is incident on the irregularity

slab and then received at a receiver located near the ground at point (ρ, z) where ρ is the transverse coordinate. These irregularities form scatterers that cause the incoming wave to experience forward scattering. It is also assumed that the irregularity slab is characterized by a dielectric permittivity with a random function of position and time (Strohbehn, 1968; Wernik et al., 1980; and Zernov, 2006). This can be represented as:

$$\varepsilon(\vec{r},t) = \langle\varepsilon\rangle\left[\varepsilon_o(\vec{r}) + \varepsilon_1(\vec{r},t)\right] \tag{3.9}$$

The simplest case will be considered, where $\varepsilon_o(\vec{r})$ will be set to unity and the irregularity assumed to be of large scale i.e. $kl_c \geq 2\pi$ which implies that $kl_c >> 1$.

$$\varepsilon = \langle\varepsilon\rangle\left[1 + \varepsilon_1(\vec{r},t)\right] \tag{3.10}$$

where $\langle\varepsilon\rangle$ is the ensemble average of the background dielectric permittivity which for the ionosphere is given by:

$$\langle\varepsilon\rangle = (1 - \omega_{po}^2/\omega^2)\varepsilon_o \tag{3.11}$$

We must be reminded that eq. (3.11) is related to the refractive index, n of the background ionosphere by, $n^2 = 1 - \omega_{po}^2/\omega^2$ (Ratcliffe, 1972; Beer, 1976; Budden, 1985).

The dielectric permittivity in terms of three-dimensional (3-D) ionospheric electron density can be expressed for cold collision-less plasma as:

$$\varepsilon = 1 - \frac{e^2\left[N_o(\vec{r}) + \Delta N(\vec{r},t)\right]}{m\varepsilon_o\omega^2} \tag{3.12}$$

In eq. (3.12) the full electron density has been written in two parts, the first part $(N_o(\vec{r}))$ is slowly varying in space distribution in the background ionosphere while

the second term $(\Delta N(\vec{r},t))$ represents fluctuations of the electron density in space and time, though the time dependence is assumed to be slowly varying. The term is therefore taken as a zero-mean random function, in other words, it fulfils:

$$\langle \Delta N(\vec{r},t) \rangle = 0 \tag{3.13}$$

Physically speaking, the fluctuation term $(\Delta N(\vec{r},t))$ in eq. (3.12) is statistically inhomogeneous because the extent of fluctuations is factored on the magnitude of the electron density of the background plasma. It is therefore more physically relevant to re-write the equation as:

$$\varepsilon = 1 - \frac{e^2 N_o(z)}{m\varepsilon_o \omega^2} - \frac{e^2 N_o(z)}{m\varepsilon_o \omega^2} \frac{\Delta N(\vec{r},t)}{N_o(z)} \tag{3.14}$$

Eq. (3.10) can be substituted into eq. (3.15) to yield:

$$\varepsilon = \langle \varepsilon \rangle \left[1 - \frac{\omega_{po}^2/\omega^2 \left[\Delta N(\vec{r},t)/N_o(z)\right]}{(1-\omega_{po}^2/\omega^2)} \right] \tag{3.15}$$

Comparing eqs. (3.10) and (3.15), the fluctuating part of the dielectric permittivity can be expressed as:

$$\varepsilon_1(\vec{r},t) = -\frac{(\omega_{po}^2/\omega^2)\left[\Delta N(\vec{r},t)/N_o(z)\right]}{1-(\omega_{po}^2/\omega^2)} \tag{3.16}$$

ω_{po} is the angular plasma frequency of the background electrons, with density, $N_o(z)$, and ω is the angular frequency of the incident electromagnetic wave. ε_o is the permittivity of free space, e is the electronic charge and m is the mass of an electron. The background electron density $N_o(z)$ is a function of height with a typical scale height, H. The percentage electron density fluctuation $(\Delta N(\vec{r})/N_o(z))$ is assumed to be homogeneous random field with a typical inner scale size l_o. For

ionospheric propagation, the scale height is comparatively large when compared to the inner scale size of the irregularities and the inner scale size is in turn comparatively large when compared to the wavelength of the incident wave (Liu and Yeh, 1982; and Kintner and Ledvina, 2005). This suggests that the most significant scattering process in the ionosphere under radio propagation is the forward scattering. As seen in eq. (3.16), the level of fluctuations experienced by the background ionosphere for a wave propagating at an angular frequency, ω increases with refractive effects.

In many cases, wave propagation through the ionosphere can be described by the Helmholtz's equation for a component of the electromagnetic field (Liu et al., 1974a, 1974b; Wernik et al., 1980; Kintner and Ledvina, 2005; and Zernov, 2006).

$$\nabla^2 E + k^2 \varepsilon(\vec{r},t) E = 0 \tag{3.17}$$

$$\nabla^2 E + k^2 \left[1 + \varepsilon_1(\vec{r},t)\right] E = 0, \quad 0 < z < L \tag{3.18}$$

The solution of eq. (3.18) is of the form expressed below, which is a wave propagating in the negative z-direction.

$$E = u(\vec{r}) \exp(-jkz) \tag{3.19}$$

where $u(\vec{r})$ is the complex amplitude of the wave and the time dependent term has been omitted.

Eq. (3.18) is a partial differential equation with random coefficient that produces a suitable solution upon which scintillation theories can be based. Unfortunately, the exact solution of eq. (3.18) seems impossible, hence, there is need to settle for series of approximations to achieve an applicable results (Sreenivasiah et al., 1976; and Yeh and Liu, 1977). Eq. (3.19) can be used to

modify eq. (3.18) in terms of the complex amplitude. Here, the intention is to adopt a simple case of an incident plane wave that is propagating along the z axis of the parabolic equation such that the transverse parts of the axes are suppressed in the process of solving the second-order moment equation (Bitjukov et al., 2003):

$$\frac{\partial^2 u}{\partial z^2} - 2jk\frac{\partial u}{\partial z} + \nabla_\perp^2 u + k^2 \varepsilon_1(\vec{r},t)u = 0 \qquad (3.20)$$

The order of magnitude of the first two terms in eq. (3.20) can be approximated as follows:

$$-2jk\frac{\partial u}{\partial z} \sim -2jk\frac{u}{l_\varepsilon}$$

$$\frac{\partial^2 u}{\partial z^2} \sim \frac{u}{l_\varepsilon^2}$$

Therefore, the approximation of these two terms is,

$$-2jk\frac{\partial u}{\partial z} + \frac{\partial^2 u}{\partial z^2} \sim -2jk\frac{\partial u}{\partial z}\left(1 - \frac{1}{2jkl_\varepsilon}\right) \qquad (3.21)$$

From eq. (3.21), the contribution of the second order derivative of u is small for the large scale irregularities and therefore negligible. Then, we can finally obtain the approximate eq. (3.22).

$$-2jk\frac{\partial u}{\partial z} + \nabla_\perp^2 u + k^2 \varepsilon_1(\vec{r},t)u = 0, \qquad 0 < z < L \qquad (3.22)$$

This approximation that assumes the neglect of the second order derivative in z is known as the *quasi-optics or parabolic equation approximation*. The physical implication of this is that the wave is not expected to undergo reflection on the z-axis. Eq. (3.22) is the so-called parabolic equation approximation whose derivations in terms of moments are used for scintillation theories and shall be taken as the

starting point for the problem. $\nabla_{\perp}^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$ (i.e. the transverse Laplacian operator).

If eq. (3.19) is re-written so that the time-dependent term is retained, then,

$$E(\vec{r},t) = u(\vec{r},\omega)\exp(j\omega\left\{t - \int_0^z k dz\right\}) \qquad (3.23)$$

where $u(\vec{r},\omega)$ is the complex amplitude as a function of position and angular frequency, represented as, $u(\vec{r},\omega) = \exp\Psi(\vec{r},\omega)$. Under the usual assumption of neglecting the depolarization effect of the field and adopting Wentzel-Kramer-Brillouin (WKB) solution, the complex amplitude of the wave approximately satisfies the parabolic equation. Eq. (3.23) is a monochromatic wave equation. In its stead, a non-monochromatic wave may be introduced by assuming that a signal with Fourier amplitude $A(\omega)$ is incident on the irregularity slab and later arrives at the receiver. Then, the signal can be represented by:

$$P(\vec{r},t) = \int_{-\infty}^{+\infty} A(\omega)u(\vec{r},\omega)\exp((j\omega\left\{t - \int_0^z [\langle\varepsilon(\zeta)\rangle]^{\frac{1}{2}} d\zeta/c\right\}))d\omega \qquad (3.24)$$

where $k = \omega(\langle\varepsilon\rangle)^{\frac{1}{2}}/c$ $(k^2 = k_{pm}^2\langle\varepsilon\rangle = (\omega/c)^2\langle\varepsilon\rangle)$ and ζ is the normalized coordinate for the z-axis.

Starting from the parabolic eq. (3.22), it is possible to derive a closed set of equations for the statistical moments of the field u, by using characteristic functional derivatives (see Lee, 1974 and references therein for details). Introducing the principle of general moments to the complex amplitude $u(\vec{\rho},z,k)$ where the frequency of the wave is explicitly represented by the argument k (the wave-number). For convenience, it is better to express the field of the complex amplitude

as function of two arguments, hence, $\vec{s} = (\vec{\rho}, k)$ is introduced so that the field can now be expressed as $u(\vec{s}, z)$. It is worthy to note that the functions under considerations are complex random functions and are statistically $2n$-dimensional. This implies that two correlation functions will be dealt with simultaneously.

From eq. (3.24), it is easy to note that the contributory effect of the background dispersive properties on the signal is mainly in the exponential factor while the effect due to random scattering falls within the complex amplitude $u(\vec{r}, \omega)$ which satisfies the parabolic equation. In essence, statistical properties (intensity fluctuation, probability density, correlation etc) of signal $P(\vec{r}, t)$ can be defined upon the statistical characteristics of the complex amplitude.

3.2.2 Derivations and Solution of the moment Functions

The *mnth*-order moment of the field is defined by Lee (1974) as:

$$\Gamma_{mn}(z, \vec{s}_1, \vec{s}_2 \ldots\ldots\ldots, \vec{s}_m; \vec{s}_1', \ldots\ldots\ldots, \vec{s}_n') = \langle u_1 u_2 \ldots\ldots\ldots u_m u_1'^{*} \ldots\ldots\ldots u_n'^{*} \rangle \quad (3.25)$$

where $u_i = u(\vec{s}_i, z)$ and $u_i' = u(\vec{s}_i', z)$.

Lee (1974) and, Yeh and Liu (1982) showed that the general moment for the field satisfies the following equation, where k_{po} is the wave-number of the background plasma.

$$\frac{\partial}{\partial z}\Gamma_{m,n}(z, \vec{s}_1 \ldots\ldots\ldots \vec{s}_m; \vec{s}_1' \ldots\ldots\ldots, \vec{s}_n') = -\frac{j}{2}\left[\frac{\nabla_1^2}{k_1} + \ldots\ldots\ldots\frac{\nabla_m^2}{k_m} - \frac{\nabla_1'^2}{k_1'} - \ldots\ldots\ldots\frac{\nabla_n'^2}{k_n'}\right]\Gamma_{m,n}$$

$$-\frac{k_{po}^4}{8}\left[\sum_{i=1}^{m}\sum_{j=1}^{m}\frac{A_{\Delta N}(\vec{\rho}_i - \vec{\rho}_j)}{k_i k_j} - \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{\left[A_{\Delta N}(\vec{\rho}_i - \vec{\rho}_j') + A_{\Delta N}(\vec{\rho}_j - \vec{\rho}_i')\right]}{k_i k_j'} + \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{A_{\Delta N}(\vec{\rho}_i' - \vec{\rho}_j')}{k_i' k_j'}\right]\Gamma_{m,n}$$

(3.26)

$\nabla_i^2 = \partial^2/\partial x_i^2 + \partial^2/\partial y_i^2$ and $\nabla_j'^2 = \partial^2/\partial x_j'^2 + \partial^2/\partial y_j'^2$. Eq. (3.26) is the general set of equations describing the behaviour of the higher statistical moments of the scintillating signal. This set of equations was first used to develop a multiple scattering scintillation theory for the ionosphere by Liu *et al.* (1974a). For a condition where $A_{\Delta N}(\cdot)$ in the last term is set to zero in eq. (3.26), the equation still remains valid but now defines statistical characteristics of the signal outside the irregularity slab i.e. for $z > L$.

For the purpose of this work, attention shall be on the two second-order moment functions that will be defined as:

$$\Gamma(\vec{\rho}, z, \omega_1, \omega_2) = \langle u(\vec{\rho}_1, z, \omega_1) u^*(\vec{\rho}_2, z, \omega_2) \rangle = \left\langle \exp \Psi(\vec{\rho}_1, z, \omega_1) \cdot \exp \Psi^*(\vec{\rho}_2, z, \omega_2) \right\rangle \equiv \langle u_1 u_2^* \rangle \tag{3.27}$$

$$\bar{\Gamma}(\vec{\rho}, z, \omega_1, \omega_2) = \langle u(\vec{\rho}_1, z, \omega_1) u(\vec{\rho}_2, z, \omega_2) \rangle = \left\langle \exp \Psi(\vec{\rho}_1, z, \omega_1) \cdot \exp \Psi\ (\vec{\rho}_2, z, \omega_2) \right\rangle \equiv \langle u_1 u_2 \rangle \tag{3.28}$$

The wave-number in the argument has been replaced by the two angular frequencies ω_1 and ω_2. Γ and $\bar{\Gamma}$ in eqs. (3.27) and (3.28) denote symmetric and anti-symmetric two-frequency mutual coherence functions respectively. ω_1 and ω_2 are now the two angular frequencies that satisfy the mutual coherence functions. The statistical characteristics of signal $P(\vec{r}, t)$ in eq. (3.24) can be derived using quantities Γ and $\bar{\Gamma}$. The intensity of a pulse at the receiving points is defined as $|P(\vec{r}, t)|^2$, and can be further represented in frequency domain as the ensemble average of the product of the signal and its conjugate:

$$\langle P(t) P^*(t) \rangle = \int \int_{-\infty}^{+\infty} ((A(\omega_1) A^*(\omega_2) \Gamma(0, z, \omega_1, \omega_2)) \cdot \exp(j\omega_1 \left\{ t - \int^z \left[\langle \varepsilon(\zeta) \rangle \right]^{\frac{1}{2}} d\zeta / c \right\}$$

$$-j\omega_2\left\{t-\int_0^z[\langle\varepsilon(\zeta)\rangle]^{\frac{1}{2}}d\zeta/c\right\})d\omega_1 d\omega_2 \qquad (3.29)$$

In eq (3.29), the integral of the type of eq. (3.24) is now integrated over two frequency domains with the mutual coherence function appearing naturally. In constructing the mutual coherence function, we must account for at least second order moment approximation of the complex amplitude. The mutual coherence function is the effect due to the ionospheric electron density fluctuations. In the absence of fluctuations, the mutual coherence function in eq. (3.29) is set to unity.

Another important assumption is that under forward scattering, the field at any height z obeys the property of "dynamic causality", that is, the field at any height z depends on the property (fluctuating dielectric permittivity) of the medium only in the region that lies before that height ($z_2 < z$). Mathematically, this is expressed by defining the correlation scale of the fluctuating dielectric permittivity of the medium in z direction in terms of dirac-delta function in z. This approximation is known as the *Markov random process approximation* (Lee, 1974; Yeh and Liu, 1982; and Zernov, 2006). The physical implication of this assumption is that, if the wave propagates in a distance much greater than the longitudinal correlation distance of the fluctuating dielectric permittivity of the random medium, the field will vary only in a very minute correlation scale of the fluctuating dielectric permittivity in z direction.

$$B_\varepsilon(\vec{\rho}-\vec{\rho}_2, z-z_2) = A_\varepsilon(\vec{\rho}-\vec{\rho}_2)\delta(z-z_2) \qquad (3.30a)$$

Therefore, under Markov random process approximation, the irregularity correlation function and the irregularity spectrum have a Fourier transform pair of the form:

$$A_\varepsilon(\vec{\rho}) = \int_{-\infty}^{+\infty} B_\varepsilon(\vec{\rho},\zeta)d\zeta = 2\pi\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\Phi_\varepsilon(\vec{k}_\perp,0)\exp(j\vec{k}_\perp\cdot\vec{\rho})d^2k_\perp \qquad (3.30b)$$

(Yeh and Liu, 1982)

where $A_\varepsilon(\vec{\rho})$ is the two-dimensional Fourier transform of the irregularity spectrum, $B_{\Delta N}(\vec{\rho},\xi)$ is the correlation function for the irregularity slab, $\Phi_{\Delta N}(\vec{k}_\perp,0)$ is the power spectrum, defined in inverse power law as, $\Phi(k_\perp,0) \propto \left(1+k^2{}_\perp/K^2\right)^{-p/2}$, p is the spectra index and $\vec{k}_\perp$ is the wave-number for the transverse coordinate (Ishimaru, 1977).

From eq. (3.26), the following set of equations can be derived for two frequencies.

$$\frac{\partial\Gamma}{\partial z} + j\frac{\left[k_2\nabla_\perp^2 - k_1\nabla_\perp'^2\right]\Gamma}{2k_1k_2} + \frac{k_{po}^4}{8}\left[\left(\frac{1}{k_1^2}+\frac{1}{k_2^2}\right)A_{\Delta N}(0) - \frac{2}{k_1k_2}A_{\Delta N}(\vec{\rho})\right]\Gamma = 0 \qquad (3.31)$$

For convenience, the ionosphere is assumed to be isotropic, hence, it can be assumed that $\nabla_\perp^2 = \nabla_\perp'^2$.

$$\frac{\partial\Gamma}{\partial z} + j\left(\frac{k_2-k_1}{2k_1k_2}\right)\nabla_\perp^2\Gamma + \frac{k_{po}^4}{8}\left[\left(\frac{k_2^2+k_1^2}{k_1^2k_2^2}\right)A_{\Delta N}(0) - \frac{2}{k_1k_2}A_{\Delta N}(\vec{\rho})\right]\Gamma = 0 \qquad (3.32)$$

Similarly, for the anti-symmetric function,

$$\frac{\partial\overline{\Gamma}}{\partial z} + j\left(\frac{k_2-k_1}{2k_1k_2}\right)\nabla_\perp^2\overline{\Gamma} + \frac{k_{po}^4}{8}\left[\left(\frac{k_2^2+k_1^2}{k_1^2k_2^2}\right)A_{\Delta N}(0) + \frac{2}{k_1k_2}A_{\Delta N}(\vec{\rho})\right]\overline{\Gamma} = 0 \qquad (3.33)$$

where $k_1 = \omega_1\left[\langle\varepsilon(\omega_1)\rangle\right]^{\frac{1}{2}}/c$ and $k_2 = \omega_2\left[\langle\varepsilon(\omega_2)\rangle\right]^{\frac{1}{2}}/c$. Analytically, approximate solution of eq. (3.32) will assume the form:

$$\Gamma = \exp\left(-\frac{k_{po}^4}{8}\left(\left[\frac{k_2^2+k_1^2}{k_1^2k_2^2}\right]A_{\Delta N}(0)z+\phi\right)\right) \qquad (3.34)$$

For numerical simplicity, it will be convenient to introduce new variables to transform eqs. (3.32) and (3.33) into dimensionless forms in terms of normalized coordinates (Liu and Yeh, 1975).

$$r = \rho/l_o\,, \quad \xi = x/l_o\,, \quad \eta = y/l_o\,, \quad \zeta = z/k_o l_o^2 \tag{3.35}$$

$$\omega_1 = \omega_o(1-X)\,, \quad \omega_2 = \omega_o(1+X)\,, \quad k_o = \omega_o/c \tag{3.36}$$

After some mathematical manipulations (see Appendix B), eqs. (3.32) and (3.33) transform to:

$$\frac{\partial\Gamma}{\partial\zeta} + jB_1\nabla_\perp^2\Gamma + \left(\frac{1}{4}\right)\left(\frac{\omega_p}{\omega_o}\right)^4 \frac{k_o^3 l_o^3}{l_o} \frac{\left\langle\left(\Delta N/N_o^2\right)^2\right\rangle}{\left\langle\left(\Delta N/N_o^2\right)^2\right\rangle} \cdot$$

$$\left[\frac{\left[1+X^2-\left(\omega_p/\omega_o\right)^2\right]}{D^2} A_{\Delta N}(0) - \frac{1}{D} A_{\Delta N}(\xi,\eta)\right]\Gamma = 0 \tag{3.37}$$

$$\frac{\partial\overline{\Gamma}}{\partial\zeta} + jB_1\nabla_\perp^2\overline{\Gamma} + \left(\frac{1}{4}\right)\left(\frac{\omega_p}{\omega_o}\right)^4 \frac{k_o^3 l_o^3}{l_o} \frac{\left\langle\left(\Delta N/N_o^2\right)^2\right\rangle}{\left\langle\left(\Delta N/N_o^2\right)^2\right\rangle} \cdot$$

$$\left[\frac{\left[1+X^2-\left(\omega_p/\omega_o\right)^2\right]}{D^2} A_{\Delta N}(0) + \frac{1}{D} A_{\Delta N}(\xi,\eta)\right]\overline{\Gamma} = 0 \tag{3.38}$$

where the transverse Laplacian $\nabla_\perp^2$ now stands for $\partial^2/\partial\xi^2 + \partial^2/\partial\eta^2$, and,

$$D = \alpha\beta(1-X^2) \tag{3.39}$$

$$B_1 = \left[(\alpha+\beta)X + \beta - \alpha\right]/2D\,. \tag{3.40}$$

$$B_2 = \left[\alpha+\beta+(\beta-\alpha)X\right]/2D \tag{3.41}$$

$$\alpha = \left[1-\left\{\left(\omega_p/\omega_o\right)^2/(1-X)^2\right\}\right]^{\frac{1}{2}} \tag{3.42}$$

$$\beta = \left[1-\left\{\left(\omega_p/\omega_o\right)^2/(1+X)^2\right\}\right]^{\frac{1}{2}} \tag{3.43}$$

B_3 and C can also be expressed as,

$$B_3 = \left|1 + X^2 - (\omega_p/\omega_o)^2\right|/D^2 \tag{3.44}$$

$$C = (1/4)(\omega_p/\omega_o)^4 k_o^3 l_o^3 \left\langle \left(\Delta N/N_o^2\right)^2 \right\rangle \tag{3.45}$$

If eqs. (3.44) and (3.45) are substituted into eqs. (3.37) and (3.38), then,

$$\frac{\partial \Gamma}{\partial \zeta} + jB_1 \nabla_\perp^2 \Gamma + \frac{C}{l_o \left\langle \left(\Delta N/N_o^2\right)^2 \right\rangle}\left[B_3 A_{\Delta N}(0) - \frac{1}{D} A_{\Delta N}(\xi,\eta)\right]\Gamma = 0 \tag{3.46}$$

Define,

$$G(0) = A_{\Delta N}(0)/l_o \left\langle \left(\Delta N/N_o^2\right)^2 \right\rangle \text{ and } G(\xi,\eta) = A_{\Delta N}(\xi,\eta)/l_o \left\langle \left(\Delta N/N_o^2\right)^2 \right\rangle \tag{3.47}$$

Substituting eq. (3.46) into eq. (3.47),

$$\frac{\partial \Gamma}{\partial \zeta} + jB_1 \nabla_\perp^2 \Gamma + \left[CB_3 G(0) - \frac{C}{D} G(\xi,\eta)\right]\Gamma = 0 \tag{3.48}$$

Similarly,

$$\frac{\partial \overline{\Gamma}}{\partial \zeta} + jB_2 \nabla_\perp^2 \overline{\Gamma} + \left[CB_3 G(0) + \frac{C}{D} G(\xi,\eta)\right]\overline{\Gamma} = 0 \tag{3.49}$$

Eqs. (3.48) and (3.49) will have approximate solutions of the form earlier expressed in eq. (3.34).

$$\Gamma = \exp-\left(CB_3 G(0)\zeta + \phi\right) = \exp-\left(CG(0)B_3\zeta\right) \cdot \exp-\phi \tag{3.50}$$

$$\overline{\Gamma} = \exp-\left(CB_3 G(0)\zeta - \phi\right) = \exp-\left(CG(0)B_3\zeta\right) \cdot \exp+\phi \tag{3.51}$$

If $\Gamma_1 = \exp-\phi$ and $\Gamma_2 = \exp+\phi$, then,

$$\Gamma = \Gamma_1 \exp-\left(CB_3 G(0)\right)\zeta \tag{3.52}$$

$$\overline{\Gamma} = \Gamma_2 \exp-\left(CB_3 G(0)\right)\zeta \tag{3.53}$$

Considering the geometry of the problem, the irregularity slab has a height from $z = 0$ to $z = L$. Therefore, the statistical properties of the traversing signal through the irregularity slab can be defined within these limits.

$$\therefore \Gamma = \Gamma_1 \text{ and } \overline{\Gamma} = \Gamma_2 \quad , z = 0 \tag{3.54}$$

$$\Gamma = \Gamma_1 \exp -CG(0)B_3 L/k_o l_o^2 \text{ and } \overline{\Gamma} = \Gamma_2 \exp -CG(0)B_3 L/k_o l_o^2 \quad , z = L = \zeta_o \tag{3.55}$$

The set of eq. (3.54) has less physical implication to the problem and can therefore be suppressed and the arguments of the function are now taken as $\vec{\rho}, z, \omega_o$ and X such that:

$$\Gamma(\vec{\rho}, z, \omega_o; X) = \Gamma_1 \exp -CG(0)B_3 L/k_o l_o^2 \tag{3.56}$$

$$\overline{\Gamma}(\vec{\rho}, z, \omega_o, X) = \Gamma_2 \exp -CG(0)B_3 L/k_o l_o^2 \tag{3.57}$$

X is the fractional frequency separation at a given receiving point and ω_o is the wave frequency within the irregularity slab. With eqs. (3.50) and (3.51), eqs. (3.48) and (3.49) will reduce to:

$$\frac{\partial \phi}{\partial \zeta} + jB_1\left[\nabla_\perp^2 \phi - (\nabla_\perp \phi)^2\right] - \left[\frac{C}{D} G(\xi, \eta)\right] = 0 \tag{3.58a}$$

Satisfying the initial condition $\phi(\rho, 0) = 0$ (see Appendix C). Eq. (3.58a) is related to Γ_1 as:

$$\frac{\partial \Gamma_1}{\partial \zeta} + jB_1\left[\nabla_\perp^2 \Gamma_1 - (\nabla_\perp \Gamma_1)^2\right] - \left[\frac{C}{D} G(\xi, \eta)\right]\Gamma_1 = 0 \tag{3.58b}$$

The non-linear term, $(\nabla_\perp \Gamma_1)^2$ in eq. (3.58 a & b) is small and can therefore be neglected.

$$\frac{\partial \Gamma_1}{\partial \zeta} + jB_1 \nabla_\perp^2 \Gamma_1 - \left[\frac{C}{D} G(\xi,\eta)\right]\Gamma_1 = 0 \tag{3.59}$$

Similarly,

$$\frac{\partial \Gamma_2}{\partial \zeta} + jB_2 \nabla_\perp^2 \Gamma_2 + \left[\frac{C}{D} G(\xi,\eta)\right]\Gamma_2 = 0 \tag{3.60}$$

Eqs. (3.59) and (3.60) can be solved by assuming an approximate solution of the form:

$$\Gamma_1 = \exp\left[\frac{jCG(\xi,\eta)\zeta\Gamma_3}{B_1 D}\right] \tag{3.61}$$

$$\Gamma_2 = \exp\left[\frac{-jCG(\xi,\eta)\zeta\Gamma_4}{B_2 D}\right] \tag{3.62}$$

Also with eqs. (3.61) and (3.62), eqs. (3.58) and (3.59) further reduce to:

$$\frac{-j}{B_1}\frac{\partial \Gamma_3}{\partial \zeta} + \nabla_\perp^2 \Gamma_3 = 0, \quad \zeta > \zeta_o \tag{3.63}$$

$$\frac{-j}{B_2}\frac{\partial \Gamma_4}{\partial \zeta} + \nabla_\perp^2 \Gamma_4 = 0, \quad \zeta > \zeta_o \tag{3.64}$$

Eqs. (3.63) and (3.64) are the mutual coherence functions for $z > L$ in the region below the irregularity slab, satisfying the initial condition, $\Gamma_3(\rho,0) = 1$. Using method of separation of variables,

$$\Gamma_3 = \exp jB_1 k_\perp^2 \zeta \tag{3.65}$$

Referring to the geometry of the problem in figure 1, eq. (3.65) has two physical interpretations as regards to the characteristics of the wave within and outside the irregularity slab. If $\exp jB_1 k_\perp^2 \zeta$ is the solution for the entire ray path i.e. from $z = 0$ to $z = z$, and $\exp jB_1 k_\perp^2 (\zeta - \zeta_o)$ is the solution for the wave outside the

irregularity slab i.e. $z > L$. It will be logical to assume that the solution within the irregularity slab is $\exp jB_1k_\perp^2(\zeta-\zeta_o)-\exp jB_1k_\perp^2\zeta$. Eq. (3.65) can then be written as,

$$\Gamma_3 = \exp jB_1k_\perp^2(\zeta-\zeta_o)-\exp jB_1k_\perp^2\zeta \tag{3.66}$$

Similarly,

$$\Gamma_4 = \exp jB_2k_\perp^2(\zeta-\zeta_o)-\exp jB_2k_\perp^2\zeta \tag{3.67}$$

Substituting eqs. (3.66) and (3.67) into (3.61) and (3.62),

$$\Gamma_1 = \exp\left[\{jCG(\xi,\eta)/B_1D\}\cdot\left(\exp(jB_1k_\perp^2(\zeta-\zeta_o))-\exp(jB_1k_\perp^2\zeta)\right)\right] \tag{3.68}$$

$$\Gamma_2 = \exp\left[\{jCG(\xi,\eta)/B_1D\}\cdot\left(\exp(jB_2k_\perp^2(\zeta-\zeta_o))-\exp(jB_2k_\perp^2\zeta)\right)\right] \tag{3.69}$$

Under Markov random process approximation,

$$A_c(\vec{\rho}) = 2\pi\int\int_{-\infty}^{+\infty}\Phi_\varepsilon(\vec{k}_\perp,0)\exp j\vec{k}_\perp\cdot\vec{\rho}d^2k_\perp$$

From eq. (3.47),

$G(\xi,\eta) = A_{\Delta N}(\xi,\eta)/l_o\left\langle\left(\Delta N/N_o^2\right)^2\right\rangle$ then,

$$\Gamma_1 = \exp-j2\pi C\Big/B_1Dl_o\left\langle\left(\Delta N/N_o^2\right)^2\right\rangle\int\int_{-\infty}^{+\infty}\Phi_{\Delta N}(k_\perp,0)\left\{\exp jB_1k_\perp^2(\zeta-\zeta_o)-\exp jB_1k_\perp^2\zeta\right\}\cdot \exp j\vec{k}_\perp\cdot\vec{\rho}d^2k_\perp \tag{3.70}$$

$$\Gamma_2 = \exp-j2\pi C\Big/B_2Dl_o\left\langle\left(\Delta N/N_o^2\right)^2\right\rangle\int\int_{-\infty}^{+\infty}\Phi_{\Delta N}(k_\perp,0)\left\{\exp jB_2k_\perp^2(\zeta-\zeta_o)-\exp jB_2k_\perp^2\zeta\right\}\cdot \exp j\vec{k}_\perp\cdot\vec{\rho}d^2k_\perp \tag{3.71}$$

Eqs. (3.70) and (3.71) are substituted into eqs. (3.56) and (3.57) to obtain the required two-frequency mutual coherence functions.

$$\Gamma(\vec{\rho},z,\omega_o,X) = \exp(-CG(0)B_3\zeta_o + j2\pi C\Big/B_1Dl_o\left\langle\left(\Delta N/N_o^2\right)^2\right\rangle\int\int_{-\infty}^{+\infty}\Phi_{\Delta N}(k_\perp,0)\cdot \left\{\exp jB_1k_\perp^2(\zeta-\zeta_o)-\exp jB_1k_\perp^2\zeta\right\}\cdot\exp j\vec{k}_\perp\cdot\vec{\rho}d^2k_\perp \tag{3.72}$$

$$\overline{\Gamma}(\vec{\rho},z,\omega_o,X)=\exp(-CG(0)B_3\zeta_o-j2\pi C/B_2Dl_o\left\langle\left(\Delta N/N_o^2\right)^2\right\rangle\int\int_{-\infty}^{+\infty}\Phi_{\Delta N}(k_\perp,0)\cdot$$

$$\left\{\exp jB_2k_\perp^2(\zeta-\zeta_o)-\exp jB_2k_\perp^2\zeta\right\}\cdot\exp j\vec{k}_\perp\cdot\vec{\rho}d^2k_\perp \quad (3.73)$$

The transverse coordinate wave-number term has a slowly varying factor in the z-direction of the integrand and can be held constant while attention is shifted on the phase vector factor $\vec{k}_\perp\cdot\vec{\rho}$ in $\exp j\vec{k}_\perp\cdot\vec{\rho}$ of the integrand as we integrate over one of the transverse axes. The integral in terms of the phase vector factor is identifiable with the zero-order Bessel function such that in polar coordinates,

$$\int_{-\infty}^{+\infty}d\vec{k}_\perp\exp j\vec{k}_\perp\cdot\vec{\rho}=\int_0^{2\pi}d\varphi\int_0^{k_\perp}k_\perp\exp(jk_\perp\rho\sin\varphi)dk_\perp=2\pi k_\perp J_o(k_\perp\rho)/\rho \quad (3.74)$$

where $J_o(k_\perp\rho)$ is the zero-order Bessel function.

Therefore,

$$\Gamma(\vec{\rho},z,\omega_o,X)=\exp(-CG(0)B_3\zeta_o+j4\pi^2C/B_1Dl_o\left\langle\left(\Delta N/N_o^2\right)^2\right\rangle\int_{-\infty}^{+\infty}k_\perp\Phi_{\Delta N}(k_\perp,0)\cdot$$

$$\left\{\exp jB_1k_\perp^2(\zeta-\zeta_o)-\exp jB_1k_\perp^2\zeta\right\}\cdot J_o(k_\perp\rho)dk_\perp/\rho \quad (3.75)$$

$$\overline{\Gamma}(\vec{\rho},z,\omega_o,X)=\exp(-CG(0)B_3\zeta_o-j4\pi^2C/B_2Dl_o\left\langle\left(\Delta N/N_o^2\right)^2\right\rangle\int_{-\infty}^{+\infty}k_\perp\Phi_{\Delta N}(k_\perp,0)\cdot$$

$$\left\{\exp jB_2k_\perp^2(\zeta-\zeta_o)-\exp jB_2k_\perp^2\zeta\right\}\cdot J_o(k_\perp\rho)dk_\perp/\rho \quad (3.76)$$

3.3 Simulation of GPS User's Position Fixing

This section presents basic concepts on how GPS receiver determines its position. Under a clear view, at any time and any location on the earth, a GPS receiver should have a direct line of sight (LOS) and be able to receive signals from four to twelve satellites (Carrano, 2007). Most often, a GPS satellite will receive

signals from more than four satellites. Since four satellites are the minimum required to obtain the user position, this arrangement provides the user position at any time and any location. For a constellation with 5° elevation mask angle, more than 80% of the times, seven or more satellites are in view (Kintner, 1999).

3.3.1 GPS Pseudorange Derivation

GPS satellites are configured primarily to provide the user with the capacity of determining his position in terms of latitude, longitude and altitude (Hofmman-Wellenhof et. al, 1997; and Smith and Davies, 2000). This is accomplished by the simple re-section process using the distances between the receiver and the satellites.

A GPS satellite at coordinate (x, y, z) transmits pulse at time t_{1T} to be received by a receiver that is located on a sphere centred at the coordinate of the satellite at time t_{1R}. The distance of the receiver from the satellite can therefore be expressed as $r_1 = c(t_{1R} - t_{1T})$. Similarly, a second satellite at coordinate (x_2, y_2, z_2) transmits pulse at time t_{2T} to be received by a receiver at time t_{2R}, such that the distance of the receiver from the satellite now becomes $r_2 = c(t_{2R} - t_{2T})$. For a receiver to observe these two satellites simultaneously, the user must be located on both spheres. By implication, the user location is restricted to the intersection of the two spheres, a circle (Kintner and Ledvina, 2005). Next, a third satellite at location (x_3, y_3, z_3) transmits pulse that is received and restricts the user on a third sphere. However, for simultaneous observation of the three satellites by a single receiver, the receiver must be located on the intersection of the three spheres that defines the coordinate of each satellite. The intersection leads to a set of three simultaneous

quadratic equations that must be solved within the receiver (Hofmman-Wellenhof et. al, 1997; and Tsui, 2000).

GPS receivers apply a slightly different technique. They typically use an inexpensive crystal clock which is set approximately to GPS time. Thus, the clock of the ground receiver is offset from the true GPS time, and because of this offset, the distance measured to the satellite from the receiver is slightly varied from the "true" range. The receiver overcomes this by simultaneous measurement of the distances (pseudoranges) to four satellites as against three satellites earlier mentioned, and the fourth unknown variable becomes the receiver clock bias (Hofmman-Wellenhof et. al, 1997).

The pseudorange can be defined as:

$$P = \rho + c(\delta_S - \delta_R) + \delta_{error} \qquad (3.77)$$

where ρ is the true range and δ_{error} is the error term. The pseudorange has two primary error sources namely, receiver clock offset and satellite clock offset. In addition to these errors, the error term, δ_{error} comprising (i) ionospheric delay error (ii) trospospheric delay error, (iii) multipath effects, play very vital role as sources of challenges to transionospheric radio propagation (Tsui, 2000). The satellite clock error is monitored by the control segment and modeled with three coefficients that are included in the navigation message so that the satellite clock offset can be calculated within the receiver. The receiver clock offset is determined in the process of navigation.

The real range ρ is the distance from the ith satellite to the receiver. Denoting satellite position as (x^i, y^i, z^i) and the receiver position as (x, y, z). Recall that the

satellite position is determined by the receiver from the ephemerides in the navigation message (Kintner and Ledvina, 2005). If the error term is neglected, eq. (3.77) can therefore be expressed as;

$$P^i - c\delta_S^i = \rho_i + c\delta_R \tag{3.78}$$

where,

$$\rho_i = \sqrt{(X^i - X)^2 + (Y^i - Y)^2 + (Z^i - Z)^2} \tag{3.79}$$

It is obvious that the right side of eq. (3.78) contains the four unknown variables X, Y, Z and δR. Hence, to solve for these variables, a minimum of four satellite links are required to yield four equations (for valid navigation solution) (Kintner, 1999; and Thomas et al., 2001). Eq. (3.79) being nonlinear can only be solved by numerical root finding technique. Multi-dimensional Newton-Raphson method with a reasonable guess of the initial receiver position provides excellent numerical approach for solving eq. (3.79) (Kintner and Ledvina, 2005). For an initial guess defined by X_0, Y_0, Z_0.

The initial guess for the true range is:

$$\rho_0^{(i)} = \sqrt{(X^i - X_0)^2 + (Y^i - Y_0)^2 + (Z^i - Z_0)^2} \tag{3.80}$$

3.3.2 Numerical Technique

Numerically, one of the general methods normally adopted in solving a function defined by $f(x)=0$ is the Newton-Raphson method. Starting from first principle, first-order Taylor expansion can be algebraically solved to arrive at this method,

$$f(X) \approx f(X_n) + (X_{n+1} - X_n) f'(X_n) = 0 \tag{3.81}$$

Therefore,

$$X_{n+1} = X_n - f(X_n)/f'(X_n) \qquad (3.82)$$

Eq. (3.82) is the Newton-Raphson formula. It implies that Newton-Raphson method is a solution of first-order Taylor expansion. In three-dimensional form, eq. (3.81) can be written as,

$$f(X_{n+1},Y_{n+1},Z_{n+1}) \approx f(X_n,Y_n,Z_n) + \Delta X \frac{\partial f}{\partial X}\bigg|_{X_n,Y_n,Z_n} + \Delta Y \frac{\partial f}{\partial Y}\bigg|_{X_n,Y_n,Z_n} + \Delta Z \frac{\partial f}{\partial Z}\bigg|_{X_n,Y_n,Z_n} = 0 \qquad (3.83)$$

where,

$$\Delta X = (X_{n+1} - X_n), \Delta Y = (Y_{n+1} - Y_n) \text{ and } \Delta Z = (Z_{n+1} - Z_n). \qquad (3.84)$$

Unlike eq. (3.81), eq. (3.83) cannot be easily solved because there are three unknowns in one linear equation. In solving this equation, we must establish a set of three linear equations that must be solved simultaneously. As defined in eq. (3.79), $\rho^i = \sqrt{(X_i - X_0)^2 + (Y_i + Y_0)^2 + (Z_i + Z_0)^2}$. Following the principle established in eq. (3.83), we can expand eq (3.79), such that,

$$\rho^i \cong \rho_0^i + \Delta X \frac{\partial \rho^i}{\partial X}\bigg|_{X_0,Y_0,Z_0} + \Delta Y \frac{\partial \rho^i}{\partial Y}\bigg|_{X_0,Y_0,Z_0} + \Delta Z \frac{\partial \rho^i}{\partial Z}\bigg|_{X_0,Y_0,Z_0} = 0 \qquad (3.85)$$

Substituting eq.(3.85) into eq. (3.79) and rearranging, then;

$$\rho^i - \rho_0^i = \frac{\Delta X(X_i - X_0)}{\rho^i} + \frac{\Delta Y(Y_i - Y_0)}{\rho^i} + \frac{\Delta Z(Z_i - Z_0)}{\rho^i} + c(\delta_S^i - \delta_R) \qquad (3.86a)$$

Assuming the user bias error is represented as E and the change in pseudorange $\rho^i - \rho_0^i - c\delta_s^i$ as b_i. We can then re-write eq. (3.86a) as;

$$a_{i1}\Delta X + a_{i2}\Delta Y + a_{i3}\Delta Z + E = b_i \qquad (3.86b)$$

where $a_{i1} = \frac{X_i - X_0}{\rho^i}$, $a_{i2} = \frac{Y_i - Y_0}{\rho^i}$ and $a_{i3} = \frac{Z_i - Z_0}{\rho^i}$ (3.87)

For a simple case whereby we have four satellites in the orbital plane, eq. (3.86b) can be expressed in matrix form as;

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & -1 \\ a_{21} & a_{22} & a_{23} & -1 \\ a_{31} & a_{32} & a_{33} & -1 \\ a_{41} & a_{42} & a_{43} & -1 \end{bmatrix} \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \\ E \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \tag{3.88}$$

The solution of eq. (3.88) is,

$$\begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \\ E \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & -1 \\ a_{21} & a_{22} & a_{23} & -1 \\ a_{31} & a_{32} & a_{33} & -1 \\ a_{41} & a_{42} & a_{43} & -1 \end{bmatrix}^{-1} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \tag{3.89}$$

Eq. (3.89) is an approximate solution of eq. (3.88). In order to achieve a greater accuracy, an iterative approach is required whereupon the most recent solution becomes the initial guess and the entire process is repeated until the desired accuracy is obtained. This is usually done by setting a pre-determined threshold value for the iteration. A quantity defined in eq. (3.90) below is often set aside for this purpose (Tsui, 2000).

$$\delta v = \sqrt{\Delta X^2 + \Delta Y^2 + \Delta Z^2 + E^2} \tag{3.90}$$

The value of E in eq. (3.90) is usually very small and can be sometimes neglected. The discussion so far is on four satellite system. In a situation whereby we have more than four satellites, some modifications must be made on the approach above such that,

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & -1 \\ a_{21} & a_{22} & a_{23} & -1 \\ a_{31} & a_{32} & a_{33} & -1 \\ a_{41} & a_{42} & a_{43} & -1 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ a_{n1} & a_{n2} & a_{n3} & -1 \end{bmatrix} \cdot \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \\ E \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ . \\ . \\ . \\ b_n \end{bmatrix} \tag{3.91}$$

In a concise form, $b = a\Delta l$, where b and Δl are vectors. Matrix-a in eq. (3.91) is obviously not a square matrix, therefore, $\Delta l \neq a^{-1}b$. The problem is over-determined. In situations like this whereby we have more equations than unknowns in a set of linear equations, the least squares method is usually employed to find the solutions (Kintner and Ledvina, 2005; and Tsui, 2000). This is done by using the pseudo-inverse of matrix-a to obtain the solution (Tsui, 2000).

Therefore,

$$\Delta X = \left[a^T a\right]^{-1} a^T \Delta b \tag{3.92}$$

where a^T is the transpose of the a-matrix. From eq. (3.91), values of ΔX, ΔY and ΔZ can be obtained. On a general note, the least squares method provides better accuracy for GPS ranging than when just four satellites are used because more data are involved (Hofmann-Wellenhof et al., 1997; and *Özlüdemir*, 2004).

For better practical applications, it is often necessary to convert the determined user position in cartesian coordinate system to the corresponding spherical coordinate system and finally to the earth coordinate system (latitude (L), longitude (l), and altitude (h)).

Defining,

$$r = \left[X_u^2 + Y_u^2 + Z_u^2\right]^{\frac{1}{2}} \tag{3.93}$$

$$L_c = \tan^{-1}\left(\frac{Z_u}{\left[X_u^2 + Y_u^2\right]^{\frac{1}{2}}}\right) \tag{3.94}$$

$$l = \tan^{-1}\left(Y_u / X_u\right) \tag{3.95}$$

where L_c is the geocentric latitude. In real sense, the surface of the earth is not a perfect sphere, therefore, there is need to take the shape of the earth into consideration. The geodetic latitude, L can be calculated from L_c by the equation stated below.

$$L = L_c + e_p \sin 2L \tag{3.96}$$

where e_p is the ellipticity of the earth.

The altitude can be evaluated by,

$$h = \left[X_u^2 + Y_u^2 + Z_u^2\right]^{\frac{1}{2}} - a_e(1 - e_p \sin^2 L) \tag{3.97}$$

These parameters are the desired user position in the earth coordinate system

3.3.3 Computational Algorithm

An arbitrary initial guess user position and receiver clock bias are chosen in an earth centre earth fixed condition (Tsui, 2000). The satellite positions are given in matrix format. The pre-measured pseudorange and satellite clock bias are also given as vectors. Eq. (3.83) is thereafter used to evaluate the pseudorange ρ_i. The difference between pre-measured pseudorange and the sum of satellite clock offset and determined pseudoranges is evaluated as $d\rho_i$. In the computation, a simple case where the error term is neglected is considered.

The second part of the program carefully generates three-column elements matrix by solving equation (3.87). After this, the fourth column which is -1 throughout is introduced to the first three columns such that the matrix now becomes $n \times 4$ matrix. The reason for making the fourth column negative ones is born out of the fact that the fourth unknown variable to be determined is the receiver clock offset which has the same value for all satellite signals and pseudoranges (Kintner and Ledvina, 2005). The $n \times 4$ matrix is thereafter transposed to obtain a new $4 \times n$ matrix. The newly transposed matrix is multiplied by the former matrix to obtain an invertible 4×4 square matrix. This newly transposed matrix is also used to multiply $d\rho_i$ to obtain a 4-row vector. From this point, we employed Guassian elimination technique to determine the four unknown variables. The 4-row vector then augments the 4×4 square matrix such that a new 4×5 matrix is obtained. The entire process follows least square method (Kintner, 1999: Tsui, 2000; and *Özlüdemir*, 2004).

Having obtained the four unknown variables, there is need to update the corrected coordinates to obtain the desired accuracy. This solution being an approximation requires an iterative approach whereupon the most recent solution becomes the initial guess and the entire process is repeated until the desired accuracy is obtained (Kintner and Ledvina, 2005). An error bound defined by the square root of the sum of the square of each of the first three determined unknown variables is set to cater for unwanted divergence during iterations. If the error bound is lesser than pre-set threshold value, iteration ceases, otherwise, the process continues until the condition is attained.

The first three values obtained are the user position in Cartesian coordinate system while the fourth value represents the receiver clock offset measured in unit of distance. Under practical applications, the user position must be transformed from Cartesian coordinate system to the earth coordinate system and the receiver clock offset must be divided by the speed of light to obtain the offset in the unit of time. To convert the user position to the practical earth coordinate system, it must be initially converted to spherical coordinate system which transforms the determined values in terms of geodetic latitude, longitude and altitude. The earth not being a perfect sphere requires that the user geodetic latitude be converted to the practical flat earth geometry so that the user position can be recorded in true latitude, longitude and altitude (Tsui. 2000).

The FORTRAN code for the computation and a sample output is shown in appendix A4. For a sample running of the code for eleven satellites, we chose the initial guess for the user position and receiver clock offset as zeros. The satellite parameters are shown in Table 3.1.

Table 3.1: The Satellite parameters for the Simulation

SV	Satellite Position coordinates	Satellite Pseudoranges (m)	Satellite Clock offset (ns)
SV1	20192000, 20305000, 21000000	43008000	8000000
SV2	20300000, 20195000, 20200000	44075400	9000000
SV3	20500000, 20192000, 20305000	45900000	13200000
SV4	20500000, 20305000, 20192000	46001000	10000000
SV5	20401000, 20305000, 20200000	44058010	8540000
SV6	20305000, 20192000, 20300000	45695400	9810000
SV7	20603000, 20200000, 20192000	46750500	8000000
SV8	20305000, 20185000, 20300090	45980000	12500000
SV9	20600000, 20195000, 20200200	43985000	8850000
SV10	20510000, 20192000, 20003000	46700000	12200000
SV11	20401000, 20205000, 20300000	45050000	10000000

The threshold bound to regulate the iteration was chosen as 1000.

3.4 Measurements of Some Ionospheric Observables

3.4.1 Equipment and Data Resources

Our previous investigations have exclusively dwelt on numerical modeling. In this section, our interest lies in the measurements of ionospheric parameters with the aid of GPS Ionospheric Scintillation and Total Electron Content (TEC) Monitors (GISTMs) located at three stations within the equatorial (Appleton) anomaly. These stations are; Nairobi (KENYA), Ascension Island and Garmin 480 WAAS certified GPS receiver for aircraft navigation with vertical guidance located at Sao Jose dos Campos (BRAZIL), all under the coordination of US Air Force Research Laboratory.

Each GISTM comprises Ashtech Z-12 11-channel dual frequency (L1 (1.5754 GHz) and L2 (1.2276 GHz)) coarse acquisition (C/A) code receiver version with OEM4 card (details on the receiver hardware are provided in Van Dierendonck et al., 1993; and Van Dierendonck 2005). The receiver is modified with a software (firmware), and is capable of tracking up to fourteen GPS satellites simultaneously. It monitors amplitude scintillation (at 50 Hz rate), and computes S_4 index (standard deviation of the receiver power normalized by its mean value (Briggs and Parkin, 1963; and Van Dierendonck et al., 1993)), and the code carrier divergence (at 1 Hz rate) for each satellite being tracked on L1 frequency. The TEC from the combined L1 and L2 pseudoranges and carrier phase measurements are also recorded (Carrano, 2007). The recorded data are stored in two CD-ROMs that accompanied the receiver. These CD-ROMs are: Ashtech Z-12 CD-ROM (for logging raw

observables) and a GPS Silicon Valley (GSV) CD-ROM (for logging extracted binary codes data into ASC II format) (Van Dierendonck, 2005).

The raw amplitude measurements are detrended (normalized) with sixth-order (pole) Butterworth low-pass filter output (with a receiver user-specified cut off frequency: in our case, 0.1 Hz, 3dB are used). The total amplitude scintillation index S_4, which includes S_4 due to effects of ambient noise (multipath) is computed over 60-second interval. The corrected S_4 index is also computed by averaging the signal-to-noise density (C/No) over the same 60-second interval (Van Dierendonck et al., 1993). In all the cases presented in this work, only the signals coming from the satellites with an elevation angle greater than 30° were taken into consideration for TEC measurements (Carrano and Groves, 2006).

Other data logged by the system include the location geographic coordinates, the receiver's signal lock status, satellite's angle of view (azimuth and elevation), and pseudorandom noise code (PRN) (Carrano, 2007). The receiver firmware provides opportunity to log-transfer ionospheric parameters in real-time via low bandwidth internet connections amongst coordinated remote stations (Carrano and Groves, 2006; and Carrano, 2007).

Overall, three different data sets were considered. The first set of data comprised of measurements of TEC at Nairobi, Kenya station for three consecutive days in December, 2008. At the Ascension Island and Sao Jose dos Campos sites, impacts of scintillation on carrier signals were considered. Lastly, we looked at and analyzed spread F measurements from Jacamarca on a phenomenal day.

3.5 GPS and Aviation Safety

In recent times all over the world, the volume of air traffic in the sky has increased tremendously leading to over-bearing impact on existing infrastructure. Many at times, controllers are left with the only option of managing traffic on first come, first served basis: stack and hold airplanes at low altitudes on arrivals, or at times diverted, only to be given slots to land in turns. The consequential effects of this pattern are: fuel wastage, environmental pollution, and ineffective flight scheduling. Another worrisome incidents in the past are the situations whereby airplanes are declared missing in an airspace, only to be located after many days when all occupants have perished, or in some worst situations, not seen at all, even when seen on time, mobilization and coordination of emergency resources is often very poor. Our intention is to use GPS-based initiatives to address these problems.

3.5.1 Airspace Grid Referencing

In this section, a case study of Nigerian airspace is adopted with the nation's twenty one airport applied. Figure 3.2 shows a model of Nigerian airspace into thirty grids that should be strategically located in all control towers across the country. The intention is to have precise reference to any part of the airspace at any point in time in a manner that will enhance swift response to aeronautical emergencies, and at the same time establish operational framework that will enhance efficient flight routing.

The application of the grid calibration is based on vertical and horizontal coordinate system. The vertical coordinate is scaled in alphabets while the horizontal coordinate is scaled in numerals. With this calibration, it can be clearly seen that Jos airport is located on grid C4 while Yola airport is on D6. As airplanes coast on the

sky, the airport that has control over the airplane can easily see and know the location of the airplane until it is handed over to another controller in another airport as the airplane cruises over. This process continues until the airplane reaches its final destination. Assuming a flight schedule from Lagos to Maiduguri has problem at a point marked X and finally crashed, the point can be easily referred to as grid B3 so that aeronautical search and rescue teams can be dispatched from neighbouring airport(s) (in this case, Abuja and Markudi) to the scene (Fire Service Manual, 2006).

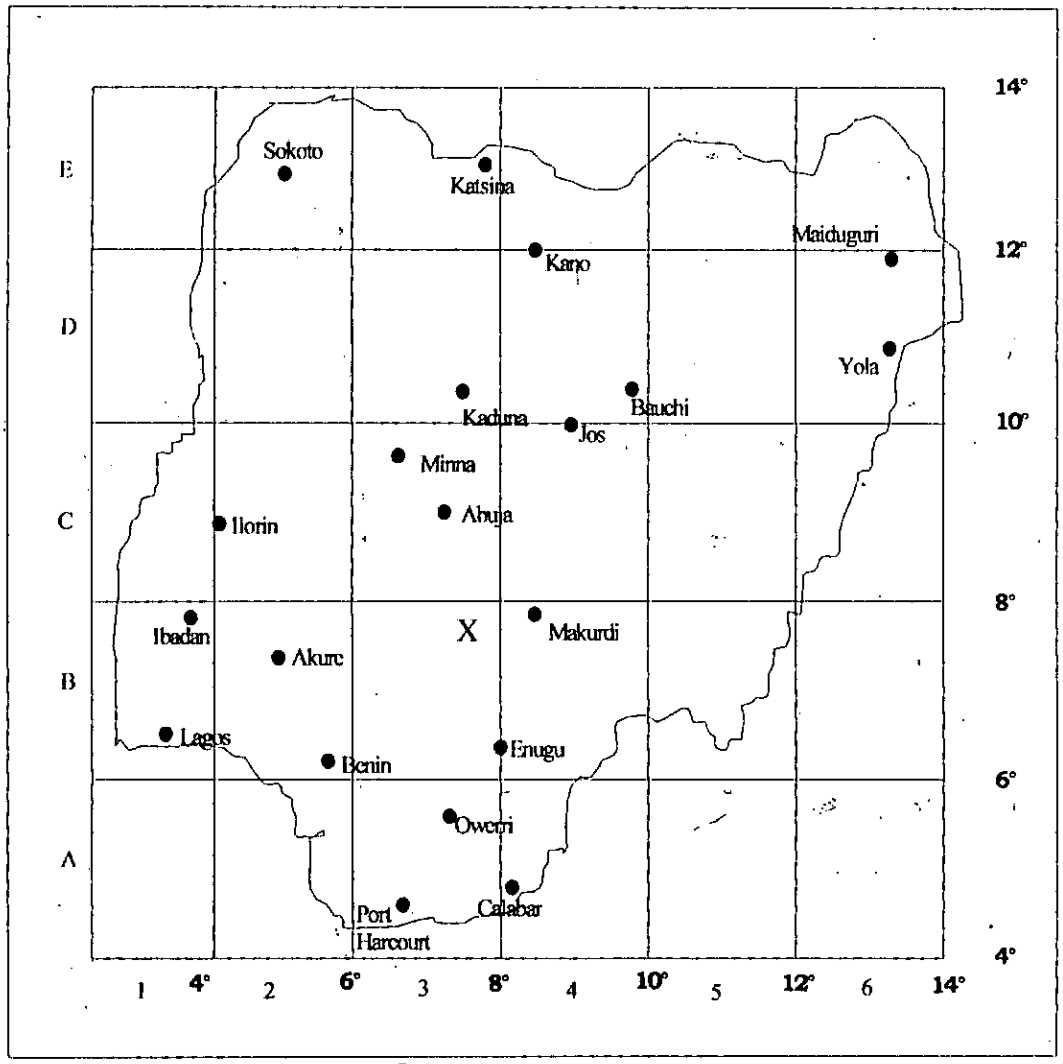


Figure 3.2: A Grid Model of Nigerian Airspace

CHAPTER FOUR

4.0 RESULTS

This chapter presents analyses of the acquired results. Results for each component of the numerical modeling are analyzed and discussed separately. The discussion opens with a model to explain the growth and evolution of ionospheric irregularities and another on the effects of ionospheric irregularities on GPS infrastructure, and what can be done to mitigate or totally eliminate such effects.

Results of some measured ionospheric observables at some stations within the equatorial anomaly region were presented and discussed. For clearer elucidation of the working principles of GPS receiver, a numerical simulation of how it (GPS receiver) determines the position of its user was presented and explained. Finally, how a GPS-driven technique can be used to enhance aviation safety, with interest in Nigeria environment was also discussed.

4.1 Ionospheric Irregularities Growth and Evolution Map

The numerical evaluation of eq. (3.8) from the nonlinear model, using iterative process yields the map shown in Figure 4.1. Eq. (3.8) defines the gradient of ionospheric irregularities drift (Basu, 1978), and p in eq. (3.8) is the parameter related to the strength of the ionospheric irregularities. This parameter was varied from 0.1 to 2.0 at incremental steps of 0.1. For each p parameter, ten thousand iterations at an initial guess of 0.01 were carried out with the last hundred retained and the first nine thousand nine hundred discarded. Finally, the gradient of the irregularities drift was plotted against p-parameter.

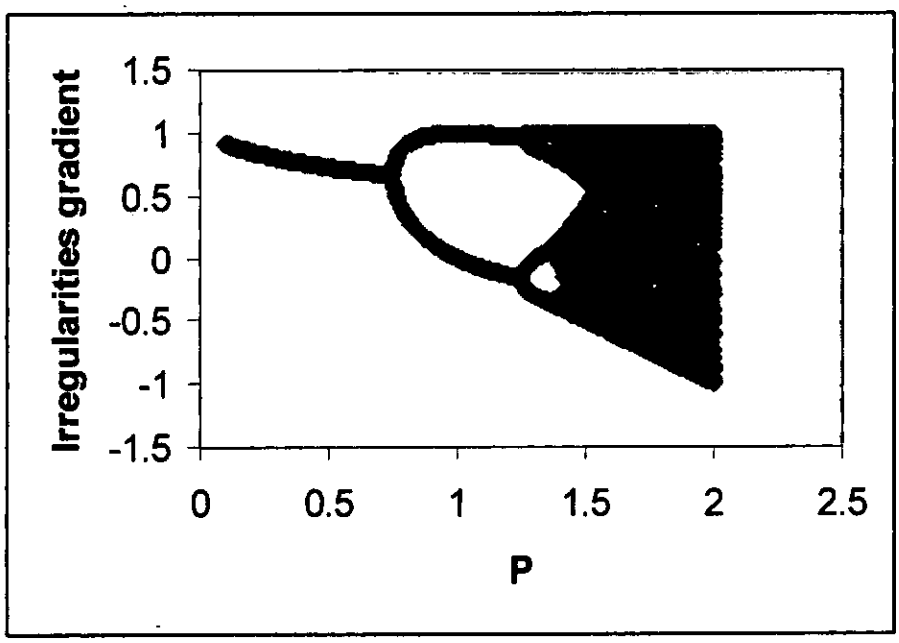


Figure 4.1: Map showing Nonlinear Regime of the Evolution of Ionospheric Irregularities

4.2 Phase Screen Model and Mutual Coherence Functions for two different Ionospheric Thicknesses

The numerical technique used in solving integrals of eqs. (3.75) and (3.76) is the trapezoidal rule. Some basic features that are believed to be realistic based on some recent *in-situ* measurements are included in the numerical evaluations. These measurements indicate that the power spectrum for the irregularities follows inverse power-law dependence in wave number, that the electron density fluctuation falls between 1% and 20% and that the irregularity slab may extend to several hundred kilometers (Rufenach, 1972; Dyson et al., 1974; Liu and Yeh, 1975; Gherm et al., 2000, 2005; Ray et al., 2006). In the computation, the ionospheric irregularity is assumed to be characterized by the inverse power-law spectrum of the form $\Phi(k_{\perp},0) \propto \left(1+k^2{}_{\perp}/K^2\right)^{-p/2}$, where p is the spectral index which for most practical

situations is of the order of 3 – 4, and has been arbitrarily taken as 4 for the purpose of this computation. K is related to the irregularity inner scale size l_o by $K = 2\pi/l_o$ (Zernov, 2006). Small values of $k_\perp$, of the order of 0.001 and transversal distance of 300 metres are used for the computation.

Normalized parameters C, ζ_o, ζ etc are also used in the computation. The set of values $C = 0.498$, $\zeta_o = 0.212$, $\zeta = 1.484$; $C = 3.872$, $\zeta_o = 0.017$, $\zeta = 0.116$, corresponding to ionospheric parameters L=50km, $l_o = 300m$, $f_p = 10MHz$, $\left\langle (\Delta N/N_o)^2 \right\rangle^{1/2} = 1 - 20\%$ and z=350km for wave frequencies f=125 MHz and f=1.6 GHz respectively. The choice of plasma frequency as high as 10 MHz is used to simulate a worst propagation condition when the ionosphere is highly disturbed (sufficient to impose scintillations on traveling signal). For the height z=350km, this value coincides with the F2-peak height, which of course, has practical importance in the study of radio propagation through the ionosphere. The operating frequencies are taken as 125 MHz and 1.6 GHz so that the effect of fluctuating ionosphere (scintillation) on two different propagation bands (VHF and L-band) can be investigated. Figures 4.2 and 4.3 show the result from the computational evaluations.

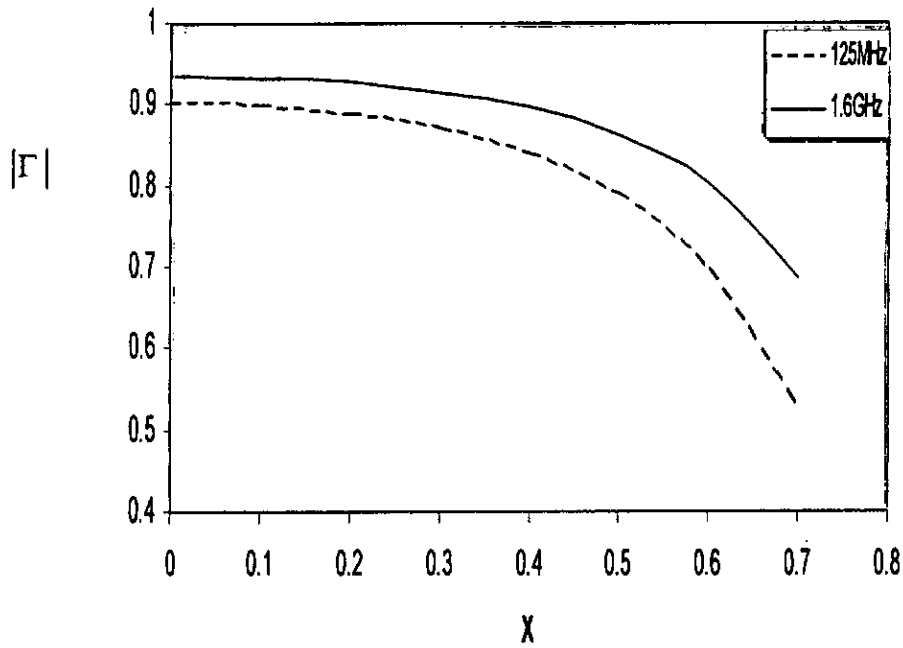


Figure 4.2: Mutual Coherence Function as a Function of Fractional Frequency Separation, at 125 MHz and 1.6 GHz Operating Frequencies.

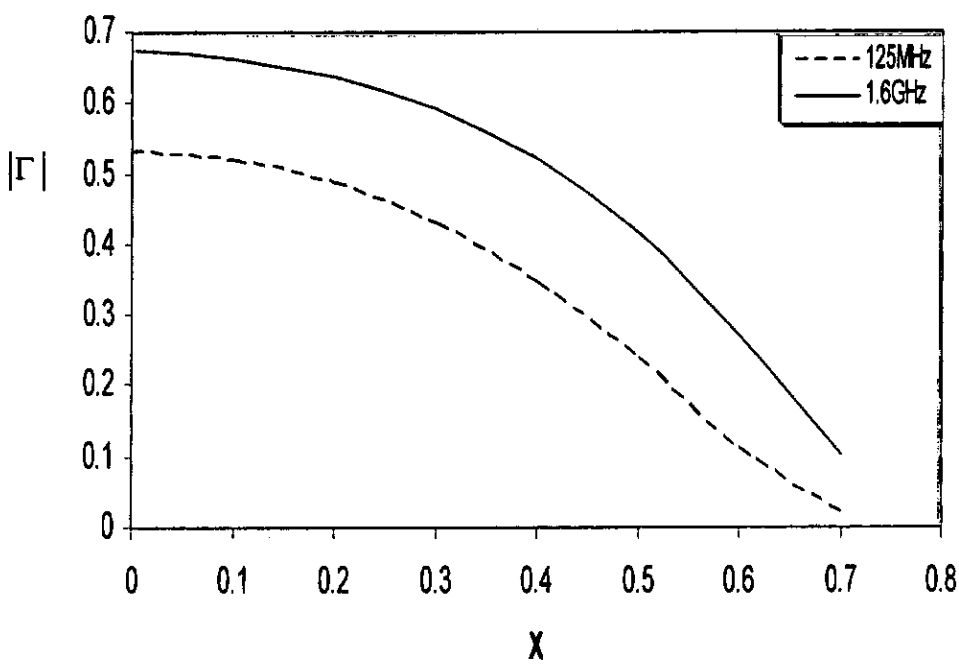


Figure 4.3: Mutual Coherence Function as a Function of Fractional Frequency Separation, at 125 MHz and 1.6 GHz Operating Frequencies for an Increased Slab Thickness.

4.3 Experimental Observation of Ionospheric Data

4.3.1 Spread F Records

The records of Spread F measured at Jacamarca (PERU) were presented. These records were obtained by the Jacamarca radar on the 17th October, 1994. The records have been calibrated to observe the behaviour of the ionosphere in terms of Spread F irregularities at different altitudes and local time, and also coded for the interpretation of carrier-to-noise ratio (C/No), i.e the C/No that the ionosphere at these altitudes and time will impose on transient signals. When the impact of this carrier-to-noise ratio is too severe, navigation service delivery using GPS-based infrastructure may be jeopardized. These records are shown in Figure 4.4.

4.3.2 Total Electron Content (TEC)

Figures 4.5 (a, b and c) show verticalized total electron content (TEC) measurements collected by Ashtech Z-12 GPS receiver that is located at Nairobi, Kenya. These records were observed for three consecutive days in December, 2008. With these records, one can conveniently observe ionospheric electron density gradients distribution at different hours of the day. Ionospheric irregularities are fluctuations that are often superimposed upon the background (smooth) electron density distribution. When the total electron content is large, there is the potentiality that the ionospheric irregularities will be more intense, and density fluctuations will be large if there is abundance of density in the background. When this happens, the ionosphere is compelled to impact negatively on GPS-based navigation system, but this impact becomes consequential when it stresses the receiver, so that it loses lock on too many satellite channels. These records will be interpreted and analyzed later.

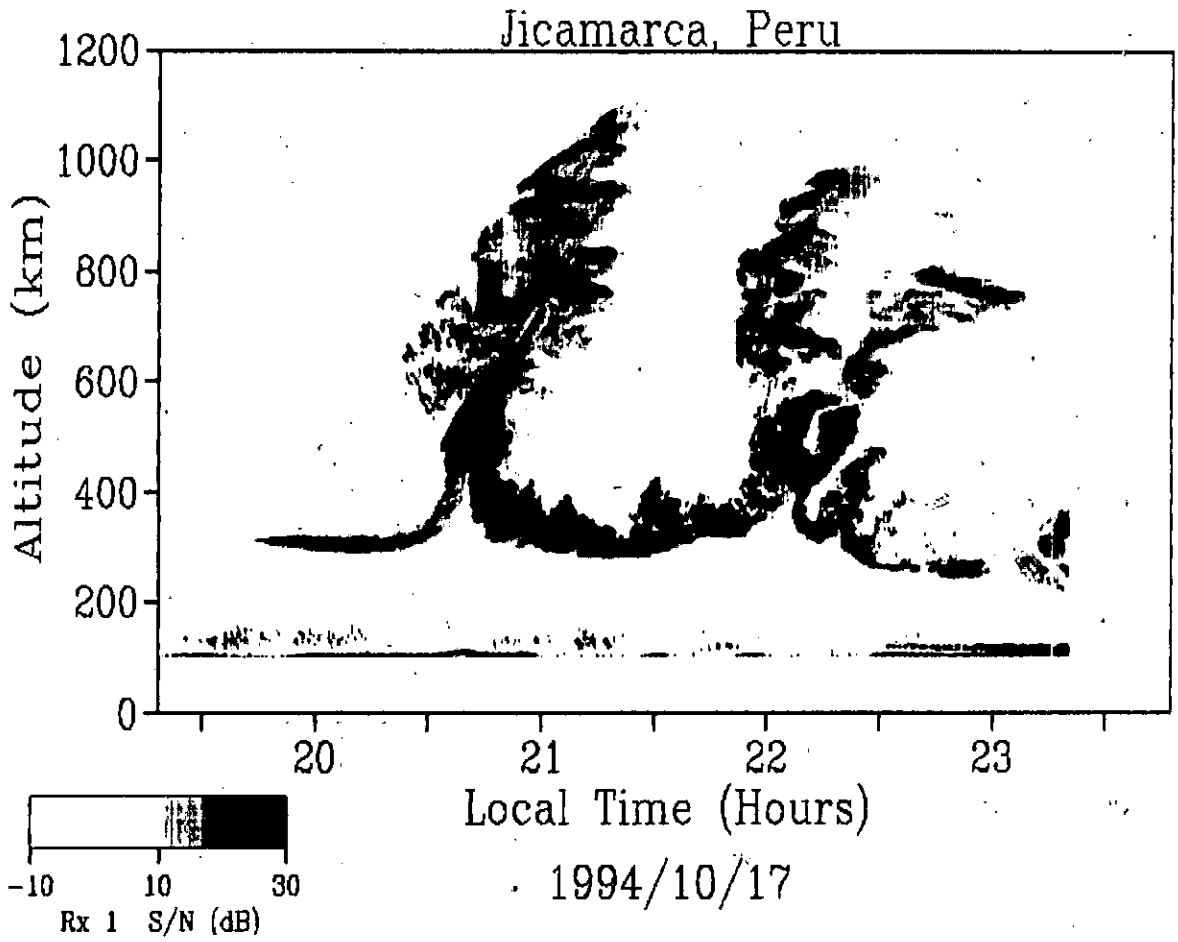


Figure 4.4: Spread F Occurrence on the 17th October, 1994 at Jacamarca, Peru.

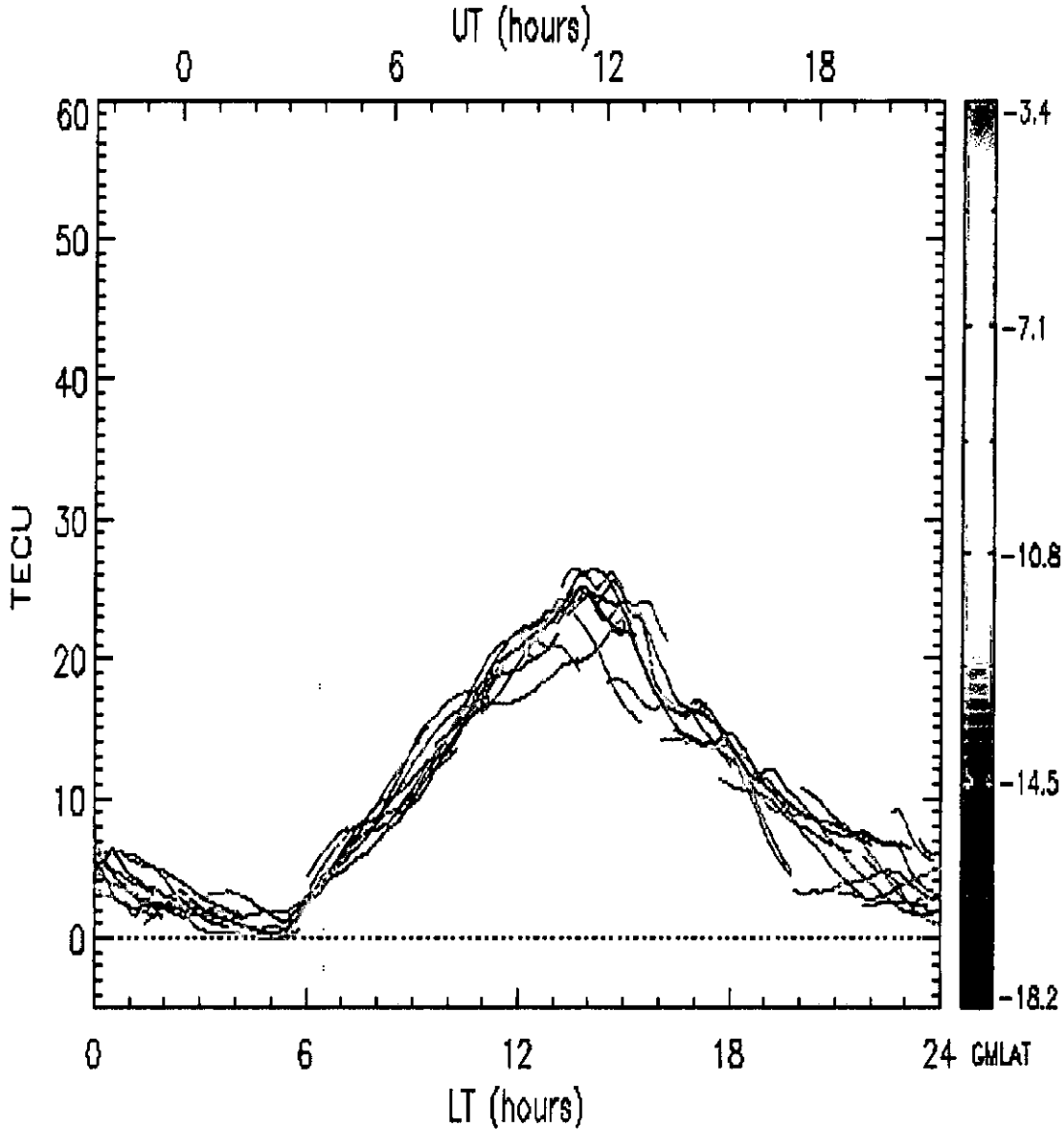


Figure 4.5 (a): Verticalized TEC mesurements collected by Ashtech Z-12 GPS receiver at Nairobi, Kenya on the 1st December, 2008.

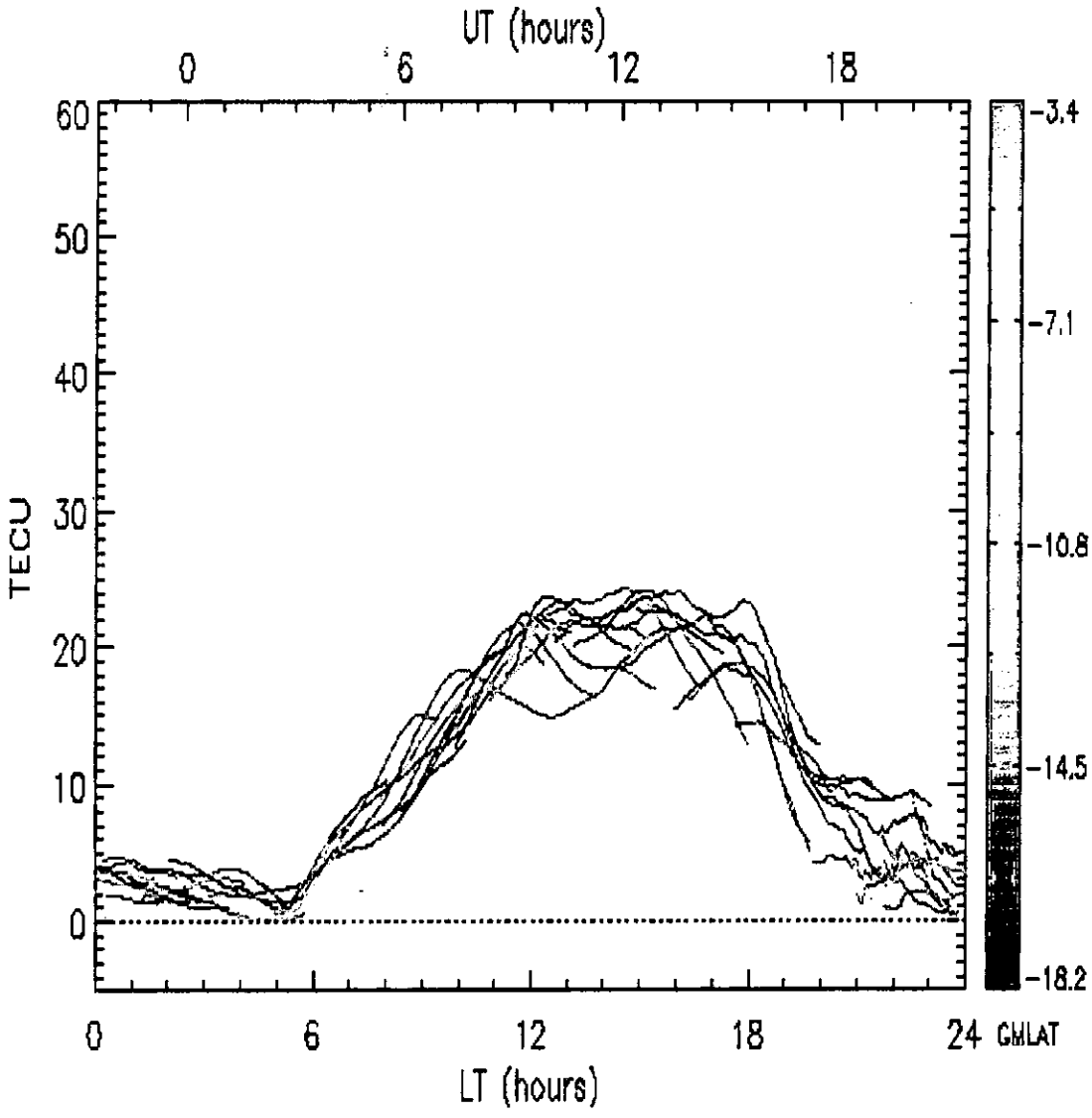


Figure 4.5 (b): Verticalized TEC measurements collected by Ashtech Z-12 GPS receiver at Nairobi, Kenya on the 2nd December, 2008.

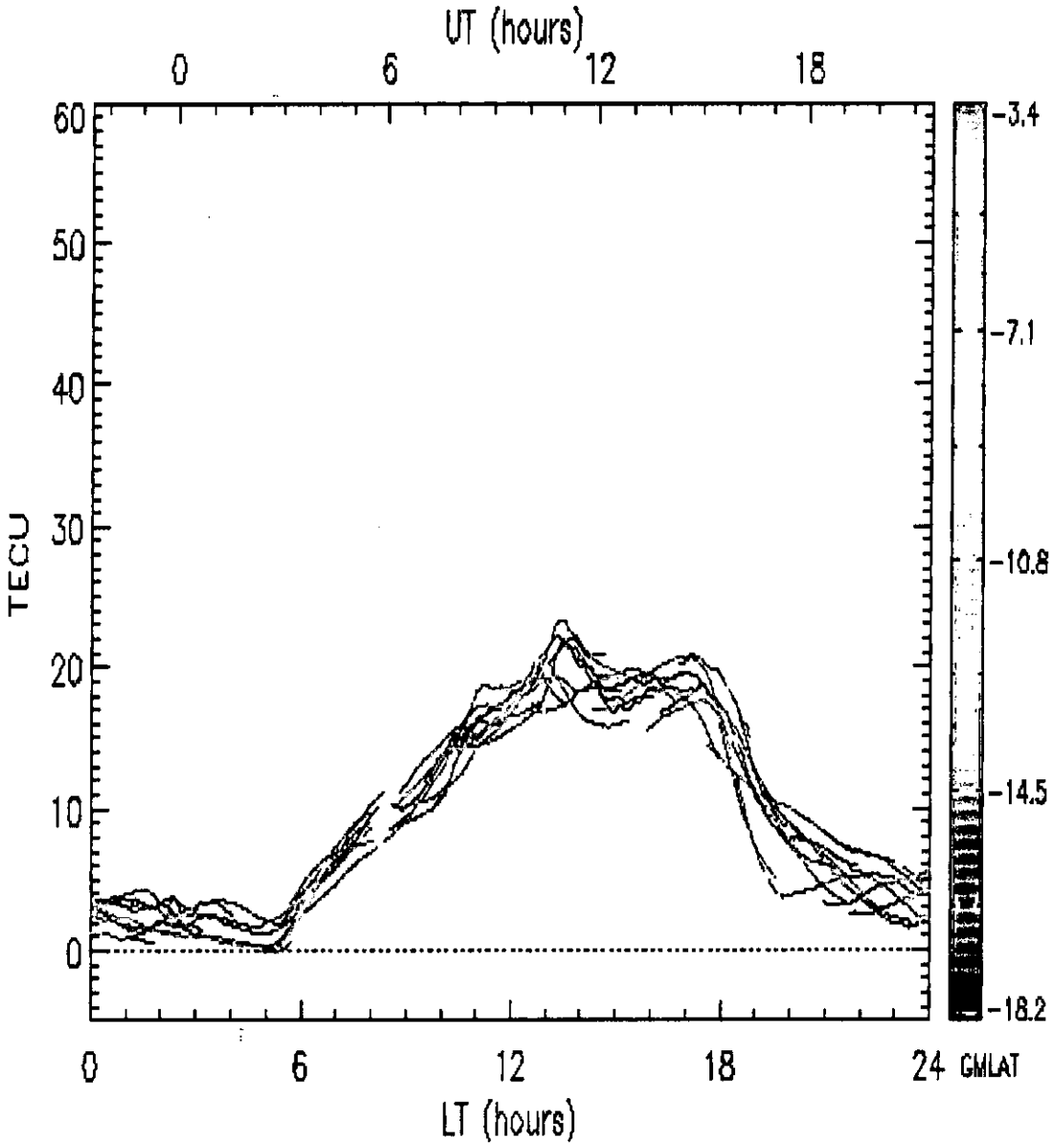


Figure 4.5 (c): Verticalized TEC mesurements collected by Ashtech Z-12 GPS receiver at Nairobi, Kenya on the 3rd December, 2008.

4.3.3 Carrier-to-Noise Ratio (C/No)

This subsection presents carrier-to-noise ratio (C/No) measured using Ashtech Z-12 GPS receiver at Ascension Island (South Atlantic Ocean) on the 16th March, 2002 (PRN 08) and 18th March, 2001 (eight PRNs) respectively, i.e. during solar maximum scintillation event. Also, measurement of carrier-to-noise ratio (C/No) at Sao Jose dos Campos, Brazil on the 5th/6th December, 2006 (PRN 29), during solar minimum scintillation event was presented. Impacts of carrier-to-noise (C/No) induced by the ionosphere on transient signals become consequential when the amplitude scintillation index is near or equal to unity.

The measured carrier-to-noise ratio (C/No) is a measure of the modulation effect that the ionosphere will impose on signals. The records of 18 March, 2001 were similar to that of 16 March, 2002, but expanded for short duration of 100 seconds as opposed to normal after sunset to midnight measurement. These records are shown in Figures 4.6-4.8. This is to make observation of all satellites in view (all PRNs) possible for the 100 seconds to see how many satellites lose lock to deep fading (C/No greater than 22dB-Hz), and the corresponding time for their re-acquisition.

When too many satellites' locks are lost simultaneously for a protracted time, severe threat will be issued on GPS-based navigation service delivery. In other words, a quick re-acquisition of lost signal's lock by GPS receiver for aviation applications is necessary to sustain service continuity and availability. These records will be discussed in details later.

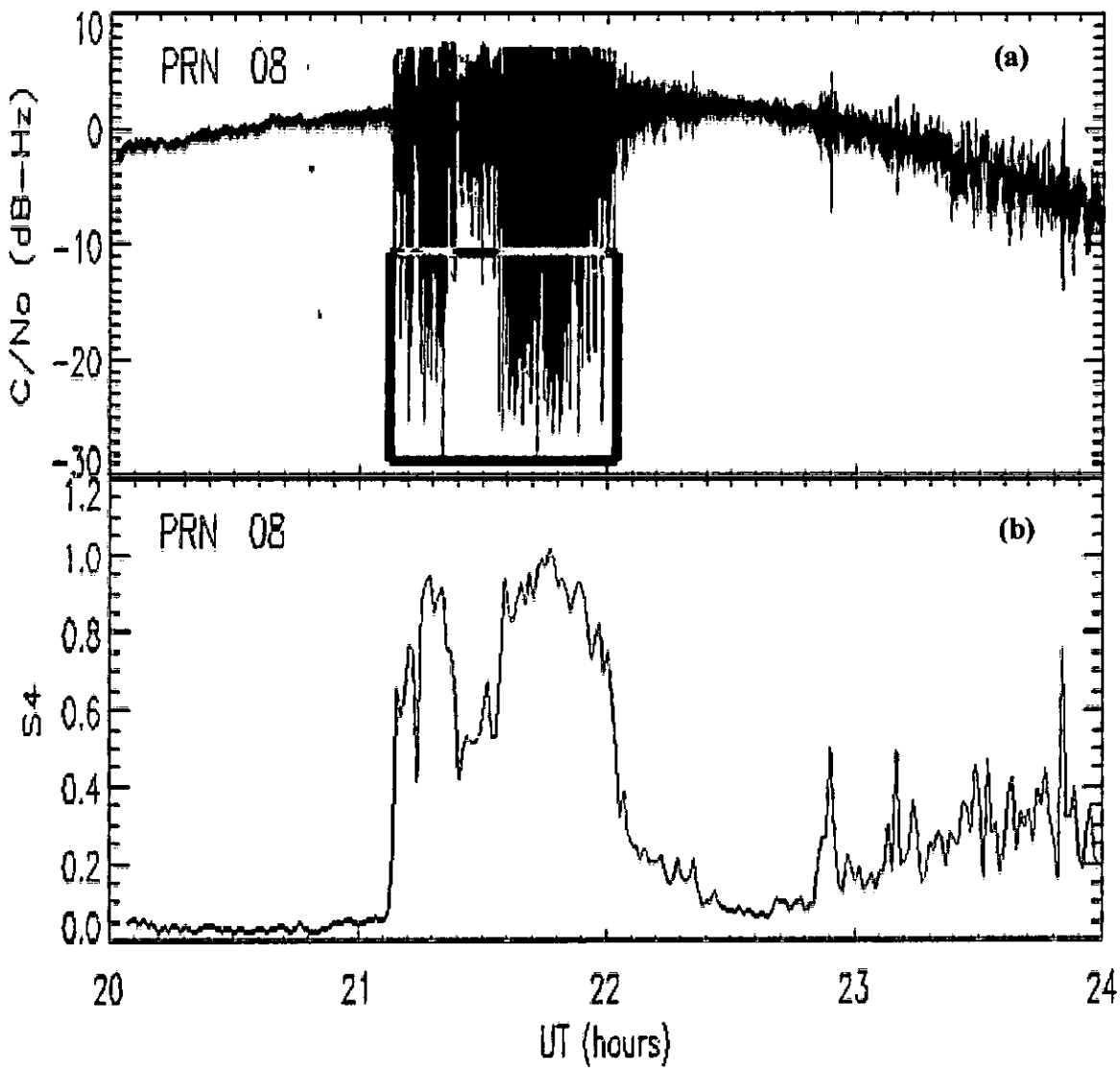


Figure 4.6: (a) Top panel: Carrier (signal) to noise ratio, **(b) bottom panel:** corresponding scintillation index records during deep scintillation event at Ascension Island on the 16th March, 2002.

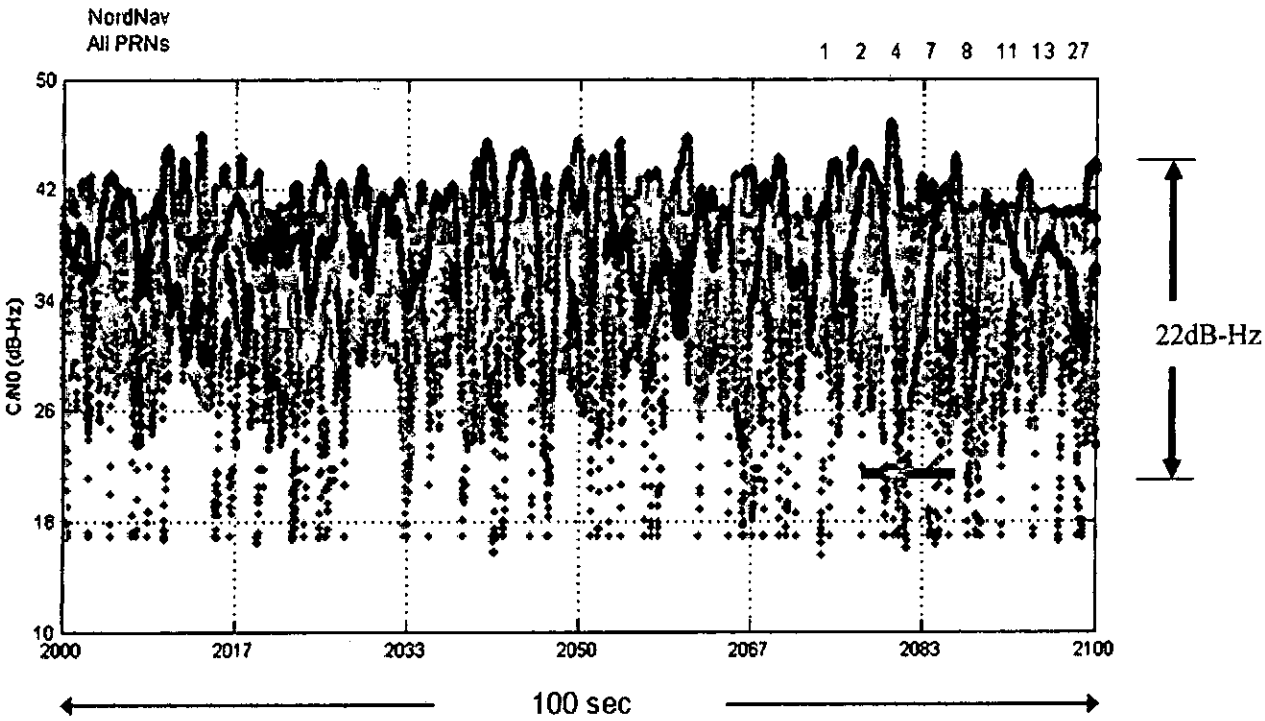


Figure 4.7: Carrier (signal)-to-noise ratio during deep scintillation event at Ascension Island on the 18th March, 2001 for 100 seconds time duration for all Eight Satellites in View.

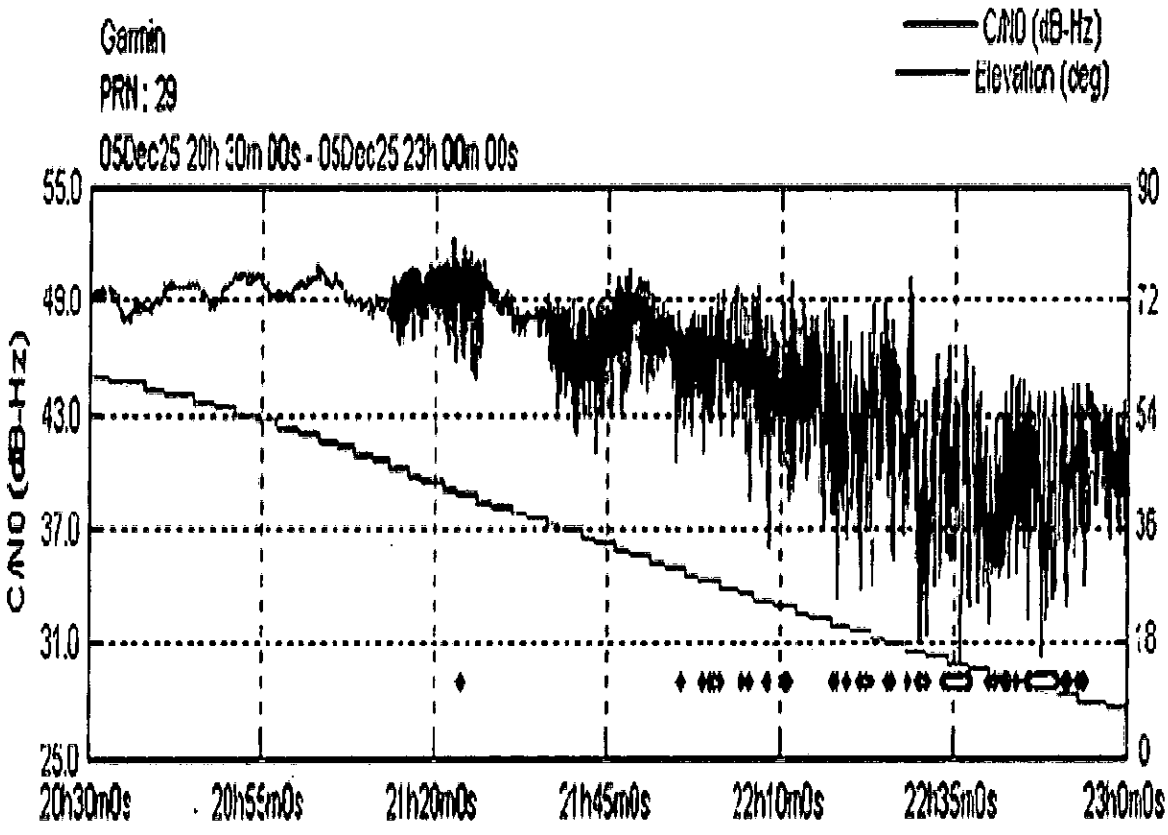


Figure 4.8: Carrier-to-noise (C/No) ratio of PRN 29 during scintillation event at Sao Jose dos Campos, Brazil on the 5th/6th December, 2006.

4.3.4 Receiver Re-acquisition Time

This section considers the satellite channels availability (%) against re-acquisition time (seconds) during severe scintillations. The data were based on the time it takes to simultaneously acquire back specific number of lost satellite channels. In other words, the percentage of the number of satellite channels (sufficient for navigation solution) that can be re-acquired, having been earlier lost to scintillations, at a predefined time (from one second to twenty seconds) for a given epoch. The Validated ICAO Standards and Recommended Practices for GNSS-2000 (ICAO, 2000) recommended

that in the space of twenty minutes, the availability of four or more satellites for GPS receivers for aviation applications must be 99.999% to 100%. This forms the basis for our analysis.

Whenever a deep fade (greater than 22dB-Hz) occurs for a satellite channel, the receiver is caused to lose such channel at a predefined re-acquisition time. For instance, out of the entire available satellites (each known by its PRN), how many are lost and re-acquired back at a predefined time window, say 20 seconds. Assuming the available satellites at the epoch are ten, the percentage availability for say 2 seconds re-acquisition time is the percentage of the satellites that were lost and re-acquired within this time frame. Therefore, the percentage that 4 or more satellites will be available for navigation operations at a pre-defined time (say 2 seconds) is the number of times that four or more satellites from the entire number of available satellites for the epoch will be re-acquired within 2 seconds having been initially lost to deep fading. If within 2 seconds, despite the deep fading, 4 or more satellites are observed at all times by the receiver, the availability under this condition is 100%. Table 4.1 shows the result of such observations, and it is graphically represented in Figure 4.9.

Table 4.1: Receiver Reacquisition Time and Satellite Channels Availability

Re-acquisition Time (sec)	ICAO Availability Requirements (%)	4 or more SV Availability (%)	5 or more SV Availability (%)
1.0	100.000	100.000	100.000
2.0	100.000	100.000	100.000
3.0	100.000	100.000	100.000
4.0	100.000	100.000	100.000
5.0	100.000	100.000	100.000
6.0	100.000	100.000	99.999
7.0	100.000	100.000	99.998
8.0	100.000	100.000	99.980
9.0	100.000	100.000	99.980
10.0	100.000	100.000	99.970
11.0	100.000	99.999	99.500
12.0	100.000	99.999	98.800
13.0	100.000	99.999	98.000
14.0	100.000	99.999	97.000
15.0	100.000	99.960	96.000
16.0	100.000	99.500	95.000
17.0	99.999	99.000	94.500
18.0	99.998	98.300	94.000
19.0	99.980	97.800	93.600
20.0	99.970	97.000	93.000

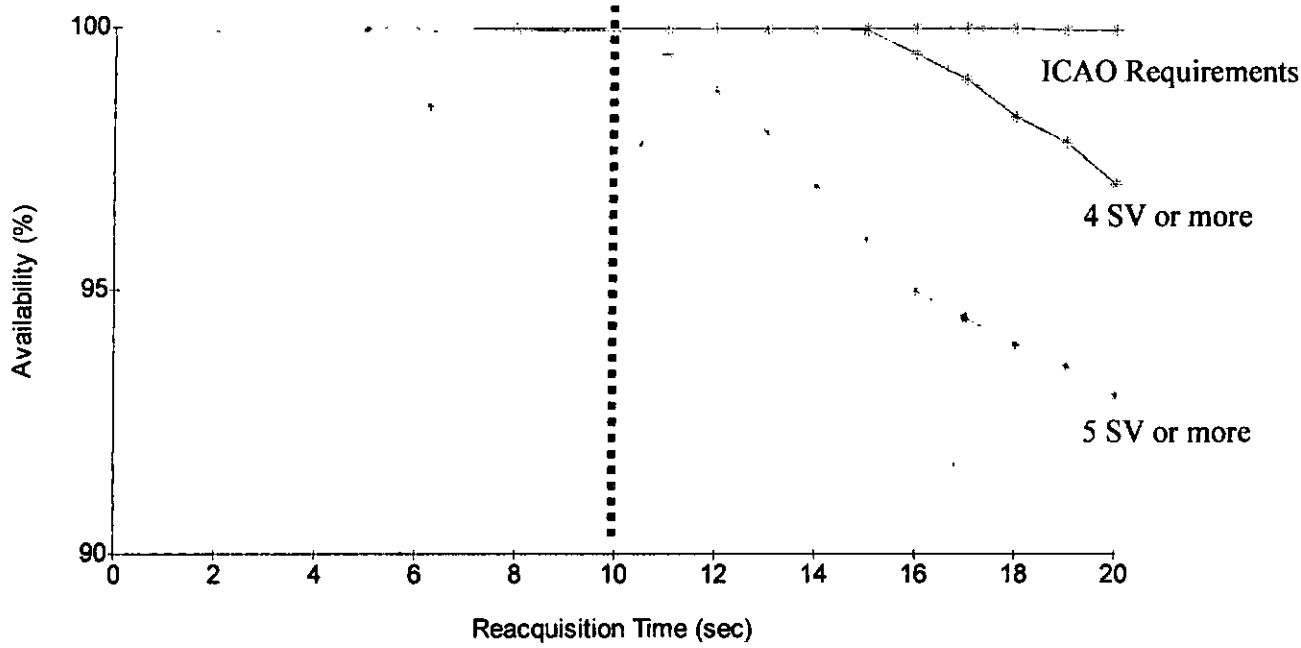


Figure 4.9: Satellite Availability against the corresponding Re-acquisition Time.

4.4 Working Operations of GPS Receiver Position Fixing

The numerical technique has a starting point that is based on multi-dimensional Newton-Raphson method. As we mentioned earlier, a simple case of a receiver tracking only four satellites can be solved exactly. On the other hand, if many satellites are being tracked, then any solution to a subset of four pseudoranges will not be solutions to the remaining pseudorange equations (Kintner, 1999). To further elucidate the operational framework of a GPS receiver system, software approach was used to explain how GPS receiver determines the position of its user. FORTRAN code was developed for the computation of the user's position with some surrogate satellite parameters to run the code. At the end, the sample output shown in appendix A4 was obtained. For a sample running of the code for eleven satellites, we chose the initial

guess for the user position and receiver clock offset as zeros. The satellite parameters for the simulation exercise are discussed in section 3.3.3 (Table 3.1).

Finally, safety mechanism was provided to enhance aviation safety in Nigeria using GPS-based technology.

CHAPTER FIVE

5.0 DISCUSSIONS

5.1 Nonlinear Dynamics Model

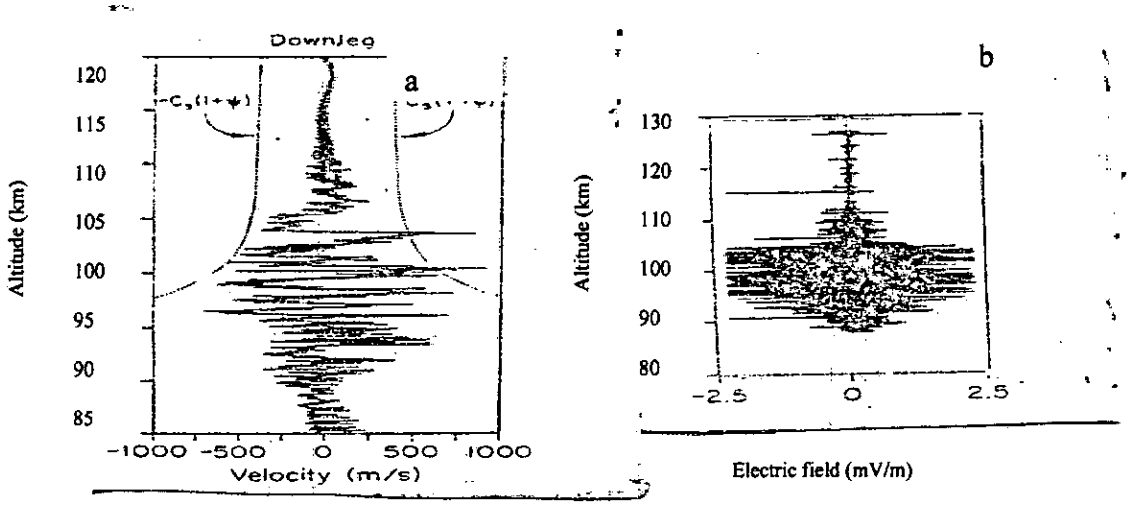
Vertical upwelling of $\mathbf{E} \times \mathbf{B}$ drifts of ions and electrons in a region of plasma depletion through nonlinear dynamics processes are responsible for irregularities generation in the ionosphere (Hudson, 1978), and they grow and evolve following bifurcation patterns (Huang, 1975).

Looking at Figure 4.3 closely, at the initial stages of growth (from $p = 0.1$ through to $p \approx 0.8$), the gradient of the irregularities follows a linear process. As the value of p increases, bifurcations start setting in, defining the proper route towards nonlinear processes. As p increases further, the irregularities also grow, further increasing the number of bifurcations, thereby leading to a situation whereby the bifurcations become ubiquitous. At this point, the irregularities are well clustered together, forming the nonlinear regime of the gradient, which attains maximum values of ± 1 at $p = 2$.

It is important to remind us once again that in our numerical computations, a more general perspective has been followed by setting the constant coefficient of eq. (3.8) to unity. Under practical situation, this coefficient depends on the seeding source of irregularities, where real maximum value(s) of the irregularities will be attained. However, our choice of this coefficient did not in any form affect the physical features of the equation because whatever value of the coefficient chosen, the features of the map in Figure 4.1 remain the same, and the analysis also therefore remains the same but for a matter of factor.

Comparing these results with experimental measurements by Pfaff et al. (1982; 1988) and Rama Rao et al. (2006) (see Figures 5.1 a, b and c, and 5.2 a and b), it was established that the growth of ionospheric electric field under HF propagation using Langmuir probe, vertical drift velocities associated with the horizontal electric fields in the ionosphere and scintillation effects occasioned by the presence of ionospheric irregularities follow nonlinear processes of evolution just like the present model. Though, it was observed that in measurements by Pfaff et al. (1982; and 1988), the irregularities evolve in a descending order as contrary to that of the present study which follows ascending order. The reason for this is not far-fetched, the p-parameter, which is related to the strength of the irregularities, has been used in inverse form during the numerical computations just for numerical convenience. In the measurement by Rama Rao et al. (2006), the records were taken at night using two propagation bands (VHF: 244 MHz and L-band: 1.5 GHz). Their measurements are observed to grow nonlinearly with observable windows at some irregular intervals. The nonlinearity was also enhanced for the VHF propagation in comparison with the L-band propagation under the same ionospheric irregularities condition, which is of course expected.

Amplitude scintillation index S_4 depends on irregularities density fluctuations (Liu et al., 1974; Yeh et al., 1975; 1982; and Rama Rao et al., 2006). As the value of scintillation index reaches unity ($S_4 = 1$: 100% occurrence), extremely strong scintillation effect will be experienced.



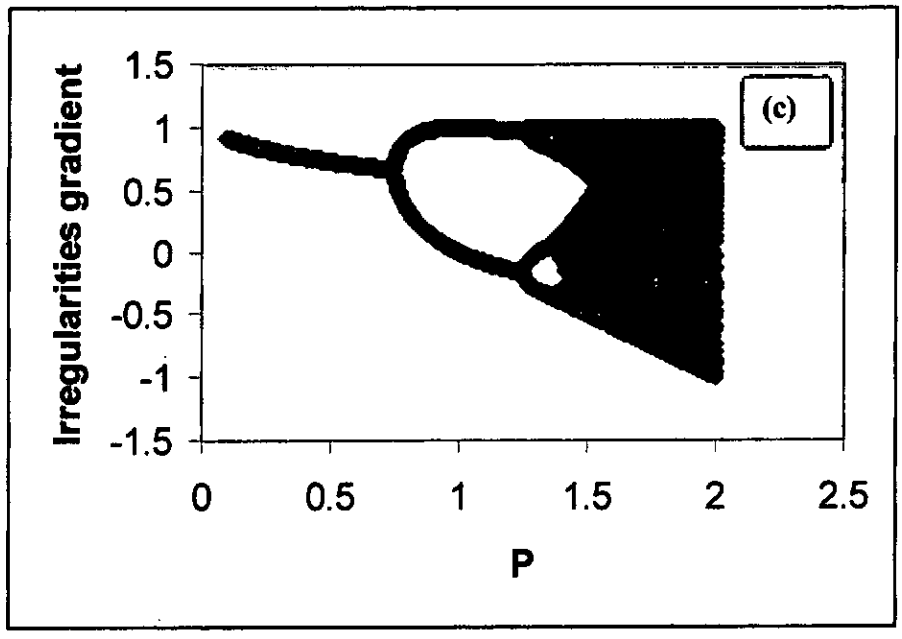
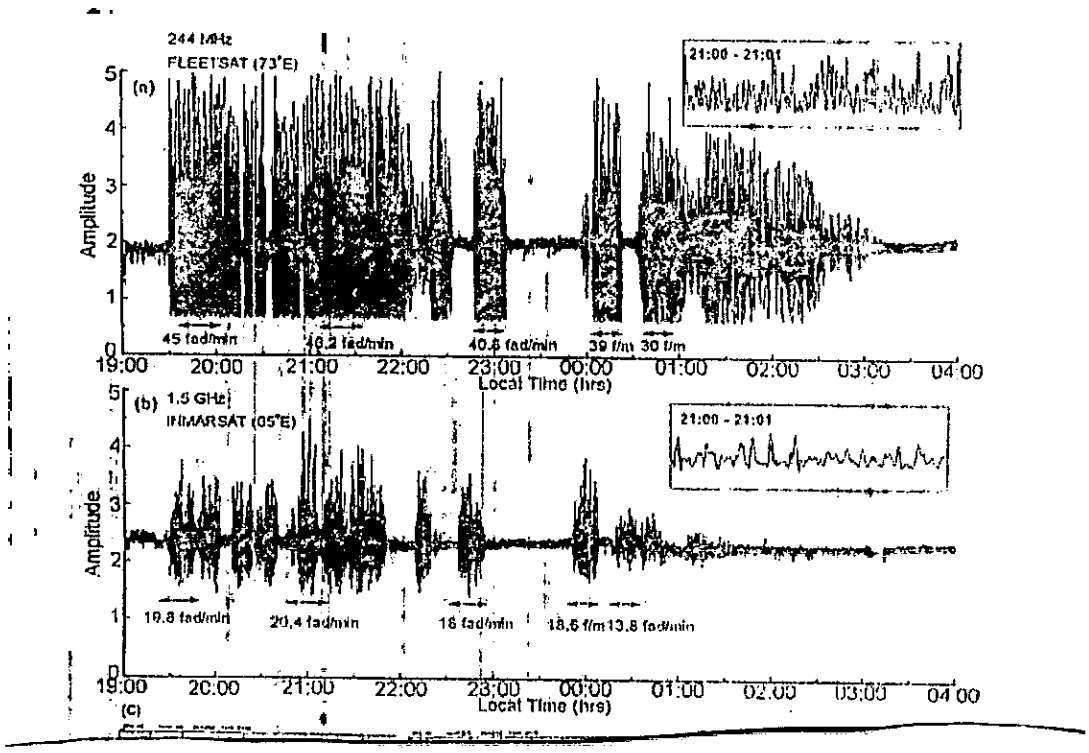


Figure 5.1: (a) Vertical Drift Velocities Associated with the Horizontal electric Fields (After Pfaff et al., 1988), (b) Variation of electric Field Data with Altitude Using Langmuir Probe (After Pfaff et al., 1982), and (c) The model on which (a) and (b) were compared.

a



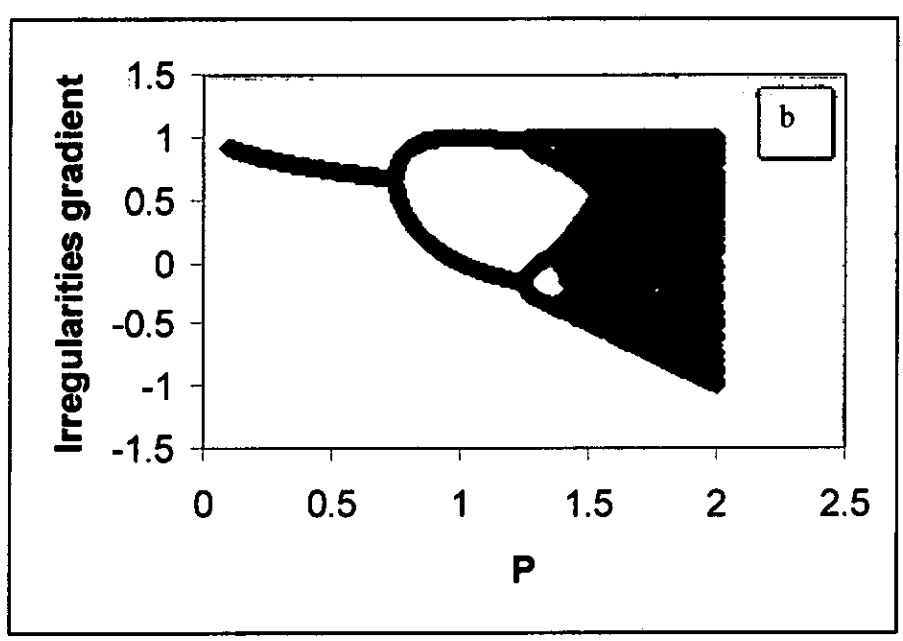


Figure 5.2: Scintillation Records of 244MHz and 1.5GHz Signals on 26 March, 2004 in India (After Rama Rao et al., 2006) (b) The model on which (a) was compared.

This becomes strong or just moderate for $S_4 \geq 0.5$. S_4 value of approximately 0.8 would be sufficient to cause loss of lock to GPS especially those with wide band receivers (Aquino et al., 2003). In area of practical applications, the irregularities density gradient of ± 1 at $p = 2$ depicts a condition for very strong scintillation with $S_4 = 1$. Under this condition, severe havoc is expected on satellite communication and navigation systems, though, as we must have seen in the measurements by (Rama Rao et al., 2006), the extent of severity depends on the signal frequency. At $p = 0.75$, a route to nonlinear regime of the irregularities gradient has just been reached. Here, the scintillation effect on signals will be weak and tolerable loss of signal may be experienced. Below this p-parameter, the growth of the irregularities gradient is linear. Under this condition, scintillation effect becomes so infinitesimal.

5.2 Ionospheric Irregularities Thickness and Their Impacts on Transionospheric Radio System

Figures 4.2 and 4.3 show the graphs of the mutual coherence function $|\Gamma|$ against the fractional frequency separation X, for the two operating frequencies, 125 MHz and 1.6 GHz (range for GPS frequency) at two different thicknesses of the irregularity slab. All the graphs have a common feature of decrease of the mutual coherence function with increasing fractional frequency separation. Specifically in Figure 4.2, there is a sharp descent starting from a value of 0.7 for the mutual coherence function and a correlation value in the neighbourhood of 60% for the 125 MHz operating frequency. This value rises to above 70% for the 1.6 GHz operating frequency. As mentioned earlier, a unit value for the mutual

coherence function depicts normal undisturbed ionosphere where there is no fluctuation. In all the graphs, the mutual coherence function starts with a value lower than unity which literally translates to fluctuating ionospheric conditions. As seen in Figure 4.2, the mutual coherence function has a value of about 0.9 for the least fractional frequency separation (X=0) at 125 MHz, while a value close to unity is attained at 1.6 GHz, hence minimal fluctuation is expected at this operating frequency. If the thickness of the irregularity slab is increased to about six times its original value, the value of the mutual coherence function drops to about 0.5 for the same fractional frequency separation due to the enhanced fluctuations caused by the increased thickness (see Figure 4.3). This value also dropped to 0.7 for 1.6 GHz operating frequency under the same condition.

In Figures 4.3, the thickness of the irregularity slab has increased to about six times the original thickness. Here, $\zeta_o = 1.272$ and 0.102 for operating frequencies f=125 MHz and 1.6 GHz respectively. Because more scattering is expected due to the increased number of wave scatterers, the signals decorrelate rapidly with loss of coherency in bandwidth when the fractional frequency separations are 40% and 60% for the operating frequencies f=125 MHz and 1.6 GHz respectively (see Figures 4.3). In other words, the tolerable bandwidth has been reduced, as contrary to the initial slab thickness where broad bandwidth is guaranteed. For instance, using a centre frequency of 1.6 GHz under that moderate condition will still accommodate fractional frequency separation of over 70%. This is also corroborated by the measurements by Rama Rao et al. (2006) (Figure 5.2): that is, the severity of scintillation impact depends on the frequency of propagation.

It is worthy to mention that the fractional frequency separation, X, at which the mutual coherence function almost disappears or decreases to a specific level is referred to as the *coherence bandwidth,* and it is a measure of how two waves transmitted simultaneously at different frequencies are correlated. The inverse of the coherence bandwidth is the pulse broadening (Green, 1981).

5.3 Observed Ionospheric Data

Spread F irregularities are the most common F-region irregularities with the potential consequences on GPS-based technologies. These are turbulent processes and plasma instabilities that grow and evolve in the night-time equatorial ionosphere (Fromouw, et al., 1978).

In the past, authors such as Calvert and Cohen (1961) and Woodman and La Hoz, (1976) parameterized spread F records in terms of altitude, time evolution, scale sizes, vertical and zonal drifts and magnetic aspect angle dependence of 3m. In Figure 4.4, spread F records that were measured at Jacamarca on the 17th October, 1994 are shown. These records indicated the evolution of spread F irregularities to altitude of the order of several hundreds kilometres. Between 2030 and 2200 hrs, the irregularities spread uniformly within the F-region height at the carrier-to-noise ratio (C/No) of about 22dB to 24dB. This is the measure of the modulation that the irregularities will impose on the transient GPS signals. GPS receivers are designed to maintain lock under scintillation depth of 22dB; this implies that scintillation activity higher than this value may impose unfavourable consequences on the performance of GPS-based navigation systems. Valladares et al. (2001) posited that extension of ionospheric irregularities to altitudes of 500km and above signifies that the

irregularities are of small scale sizes, which is more problematic to GPS signals. This phenomenon occurs between 2030 and 2100hrs; 2200 and 2230hrs.

In a similar interpretation, Figure 4.6 shows the raw carrier-to-noise ratio (C/No), and the corresponding scintillation intensity index measured with an Ashtech Z-12 receiver at Ascension Island on the 16th March, 2002 for PRN 8 satellite. We must be reminded that this date is equinoctial date during high solar period, where scintillation event is expected to attain maximum occurrence, hence, a very good campaign period for observation. The rapid fluctuations in the C/No between 2105 and 2205hrs indicate modulation effect (scintillation) imposed on the signal by the ionospheric irregularities. Within this window, the C/No is about 22-24dB over the mean average (indicated by the red rectangle), and the corresponding scintillation is 1.0 and above, which portends danger to GPS service as earlier explained. These figures further confirmed that scintillation is a post sunset event (Carrano, 2007).

In another observation, data were collected by Garmin receiver at Sao Jose dos Campos in Brazil between December 2005 and January 2006. This receiver is certified for aircraft navigation with vertical guidance; hence, its performance evaluation during scintillation is essential to assess the impact of scintillation on GPS navigation.

These records (Figure 4.8) show carrier-to-noise density ratio (C/No) of PRN 29 recorded by WAAS Garmin receiver during scintillation event. C/No fluctuation was as high as 20dB. Red points on the figure indicate that phase lock of this satellite signal was lost at corresponding epochs. It is observed also that the receiver lost high elevation satellites under strong scintillation in addition to low elevation satellites.

Loss of high elevation satellites due to scintillation poses serious hazard to GPS navigation.

GPS avionics on aircraft require at least four satellites with good geometry for correct navigation solution. Hence, loss of carrier information due to scintillation could be hazardous to GPS navigation if significant number of satellite signals was lost simultaneously. Seo et al. (2007) concluded that even under intense scintillation, the percentage that two satellites will be lost simultaneously is 0.5%, which is still tolerable for navigation applications. However, this conclusion may be invalid if the scintillation patches cover large portion of the sky at the same time.

On close observation, the analysis of Figure 4.9 indicates the percentage that a fixed GPS receiver during scintillation period in solar maximum can reacquire successfully four or more satellites channels in the first fifteen second before slight departure from the ICAO's requirements for availability, 99.999% to 100% (ICAO, 2000; and Van Dyke, 2001). At the twentieth second, the availability comes to 97%. Similarly, when five or more satellites were considered, the ICAO requirements were satisfied for the first ten seconds before departure is experienced, and this departure drops to about 93% at the twentieth second. Low percentage (less than 99%) of satellite channels availability may hinder the capability of the receiver to calculate navigation solution under scintillation effect, especially, when this effect overrides the availability of at least four satellites with good geometry (which is the least condition for correct navigation solution).

Overall, for the first ten seconds, the fixed GPS receiver shows excellent performance of re-acquiring five or more satellites channels successfully with 100%

availability. This depicts the benefit of short re-acquisition time. Although this investigation revealed that losing track on too many satellites by a receiver may be rare. However, in order to obviate aviation disaster in event of unexpected simultaneous deep signal fading on aviation GPS receivers, though, conservatively expensive to achieve because of significant tradeoff between antenna gain and directivity; it is suggested that intended GPS receivers for transequatorial aviation applications must be enhanced to cope with fading of over 25dB-Hz, and short re-acquisition time. Synthetic aperture type antenna may be excellent for such receivers' design. Precisely, re-acquisition of less than or not more than 10 seconds is needed for aviation applications (as shown by the green dash line boundary in Figure 4.9). This is in agreement with the ICAO recommendation of 99.999 to 100% availability, and 10 seconds time to alert for approach with vertical guidance-I (APV-I) and 6 seconds time to alert for approach with vertical guidance-II (APV-II) (ICAO, 2000, Van Dyke, 2001). This is the same for CAT I precision approach (instrument guidance down to decision height of 200 ft (60m)), and just 2 seconds for CAT II/III precision approaches (instrument guidance down to decision height of 100 ft (30m)) (Viarrsson et al., 2001).

The records shown in Figure 4.7 explain the carrier-to noise ratio outputs for all eight satellites in view for a GPS receiver located at the Ascension Island during severe scintillation effect (S_4 index greater or equal to unit), in a year of solar maximum on the 18th March 2001. This observation was carried out at a short period (just 100 seconds) to observe the possibility of simultaneous loss of satellite channels to fading. Frequent deep signal fading of almost all the eight satellites in view are

observed. The question is that, is there any possibility of losing all eight satellite channels (PRNs; 1, 2, 4, 7, 8, 11, 13, 17) or more than 50% of the eight channels at the same time?

As it has been mentioned earlier, deep signal fading of over 22dB-Hz can break the GPS receiver's carrier tracking lock to a satellite channel (Brown et al., 1998). The lost channel cannot be used for position calculation until it is re-acquired back to sustain continuity. If the lost channels are just one or two, and provided that the GPS receiver can still view at least four satellites with good geometry conveniently, such loss may not portend serious consequence for navigation service. Otherwise, if many channels are lost simultaneously, thereby allowing the receiver to view less than four channels, error is bound to occur in the navigation solution. This portends great risk for aviation applications. Protracted deep fades may cause over-riding continuous resets of carrier smoothing filter, and this often shortens effective smoothing times (Seo et al., 2007). Less smoothing leads to higher noise levels, and lower availability of navigation service.

From the analysis of Figure 4.7, almost all the satellite channels in view by the receiver suffer deep fading, though not simultaneously. This implies that the receiver did not fail to track multiple satellites simultaneously. Deep signal fading of different satellites do not usually happen at the same time. Hence, if a receiver can re-acquire a lost satellite before it losses another, simultaneous loss will definitely be averted, and consequently reduce the impact of scintillation on GPS navigation.

Figures 4.5 (a, b and c) show the measurement of total electron content (TEC) by an Ashtech Z-12 GPS receiver located at Nairobi, Kenya. These records were

observed for three consecutive days (1st, 2nd, and 3rd) in December, 2008. The three plots have similar features with the maximum crest between 1200LT and 1600LT, and trough between midnight and 0600LT. This is expected because the sun is just about to set at this period, therefore, production of electron for the enhancement of ionospheric electron density is low. On the average, less than 5 TECU was observed for the three days during these hours. Similarly, the crest observed between 1200LT and 1600LT (25 TECU on the average) was due to the sufficient solar energy that radiates on the ionosphere to enhance electron production.

This analysis validates Appleton's postulation that maximum electron production is expected when the sun is overhead (Ratcliffe, 1972). After this time (1200LT-1600LT), the TEC begin to decay to another minimum valley at the next midnight.

5.4 Simulation of GPS Position Fixing

With the inputs parameters earlier shown for the simulation of the user position for eleven satellites in view by the receiver, the user position obtained are; on latitude 007.0°, longitude 015.4° and altitude 8.94393×10^{7} m.

Following a least squares approach, our code provides means of determining the unknown four user position variables for a receiver with clear view of many satellites. Under practical application sense, GPS receivers always have six to eight satellites in view. The maximum that can be seen at any instant in time for a receiver with a clear view to the horizon is twelve satellites (Carrano, 2007). By implication, this leads to superfluous linear systems which needs optimization technique for unique solution (see appendix A3 for details on the optimization technique (Least Squares Method)).

Tracking more than four satellites by a receiver has some obvious advantages. The probability of maintaining track on at least four satellites is enhanced for a receiver antenna with translating obstacles across its view, even when track is lost on some of the satellites. In addition, the error in navigation solution is minimized when a receiver tracks more than four satellites (Tsui, 2000).

To further elucidate the concept, we simulate the case of a GPS receiver that is simultaneously observing seven satellites. The satellite input parameters for the code have been earlier discussed under computational algorithm. With these parameters, the code fixed the user position on latitude 007.0°, longitude 015.4° and altitude 8.94393×10^7 m, geographically; this location falls very close to Nigerian airspace, albeit in Cameroon. This location has practical relevance to ionospheric modeling community in that it falls within the Appleton anomaly region where scintillation has been known to constitute threat to communication and navigation systems on daily basis. Many GPS receiver's network has been set up across the equatorial anomaly region starting from South America region through to the Southern Asia/Northern Australia region, but unfortunately, the distributions of these receivers is sparse in the Africa region. Considering this region for distributions of this equipment will enhance research efforts that can consequently put our findings on a firmer footing.

5.5 Applications of Study to Aviation Safety

The essence of air traffic management (ATM) system is to provide guidance, separation, coordination and control of aircraft movements. As mentioned in the introduction, the need to control large traffic has subjected ATM systems all over the world to some limitations that usually necessitate stacking and holding of airplanes on

arrivals. The consequence of this is excess fuel burnt and emissions. IPCC (1995) pointed out that if these limitations could be addressed, which of course, is one of the cardinal objectives of this study, aircraft fuel burnt and emissions will be reduced in the range of 6 to 12 %.

As we mentioned earlier, GPS receivers have capability of determining the user's position in terms of geographic latitude, longitude and altitude. Before application of the technique presented herein, a prerequisite condition is that all airplanes plying the airspace must have GPS antenna/receivers mounted on them and that such receiver will be capable of encoding its position via radio infrastructure, and retransmits it to a visual display unit at a control tower that has jurisdiction over the airplane, so as to ascertain its position on the grid calibration at any point in time.

With this concept, the position of an airplane can be easily known in the sky. Consequently, effective traffic control can be guaranteed when there is need to control large traffic. With GPS-based control mechanism, efficient multilateral separation of airplanes can be sustained (Burnley, 2008). Another important benefit of the concept is that, it is environmentally friendly. Airplanes can be flown at continuous (green) descent/approaches: in a progressive down-stair manner at/or near idle power with enhanced thrust each time the airplanes level off, thereby reducing environmental pollution (emissions and noise) (Standar, 2008). Also business-driven operations can be attained using this technique: In achieving this, controllers and pilots must continuously iterate flight landing time by synchronizing system's time with reference to that of GNSS satellites so as to know accurately the airplane landing time. With this, airline operators can achieve effective flight scheduling.

In the area of emergency management, the technique provides an easy referencing of any location within flight routes of the airspace. As briefly mentioned in the airspace grid referencing (see Figure 3.1), any particular grid where there is a problem can be easily referred so that emergency response agencies can be dispatched to reach the scene at minimum time. On arrival at the specified grid and the distressed airplane is not quickly located, a specified radius of search has at least been establish upon which the airplane will be located on thorough searching. Similarly, GPS receivers can be mounted on emergency vehicles as earlier explained for airplanes, for efficient coordination of resources at emergency scenes under a separately located control point where a visual display unit will be used for coordination (Hofmann-Wellenhof et al., 1997). Any emergency vehicle that is approaching obstacle/hitch can be redirected to en-route via a safer path (Fire Service Manual, 2006).

CHAPTER SIX

6.0 SUMMARY AND RECOMMENDATIONS

This thesis presented the role that the transionospheric radio (GPS) system plays in enhancing aviation safety. Ionospheric irregularities impacts severely on transionospheric radio (GPS) system. In a view to give adequate explanation of the formation and evolution of these irregularities and their consequences on transionospheric radio (GPS) system, two numerical modeling efforts came to play: (i) Nonlinear dynamics model, and (ii) Phase Screen Model. The mitigation strategies for these potential consequences on transionospheric radio (GPS) systems were proffered. Analyses of some ionospheric observables that were measured at some stations within the equatorial (Appleton) anomaly showed that scintillation is a post sunset event, and that when intense, it could be hazardous to GPS-based air navigation service delivery, by imposing suppression on the carrier signals. Finally, the importance of GPS-based technology in enhancing aviation safety was discussed.

The nonlinear dynamics model was used to establish the features of ionospheric irregularities growth and evolution. The investigation revealed that ionospheric irregularities grow and evolve via bifurcation patterns, and that nonlinear dynamics processes are responsible for the formation of ionospheric irregularities through vertical upwelling of $\mathbf{E}\times\mathbf{B}$ drifts of ions and electrons in region of plasma depletion, which are in agreement with the conclusions drew by Calvert and Cohen (1961), and Hudson (1978).

From this approach, a map of ionospheric irregularities growth and evolution was obtained via numerical iterations. In the nonlinear regime of the map, the density irregularities gradient attained maximum values of ± 1 which was analyzed to be corresponded to extremely strong scintillation condition where the amplitude scintillation index, $S_4 = 1$ (100% occurrence) which may portend significant hazard to transionospheric radio system (GPS) service availability and continuity, especially when impacts are simultaneously on too many satellites. For $S_4 = 0.5$; representing the neighbourhood of the commencement of bifurcations, moderate scintillation may be expected at this point, and at points before, the traversing signal will undergo weak scintillations.

The map was compared with experimental measurements. Overall, it shows good agreement with experimental measurements. In the simulation, only density irregularities gradients with a factor that is related to the strength of the irregularities were varied, but at the end, the study added along-side with the analyses, some relevant parameters such as total electron content (TEC), and scintillation index (S_4) (Briggs and Parkin, 1963; and Fremouw et al., 1978).

The second numerical model was based on principles of optics, whereby the ionosphere was assumed to be bounded below by a thin screen (phase screen), so that traveling radio wave through it suffers diffraction patterns.

The correlation between frequencies of propagation in a fluctuating transionospheric link has been investigated using mutual coherence functions. Markov random process and parabolic equation approximations were used to derive the two frequency two position mutual coherence functions with effects of

scattering taken into cognizance (Gherm and Zernov, 1998). The results revealed that the level of coherency maintained by transmitting radio signals through the ionosphere depends strongly on the centre frequency of propagation, which was in consonance with experimental measurements by Rama Rao et al. (2006). For instance, good coherency is maintained at the L-band propagation frequency (GPS frequency), relative to the VHF band frequency under the same ionospheric conditions.

Another important inference is that the propensity for forward scattering, which impacts fluctuations on the traversing signals, is enhanced by the increase in the thickness of the irregularity slab. Under these conditions, signals are rapidly decorrelated and loss of coherency in bandwidth is experienced. This implies that the sustenance of coherence bandwidth depends on the concrete conditions of propagation.

Analyses of relevant ionospheric parameters, measured at different stations (within the equatorial anomaly) at different times and seasons indicated that the impacts of scintillation on GPS-based navigation facilities can only be hazardous when too many satellites' signals are lost to deep fading (greater than 22dB-Hz) at exact the same time, to the extent that the receiver can no more view up to four satellites at good geometry for proper navigation solution.

This situation was not noticed throughout the investigations. This implies that, even under severe scintillation conditions like the records considered in this study, the hazardous situation may be rare. However, in order to obviate aviation disaster in event of unexpected simultaneous deep signal fading on aviation GPS receivers,

though, conservatively expensive to achieve because of significant tradeoff between antenna gain and directivity, we suggested that intended GPS receivers for transequatorial aviation applications should be enhanced to cope with fading of over 25dB-Hz, and short re-acquisition time of less than 10 seconds. This is necessary for the overall reliability and robustness of satellite communication and navigation systems.

Thereafter, the underlying principles on which GPS operations are based were looked into, especially as on how it determines the position of its user. Ordinarily, under clear view, at any time and any location on the earth, GPS receivers can observe four to twelve satellites (Van Dierendonck, 2005; and Carrano, 2007). In event whereby more than four satellites are being observed, the study provided a code that reduces superfluous unknown variables to a system of four variables by multidimensional Newton-Raphson technique, and thereafter solved by Gaussian elimination approach. At the end, a case study of a receiver tracking eleven satellites simultaneously was simulated, to further elucidate the operational principles of GPS receiver.

For applications in aviation industry, a GPS-driven airspace availability optimization technique whereby airplanes can be flown efficiently by controllers at (a) minimum multilateral separations, in other words, flying many airplanes simultaneously, (b) continuous descent (green) approaches, in progressive down-stair technique at/or near idle power and enhance thrust at level off points, thereby reducing fuel burnt (emissions), and consequently reducing environmental impacts, (c) more business-driven opportunities; with GPS, the exact time of airplane

landing can be predicted and as such, attaining effective flight scheduling. Also, the concept brought insights into how a robust aeronautical search and rescue (SAR) mechanisms can be designed for each state of the world.

6.1 Significant Contributions to Knowledge

This thesis has contributed to knowledge in the following ways:

1. Previous studies indicated the nonlinearity of ionospheric irregularities spectrum. In the present study, a numerical approach has been used to develop this nonlinearity via 'nonlinear dynamics model' so as to establish adequate explanations of the growth and evolution of these irregularities
2. This study concluded that sustenance of radio transmission coherency through the ionospheric irregularities depends on the thickness of the irregularities and the frequency of propagation.
3. Previous studies emphasized potential consequences of scintillation on GPS system. Analyses of data from the present study revealed that the effects of such scintillation can only be hazardous to GPS systems when too many satellites' channels are simultaneously lost to scintillation (fading), which was rare during observations. However, in order to guarantee safety under anomalous conditions, the study suggested that intended GPS receivers for transequatorial aviation applications should be enhanced to cope with fading of over 25dB-Hz, and short re-acquisition time of less than 10 seconds.
4. This study provided numerical insights (simulations) into operational principles of GPS receiver, especially on how it determines the position of its user in a manner to assist future modeling efforts.

5. This study provided GPS-driven techniques for the enhancement of aviation safety. These techniques are presented in two categories: GPS-driven airspace availability optimization technique and a robust aeronautical search and rescue (SAR) mechanisms.

6.2 Limitations and Future Directions

1. Large volume of multi-station data from equatorial anomaly longitudes will help to define firmer climatology for better assimilation and predictions. Also, the data used in the analyses are from fixed/stationed GPS receivers, future efforts should validate the conclusions from this thesis by taking data from receivers on moving platforms (aircrafts) to observe Doppler effects on measured data.
2 The effort in applying GPS to enhance aviation safety is still hypothetical, and as such subject to practical ramification by the International Civil Aviation Organization (ICAO). Therefore, when definitive footing is established, this study should be applied to encompassed the entire globe. Another caution is that conventional radar should still remain as backup to GPS-driven technique in case of impacts that may arise from unforeseen ionospheric variability (Groves et al., 1997) or men of malicious intents.

REFERENCES

Aarons, J. (1982). Global Morphology of Ionospheric Scintillations, *Proc. IEEE*, **70** (4), pp. 360 – 378.

Aarons, J., J. P. Mulley, H. E. Whitney, and E. M Mackenzie (1980). The Dynamics of Equatorial Irregularities Patch Formation, Motion and Decay, *J. Geophys. Res.*, **85**, pp. 139 – 378.

Aquino, M. H. O., S. Waugh, A. Dodson, T. Moore and S. Skone (2002). GPS Based Ionospheric Scintillation Monitoring, Institute of Engineering Surveying and Space Geodesy (IESSG), University of Nottingham, Nottingham, UK and Department of Geomantic Engineering, University of Calgary, Calgary, Canada.

Aquino, M. H. O., F. S. Rodrigues, A. Dodson, T. Moore and S. Waugh (2003). Results of Statistical Analysis of GPS Ionospheric Scintillation Data in Northern Europe, 13-15 October, 2003, Matera, Italy, Institute of Engineering Surveying and Space Geodesy (IESSG), University of Nottingham, Nottingham, UK.

Basu, Su. (1978). Ogo 6 Observations of Small-Scale Irregularities Structures Associated with Subtrough Density Gradients, *J. Geophys Res.*, **83**, pp.182 – 190.

Basu, S. and M. C. Kelly (1979). A Review of Recent Observations of Equatorial Scintillations and their Relationship to Current Theories of F-Region Irregularities Generation, *Radio Sci.*, **14** (3), pp. 471 – 485.

Basu, Su., S. Basu, C. E Valladres, H. C. Yeh, S. Y. Su, E. MacKenzie, P. J. Sultan, J. Aarons, F. J. Rich, P. Doherty, K. M. Groves, and T. W. Bullett (2001). Ionospheric Effects of major Magnetic Storms during International

Space Weather Period of September and October 1999: GPS Observations, VHF/UHF Scintillatios, and in-situ Density Structure at Middle and Equatorial Latitudes, *J. Geophys Res.*, **106**(A12), pp.30389 – 30413.

Beer, T. (1976). The Aerospace Environment, Wykeham Publications (London) Ltd, UK, Ch. 4.

Bitjukov, A. A., V. E Gherm, and N. N. Zernov (2003). Quasi-classic Approximation in Markov's Parabolic Equation for Spaced Position and Frequency Coherency, *Radio Sci.*, **38**, 1021, doi: 10.1029/2002RS002714, pp. 41 – 46.

Booker, H. G. (1958). The Use of Radio Stars to Study Irregular Refraction of Radio Waves in the Ionosphere, *Proceedings of the IRE*, pp. 315-320.

Briggs, B. H. and I. A. Parkin (1963). On the Variation of Radio Star and Satellite Scintillations with Zenith Angle, *J. Atmosph. Terr. Phys.*, **25**, pp. 339-365.

Brown, A., E. Holm, and K. M. Groves (1998). GPS Ionospheric Scintillation Measurements Using a Beam Steering Antenna Array for Improved Signal/Noise, Proc. of the 54th annual Meeting of the Institute of Navigation.

Budden, K. G. (1985). The Propagation of Radio Waves, Cambridge, Great Britain, Ch. 4.

Burnley, J (2008). Global Position, *Future Airport*, **1**, pp. S9-S13.

Burrel, A. G., N. A. Banito, and C. S. Carrano (2006). Total Electron Content Processing from GPS Observations to Facilitate Ionospheric Modeling, *Proceedings of the American Meteorological Society*, Atlanta, GA, pp. 1-10.

Calvert, W. and R. Cohen (1961). The Interpretation and Synthesis of Certain Spread F Configurations Appearing on Equatorial Ionograms, *J. Geophys Res.*, **66**, pp.3125 – 3140.

Carrano, C. S. (2007). GPS-SCINDA: A Real-time GPS Data Acquisition and Ionospheric Analysis System for SCINDA, Scientific Report No. X, Air Force Research Laboratory (ARFL), Space Vehicle Directorate, 29 Randolph Rd, Air Force Materiel Command, Hanscom Air Force Base (AFB), MA, pp. 1-59.

Carrano, C. S. and K. M. Groves (2006). The GPS Segment of the AFRL\ SCINDA Global Network and the Challenges of Real-time TEC Estimation in the Equatorial Ionosphere, *Proceeding of the Institute of Navigation (ION)*, NTM, Monterey, CA, pp. 1036-1047.

Carpenter, D. L. (1966). Whistler Studies of the Plasmapause in the Magnetosphere, I. Temporal Variations in the Position of the Knee and some Evidence of Plasma Motion near the Knee, *J. Geophys. Res.*, **71**, pp. 693.

Chapman, S. (1931). The Absorption and Dissociative or Ionizing Effects of Monochromatic Radiation in an Atmosphere on a Rotating Earth, *Proc. Phys. Soc. London*, **43**(26), pp. 484.

Dandekar, B. S. and K. M. Groves (2004). Using Ionospheric Scintillation Observations for Studying the Morphology of Equatorial Ionospheric Bubbles, *Radio Sci.*, **39**, RS3010, doi: 10.1029/2003RS003020, pp.1 – 16.

Davies, K. (1966). Ionospheric Radio Propagation, Dover Publication Inc., NY 10014, USA.

Davies, K. (1990). Ionospheric Radio, Peter Peregrinus Ltd., London, UK.

Davies, K. and E. K. Smith (2000). Ionospheric Effects on Land Mobile Satellite Systems, Department of Electrical and Computer Engineering, University of Colorado, Colorado, USA.

de-Paula, E. R., F. S. Rodrigues, K. N. Iyer, I. J. Karton, M. A. Abdu, P. M. Kintner, B. M. Ledvina, and H. Kil (2003). Equatorial Anomaly Effects on GPS Scintillations in Brazil, *Adv. Space Res.*, **31**(3), pp. 749 – 754.

Doherty, P. H, T. Dehel, J. Klobuchar, S. H. Delay, S. Datta-Barua, E. R de-Paula, and. F. S Rodrigues (2002). Ionospheric Effects on Low-Latitude Space Based Augmentation Systems. *ION, GPS-2002*, Portland, OR, USA, p. 1321.

Doherty, P. H, S. H. Delay, C. E Valladares, and J. Klobuchar (2004).. Ionospheric Scintillation Effects on GPS in the Equatorial and Auroral Regions. Navigation: *Journal of the Institute of Navigation*, p. 235.

Dyson, P. L, J. P. McClure, and H. B. Hanson (1974). In-situ Measurements of Amplitude and Scale Size Characteristics of Ionospheric Irregularities, *J. Geophys. Res.*, **79**, pp. 1497 – 1502.

Elkins, T. J. and M. D. Papagiannis (1969). Measurement and Interpretation of Power Spectrum of Ionospheric Scintillation at a Sub-Auroral Location, *J. Geophys. Res.*, **74**(16), pp. 4105 – 4115.

Fesen, C. G, G. Crowley, R. G. Roble, A. D. Richmond, and B. G. Fejer (2000). Simulation of the Pre-reversal Enhancement in the low Latitude Vertical Ions Drifts, *Geophs. Res. Lett.*, **27**(13), pp. 1851-1854.

Fire Service Manual (2006). Fire Service Technology, Equipment and Media, HM Fire Service Inspectorate Publications Section, The Stationery Office (TSO),

London, Great Britain.

Fremouw, E. J., E. L. Leadabrand, R. C. Livingstone, M. D. Cousins, C. L. Rino, B. C. Fair, and R. A. Long (1978). Early Results from the DNA Wideband Satellite Experiment: Complex-signal Scintillation, *Radio Sc.*, **13**(1), pp. 167-187.

Gherm, V. E. and N. N. Zernov (1998). Scattering Function of the Fluctuating Ionosphere in the HF Bands, *Radio Sci.*, **33**(4), pp. 1019 – 1033.

Gherm, V. E., N. N. Zernov, S.M. Radicella, and H.J. Strangeways (2000). Propagation Model for Signal Fluctuations on Transionospheric Radio Links, *Radio Sci.*, **35**(5), pp. 1221 – 1232.

Gherm, V. E., N. N. Zernov, and H.J. Strangeways (2005). Propagation Model for Transionospheric Fluctuating Paths of Propagation: Simulator of the Transionospheric Channel, *Radio Sci.*, **40**, RS1003, doi: 10.1029/2004RS003097, pp. 1 – 9.

Gonzalez, W. D., B. T. Tsurutani, and A. L. C. Gonzalez (1999). Interplanetary Origin of Geomagnetic Storms, *Space Sc., Rev*, **88**, pp. 529-562.

Green, D. C. (1981). Radio Systems Technology, Pitman Press, Bath, Great Britain, ch 8.

Grewal, G. S. (1980). Numerical Methods in Engineering and Science, Khana Publishers, Nai-Parak, Delhi India.

Groves, K. M., S. Basu, E. J. Weber, M. Smitham, H. Kuenzler, C. E. Valladares, R. Shecham, E. MacKenzie, J. A. Secan, P. Ning, W. J. McNeill, D. W. Moonan, and M. J. Kendra (1997). Equatorial Scintillation and System

Support, *Radio Sci.*, **32**(5), pp. 2047 – 2064.

Gupta, B. D. (1998). Mathematical Physics, 2nd Revd edn., Vikas Publishing House, PVT Ltd, New-Delhi, India.

Hofmann-Wellenhof, B., H. Lichtenegger, and J. Collins (1997). Global Positioning System: Theory and Practice, 4th Revd edn, Spriger-Verlag Wien, New York, USA.

Huang, C. (1975). The Travelling Bifurcation of the Equatorial F2 Layer, *Radio Sci.*, **10**, pp. 507 – 516.

Hudson, M. K. (1978). Spread F Bubbles: Nonlinear Rayleigh-Taylor Model in Two Dimensions, *J. Geophys Res.*, **83**, pp.3189 – 3194.

Intergovernmental Panel on Climatic Change (IPCC), (1995). Aviation and Global Atmosphere, A Special Report of IPCC Working Groups I and III in collaboration with the Scientific Assessment Panel to the Montreal Protocol on substances that deplete the Ozone layer [Penner, J. E (*University of Michigan*), D. H. Lister (*Defence Research & Evaluation Agency*), D. J. Griggs (*UK Meteorological Office*), D. J. Dokken (*University Corporation for Atmosp. Res.*), and M. McFarland (*DuPont Fluoroproducts*) (Eds.)].

Ishimaru, A. and S. T. Hong (1975). Multiple Scattering Effects on Coherent Bandwidth and Pulse Distortion of a Wave Propagating in a Random Distribution of Particles, *Radio Sci.*, **10**(6), pp. 637 – 644.

Ishimaru, A. (1977). Theory and Application of Wave Propagation and Scattering in Random Media, *Proc. IEEE*, **65**(7), pp. 1030 – 1060.

Jordan, E. and K. G. Balmain (1978). Electromagnetic Waves and Radiating

Systems, Prince-Hall of India Private Ltd, New Delhi, India.

Jordge, V. J. and J. S. Eugene (1998). Classical Dynamics, Cambridge University Press, Cambridge, UK.

Kelly, M. C., K. D. Baker, J. C. Ulwock, C. L. Rino, and M. J. Baron (1980). Simultaneous Rocket Probe, Scintillation, and Incoherent Scatter Radar Observations of Irregularities in the Auroral Zone Ionosphere, *Radio Sc.*, **15**(3), pp. 491-505.

Kelly, M.C. (1989). The Earth's Ionosphere, Academic Press Inc., London.

Keskinen, M. J. and S. L. Ossakow (1983). Theories of High-latitude Ionospheric Irregularities: A Review, *Radio Sc.*, **18**(6), pp. 1077-1091.

Kintner, P. M., H. Kil, T. L. Beach, and. E. R. de-Paula (2001). Fading Time Scale Associated with GPS Signals and Potential Consequences, *Radio Sc.*, **36**, pp. 731-743.

Kintner, P. M. (1999). Global Positioning System: Theory and Design, School of Electrical Engineering, Cornell University, Ithaca, NY, USA.

Kintner, P. M. and B. M. Ledvina (2005). The Ionosphere, Radio Navigation and Global Navigation Satellite Systems, *Adv. Space Res.*, **35**, pp. 788 – 811.

Koster, J. R. (1963). Some Measurements on the Sunset Fading Effects, *J. Geophys. Res.*, **68**(9), pp. 2571 – 2578.

Kreyszig, E. (2005). Advanced Engineering Mathematics, 8th edn., John Wiley & Sons (Asia) Pte Ltd., Singapore.

Lawrence, R. S. (1958). An Investigation of the Perturbations Imposed Upon Radio Waves Penetrating the Ionosphere, *Proceedings of the IRE*, pp. 315-320.

Lee, L.C. (1974). Wave Propagation in a Random Medium: A Complete Set of the Moment Equations with Different Wavenumbers, *J. Math. Phys. (NY)*, **15**(9), pp. 1431 -1435.

Leitinger, R. (1989). The Importance of the Ionosphere in Modern Satellite Communications, A paper presented on Theoretical and Experimental Radio-Propagation Science at the International Centre for Theoretical Physics, Trieste, Italy, 6 – 24th Feb, 1989.

Li, G., B. Wing, W. Wan, and B. Zhao (2006). Observations of GPS Ionospheric Scintillations Over Wuhan During Geomagnetic Storms, *Ann. Geophs.*, **24**, pp. 1581-1590.

Liu, C. H., A. W. Wernik, K.C. Yeh, and M.Y. Youakim (1974a). Effects of Multiple Scattering on Scintillation of Transionospheric Radio Signals, *Radio Sci.*, 9(6), pp. 599 – 607.

Liu, C. H., A. W. Wernik, and K.C. Yeh (1974b). Propagation of Pulse Trains through a Random Medium, *IEEE Trans. Antenn. Propagat.*, 9(AP-22), pp. 624 – 627.

Liu, C. H. and K.C. Yeh (1975). Frequency and Spatial Correlation Functions in a Fading Communication Channel through the ionosphere, *Radio Sci.*, **10**(12), pp. 1055 – 1061.

Makela, J. J., M. C. Kelly, J. J. Sojka, X. Pi, and A. J. Mannucci (2001). GPS Normalization and Preliminary Modeling Results of Total Electron Content During a Mid-latitude Space Weather Event, *Radio Sc.*, **36**(2), pp. 351-361.

Manucci, A. J., B. D. Wilson, D. N. Yuah, C. H. Ho, U. J. Lindquister, and T.

F. Runge (1998). A Global Mapping Technique for GPS-derived Ionospheric Total Electron Content Measurements, *Radio Sc.*, **33**(3), pp. 562-582.

Matsushita, S. (1963). Equatorial Ionospheric Variations during Geomagnetic Storms, *J. Geophy. Res.*, **68**(9), pp. 2595-2599.

Mazaudier, C. A. (2008). Electric Current Systems in the Earth's Environment, *International Heliophysical Year (IHY) Workshop*, 10-22 November, Enugu, Nigeria.

McClure, J. P. (1964). The Height of Scintillation-Producing Ionospheric Irregularities in the Temperate Latitudes, *J. Geophys. Res.*, **69**(13), pp. 2775 – 2780.

McClure, J. P., W. B Hanson, and J. H. Hoffman (1977). Plasma Bubbles and Irregularities in the Equatorial Ionosphere, *J. Geophys. Res.*, **82**, pp. 2650.

Okeke, F. N. (2005). Geomagnetic Variations in the Ionosphere, Department of Physics and Astronomy, University of Nigeria, Nssuka, Nigeria.

Ostgaard, N. (2008). Energy Flow into the Magnetosphere-Ionosphere System, *International Heliophysical Year (IHY) Workshop*, 10-22 November, Enugu, Nigeria.

Oyinloye, O. (1987). Radio Wave Propagation in the Earth's Atmosphere, University of Ilorin Press, Ilorin, Nigeria, pp. 86-103.

Ozludemir, M. T. (2004). The Stochastic Modeling of GPS Observations. *Turkish J. Eng. & Env. Sc.*, **28**, pp. 223-231.

Panofsky W. K. H. and M. Phillips (1971). Classical Electricity and Magnetism, 2nd edn., Addison-Wesley Publishing Coy, Massachusetts, USA.

Pfaff, R. F., M. C. Kelly, B. G. Fejer, N. C. Maynard, and K. D. Baker (1982). In-situ Measurements of Waves Electric Fields in the Equatorial Electrojet, *Geophys Res., Lett.*, pp. 688.

Pfaff, R. F., M. C. Kelly, B. G. Fejer, E. Kudeki, and K. D. Baker (1988). Electric Field and Plasma Density Measurements in the Strongly-Driven Daytime Equatorial Electrojet-2. Two Stream Waves, *J. Geophys Res.*, **92**, pp. 13597.

Rama Rao, P. V. S., S. Tulasi Ram, S. Gopi Krishna, K. Niranjar, and D. S. V. V. D. Prasad (2006). Morphology and Spectral Characteristics of L-band and VHF Scintillations and their Impact on Transionospheric Communications, *Earth Planet Space*, **58**, pp. 895 – 904.

Ray, S., A. Paul, and A. Das Gupta (2006). Equatorial Scintillations in Relation to the Development of Ionization Anomaly, *Ann. Geophys.*, **24**, pp. 1429 – 1442.

Ratcliffe, J. A. (1972). An Introduction to the Ionosphere and Magnetosphere, Cambridge University Press, Cambridge, Great Britain, Ch. 8.

Rino, C. L., V. H. Gonzalez, and A. R. Hessing (1981). Coherence Bandwidth Loss in Transionospheric Radio Propagation, *Radio Sci.*, **16**(2), pp. 245 – 255.

Rishbeth, H., A. J. Lyon, and M. Peart (1963). Diffusion in the Equatorial F Layer, *J Geophs. Res.*, **68**(9), pp. 2559-2569.

Rishbeth, H. and O. K Garriot (1969). Introduction to Ionospheric Physics, Academic Press, NY, USA.

Rufenach, C. L. (1972). Power-Law Wavenumber Spectrum Deduced from Ionospheric Scintillation Observations, *J. Geophys. Res.*, 77(25), pp. 4761 –

4772.

Seo, J., T. Walter, E. Marks, T. Y. Chiou, and P. Enge (2007). Ionospheric Scintillation Effects on GPS Receivers during Solar Minimum and Maximum, *International Beacon Satellite Symposium*, 11-15 June, Boston, MA, USA.

Sreenivasiah, I. A. Ishimaru, and S.T. Hong (1976). Two Frequency Mutual Coherence Function and Pulse Propagation in a Random Medium: An Analytic Solution to the Plane Wave Case, *Radio Sci.*, **11**(10), pp. 775 – 778.

Serafim, P. E. (1980). Effect of Shape of Plasma Density Power Spectra on Ionospheric Scintillations, *Radio Sci.*, **15**(5), pp. 1031 – 1043.

Shimazaki, T. (1964). Nighttime Variations of F Region Electron Density Profiles at Puerto Rico, *J. Geophs. Res.*, **69**(13), pp. 2781-2797.

Singleton, D. G. (1968). The Morphology of Spread-F Occurrence over Half a Sunspot Cycle, *J. Geophys. Res.*, **73**(1), pp. 295– 308.

Singleton, D. G. (1977). The Reconciliation of an F-Region Irregularity Model with Sunspot Cycle Variations in Spread-F Occurrence, *Radio Sci.*, **12**(1), pp. 107– 113.

Standar, M (2008). A Green Approach, *Future Airport*, **1**, pp. S31-S32.

Strohbehn, J. W. (1968). Line-of-sight Wave Propagation through the Turbulent Atmosphere, *Proc. IEEE*, **56**(8), pp. 1301 – 1318.

Thomas, R. M., M. A. Cervera, K. Eftaxiadis, S. L. Manurung, S. Saroso, L. Effendy, A. G. Ramli, W. Salwa Hassan, H. Rahman, M. N. Dalimni, K. M. Groves, and Y. Wang (2001). A Regional GPS Receiver Network for Monitoring Equatorial Scintillation and Total Electron Content, *Radio Sc.*,

36(6),pp. 1545-1557.

Tsui J. B. (2000). Fundamentals of Global Positioning System Receiver: A Software Approach, John Wiley & Sons Inc, New York.

Validated International Civil Aviation Organization (ICAO) Standards and Recommended Practices for Global Navigation Satellite System (GNSS)-2000.

Valladares, C. E., S. Basu, K. Groves, M. P. Hagan, D. Hysell, A. J. Mazzella Jr., and R. E. Sheehan (2001). Measurements of the Latitudinal Distributions of Total Electron Content During Equatorial Spread F Events, *J. Geophys. Res.*, **106**(A12), pp. 29133-29152.

Valladares, C. E., J. Villalobos, R. Sheehan and M. P. Hagan (2004). Latitudinal Extension of Low-latitude Scintillations Measured with a Network of GPS Receivers, *Ann. Geophys.*, **22**, pp.3155-3175.

Valladares, C. E., and P. H. Doherty (2009). The Low-latitude Ionospheric Sensor Network (LISN), *Proceedings of the Institute of Navigation-ITM,* 26-28 Jan., Anaheim, CA, USA.

Van Dierendonck, A. J. (2005). GSV4004B GPS Ionospheric Scintillation and TEC Monitor (GISTM) User's Manual.

Van Dierendonck, A. J., J. Klobuchar, and Q. Hua (1993). Ionospheric Scintillation Monitoring Using Commercial Single Frequency C/A Code Receivers, *Proccedings of the Institute of Navigation (ION)*, GPS-93, Allexandria, Va, pp. 1333-1343.

Van Dyke, K. L. (2001). Use of Standalone GPS for Approach with Vertical Guidance. *Proceedings of the Institute of Navigation (ION) NTM-2001*, 22-24

Jan., Long Beach, CA, p. 301-309.

Viarrsson, L, S. Pullen, G. Green, and P. Enge (2001). Satellite Autonomous Integrity Monitoring and its Role in Enhancing GPS User's Performance. Department of Aeronautics and Astronautics, Stanford University, USA.

Wernik, A. W., C. H. Liu, and K.C. Yeh (1980). Model Computations of Radio Waves Scintillation caused by Equatorial Ionospheric Bubbles, *Radio Sci.*, **15**(3), pp. 559 – 572.

Wernik, A. W., J. A. Secan, and E. J. Fremouw (2003). Ionospheric Irregularities and Scintillation, *Adv. Space Res.*, **31**(4), pp. 971 – 981.

Whitney, H. E. and S. Basu (1977). The Effect of Ionospheric Scintillation on VHF/UHF Satellite Communications, *Radio Sci.*, **12**(1), pp. 123 – 133.

Woodman, R. F. and C. La Hoz (1976). Radar Observations of F Region Equatorial Irregularities, *J. Geophs. Res.*, **81**(31), pp. 5447-5466.

Yeh, K.C. and C. H Liu (1977). An Investigation of Temporal Moments of Stochastic Waves, *Radio Sci.*, **12**(5), pp. 671 – 680.

Yeh, K.C., C. H. Liu, and M.Y. Youakim (1975). A Theoretical Study of the Ionospheric Scintillation Behaviour caused by Multiple Scattering, *Radio Sci.*, **10**(1), pp. 97 –106.

Yeh, K.C and C. H. Liu (1982). Radio Wave Scintillations in the Ionosphere, *Proc. IEEE*, **70**(4), pp. 324 – 360.

Zernov, N. N. (2006). Propagation of Radio Waves in a Fluctuating Ionosphere: Ionospheric Scintillations, A paper presented at a Workshop on the Future of Ionospheric Research for Satellite Navigation and Positioning: Its Relevance for

Developing Countries, ICTP, Trieste, Italy, 4 – 15th Dec., 2006.

APPENDICES

Appendix A: Numerical Techniques

Appendix A1: Trapezoidal Rule

```
      Program Trapez
C     Program to solve for an integral of a function using
C     Trapezoidal Rule
C     Program written by Akala A.O on the 28th Nov., 2006
      Real Integ
      Y(X)=f(X)
      Write (*,*)'Enter the values of Xo,Xn,No. of subintervals'
      Read (*,*)Xo,Xn,N
      H=(Xn-Xo)/N
      Sum=Y(Xo)+Y(Xn)
      Do 10 I=1,N-1
C     Adding all values of Y together
      Sum=Sum+2.0*Y(Xo+(I*H))
C     Multiplying the sum by the step and then divided by 2.0
      Integ=(H/2.0)*sum
 10    Continue
       Write (*,20)Integ
 20     Format (5x,'The integral=',F7.3)
         Stop
         End
```

Appendix A2: Iterative Method for Difference Equation.

```
      Program equat
C     Program to solve difference equation using iterative method
C     Program written by AKALA A.O on the 12th Aug.,2007
      Open(Unit=1,File='Res1.out',Status='New')
       Write(*,*)'Enter the Number of Iterations'
        Read(*,*)N
C     Values of p are generated from 0.1 to 2.0
       p=0.099+(0.001*J)
          X=0.01
 10   Do 45 J=1,1901
 15   Do 40 I=1,N
       p=0.099+(0.001*J)
        X=1.0-(p*(X**2))
C    Only last hundred itrations are retained for every value of p
       If (I.GE.(N-100.0))Then
        Write (*,32)p,X
        Write (1,35)p,X
        Else
        End if
 32    Format (5x, F5.3, 2x, F9.7)
 35    Format (5x, F5.3, 2x, F9.7)
 40   Continue
 45   Continue
      Stop
      End
```

Appendix A3: Least Squares Method (Pseudo-inverse Application)

An over-determined linear equation is defined in matrix notation as,

$$A\hat{\mathbf{x}} \approx \mathbf{b} \qquad \text{(A4.1)}$$

where A is $m \times n$ matrix $(m > n)$, $\hat{\mathbf{x}}$ is an unknown n-dimensional parameter vector, and $\mathbf{b}$ is a known measured vector. Our interest is to find the solution of eq. (A4.1). A being a non-square matrix makes direct exact solution of eq. (A4.1) impossible. Under this condition, the number of equations (m) is more than the number of variables (n). To determine unique solutions for such equations, an optimization technique is used. To this end, we define a residual $\mathbf{r}$, of eq. (A4.1) as;

$\mathbf{r} = A\mathbf{x} - \mathbf{b}$.

The Euclidean norm squared of the residual is minimized to attain an optimization procedure. The Euclidean norm squared of the residual is defined as,

$$\|\mathbf{r}\|^2 = \sum_{i=1}^{m} \left([A\mathbf{x}]_i - \mathbf{b}_i \right)^2 \quad i = 1,2,3,.....m \qquad \text{(A4.2)}$$

where $[A\mathbf{x}]$ represents the i-th component of vector $A\mathbf{x}$. We can simplify eq. (A4.2) further by adopting the principle of least squares. In n-dimensional Euclidean Space, the squared norm of a is $a^T a$, (Kreyzsig, 2005), where a^T is the transpose of a.

$$\|\mathbf{a}\|^2 = (\mathbf{a},\mathbf{a}) = a^T a \qquad \text{(A4.3)}$$

Equation (A4.2) can therefore be re-written in form of eq. (A4.3) as,

$$\|r\|^2 = (A\mathbf{x} - \mathbf{b})^T (A\mathbf{x} - \mathbf{b}) = (A\mathbf{x})^T (A\mathbf{x}) - \mathbf{b}^T A\mathbf{x} - (A\mathbf{x})^T \mathbf{b} + \mathbf{b}^T \mathbf{b} \qquad \text{(A4.4a)}$$

but, $\mathbf{b}^T A\mathbf{x} = (A\mathbf{x})^T \mathbf{b}$, therefore,

$$(A\mathbf{x} - \mathbf{b})^T (A\mathbf{x} - \mathbf{b}) = (A\mathbf{x})^T (A\mathbf{x}) - 2(A\mathbf{x})^T \mathbf{b} + \mathbf{b}^T \mathbf{b} \qquad \text{(A4.4b)}$$

The minimum value for the Euclidean norm squared of the residual can then be determined when its derivative with respect to $\mathbf{x}$ assumes value zero, that is,

$$\frac{d}{d\mathbf{x}}\left[(A\mathbf{x})^T(A\mathbf{x}) - 2(A\mathbf{x})^T\mathbf{b} + \mathbf{b}^T\mathbf{b}\right] = 2A^TA\hat{\mathbf{x}} - 2A^T\mathbf{b} = 0 \qquad \text{(A4.5a)}$$

From eq. (A4.5a), the minimizing vector $\hat{\mathbf{x}}$ is the solution of the normal equation. Thus,

$$A^TA\hat{\mathbf{x}} = A^T\mathbf{b} \qquad \text{(A4.5b)}$$

Vector $\hat{\mathbf{x}}$ can be easily solved by multiplying $(A^TA)^{-1}$ by $A^T\mathbf{b}$. The product (A^TA) results to an invertible symmetric square matrix unlike matrix-A.

$$\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b} \qquad \text{(A4.6)}$$

where $(A^TA)^{-1}A^T$ is the pseudo-inverse of matrix-A (Tsui, 2000).

Appendix A4: Software Approach to GPS User's Position Fixing

```
      Program Range
C     Program written by AKALA AO on the 20th Jan., 2008
      Real lat, long, L, Lc,lpe
      Parameter (N=N)
     Dimension S(N,3),Rg(N),R(3),a(N,4),Pr(N),dl(4),dRg(N),b(N,4)
      Dimension pinv(N),c(N,4),D(N),dx(4),bias(N)
      Open(unit=1,file='sat1.out',status='new')
C   Enter the user guess position coordinates and the clock error
      Write(*,*)'Enter user guess positions and clock error'
      Read(*,*)X0,Y0,Z0,b0
        R(1)=X0
        R(2)=Y0
        R(3)=Z0
        bu=b0
C     Enter the satellite position in matrix format
      Write(*,*)'enter the satellites positions'
      Read(*,*)((S(i,j),j=1,3),i=1,N)
C   Enter the premeasured values of pseudoranges for each
C   satellite
        Write(*,*)'Enter pseudorange values'
       Read(*,*)(Pr(i),i=1,N)
        Write(*,*)'Enter satellite bias values'
       Read(*,*)(bias(i),i=1,N)
  5     Do 10 i=1,N
  8    Rg(i)=((S(i,1)-R(1))**2+(S(i,2)-R(2))**2+(S(i,3)-
     * R(3))**2)**0.5
```

```
 10   Continue
        write(*,*)(i,Rg(i),i=1,N)
C     ************************************************************
C     Generating the first three columns for the a-matrix
 12     Do 20 j=1,3
 14     Do 15 i=1,N
        a(i,j)=(S(i,j)-R(j))/Rg(i)
 15     continue
 20      continue
C      Generating the fourth column with ones-element
        Do 22 i=1,N
        a(i,4)=-1
 22       continue
C      Determining the product of transpose of matrix A and A
         do 39 j=1,4
         do 35 i=1,4
         C(i,j)=0.0
 35        continue
           do 40 i=1,N
          do 36 k=1,N
         C(i,j)=c(i,j)+(a(k,j)*a(k,i))
 36        continue
 40         continue
 39        continue
C     ************************************************************
C       Determining the change in range
        Do 44 i=1,N
        dRg(i)=pr(i)-Rg(i)-bias(i)
 44       Continue
             Do 46 i=1,N
           D(i)=0.0
 46          continue
           Do 49 i=1,4
           Do 47 k=1,N
           D(i)=D(i)+a(k,i)*dRg(k)
 47          continue
 49          continue
C     ************************************************************
C       Determining the augmented matrix
       Do 52 j=1,5
        c(i,5)=D(i)
 52       continue
C       Solving the augmented matrix with Gaussian Elimination
         Do 54 j=1,3
         Do 54 i=j+1,4
         T=c(i,j)/c(j,j)
         Do 54 k=1,5
          c(i,k)=c(i,k)-c(j,k)*T
 54         continue
         dx(4)=c(4,5)/c(4,4)
C     ********Starting the back substitution procedure**********
         Do 58 i=3,1,-1
         sum=0.0
         Do 57 j=i+1,4
         dx(4)=c(4,5)/c(4,4)
          Sum=sum+c(i,j)*dx(j)
          dx(i)=(c(i,5)-sum)/c(i,i)
```

```
 57        Continue
 58        Continue
            bu=bu+dx(4)
 61           Do 62 k=1,3
            R(k)=R(k)+dx(k)
 62    Continue
            error=((dx(1)**2)+(dx(2)**2)+(dx(3)**2))**0.5
            Write(*,*)'Enter the threshold value'
            Read(*,*)threshold
            if(error.LT.threshold) then
            goto 5
            else
            goto 64
            end if
 64        Write(1,*)
          Write(1,*)'The user position in cartesian coordinate'
           Write(1,*)(i,dx(i),i=1,3)
            Write(1,*)dx(4)
C    **Converting the results to spherical coordinate system***
           rse=(dx(1)**2+dx(2)**2+dx(3)**2)**0.5
C    ***Determining the latitude and longitude of the user*****
            Lc=ATAN(dx(3)/(((dx(1)**2)+(dx(2)**2))**0.5))
            long=(ATAN(dx(2)/dx(1)))*(180.0/(4.0*ATAN(1.0)))
C     ***longitude in conformity to practical earth geometry****
         e=1.0/298.257223563
 65         lpe=Lc
 66         L=Lc+e*SIN(2.0*lpe)
             erro2=ABS(lpe-L)
             if(erro2.GE.1.0)then
             goto 65
             else
             lpe=L
             BB=SIN(1.0*L)**2
             end if
       lat=L*(180.0/(4.0*ATAN(1.0)))
             re=6378137.0
             alt=rse-re*(1.0-e*BB)
        Write(1,*)
       Write(1,*)'The user latitude, longitude and altitude'
        Write(1,68)lat
        Write(1,69)long
        Write(1,*)
         Write(1,*)' alt='
           Write(1,*)  alt, 'm'
 68     Format(2x,'lat=',F6.1, 'deg')
 69     Format(/2x,'long=',F6.1, 'deg')
 70          stop
           end
```

Sample Output:

```
The user position in cartesian coordinate
  1 91709040.000000   2  25230180.000000   3 11567400.000000
  74394980.000000

The user latitude, longitude and altitude
  lat= 7.0deg
```

long= 15.4deg

alt=
89439300.000000m

Appendix B: Transformation of equations (3.32) and (3.33) to (3.37) and (3.38).

It has earlier been shown in the mains that, $r = \rho/l_o$, $\xi = x/l_o$, $\eta = y/l_o$, $\zeta = z/k_o l_o^2$ $\omega_1 = \omega_o(1-X)$, $\omega_2 = \omega_o(1+X)$, $k_o = \omega_o/c$; $\langle\varepsilon(\omega)\rangle = (1-\omega_{po}^2/\omega^2)$.

Using the above expressions, it implies that,

$$k_1 = \frac{\omega}{c}\alpha(1-X) \text{ and } k_2 = \frac{\omega}{c}\beta(1-X) \tag{B1}$$

where $\alpha = \left[1-\left\{(\omega_p/\omega_o)^2/(1-X)^2\right\}\right]^{\frac{1}{2}}$ and $\beta = \left[1-\left\{(\omega_p/\omega_o)^2/(1+X)^2\right\}\right]^{\frac{1}{2}}$.

The partial derivative of the correlation function in z can be expressed in dimensionless form as,

$$\frac{\partial\Gamma}{\partial z} = \frac{1}{k_o l_o^2}\frac{\partial\Gamma}{\partial\zeta} \tag{B2}$$

Also, $\nabla_\perp^2\Gamma$ is now represented in terms of ξ and η.

$$\frac{k_2-k_1}{2k_1k_2}\nabla_\perp^2\Gamma = \frac{1}{k_o l_o^2}\frac{\left[(\alpha+\beta)X+\beta-\alpha\right]}{2D}\nabla_\perp^2\Gamma \tag{B3}$$

Defining,

$D = \alpha\beta(1-X^2)$ and $B_1 = \left[(\alpha+\beta)X+\beta-\alpha\right]/2D$, then,

$$j\left(\frac{k_2-k_1}{2k_1k_2}\right)\nabla_\perp^2\Gamma = j\frac{B_1\nabla_\perp^2\Gamma}{k_o l_o^2} \tag{B4}$$

$$k_2^2+k_1^2 = \frac{\omega_o^2}{c^2}\left[\beta^2(1+X)^2+\alpha^2(1-X^2)\right] \tag{B5}$$

$$k_1^2 k_2^2 = \frac{\omega_o^4}{c^4} D^2 \;;\; k_1 k_2 = \frac{\omega_o^2}{c^2} D \tag{B6}$$

In equation (3.44), $B_3 = \left[1 + X^2 - \left(\omega_p/\omega_o\right)^2\right]/D^2$, therefore,

$$\frac{k_{po}^4}{8}\left[\left(\frac{k_2^2 + k_1^2}{k_1^2 k_2^2}\right) A_{\Delta N}(0) + \frac{2}{k_1 k_2} A_{\Delta N}(\vec{\rho})\right]\Gamma \equiv$$

$$\left(\frac{1}{4}\right)\left(\frac{\omega_p}{\omega_o}\right)^4 k_o^2\left[B_3 A_{\Delta N}(0) + \frac{1}{D} A_{\Delta N}(\vec{\rho})\right]\Gamma \tag{B7}$$

Adding equations (B2), (B4) and (B7) together,

$$\frac{\partial \Gamma}{\partial \zeta} + jB_1 \nabla_\perp^2 \Gamma + \left(\frac{1}{4}\right)\left(\frac{\omega_p}{\omega_o}\right)^4 \frac{k_o^3 l_o^3}{l_o} \frac{\left\langle \left(\Delta N/N_o^2\right)^2 \right\rangle}{\left\langle \left(\Delta N/N_o^2\right)^2 \right\rangle} \cdot \left[B_3 A_{\Delta N}(0) - \frac{1}{D} A_{\Delta N}(\xi,\eta)\right]\Gamma = 0 \tag{B8}$$

Appendix C: Application of Wentzel-Kramer-Briollouin (WKB) Approximation (Budden, 1985, ch.7).

As expressed in equation (3.50),

$$\Gamma = \exp -\left(CB_3 G(0)\zeta + \phi\right),\ \Gamma_1 = \exp \phi \tag{C1}$$

$$\frac{\partial \phi}{\partial \zeta} \cdot e^{\phi} + jB_1(\nabla_\perp^2 \phi - (\nabla \phi)^2) \cdot e^{\phi} - \frac{C}{D} G(\xi,\eta) \cdot e^{\phi} = 0 \tag{C2}$$

If the initial condition $\phi(\rho,0) = 0$ is applied, then, equation (C2) will reduce to,

$$\frac{\partial \phi}{\partial \zeta} + jB_1(\nabla_\perp^2 \phi - (\nabla \phi)^2) - \frac{C}{D} G(\xi,\eta) = 0 \tag{C3}$$

Equation (C3) can be expressed in terms of the Γ_1 by using the differential equation below,

$$\frac{\partial \Gamma_1}{\partial \zeta} = \frac{\partial \Gamma_1}{\partial \phi} \cdot \frac{\partial \phi}{\partial \zeta} = \frac{\partial \Gamma_1}{\partial \zeta} \cdot e^{\phi} \tag{C4}$$

Therefore,

$$\frac{\partial \Gamma_1}{\partial \zeta} + jB_1(\nabla_\perp^2 \Gamma_1 - (\nabla \Gamma_1)^2) - \frac{C}{D} G(\xi, \eta)\Gamma_1 = 0 \tag{C5}$$

For a simple isotropic case, the non-linear term $(\nabla \Gamma_1)^2$ can be neglected.,

$$\frac{\partial \Gamma_1}{\partial \zeta} + jB_1 \nabla_\perp^2 \Gamma_1 - \frac{C}{D} G(\xi, \eta)\Gamma_1 = 0 \tag{C6}$$

Adopting the same process but with an initial condition $\Gamma_3(\rho, 0) = 1$,

$$\frac{\partial \Gamma_3}{\partial \zeta} + \frac{C}{D} G(\xi, \eta)\Gamma_3 = 0 \tag{C7}$$