

## Ionospheric Disturbance on GNSS Receivers and Other Systems

*Master's thesis in Radio and Space Science*

WEIHUA WANG

Department of Earth and Space Sciences  
CHALMERS UNIVERSITY OF TECHNOLOGY  
Gothenburg, Sweden 2015



THESIS FOR THE DEGREE IN MASTER OF SCIENCE

# Ionospheric Disturbance on GNSS Receivers and Other Systems

Weihua Wang



**CHALMERS**  
UNIVERSITY OF TECHNOLOGY

Department of Earth and Space Sciences  
CHALMERS UNIVERSITY OF TECHNOLOGY  
Gothenburg, Sweden 2015

Ionospheric Disturbance on GNSS Receivers and Other Systems  
Weihua Wang

© Weihua Wang, 2015.

Supervisor: Per Jarlemark, SP Technical Research Institute of Sweden  
Examiner: Jan Johansson, Department of Earth and Space Sciences

Master of Science Thesis  
Department of Earth and Space Sciences  
Research Group of Space Geodesy and Geodynamics  
Chalmers University of Technology  
SE-412 96 Gothenburg  
Telephone +46 31 772 1000

Dedicated to my parents Hanxing Wang and Fang Wang

Ionospheric Disturbances on GPS Receivers and Other Systems  
Weihua Wang  
Department of Earth and Space Sciences  
Chalmers University of Technology

## **Abstract**

Equatorial ionospheric plasma irregularities are generated at the magnetic equator areas after sunset due to plasma instabilities. Its impact on the GPS (Global Positioning System) signals, both amplitude and phase, often results in a degraded performance of the GPS receivers. Based on observational data obtained from dual-frequency GPS receivers at two triangular observational networks in south Brazil, ionospheric variability over one full year range (2013) is studied through the Total Electron Content (TEC) distribution. Small-scaled local TEC variations of the region of two observational networks are also presented. From characterization of the local TEC variations, it is thought to be primarily related to the local Equatorial Plasma Bubbles' (EPBs) activities due to the Equatorial Ionospheric Anomaly (EIA).

Keywords: GPS signals, TEC, equatorial plasma irregularities.

## **Acknowledgements**

The author would like to sincerely thank Professor Jan Johansson of Chalmers University of Technology, and Dr. Per Jarlemark at SP Technical Research Institute of Sweden for their significant and kind help in theoretical guidance, simulations, supervision and comments on this project.

Weihua Wang, Gothenburg, May 2015

# List of Abbreviations

BRDC	Broadcast Ephemeris Data
C/A-code	Coarse/acquisition code
CDMA	Code Division Multiple Access
DOP	Dilution Of Precision
EIA	Equatorial Ionospheric Anomaly
EPBs	Equatorial Plasma Bubbles
ESA	European Space Agency
IOC	Initial Operational Capability
FOC	Full Operational Capability
GLONASS	Globalnaya Navigazionnaya Sputnikovaya Sistema
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GRT	Generalized Rayleigh-Taylor
IAACs	Ionosphere Associate Analysis Centers
IBGE	Instituto Brasileiro de Geografia e Estatstica
IGS	International GNSS Service
IOC	Initial Operational Capability
IONEX format	IONospheric map EXchange ASCII format
IPP	Ionospheric Pierce Point
JPL	Jet Propulsion Laboratory
LOS	Line Of Sight
LSTIDs	Large Scale Traveling Ionospheric Disturbances
MCS	Master Control Station

MSTIDs	Medium Scale Traveling Ionospheric Disturbances
NBI	Narrow-band Interference
NASA	National Aeronautics and Space Administration
P-code	Precision code
PRN	Pseudo-random Noise
SAA	South Atlantic Anomaly
SID	Sudden Ionospheric Disturbance
SPENVIS	Space Environment Information System
STEC	Slant Total Electron Content
SV	Space Vehicles
TEC	Total Electron Content
TECU	Total Electron Content Unit
TID	Traveling Ionospheric Disturbance
TOA	Time Of Arrival
URE	User Equivalent Range Error
UT	Universal Time
VTEC	Vertical Total Electron Content

# Contents

<b>List of Figures</b>	<b>x</b>
<b>List of Tables</b>	<b>xii</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 GPS Background</b>	<b>2</b>
2.1 GPS Overview . . . . .	2
2.2 GPS Segments . . . . .	3
2.3 GPS Signals . . . . .	6
2.4 Code Pseudo-range Measurements . . . . .	8
2.5 Carrier Phase Measurements . . . . .	11
2.6 Doppler Measurements . . . . .	12
2.7 Errors and Biases . . . . .	13
<b>3 Ionosphere Background</b>	<b>16</b>
3.1 Height Characteristics . . . . .	17
3.2 Ionosphere Variation Types . . . . .	18
3.2.1 General Variations of Global Scale . . . . .	18
3.2.2 Specific Variations Around Tropical Areas . . . . .	19
<b>4 Signal Propagation through the Ionosphere</b>	<b>23</b>
4.1 Some Fundamentals of Wave Propagation . . . . .	23
4.2 Ionospheric Refraction . . . . .	25
<b>5 Methodology of Obtaining the Ionospheric Variability</b>	<b>30</b>
<b>6 Data Set</b>	<b>33</b>
<b>7 Results and Discussion</b>	<b>36</b>
7.1 swTEC Analysis . . . . .	36
7.2 Correlation Study . . . . .	39

7.3	Comparison of TEC and swTEC . . . . .	40
7.4	Conclusion . . . . .	43
7.5	Future Work . . . . .	44
<b>References</b>		<b>45</b>
<b>Appendix A Methodology for Estimating the Ionospheric Variability</b>		<b>I</b>
<b>Appendix B Observational Data Set</b>		<b>IV</b>

# List of Figures

2.1	Illustration of satellite positioning via an intersection of three spheres, where the clock icons represent orbital satellites. . . .	3
2.2	GPS constellation of 24 satellites. . . . .	3
2.3	Initial deployment of GPS control sites. . . . .	5
2.4	Major components of a GPS receiver, redrawn from [1]. . . . .	5
2.5	GPS segments illustration, redrawn from [2]. . . . .	6
2.6	An illustration of the pseudo-range measurements (Note: $\Delta t$ = signal propagation time), redrawn from [3]. . . . .	9
2.7	Major components of GPS errors and biases, redrawn from [3].	13
2.8	Good and bad satellite geometry illustration. . . . .	15
3.1	Illustration of the ionospheric layers in daytime and nighttime.	17
3.2	World geomagnetic inclination in South American regions [4].	21
3.3	Main magnetic field at Earth's surface in June 2014 [5]. . . . .	22
3.4	SPENVIS simulation of SAA at 550 km. Note that values in the index bar are logarithmic. . . . .	22
4.1	Ionospheric single layer model illustration, redrawn from [1]. .	27
5.1	An illustration of linear interpolation. . . . .	31
6.1	Map of the eight selected Brazilian observational sites. . . . .	34
7.1	Local ionospheric variability comparison for the two observational networks of a whole 2013 year. . . . .	37
7.2	Histogram plot showing the occurrence frequency of the swTEC unit. Median value: 1.4392 (top) and 0.9677 (below). . . . .	38
7.3	Whole year range correlation plot. . . . .	39
7.4	Correlation plot from April to August. . . . .	39
7.5	Illustration of plasma bubble observation. . . . .	40
7.6	Whole year TEC and swTEC comparison. . . . .	41
7.7	Spacial high-pass filtering effect can be resolved by the local observational network. . . . .	43

# List of Tables

2.1	Basic characteristic features of the GPS [1] . . . . .	4
2.2	The GPS satellite signals [1] . . . . .	8
2.3	The navigation message format [6] . . . . .	8
2.4	Sources of User Equivalent Range Errors [6] . . . . .	15
3.1	MSTIDs and LSTIDs . . . . .	20
B.1	Specific date used from IBGE website each Thursday . . . . .	IV

# Chapter 1

## Introduction

Intense ionospheric disturbance can lead to unwanted terrestrial consequences such as malfunctioning satellite communications and navigation, as well as electric power grids. As the GNSS (Global Navigation Satellite System) radio signal propagating through the atmosphere, ionospheric variation will not only affect the signal velocity, resulting in phase advance and group delay, but also bend the signal. Strong amplitude fading and phase distortion due to violent ionospheric disturbances will degrade the navigation performance a lot. Therefore, understanding the causes and mitigating the impact of the ionospheric disturbance is quite important.

Equatorial regions in Brazil can sometimes experience severe problems associated with ionospheric irregularities, mainly due to influences of the solar activities, daily and seasonal variations due to movement of the Earth, and abnormal distribution of the geomagnetic field. However, still need to further investigate, the ionospheric variations over Brazilian territory seems it may influence applications such as the calculation of the atomic time scale for satellite-based timing services and positioning at aviation traffic control.

The main purpose of this project is to study the ionospheric variability over Brazil. Characterization of the electron distribution variations, represented by the Total Electron Content (TEC), are analyzed from both temporal and spatial point of view, based on local observational data continuously recorded from geographically distributed GNSS receivers in southern Brazil. In addition, equatorial ionospheric anomaly, supposed to be closely related with activities of the Equatorial Plasma Bubbles (EPBs), has been preliminary explored and compared with outcome from the Vertical Total Electron Content (VTEC) map produced by the Ionosphere Working Group of International GNSS Service (IGS).

# Chapter 2

## GPS Background

### 2.1 GPS Overview

The satellite-based Global Positioning System (GPS) can provide precise continuous positioning, timing information and speed to any user equipped with a GPS receiver near the Earth surface. The GPS operates independent of weather conditions or geospatial restrictions as long as unobstructed line-of-sight (LOS) communication links to four or more GPS satellites are set up [3]. Basic features of the GPS are displayed in Table 2.1 [1].

The fundamental idea behind GPS is to measure the distances between a GPS receiver and several simultaneously LOS satellites. Since a user want to utilize the GPS satellite constellation to perform positioning and navigation, information about the satellite coordinates must be known in advance. Then the time of arrival (TOA) measurement, which aims to measure the signal propagation time between the satellite and the receiver, is performed. By multiplying the signal propagation time with the speed of the light, the satellite-to-receiver range can be computed. Finally, position of the receiver can be achieved by applying several TOA measurements with respect to multiple visible satellites.

Theoretically, three distances between the GPS receiver and three simultaneously visible satellites can be used to solve the user's 3D position (latitude, longitude and altitude). In this case, location of the receiver is resulting from an intersection of three spheres. Each has a radius of one receiver-to-satellite range and is centered on that particular satellite. A fourth satellite is added to solve the receiver clock offset [7]. In practice, more than four satellites are always visible for reliability and availability purposes. Figure 2.1 shows that a specific position can be determined from an intersection of the three spheres.

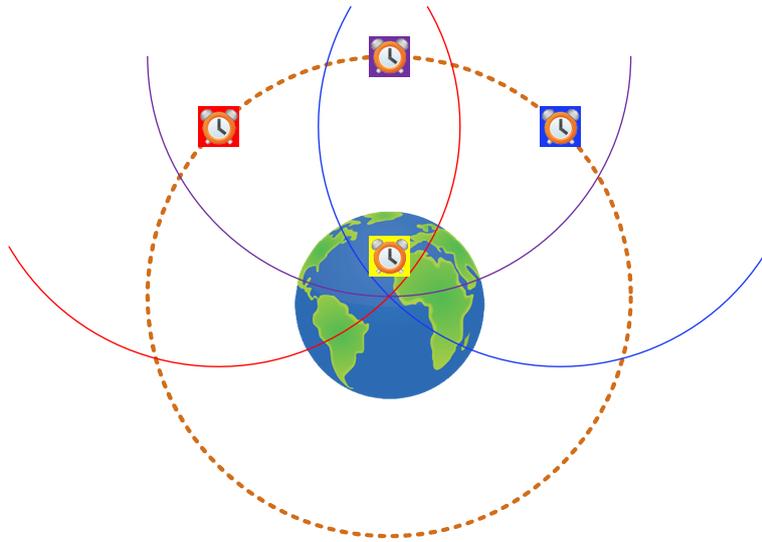


Figure 2.1: Illustration of satellite positioning via an intersection of three spheres, where the clock icons represent orbital satellites.

## 2.2 GPS Segments

GPS consists of three segments, namely the space, control and user segment. The space segment includes on-orbit operational satellites, or space vehicles (SV) in GPS jargon, and its constellation. Initially 24 active SVs, geometry of the constellation is carefully designed such that four satellites are placed in each of six inclined orbital planes. As a result, more than four GPS satellites are simultaneously visible anywhere on the ground worldwide, as long as a LOS propagation is fulfilled. This constellation was known as the initial operational capability (IOC) and is illustrated in Figure 2.2 by a MATLAB simulation. Since the GPS achieved its full operational capability (FOC)

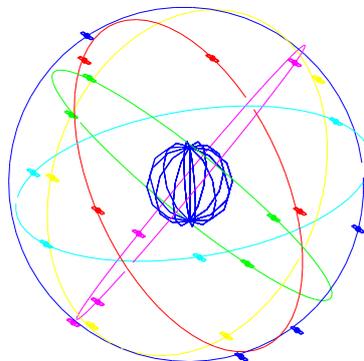


Figure 2.2: GPS constellation of 24 satellites.

Table 2.1: Basic characteristic features of the GPS [1]

orbital height	20220 km
orbital radius	26600 km
inclination	55°
period	11 h 58 m
frequencies	1575.42 MHz, 1228.60 MHz
navigation data	4D: X,Y,Z,t; velocity
availability	continuously
accuracy	15 m, 0.1 m/s
constellation	24 - 32 SV
geometry	repeating
satellite clocks	Rubidium, Cesium

after 1995 [8], the number of satellites, in the GPS constellation has always been no less than 24 SVs. Some orbital characteristics can be read from Table 2.1.

The control segment consists of a worldwide network of tracking and controlling stations. Initial deployment of the control segment composes of several components: a Master Control Station (MCS) at Schriever near Colorado Springs, an alternate MCS at Vandenberg in California, four monitor stations co-located with ground antennas at Ascension Island, Cape Canaveral, Diego Garcia and Kwajalein. Besides, two more monitor stations are located in Hawaii and Colorado Springs. Major tasks processed by the control segment includes satellite system integrity monitor and control, GPS time determination, satellite ephemeris prediction, navigation message update, relocation of functional failure satellite, and orbit maintenance. Figure 2.3 shows an initial control segment network deployment. Note that there are 17 monitoring stations nowadays worldwide for passively track all SVs in view [8].

The user segment contains all users for different application fields. Equipped with a GPS receiver and an antenna, both military and civilian users can receive the broadcast GPS signals, furthermore the user's position can be determined anytime anywhere near the Earth's surface.

According to reference [1] the main components of a GPS receiver are displayed in Figure 2.4. Signals emitted from a GPS satellite are first picked up by an antenna which must be right-hand circularly polarized. Next, the signals are down-shifted at radio frequency (RF) front end. This step is achieved

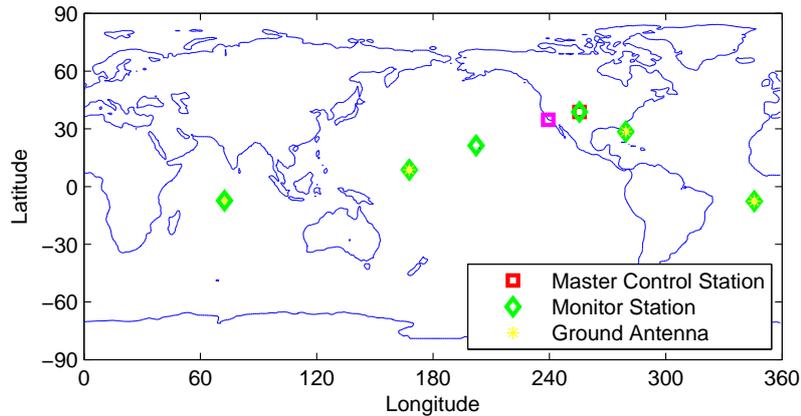


Figure 2.3: Initial deployment of GPS control sites.

by a mixer which combines the incoming RF signals with a sinusoidal signal generated from a local oscillator, resulting in intermediate frequency (IF) signals. The down-converted IF signals are immediately forwarded to a signal processor where algorithms of signal tracking and cross correlation are performed. The received signals tracking from all visible satellites, at this step, are isolated, identified by their codes, and assigned for further processing. The microprocessor controls the receiver's operation, including digital signal processing, decoding of the broadcast navigation message, control of the signal processor, on-line positions and velocities computation, and so on. The user communication part serves as a bridge between the user and the microprocessor, with common tasks such as accepting commands from the user, displaying of computing results and so on. Data logger and memory devices store the data for further post-processing. Power supply provides low voltage DC power for the entire GPS receiver.

In conclusion, the space and control segments are developed and operated by the U.S. Air Force. Each GPS receiver of the user segment utilizes the

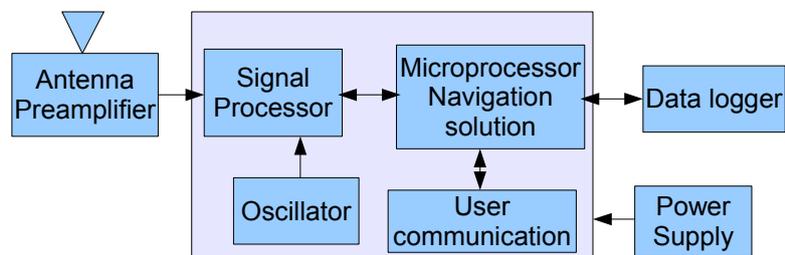


Figure 2.4: Major components of a GPS receiver, redrawn from [1].

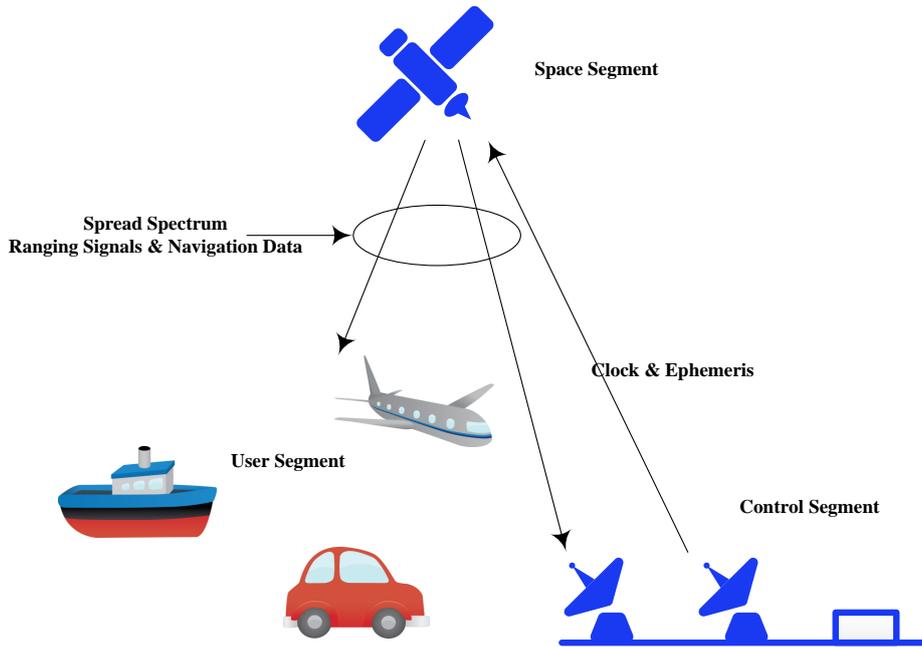


Figure 2.5: GPS segments illustration, redrawn from [2].

broadcast satellite signals to acquire its 3D position and the current time information. The three segments of GPS are collectively illustrated in Figure 2.5.

## 2.3 GPS Signals

Generally, signal emitted from each satellite includes several components: two sine waves, three digital codes and a navigation message. The L-band (radio spectrum range between 1 and 2 GHz) sine waves, also known as carrier frequencies, are coherently derived from the fundamental frequency of 10.23 MHz (see Eq.(2.1)) which is generated by the space-borne highly accurate atomic oscillator. Accessibility of the L1 and L2 carrier allows for correcting the ionospheric delay error which will be discussed in later chapters.

$$\begin{aligned}
 L1 &: 154 \times 10.23 \text{ MHz} = 1575.42 \text{ MHz} \\
 L2 &: 120 \times 10.23 \text{ MHz} = 1227.60 \text{ MHz}
 \end{aligned}
 \tag{2.1}$$

The digital codes and navigation message are transmitted by each satellite on both carriers. In order to minimize the signal interference among adjacent satellites and to save the scarce frequency resource, code division multiple access (CDMA) is used so that all GPS satellites transmit with the same frequency. Moreover, spread spectrum technique which resists against

narrow-band interference (NBI, typically emanates from radio stations or mobile phones), noise and jamming is used for security and power-saving purposes. Both the carriers and the codes are used primarily to measure the range from the satellite transmitter to the GPS receiver.

The carriers are modulated with pseudo-random noise (PRN) codes, or pseudo-random binary sequences. The noise-like property of the PRN codes result in a very low correlation with any other sequence generated by other satellites or with the sequence itself but at different time. In addition, the correlation of the PRN code with the NBI, or with the thermal noise, is also very low which strengthen the noise resistance ability. Another advantage of the PRN code lie in its periodicity, hence certain mathematical expression exists in order to generate the same PRN code. This makes satellite-to-receiver distance calculation feasible. In theory, if both the transmitter and the receiver generate exactly the same PRN code, a very high correlation between the replica sequence generated by the local GPS receiver and the transmitted sequence produced by the satellite could be achieved as long as both clocks of the receiver and the transmitter are synchronized.

Traditional GPS satellites transmit two navigation codes known as coarse acquisition code (C/A-code) and precision code (P-code). Both digital codes are in fact the PRN codes. Historically, the P-code is modulated onto both the L1 and L2 carriers, while the C/A-code is only available on the L1 carriers. Although relatively less precise compared with that of the authorized P-code, less complexity of the C/A-code enable all users to use it for range measurement. The P-code is the principal navigation ranging code. However, it is so long and complex so that a GPS receiver could not directly acquire and synchronize with it alone. In real application, the receiver would first lock onto the simple C/A-code and then, after extracting information of the current time and approximate position coordinates, synchronize with the P-code later. A third code type, known as the P(Y)-code, is utilized wherever the anti-spoofing mode of operation is required. The P(Y)-code is encrypted by multiplying the P-code with highly classified W-code [9]. This indicates that the primary users of the P(Y)-code belong to the U.S. military sections. A summary of the GPS signal types are listed at Table 2.2 [1].

With a bitrate of 50 bits per second (bps), the navigation message transmitted by every satellite includes all the necessary information about the operational satellites and the constellation, such as 'health' status, orbital elements, clock behavior and other system parameters. Additionally, an almanac is also provided which offers the approximate data for each active satellite [8]. Each complete navigation message contains five sub-frames. A brief description of each sub-frame can be read from Table 2.3 [6].

Table 2.2: The GPS satellite signals [1]

Atomic clock fundamental frequency	10.23 MHz
L1 carrier signal	$154 \times 10.23$ MHz
L1 frequency	1575.42 MHz
L1 wavelength	19.0 cm
L2 carrier signal	$120 \times 10.23$ MHz
L2 frequency	1227.60 MHz
L2 wavelength	24.4 cm
P-code frequency (chipping rate)	10.23 MHz (Mbps)
P-code wavelength	29.31 m
P-code period	266 days; 7 days/satellite
C/A-code frequency	1.023 MHz
C/A-code wavelength	293.1 m
C/A-code period	1 millisecond
data signal frequency	50 bps
data signal cycle length	30 second

Table 2.3: The navigation message format [6]

Subframe	Description
1	Satellite clock, GPS time relationship
2-3	Ephemeris (precise satellite orbit)
4-5	Almanac component (network synopsis, error correction)

## 2.4 Code Pseudo-range Measurements

When satellites-to-receiver ranges are measured, we have to address several issues regarding the GPS observables. Basically, the observations include two methods, namely the pseudo-range measurements and the carrier phase measurements. One effect is due to the Doppler shift and can be considered in measurements where accurate velocity computations are required [10].

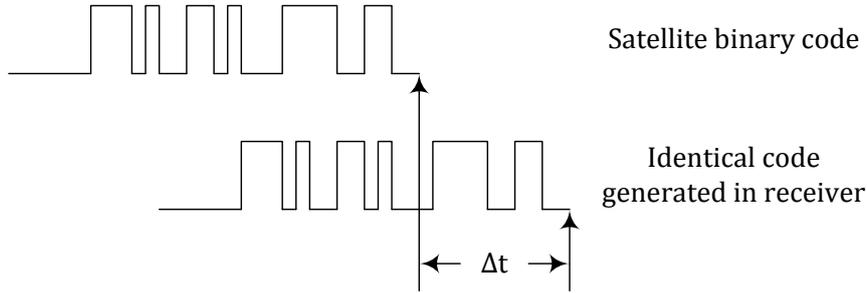


Figure 2.6: An illustration of the pseudo-range measurements (Note:  $\Delta t =$  signal propagation time), redrawn from [3].

The pseudo-range measurements determine the distance between the satellite transmitter and antenna of the receiver through measuring the GPS signal propagation time between the two.

According to [3] the idea behind the pseudo ranging is quite straightforward. First, let us assume that both the GPS satellite and the receiver are perfectly synchronized with each other at a certain instant. In other words, when the modulated signal with the PRN code is emitted from the satellite, the C/A-code or the P-code are precisely aligned with the space-borne clock. Meanwhile, an exact replica of the PRN code is generated by the receiver, accurately matched with the receiver clock. A short time later, the GPS signal is picked up by the receiver's antenna. By a maximum cross correlation analysis of the received PRN code with the local replica, the signal propagation time can be acquired. Furthermore, the satellite-to-receiver range can be computed through a multiplication of the signal propagation time with the velocity of the radio wave. A brief illustration of this ranging idea can be seen in Figure 2.6.

In real applications, both clocks can not be perfectly synchronized. Because the calculated range is not the correct range due to time biases. This is why 'pseudo' range is used. Besides the clock offsets, error sources corresponding to the propagation medium and the signal multi-path also result in deviation between measured pseudo-range and the true geometric distance. Note that the signal transmission path is slightly longer than the true geometric distance due to refraction of the atmosphere.

The mathematical derivation found in Equation (2.2) - (2.7) follows from [11]. Let us derive the pseudo-range expression under the simplest scenario, then more practical representation can be obtained by adding the error source terms. In a vacuum medium and error-free case, the pseudo-range equals the geometric satellite-to-receiver distance. An expression of the pseudo-range follows:

$$R_r^s = (t_r - t^s)C \quad (2.2)$$

where  $R_r^s$  is the pseudo-range between the receiver  $r$  and the satellite  $s$ ;  $t_r$  is the signal reception time at the receiver, while  $t^s$  is the signal emission time at the satellite transmitter; and  $C$  is the speed of light.

By introducing the clock biases, the pseudo-range can be modified to

$$R_r^s = (t_r - t^s)C - (\Delta t_r - \Delta t^s)C \quad (2.3)$$

where  $\Delta t_r$  and  $\Delta t^s$  are clock offsets for the receiver and the satellite, respectively. The satellite clock bias,  $\Delta t^s$ , can be calculated by the receiver implicitly from the GPS navigation message. More precise satellite clock information can be obtained from data centers of the IGS.

The geometric distance between the satellite and the receiver can be easily computed as

$$\rho_r^s(t_r, t^s) = \sqrt{(x_r - x^s)^2 + (y_r - y^s)^2 + (z_r - z^s)^2} \quad (2.4)$$

where  $(x_r, y_r, z_r)$  and  $(x^s, y^s, z^s)$  refer to coordinate vector of the ground receiver and the satellite, respectively. Both the coordinate vectors are functions of time, which implies that the geometric distance is indeed a function of two time variables, namely  $t_r$  and  $t^s$ . Moreover, the emission time of  $t^s$  is unknown in practice, but it can be presented in terms of the signal propagation time  $\Delta t$  as

$$t^s = t_r - \Delta t \quad (2.5)$$

The geometric distance in terms of the signal reception time and the propagation time can be expressed as

$$\rho_r^s(t_r, t^s) = \rho_r^s(t_r, t_r - \Delta t) \quad (2.6)$$

Taking various error source terms into account, the pseudo-range model can be completed with the geometric distance as well as several correction terms,

$$\rho_r^s(t_r, t^s) = \rho_r^s(t_r, t_r - \Delta t) - (\Delta t_r - \Delta t^s)C + \Delta_{ion} + \Delta_{tro} + \Delta_{mul} + \Delta_{tide} + \Delta_{rel} + \epsilon \quad (2.7)$$

where the clock biases between the receiver clock offset and the satellite clock offset are scaled by the vacuum velocity of light  $C$ . Two transmission medium induced errors due to the ionospheric effect and the tropospheric effect are denoted by  $\Delta_{ion}$  and  $\Delta_{tro}$ , respectively. The multi-path effect,  $\Delta_{mul}$ , would result in two or more propagation paths when the GPS signals approach the receiver. The Earth tidal effect, denoted as  $\Delta_{tide}$ , is responsible for displacement of the receiver due to astronomical gravity. Relativistic effects of the time dilation, gravitational frequency shift, and eccentricity effects are all included in the relativistic effect factor,  $\Delta_{rel}$ . Uncertainty in GPS receivers due to remaining systematic errors are denoted as  $\epsilon$ .

## 2.5 Carrier Phase Measurements

The carrier phase measurements utilize the phase difference between the received satellite signal and the replica carrier generated by the local receiver at the reception time. The phase term conventionally has the unit of cycle which refers to a full carrier wave. By shifting the receiver-generated phase to align with the received phase, the range would simply be the sum of the total number of full carrier cycles plus fractional cycle at the receiver [3], multiplied by the carrier wavelength. Recall from Table 2.2 that the wavelength of the L1 carrier is 19 cm which is much shorter than the wavelength of the PRN code, for example 29.31 m for the P-code. This implies that the carrier phase approach can achieve far more accurate measurements than the pseudo-range because of the short wavelength (implies high resolution) of the carrier itself.

The total number of full carrier cycles between the receiver and the satellite cannot be determined because of the initial carrier phase ambiguity. Therefore, measuring the carrier phase is equal to count the full carrier waves and fractional phase received [9].

The mathematical derivation found in Equation (2.8) - (2.15) follows from [11]. Let us first try to derive a carrier phase expression from the simplest case of a vacuum medium and an error-free scenario. The measured phase can be presented by

$$\Phi_r^s = \Phi_r(t_r) - \Phi^s(t_r) + N_r^s \quad (2.8)$$

where r and s are the receiver and the satellite, respectively.  $t_r$  refers to the GPS signal reception time.  $\Phi_r$  represents the phase term of the receiver, while  $\Phi^s$  represents the phase term of the satellite at the reception time.  $N_r^s$  refers to the phase ambiguity.

According to the fact that the received signal phase term of the satellite exactly equals to the emitted signal phase term at the satellite transmitter [12]:

$$\Phi^s(t_r) = \Phi_e^s(t_r - \Delta t) \quad (2.9)$$

where  $\Phi_e^s$  denotes the phase term at satellite emission time, and  $\Delta t$  is the GPS signal propagation time. We may rewrite Equation (2.8) as

$$\Phi_r^s(t_r) = \Phi_r(t_r) - \Phi_e^s(t_r - \Delta t) + N_r^s \quad (2.10)$$

Note that a relation among time, phase and frequency obeys

$$t = \frac{\Phi}{f} \quad (2.11)$$

The geometric distance, denoted as  $\rho_r^s(t_r, t_e^s)$ , between the satellite at the emission time  $t_e^s$  and the GPS antenna at the reception time  $t_r$  is calculated by

$$\rho_r^s(t_r, t_e^s) = \Delta t \cdot C \quad (2.12)$$

Combine Equation (2.10), (2.11) and (2.12) yields

$$\Phi_r^s(t_r) = \frac{\rho_r^s(t_r, t_e^s)f}{C} + N_r^s \quad (2.13)$$

Now taking various error effects into account, the measured carrier phase model can be completed with the geometric range between the satellite and the receiver plus or minus several rectification terms

$$\begin{aligned} \Phi_r^s(t_r) = & \frac{\rho_r^s(t_r, t_e^s)}{\lambda} - f(\Delta t_r - \Delta t_s) + N_r^s - \frac{\Delta_{ion}}{\lambda} + \\ & + \frac{\Delta_{tro}}{\lambda} + \frac{\Delta_{tide}}{\lambda} + \frac{\Delta_{mul}}{\lambda} + \frac{\Delta_{rel}}{\lambda} + \frac{\epsilon}{\lambda} \end{aligned} \quad (2.14)$$

Or

$$\begin{aligned} \lambda\Phi_r^s(t_r) = & \rho_r^s(t_r, t_e^s) - (\Delta t_r - \Delta t_s)C + \lambda N_r^s - \Delta_{ion} + \\ & + \Delta_{tro} + \Delta_{tide} + \Delta_{mul} + \Delta_{rel} + \epsilon \end{aligned} \quad (2.15)$$

where the difference between the receiver clock offset and the satellite clock offset is scaled by the vacuum velocity of light  $C$ . The ambiguity of the carrier phase (integer number) is scaled by the carrier wavelength. The atmospheric effects contain both the ionospheric and the tropospheric contribution of  $\Delta_{ion}$  and  $\Delta_{tro}$ , separately. The Earth tide effects are represented by  $\Delta_{tide}$ . The multi-path effect, the relativistic effects, and remaining errors are referred by  $\Delta_{mul}$ ,  $\Delta_{rel}$ ,  $\epsilon$ , respectively. Equation (2.15) is easy to use in real application because all terms have units of length in meter.

## 2.6 Doppler Measurements

The Doppler effect describes frequency shift of a signal due to relative motion between the transmitter and the receiver. The Doppler shift count, denoted as  $D$ , due to relative motion between the GPS satellites and a GPS receiver can be modeled in Equation (2.16) [11]. It should be noted that the Doppler shift effect will not be significant, because of the high altitude of the GPS satellites [3].

$$D = \frac{d\rho_r^s(t_r, t_e^s)}{\lambda t} - f \frac{d\beta}{dt} + \Delta_f + \epsilon \quad (2.16)$$

where  $\rho_r^s(t_r, t_e^s)$  is the geometric distance;  $\beta$  is the clock error term of  $(\Delta t_r - \Delta t^s)$ ;  $\Delta_f$  is the frequency correction term due to the relativistic effects and  $\epsilon$  is the remaining error.

## 2.7 Errors and Biases

GPS measurements are affected by several types of errors and biases. Depending on the origination of error sources, it is classified into four main aspects, namely the satellites, the receiver, the signal propagation medium and the satellite geometry effects [3]. Figure 2.7 shows the main types of errors and biases.

Errors and biases stemming from the satellites mainly include ephemeris errors and satellite clock errors. The ephemeris errors represent a discrepancy between the expected and actual orbital position of a GPS satellite. These errors can be caused by perturbation of the gravitational field or uncertainties of the solar radiation pressure [9]. Over a period of time, the space-borne atomic oscillator will drift out of stability and thus induce the satellite clock error. Thereby, performance of the satellite clock must be continually monitored by the ground control/monitor stations, and clock offsets are compensated within the navigation message.

Errors and biases coming from the receivers mainly include receiver clock errors, multi-path errors, system noise or delays. The GPS receivers use far less accurate crystal quartz clock for cost purpose, therefore, the receiver clock errors are much larger than the satellite clock errors. It can be re-

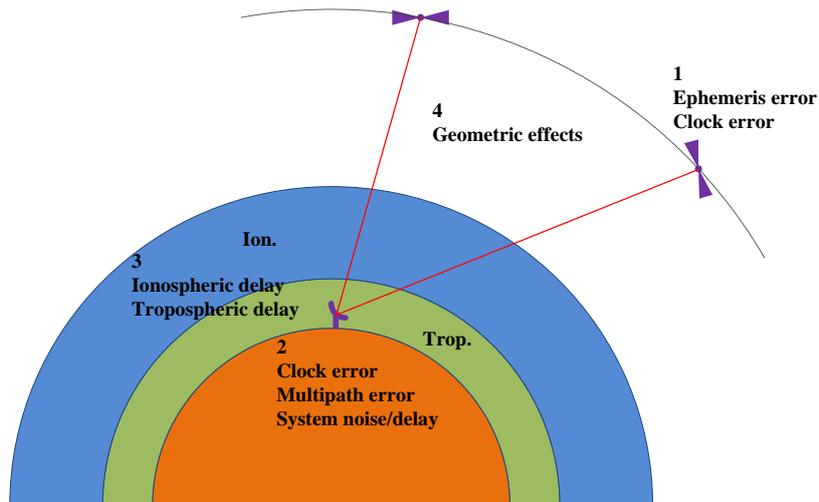


Figure 2.7: Major components of GPS errors and biases, redrawn from [3].

moved from differencing between the satellites [3], or be reduced by external accurate timing offered by local reference stations. The multi-path effects result from the possibility that more than one propagation paths could reach the GPS receiver. This error is highly depended on the surrounding environment around the receiver antenna. Except from developing good multi-path models a good countermeasure is to choose an observational site where no major reflecting objects is in the vicinity of the receiver antenna. Last but not the least, the system noise or delays contribution refer to the noise offset originating from a mismatching receiving antenna, the thermal noise from a pre-amplifier, and the delay terms ranging from transmission cable delay to intrinsic software and hardware response delay, and so on.

Errors or biases caused by the propagation medium include the ionospheric delay and the tropospheric delay. The dispersive ionosphere not only bends the GPS signal propagation path but also changes the speed as the GPS signal propagates through various ionospheric layers. The ionospheric delay is proportional to the number of electrons along the GPS signal transmission path. Properties of the ionosphere and its influence on the passing signals will be discussed in detail in later chapters. Unlike the ionosphere, the troposphere is an electrically neutral region. Therefore, the tropospheric delay is frequency-independent, compared to the frequency-dependent ionospheric delay. As the signals propagating through the troposphere, factors such as atmospheric pressure, temperature, and humidity should be taken into account when estimating the tropospheric delay [9].

Satellite geometry describes the geometric position of the GPS satellites from the receiver's point of view. Good satellite geometry can be the case when the GPS satellites are spread out over the horizon seen by the receiver, whereas bad satellite geometry can be the distribution when the visible GPS satellites are close together or even positioned almost in line from the receiver perspective. It is easily understood that satellites in a poor geometry cannot provide as much information as satellites that are in a good geometry. To illustrate this, two satellites are separated far apart in Figure 2.8 (a), thus the blue colored uncertainty area is small resulting in a good satellite geometry. In Figure 2.8 (b), however, two satellites that are closely located leads to a large uncertainty area, which corresponds to a bad satellite geometry. Dilution of precision (DOP) values are used to indicate the quality of the satellite geometry. Lower DOP value indicates a higher satellite geometry quality, and vice versa [13].

Other errors and biases mainly include the site displacement effects, the relativistic effect, hardware delays, antenna adjustment induced error, and rounding error. The site displacement effects represent periodic movements of the GPS station, which result from the solid Earth tides, ocean loading,

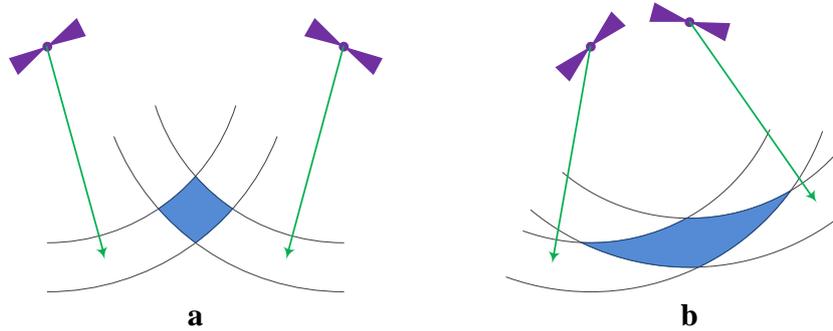


Figure 2.8: Good and bad satellite geometry illustration.

and polar motion [11]. The hardware delays can occur both at the satellite and the receiver. The satellite hardware delay refers to a time delay between signal generation by the signal generator and signal transmission by the antenna. The receiver hardware delay refers to the time difference between signal reception at the receiver antenna and the signal processing devices, for example digital correlator. Hardware delays are frequency-dependent. The difference between the delays on the L1 and L2 is called inter-frequency bias. One way to calibrate the hardware delays is to use the GPS observations collected at reference stations with precisely known coordinates [3].

A convenient way of examining the GPS measuring accuracy can resort to the User Equivalent Range Error (UERE). It refers to the error of a component in the distance from the receiver to a satellite. These UERE errors are given as  $\pm$  errors hence implying that they are zero mean errors. Major error sources and their effects in terms of the UERE are listed in Table 2.4 [6].

Table 2.4: Sources of User Equivalent Range Errors [6]

Source	Effect (m)
Signal arrival C/A	$\pm 3$
Signal arrival P(Y)	$\pm 0.3$
Ionospheric effects	$\pm 5$
Tropospheric effects	$\pm 0.5$
Ephemeris errors	$\pm 2.5$
Satellite clock errors	$\pm 2$
Multipath effect	$\pm 1$

# Chapter 3

## Ionosphere Background

Ionosphere is one part of the Earth's upper atmosphere that extends itself from about 60 km until more than 1000 km in altitude. The ionized medium is mainly produced by X-rays and Ultraviolet (UV) rays from the solar radiation. Cosmic rays, which originate from sources throughout the Milky Way and the universe, provide an incidence of charged particles for ionization to some limit extent. Neutral gas molecules within ionosphere bombarded by high energetic particles will further be stripped off one or more electrons. These 'free' electrons are negatively charged, while the molecules that lose the electrons become positively charged ions. Overall, the ionosphere is considered to be neutral because of equal amount of electrons and ions.

Propagation speed of the GNSS electromagnetic signals in the ionosphere depends on its electron density, which is typically driven by two main processes. During the day, generated electrons and ions make the plasma density raises. During the night, cosmic rays still ionize the ionosphere, although not as strong as the Sun does during the day. At night, a reverse process to ionization known as plasma recombination takes place such that the 'free' electrons are captured by positive ions to form new neutral particles. Therefore, it leads to a reduction in the electron density.

The ionosphere can bend, reflect and attenuate radio waves, thus it influences radio communication applications. Specifically, the ionosphere bounces radio waves back when radio frequencies are below about 30 MHz, enabling a long-distance communication on the Earth. At higher frequencies, like those used by the GPS satellites, however, radio waves will pass through the ionosphere, enabling radio communication between satellites and the Earth. Although the ionosphere itself occupies less than 0.1% of the total mass of the atmosphere [14], it is extremely important for mankind to study its properties.

### 3.1 Height Characteristics

The ionosphere is classified into D, E and F layers based on what wavelength of solar radiation is absorbed in that region most frequently.

D layer is the lowest layer in altitude. It does not have a fixed starting and stopping point, but includes the ionization that happens below about 90 km. Hard X-rays (wavelength smaller than 1 nm) can be absorbed at this layer due to ionization of molecular  $N_2$  and  $O_2$  [15]. Recombination is important for this layer because relative high gas density implies that the gas molecules and ions are closer together. Therefore, net ionization is low. Generally this layer has very little influence on GNSS signals.

E layer stretches up to about 150 km above the ground. Soft X-rays (wavelength between 1 and 10 nm) and far UV are absorbed here with ionization of molecular  $O_2$  [15]. This layer is characterized by irregularities in the plasma density at high latitude regions. E layer becomes weakened at night.

Topmost F layer begins above 150 km and can extends to about 1000 km. It is the region where extreme UV (wavelength between 10 and 100 nm) absorption is performed [15]. As altitude raises, light ions such as  $H_2^+$  and  $He_2^{2+}$  become dominant. Being the densest plasma layer, it effects the GNSS signals the most significant [16].

Figure 3.1 shows a general layer distribution compared in day and night. During the night, only the F and E layers remain. In detail, the D layer can disappear due to heavy recombination. Ionization in the E layer is low because primary source of the solar radiation no longer presents. In fact, the vertical structure of the E layer is primarily determined by the competitive effects of ionization and recombination. The F layer is the only layer of significant ionized and can be considered as one layer, known as the  $F_2$ , at night. During the day, the D layer is created and the E layer becomes enhanced due to much more strongly ionization. Meanwhile, the heavily ionized F layer could further develop an additional, relative weaker ionized region known as

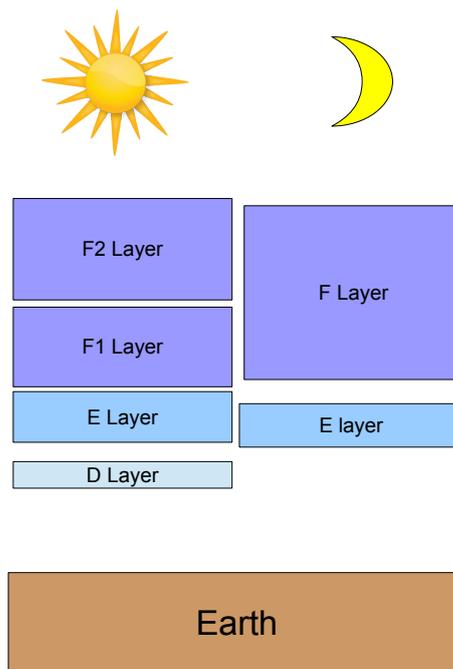


Figure 3.1: Illustration of the ionospheric layers in daytime and nighttime.

the  $F_1$  layer. The  $F_2$  layer remains both day and night can be used for high frequency radio communications [17].

## 3.2 Ionosphere Variation Types

Ionospheric density is highly depended on the amount of emitted radiation from the Sun. Therefore, the Sun and its activity or relative movement of the Earth orbiting the Sun will mainly result in ionospheric variation. There exists four main classes of regular variations, namely daily, seasonal, 11-year and 27-day variation. Here ‘regular’ implies that the variation could be periodically predicted easily. Besides regular variations, sudden-burst solar activities and character of the geomagnetic field distribution can also lead to special ionospheric phenomena, some of which are difficult to predict. The ionospheric variation in the tropical area are tightly related to ionospheric storms, travelling ionospheric disturbance, sudden ionospheric disturbance, ionospheric scintillation, geomagnetic field distribution and South Atlantic Anomaly induced influence.

### 3.2.1 General Variations of Global Scale

For the daily effect, local ionospheric variation are directly related to rotation of the Earth around its axis. In general, diurnal variation can be described as an increase of ionospheric particles right after the sunrise with a maximization at about 2 hours after the local noon, followed by slow decline in the rest of the day until the dawn.

Seasonal variations are directly related to the Earth orbiting around the Sun. Local winter hemisphere faces away from the Sun results in less received solar radiation, on the contrary, local summer hemisphere results in more received solar radiation. Seasonal variations of the D, E, and F1 layers are mainly determined by the highest angle of the Sun, thus the ionization density of these layers are highest in the summer and lowest in the winter [18]. The F2 layer, however, shows an opposite pattern. Ionization density in F2 layer reaches greatest in the winter and lowest in the summer. The seasonal variation also depends on the temperature, the colder the temperature the less effective recombination rate for molecular  $N_2$  and  $O_2$ . As a result, ionization density presents greatest in the winter [19].

The 11-year effect relates with the solar cycle. Solar activities periodic change about every 11 years with regard to the number of sunspots, flares, coronal mass ejections and the levels of solar radiation, et cetera. When solar activities strengthened, flux of short-wavelength solar radiation, ranging

from the X-rays to UV rays, booms. Besides high energetic radiation flux, associated magnetic variation interacting with the geomagnetic field can further imposes severe magnetic fields' fluctuation. Overall consequences of the enhanced solar activities are responsible for an increase of ionization density for all the ionospheric layers.

The 27-day alteration is due to rotation of the Sun around its axis. As the Sun rotates, the sunspots are visible from the Earth with a 27-day intervals. This solar induced period results in variation of the ionospheric layers with a day-to-day bias, making precise prediction of the ionization density difficult. Long time observational record shows that the fluctuation in the F2 layer are the greatest among all the layers.

### 3.2.2 Specific Variations Around Tropical Areas

The ionospheric storms are associated with solar activities such as flares and coronal mass ejections. Released from the Sun, streams of energetic particles (mainly electrons and protons) and magnetospheric energy incident upon the Earth will cause disturbances of the 'quiet-time' ionosphere [20]. The ionospheric storms tend to generate large turbulent on the ionospheric density distribution, expressed by the TEC (see below chapter), and the ionospheric current system [21]. An intense ionospheric storm can last for one day or so and would result in serious consequences such as disrupting satellite communications. Therefore, it is important to monitor or even early warning its behavior.

The traveling ionospheric disturbances (TIDs) are common ionospheric phenomena that gravity waves propagate horizontally in the neutral atmosphere [22]. It can cause perturbations on the TEC measurements. Research have shown that the solar eclipses, geomagnetic storms and solar terminator are responsible for the observed TIDs. Based on physic characteristic such as wavelength, velocity and period, TIDs are classified into medium scale TIDs (MSTIDs) and large scale TIDs (LSTIDs). The MSTIDs are thought to originate from sources of thunderstorm activity, seismic events and shears in the jet streams [23]. The LSTIDs usually generate from geomagnetic disturbances. During violent geomagnetic storms, the normally high latitude originated LSTIDs can propagate into mid-latitudes, the equatorial areas and even into the opposite hemisphere. Table 3.1 lists general properties of the two types of TIDs [24].

The sudden ionospheric disturbances (SIDs) are generated due to enhanced solar flare events. Specifically, the solar flare event produces an abnormal intense burst of hard X-rays and UV radiation, which are not absorbed by the F2, F1 and E layers, but instead causes a sudden increase in the plasma

Table 3.1: MSTIDs and LSTIDs

TID	Wavelength (km)	Horizontal phase velocity (m/s)	Period (minutes)
Medium Scaled	100 - 1000	100 - 300	< 60
Large Scaled	> 1000	100 - 300	60 - 180

density of the D layer. Such ionospheric anomalies can occur without warning and may prevail for any length of time, from a few minutes to several hours. When the SIDs occurred all ground stations facing the Sun are more or less affected. The SIDs often interferes with telecommunication systems operating in the upper medium frequency (MF) and lower high frequency (HF) ranges [25].

The ionospheric scintillation describes a manifestation of small-scale inhomogeneities in the density and movement of the free electrons in the ionosphere. It affects the GPS signals via diffraction and refraction. Therefore, the received signals present random amplitude and phase fluctuation. In severe scintillation case, navigation errors or even navigation failure may occur when the receiver's phase lock loops are unable to track loops of the GPS signals. The ionospheric scintillation generally presents its most frequent and most intense in the sub-equatorial regions, located on average about  $15^\circ$  surrounding the geomagnetic equator. Plasma bubbles are thought to be a main induction for the ionospheric scintillation in equatorial and high-latitude areas. The primary disturbance region is typically in the F-layer at altitudes between 250 and 400 km [26].

The equatorial plasma bubbles (EPBs) are ionospheric irregularities of depleted plasma density. Ionospheric turbulence and scintillation are thought to be correlate with the EPBs activities. With longitudinal scale of hundreds of kilometers [27], the EPBs most frequently begin growth in the geomagnetic equator, approximate 1 to 2 hours after the local sunset when the recombination process prevails. An initial perturbation on the bottom side of the F1 layer can rise the low density plasma vertically to the topside ionosphere by the effect of buoyancy. Regarding the movement of the EPBs, on the one hand, they accumulate to peak altitude where the density inside the bubbles are balanced to that of the surroundings, on the other hand, they extend and diffuse along the geomagnetic field lines due to pressure gradients and gravity action. However, larger amplitudes can be reached when the EPBs approach the high electron density region of the Equatorial Ionospheric Anomaly (EIA) crests which are located  $20^\circ$  north and south of the geomagnetic equator. Figure 3.2 shows a partial region of world geomagnetic

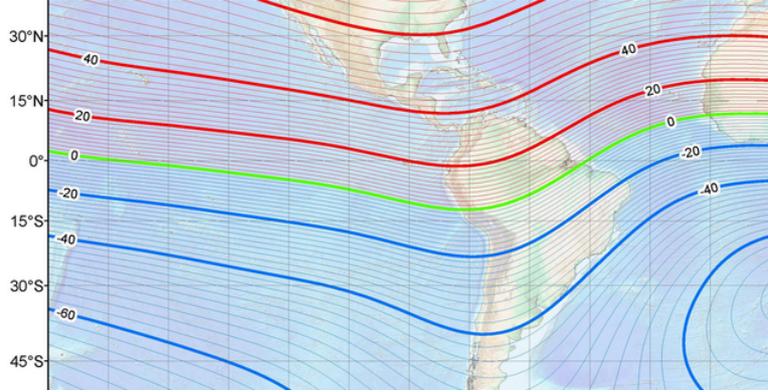


Figure 3.2: World geomagnetic inclination in South American regions [4].

inclination in 2015 where the geomagnetic latitude is clearly different from the geographic latitude.

Research have shown that possible perturbations include the non-linear evolution of the Generalized Rayleigh-Taylor (GRT) instability, growth of strong gravity waves, penetration electric fields, the TIDs and other instabilities. In the pre-midnight hours, the dominant perturbation is widely believed to be the gravity waves [27–29]. The GRT instability creates intense polarization electric fields, causing the depleted plasma inside to  $E \times B$  drift upward to the topside ionosphere [30]. As the bubbles diffuse along the geomagnetic field lines to low latitudes, they drift eastward. Velocity of the eastward EPBs slowly decrease before midnight in the local time and become dilution and vanishing thereafter.

The South Atlantic Anomaly (SAA) results from the fact that the Earth’s magnetic field, the geographic center and poles are not perfectly aligned [31]. The Earth is surrounded by a pair of concentric donut-shaped Van Allen radiation belts which trap and preserve charged particles from the solar winds. The geomagnetic field at the SAA is the weakest all over the world, seen from Figure 3.3, and the Earth’s inner Van Allen radiation belt at this region comes closest to the Earth surface dropping down to an altitude of about 200 km [32]. The result is that, for a given altitude, higher-than-usual levels of radiation can be found. Research has shown that greater TEC values at local noon in the longitude sectors over and near the SAA have been observed [33]. Developed by European Space Agency (ESA), software of Space Environment Information System (SPENVIS) provides good models for space environment analysis [34], a simulation, illustrated in Figure 3.4, shows that Brazil is mainly covered within the SAA region.

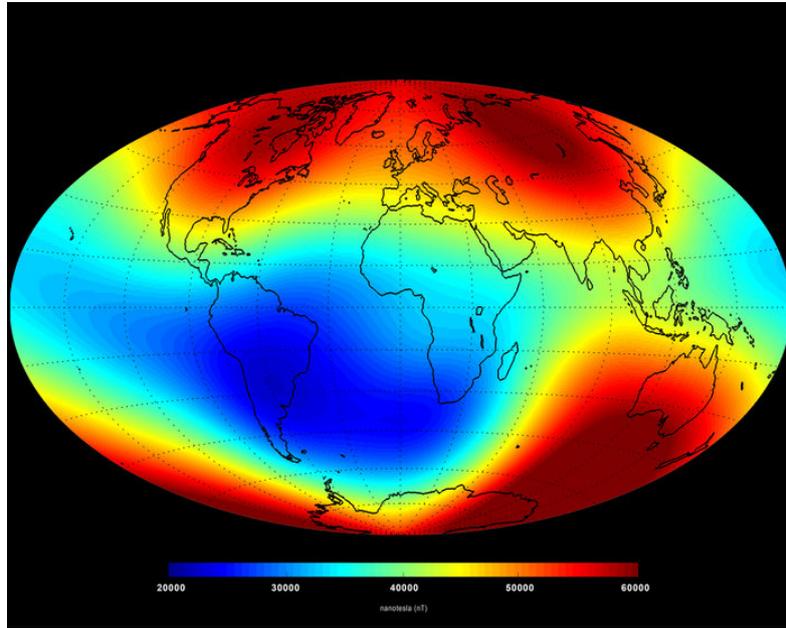


Figure 3.3: Main magnetic field at Earth's surface in June 2014 [5].

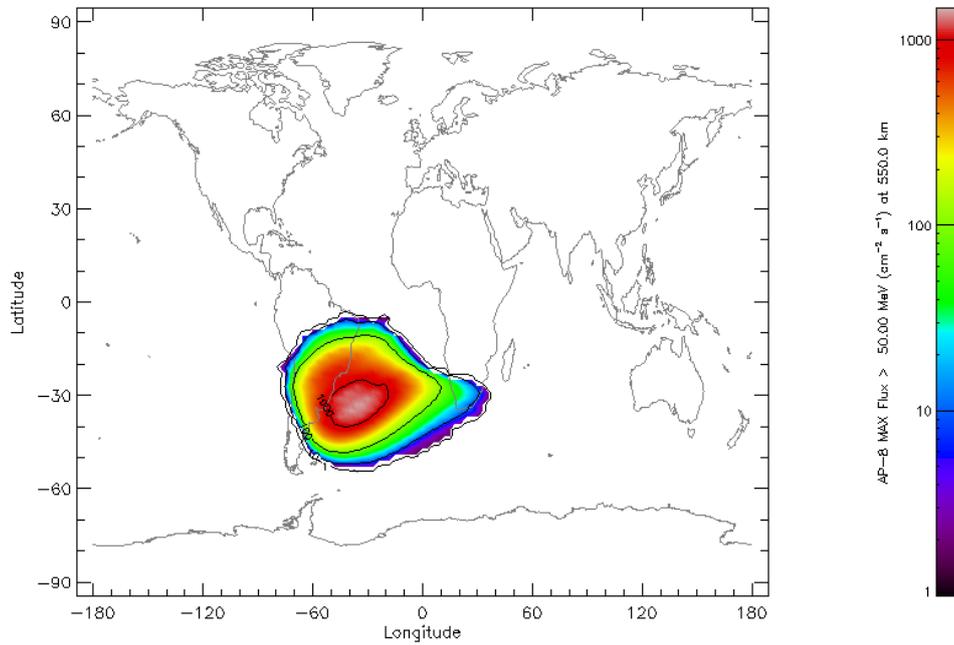


Figure 3.4: SPENVIS simulation of SAA at 550 km. Note that values in the index bar are logarithmic.

# Chapter 4

## Signal Propagation through the Ionosphere

The mathematical expressions derived below follows from [1], [35–38]. Unless otherwise indicated or cited clearly, most of the derivation can be done step by step without too much difficulty.

### 4.1 Some Fundamentals of Wave Propagation

The relation between wavelength  $\lambda$ , the frequency  $f$ , and the propagation velocity  $v$  is

$$v = \lambda \cdot f \quad (4.1)$$

Units for  $\lambda$ ,  $f$  and  $v$  are meters (m), Hertz (Hz, oscillations per second), and meters per second (m/s), respectively.

Relation between angular frequency  $\omega$  and frequency  $f$  is

$$\omega = 2\pi f \quad (4.2)$$

The phase constant or wave number  $k$  can be expressed in terms of wavelength  $\lambda$

$$k = \frac{2\pi}{\lambda} \quad (4.3)$$

The wave propagation velocity  $v$  follows with expressions that

$$v = \lambda \cdot f = \frac{\lambda}{P} = \frac{\omega}{k} \quad (4.4)$$

where  $P$ , reciprocal of the frequency, is the period of a wave.

A periodic wave can be modeled by a sinusoidal function in space and time.

$$y = A \sin 2\pi \left( \frac{t}{P} + \Phi_0 \right) \quad (4.5)$$

where  $y$  is the magnitude at time  $t$ ;  $\Phi_0$  is the initial phase of the wave at  $t$  of 0, and  $A$  is the amplitude of the wave. The phase at time  $t$  is then

$$\Phi = \frac{t}{P} + \Phi_0 \quad (4.6)$$

$2\pi\Phi$  is referred to phase angle  $\phi$ . Inserting Equation (4.2) into Equation (4.5) has

$$y = A \sin(\omega t + \phi_0) \quad (4.7)$$

The relation between time, phase and frequency follows [1]

$$t = \frac{\Phi}{f} \quad (4.8)$$

Equation (4.8) gives a fundamental relation between the phase of a periodic wave and the corresponding time reading at the satellite clock, and can be considered to be the definition equation of a clock.

Propagation velocity,  $c$ , of an electromagnetic (EM) wave in a medium of vacuum is

$$c = \frac{\lambda_{vac}}{P} = \lambda_{vac} \cdot f = \frac{\omega}{k_{vac}} \quad (4.9)$$

The propagation velocity in a non-vacuum medium is characterized by the refractive index  $n$

$$n = \frac{c}{v} = \frac{\lambda_{vac}}{\lambda} = \frac{k}{k_{vac}} \quad (4.10)$$

The propagation velocity of the EM waves in a dispersive medium depends on the frequency. From Equation (4.10), it is easy to conclude that the index of refraction (or refractivity) depends on the frequency or the wavelength in a dispersive medium.

The velocity dispersion is defined by

$$\frac{dv}{d\lambda} \quad (4.11)$$

In a dispersive medium, different propagation velocities for sinusoidal waves and groups of waves can be observed. Thus, concepts of phase velocity,  $v_{ph}$ , and group velocity,  $v_{gr}$ , are introduced.

The phase velocity is the propagation velocity for a single EM wave with uniform wavelength and it is given by,

$$v_{ph} = \lambda \cdot f \quad (4.12)$$

For GPS, the carrier waves L1 and L2 are propagating with this velocity.

The group velocity is the propagation velocity of a group of EM waves, generated by a superposition of different waves with different frequencies. It is given by

$$v_{gr} = -\frac{df}{d\lambda} \lambda^2 \quad (4.13)$$

According to [35]. This velocity has to be considered for GPS code measurements.

The relation between phase velocity and group velocity is described by Rayleigh equation as

$$v_{gr} = v_{ph} - \lambda \frac{dv_{ph}}{d\lambda} \quad (4.14)$$

Expressed in terms of the refraction index,

$$n_{gr} = n_{ph} + f \frac{dn}{df} \quad (4.15)$$

Derivation of Equation (4.14) and (4.15) can be found in [36].

## 4.2 Ionospheric Refraction

Propagation of radio signals within the ionosphere is primarily affected by ‘free’ electrons and ions. Since the ionosphere is a dispersive medium with respect to the GPS signal, refraction index of the phase can be expanded as a power series [1]

$$n_{ph} = 1 + \frac{c_2}{f^2} + \frac{c_3}{f^3} + \frac{c_4}{f^4} + \dots \quad (4.16)$$

The coefficient  $c_i$  does not depend on the carrier frequency but on the quantity of electron density  $N_e$  along the propagation range. By dropping off the

series expansion terms above the quadratic order, the phase refractive index is approximated as

$$n_{ph} = 1 + \frac{c_2}{f^2} \quad (4.17)$$

Differentiating above equation on both side gives

$$dn_{ph} = -\frac{2c_2}{f^3}df \quad (4.18)$$

Then, inserting Equation (4.17) and (4.18) into Equation (4.15) produces

$$n_{gr} = 1 - \frac{c_2}{f^2} \quad (4.19)$$

A comparison between Equation (4.17) and Equation (4.19) reveals that only an opposite sign exists for the group and the phase refractivity.

The formula of dispersion relates the refractive index  $n$  in an ionized medium with electron density  $N_e$  [37]

$$n = 1 - n_e \frac{C^2 e^2}{\pi f^2 m_e} \quad (4.20)$$

where  $e$  is the elementary charge,  $m_e$  is the mass of an electron.

An explicit derivation of index of refraction can be found in [38] as

$$n = 1 - \frac{C \cdot N_e}{f^2} \quad (4.21)$$

with  $C = 40.3$ . Equation (4.21) implies that the index of refraction is inverse proportional to the frequency squared. Reviewed from Equation (4.10), time delay of a propagation signal relates with the index of refraction. Therefore, Equation (4.21) further indicates that signals with higher frequencies are less influenced by the ionosphere.

The quadratic term coefficient of  $c_2$  can be estimated from Equation (4.21) to be

$$c_2 = -40.3N_e \quad [Hz^2] \quad (4.22)$$

Therefore, phase refractive index holds as

$$n_{ph} = 1 - \frac{40.3N_e}{f^2} \quad (4.23)$$

Equation (4.23) indicates that a rough correction can be made for the delay in signal propagation if the an a priori electron density is known. Similarly, group refractive index holds as

$$n_{gr} = 1 + \frac{40.3N_e}{f^2} \quad (4.24)$$

Combining the relationship of  $n_{gr} > n_{ph}$  with non-negative electron density  $N_e$  gives  $v_{gr} < v_{ph}$ . As a result, the group and the phase velocity are delayed and advanced, respectively. In other words, GPS code measurements are delayed and the carrier waves' phases are advanced.

To quantitatively describe the ionosphere of the Earth, the TEC is used. It is defined as the total number of electrons integrated between the receiver  $r$  and the satellite  $s$ , along a column with a cross section of one meter squared [36]. Mathematical expression is

$$TEC = \int_r^s N_e(l)dl \quad (4.25)$$

where  $N_e(l)$  denotes the varying electron density along the integration path. Usually, the TEC is measured in TEC units (TECU) where

$$1 \text{ TECU} = 10^{16} \text{ electron per } m^2 \quad (4.26)$$

Usually, the vertical total electron content (VTEC) is used instead of the slant TEC (STEC) for comparison among sets of TEC data. The relation between the VTEC and STEC are normally expressed using a mapping function.

A common used model is known as the single layer model, see Figure 4.1. In this model it is assumed that all free electrons are compressed in a thin spherical shell at a given height  $h_{ion}$  (usually in a range between 300 and 400 km). The point of intersection between the signal path and the layer is referred to as the ionospheric pierce point (IPP). Zenith angle,  $\alpha$ , is the angle spanned cross the signal path and a line extended from the center of the

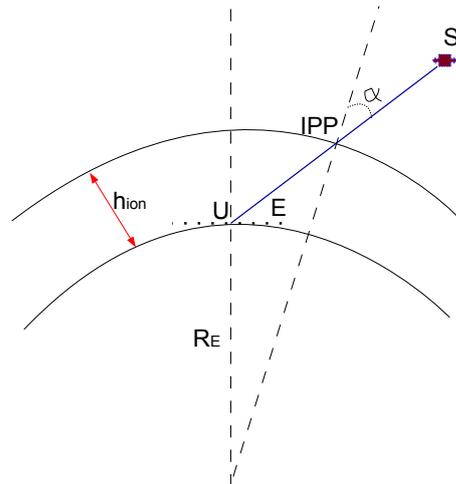


Figure 4.1: Ionospheric single layer model illustration, redrawn from [1].

Earth to the IPP. The relation between the VTEC, STEC, the elevation angle  $E$ , the given height  $h_{ion}$  and the radius of the Earth  $R_E$  can be found in Equation (4.27) and (4.28) [36]

$$VTEC = STEC \cdot \cos\alpha \quad (4.27)$$

where the zenith angle

$$\alpha = \arcsin\left(\frac{R_E}{R_E + h_{ion}} \cos E\right) \quad (4.28)$$

The signal delay due to the ionosphere, denoted as  $d_{ion}$  [m], is given by the difference between the actual signal path and the geometrical distance between the satellite and the receiver

$$d_{ion} = \int_r^s n_p ds_0 - \int_r^s ds_0 = \int_r^s (n_p - 1) ds_0 \quad (4.29)$$

Note that the true geometric range along a straight path between the satellite and the receiver can be obtained from the subtrahend item.

The phase refractive index  $n_{ph}$  from Equations (4.23) yields the phase delay as

$$d_{ion,ph} = \int_r^s \left(1 - \frac{40.3N_e}{f^2}\right) ds - \int_r^s ds_0 \quad (4.30)$$

The group refractive index  $n_{gr}$  from Equation (4.24) yields the group delay as

$$d_{ion,gr} = \int_r^s \left(1 + \frac{40.3N_e}{f^2}\right) ds - \int_r^s ds_0 \quad (4.31)$$

Both the phase delay and the group delay equations can be further simplified by integrating the first term along the geometric straight path  $ds_0$ . Review the TEC definition and substitute it into the phase and group delay,

$$d_{ion,ph} = -\frac{40.3 \cdot TEC}{f^2} \quad \text{and} \quad d_{ion,gr} = \frac{40.3 \cdot TEC}{f^2} \quad (4.32)$$

For LOS situation as shown in Figure 4.1, the zenith angle must be taken into account because the propagation path range in ionosphere varies with an altering zenith angle. By introducing the zenith angle term, the phase delay and the group delay can be modified to

$$d_{ion,ph} = -\frac{1}{\cos\alpha} \frac{40.3 \cdot VTEC}{f^2} \quad \text{and} \quad d_{ion,gr} = \frac{1}{\cos\alpha} \frac{40.3 \cdot VTEC}{f^2} \quad (4.33)$$

Only a difference in sign with respect to the above two expressions. Extracting the common components

$$d_{ion} = \frac{1}{\cos\alpha} \frac{40.3}{f^2} VTEC \quad (4.34)$$

Such notation helps to omit the subscripts of ‘ph’ or ‘gr’ by taking the correct models. It implies that the ionospheric influence for the code pseudo-range measurement is modeled by  $+d_{ion}$  and for the phase by  $-d_{ion}$ .

## Chapter 5

# Methodology of Obtaining the Ionospheric Variability

Ionospheric phenomena like plasma bubbles would result in scintillation which is thought to be primarily responsible for GPS signal distortion of phase and amplitude. Other direct consequences like loss of lock on GPS receivers and available GPS satellite number decrement can further limit the GPS tracking performance and degrade the navigational accuracy. Therefore, to analyze the temporal and spatial distribution of the plasma bubbles in Brazil is a very important issue.

The ionospheric effects can be studied by the dual-frequency ionosphere measuring model. The difference between the L1 and L2 carrier frequency of the GPS signal forms a so called L4-combination which could remove all frequency independent components. Recall the phase observable model discussed in Chapter 2.5. A simple differential combination of both phase pseudo-ranges can be formed as [11]

$$\lambda_{l1}\Phi_{l1} - \lambda_{l2}\Phi_{l2} = \rho_p(f_{l1}) - \rho_p(f_{l2}) + \lambda_{l1}N_1 - \lambda_{l2}N_2 + \lambda_{l1}\mu_{l1} - \lambda_{l2}\mu_{l2} \quad (5.1)$$

where  $\Phi_{l1}$  and  $\Phi_{l2}$  are the  $L_1$  and  $L_2$  phase pseudo-ranges (in units of cycles),  $\rho_p(f_{l1})$  and  $\rho_p(f_{l2})$  are ionospheric effects on the phase for the  $L_1$  and  $L_2$  carriers,  $N_1$  and  $N_2$  refer to the phase ambiguities of the  $L_1$  and  $L_2$  carriers,  $\mu_{l1}$  and  $\mu_{l2}$  include signal multi-path and receiver measurement noise for both carriers. The satellite L1-L2 biases are in common considered to be constant in a few hours. Variation of the receiver inter frequency biases is considered to be relatively small at constant receiver temperatures [39]. Thus, both the satellite L1-L2 biases and the receiver inter frequency biases contribution can be omitted here. Offset term of  $\lambda_{l1}N_1 - \lambda_{l2}N_2$  due to phase ambiguities remains constant, as long as no cycle slips occur which indicate a continuous of phase measurement. Consequently, an expression for the

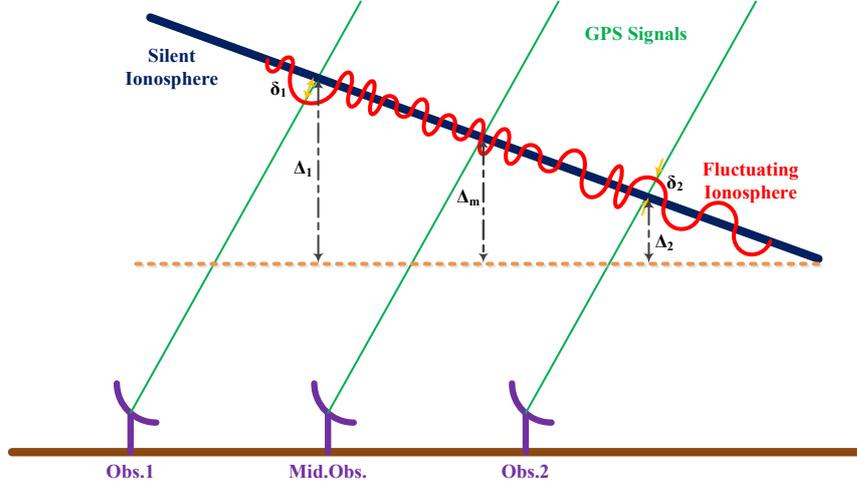


Figure 5.1: An illustration of linear interpolation.

slant ionospheric delay variations containing only error sources from signal multi-path and measurement noise is achieved.

A simple exaggerated illustration of how interpolated value and measured delay value could be applied to reflect the local ionospheric variation is shown in Figure 5.1, where only direction along the longitude is considered. Three observatories, namely observatory 1, middle observatory, and observatory 2, are equally spaced. As altitude of the GPS satellites are far more greater than the distance between adjacent observatories, green colored GPS signals, emitted from the same satellite, can assumed to be parallel received by the three sites. Here, only the ionospheric delays are considered. Delays caused by all frequency independent effects such as geometrical satellite-to-receiver distance and signal delay in the troposphere can be removed by the L4-combination. While, delays caused by other effects such as the inter frequency biases, phase ambiguity, signal multi-path and so on are omitted here to simplify the discussion.

First, let us consider an ideal case where the ionosphere remains no fluctuation, corresponding to the blue line in the Figure 5.1. Only features of long wavelengths are present in this case. In this case, the middle site sees the mean of the ionospheric delays of sites 1 and 2. In other words, the difference between measured ionospheric delay at the middle observatory (denoted as  $\Delta_m$ ) and expected ionospheric delay interpolated from other two observatories (denoted as  $\Delta_1$  and  $\Delta_2$ ) is zero.

Next, let us consider the case where the ionosphere no longer keeps stillness. At a certain moment, all the three sites see a fluctuating ionosphere,

corresponding to red curve in the Figure 5.1. At this snapshot moment, assumed that the middle observatory see the slant ionospheric delay the same as that of the none fluctuation case, however, the other two observatories clearly see different ionospheric delays greatly influenced by the undulating ionosphere. Here the notation of  $\delta_1$  and  $\delta_2$  merely imply a decrease and an increase amount of ionosphere density, respectively. Meanwhile, compared to the ionospheric delay under the none-fluctuation case, decrement ionosphere density seen by the observatory 1 leads to a shorter ionospheric delay, whereas increment ionosphere density seen by the observatory 2 leads to a longer ionospheric delay. Therefore, the difference between the measured ionospheric delay at the middle observatory and the expected ionospheric delay interpolated from the other two observatories can not be zero any more. In other words, the difference contains information about the variability of the electron content at the observatory 1 and 2 right at that snapshot moment. If many snapshot moments are involved, variability of the electron content distribution can be further acquired.

According to [40], in real application, the L4-combination is applied for the three surrounding sites at the vertices of the triangular observational network to interpolate an estimated L4 value for the inner site. The interpolation is performed via bilinear interpolation along both latitude and longitude direction. The bilinear interpolation is just an extension of linear interpolation on a two dimensional grid scale. The weighted term of each site can be determined during the bilinear interpolation procedure as long as latitude and longitude components of each sites are known in advance. Further, the difference between the interpolated value and the measured L4 value at the inner site are computed. For comparison purpose, it is reasonable to scale the difference into the zenith direction. This procedure is taken for all observed satellites in order to obtain a distribution of the ionospheric variation. Because the differences can yield both positive and negative outcomes, it makes sense to use the hourly root-mean-squared (rms) values. Finally, the obtained rms values are converted into the TEC units for analysis and comparison.

The ionospheric spatial variability due to short wave term at the local small scale is here denoted as swTEC, in order to differentiate from the long wave induced TEC values estimated by data set of the IGS Ionosphere Working Group. Both the swTEC and TEC have the same units, denoted as TECU. Note that the contributions from the multi-path and measurement noise can further be eliminated using a model presented in [41]. Detailed description about the methodology was referenced from [40] and can be read at Appendix A.

# Chapter 6

## Data Set

Data sets utilized in this project mainly includes observational data from Brazilian ground stations, navigation data from the GPS satellites and global VTEC maps data from the IGS Ionosphere Working Group.

Original observational data set can be downloaded directly from the official website of Brazilian Institute of Geography and Statistics or IBGE (Portuguese: Instituto Brasileiro de Geografia e Estatística), as it provides observational information within Brazilian territory. The selected ground stations are distributed around geographical coordinates of  $50^\circ$  West and  $22.5^\circ$  South. They are divided into two triangular observational networks surrounding one additional station each to measure the L4-combination. Figure 6.1 shows the eight observational sites. Represented by blue dots, all sites are located in the southern part of Brazil where the ground GPS receiver network is relatively dense. Triangular network 1 includes surrounding stations of ILHA, ROSA, OURI with PPTE as the centering station. Triangular network 2 includes surrounding stations of SPBO, SPJA, MGIN with EESC as the centering station. The triangular network 2 located east of the triangular network 1 is relatively smaller in scale. Such spatial difference can be visible in the characteristics of the local swTEC variation, discussed in next chapter.

Basically, observational data recorded each Thursday are chosen for a whole 2013 year study. Occasionally data are missing from one or more sites in the network. Therefore, data collected one day before or after the data-missing-Thursday are selected instead. Detail information about which date's data are chosen for both triangular networks are listed in Appendix B.

The format of the observational data has to be converted for further analysis. The first-hand downloaded observational data have type name of compressed (zipped) folder, i.e., extension of 'zip'. After decompression two type of files can be acquired, one with type of '13N', the other with '13O'. Here

only the files with '130' are preserved for further MATLAB analysis. Next, the '130' files are converted to 'Text' files via a batch program.

The daily satellite navigation messages come from the GPS broadcast ephemeris data (BRDC) files. And they can be found from the website of NASA's Archive of Space Geodesy Data. One advantage of the BRDC file lies in that it is one navigation file offered by worldwide individual observational stations and has already been merged into the non-redundant daily GPS broadcast ephemeris which can be exploited by users rather than the multiple complex individual navigation files.

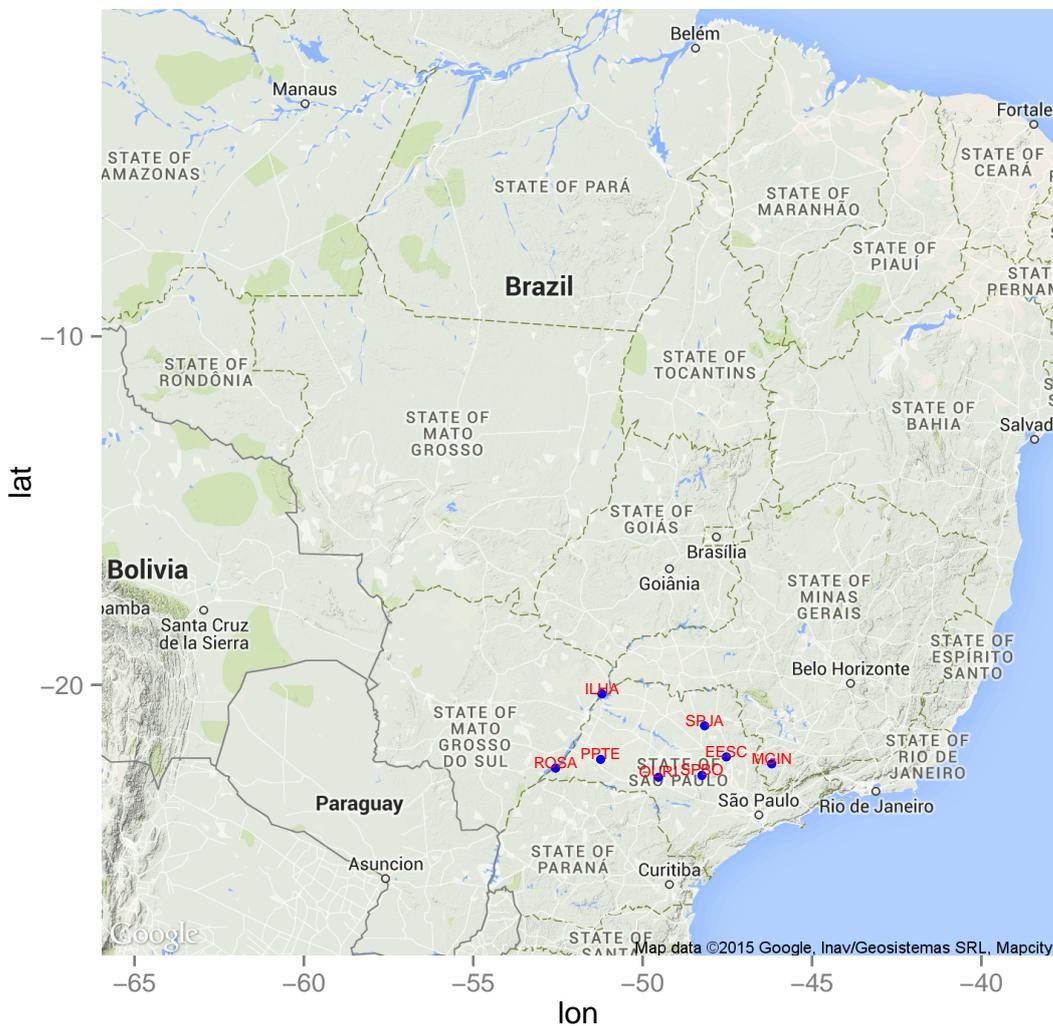


Figure 6.1: Map of the eight selected Brazilian observational sites.

three digits of 'D' and two digits of 'Y' are denoted for days and year, re-

spectively; ‘n’ indicates the satellite navigation information; ‘Z’ refers to the compact format. Here take one example, a file name of ‘brdc0160.10n.Z’ means that this navigation file is dedicated for navigation information of the 16<sup>th</sup> day of the year 2010.

Before imported into MATLAB, file format conversion needs to be performed for those original downloaded navigation files. First step is to perform a decompression operation. As a result, files with type of ‘13N’ are generated. Next step is to convert the ‘13N’ files into ‘Text’ files, which can be realized by a simple batch program.

The Ionosphere Working Group of the IGS has been continuously dedicating to produce reliable ionosphere VTEC maps since 1998. So far, there are four IGS Ionosphere Associate Analysis Centers (IAACs), and each center independently makes rapid and final VTEC maps. In this project, data set of the VTEC maps coming from the Jet Propulsion Laboratory (JPL) were used. These worldwide data set of VTEC, with a temporal-spatial resolution of 2 hours  $\times$  5°  $\times$  2.5° (corresponding to the Universal Time (UT)  $\times$  Longitude  $\times$  Latitude), are stored in IONEX format which refers to an IONospheric map EXchange ASCII format for two and three-dimensional total electron count value maps. In this project, the data set of IONEX format are extracted at 50° West and 22.5° South where the eight observational sites are located nearby. Then, averaged VTEC value of a whole 2013 year range expressed within 24 hours are calculated and further compared with the rms local ionospheric variability data (i.e. mapping to the swTEC Unit here) resulted from the two observational networks, as will be shown in next chapter Figure 7.6.

The IONEX file directly downloaded from website has a format of ‘igs-gDDD0.13i.Z’ where the three digits of ‘D’ refer to specific date information. For example, a file of ‘igs-g3530.13i.Z’ provides global ionospheric information of the 353<sup>th</sup> day of the year 2013. Before obtaining the final ‘txt’ format for subsequent MATLAB analysis, file decompression and type conversion are also required.

# Chapter 7

## Results and Discussion

### 7.1 swTEC Analysis

Ionospheric variability of the two triangular observational networks in terms of swTEC TECU for a whole year range of 2013 is shown in Figure 7.1. These swTEC values are converted from the rms delay differences calculated from the bilinear interpolation and the measured L4-combination value at each inner site. Note that the swTEC amount of the observational network 1 and 2 are colored by blue and red, respectively.

Both observational networks exhibit a similar tendency regarding the ionospheric variation. The swTEC amount of both networks show high values from October to March (corresponding to Spring and Summer of the Southern Hemisphere). Yet relative small swTEC values exist in the rest of the year. This notable seasonal discrepancy of the ionospheric variation mainly relates to the seasonal solar radiation due to the Earth orbiting around the Sun. Intense solar radiation flux in Spring and Summer can result in more active production of the ionospheric plasma irregularities or bubbles. Research has shown that activities of the plasma bubbles take place primarily from September to March in Brazil [42]. Therefore, presence or absence of the plasma bubble activities throughout the year can be used to interpret the ionospheric variation tendency shown in Figure 7.1.

The two observational networks show a relatively clear difference regarding ionospheric variation. The swTEC variations of the observational network 1 is generally higher than the network 2. In other words, being the same selected days, observational network 1 generally presents higher swTEC value than that of the observational network 2. For example, there are several days in Figure 7.1 that the swTEC are over 6 units for the observational network 1, whereas no swTEC value are greater than 6 units for the

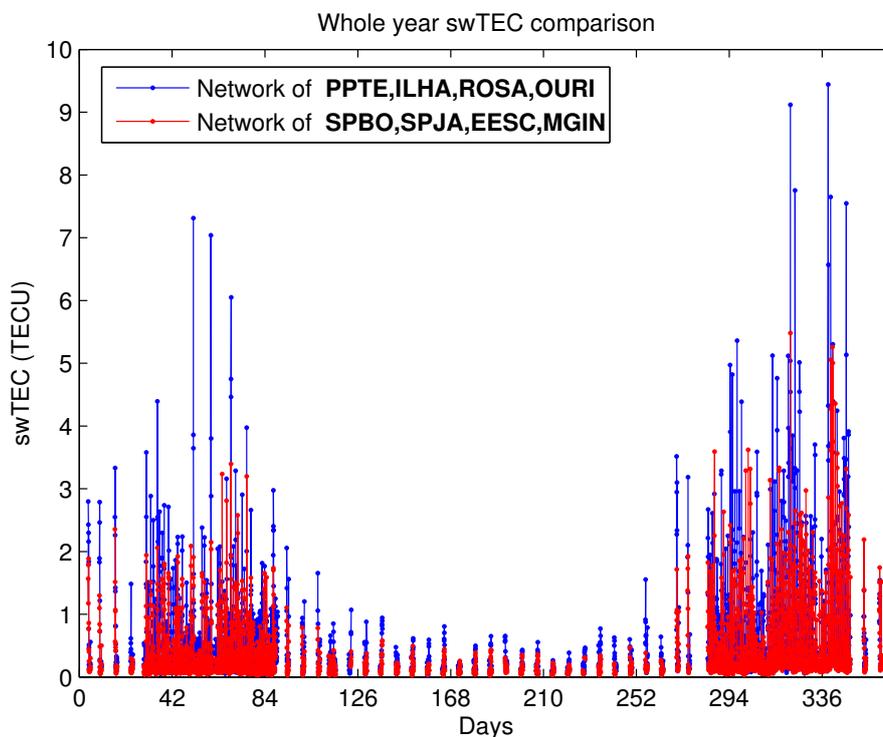


Figure 7.1: Local ionospheric variability comparison for the two observational networks of a whole 2013 year.

observational network 2. This can be more obviously viewed from histogram plot in Figure 7.2. In addition, median value for both observational networks are calculated and compared.

Possible reason for this small-scale inhomogeneities is thought to be mainly relate to the spatial filtering effect. The relative larger spatial scale of the observational network 1 indicates a less high-pass filtering ability in spatial as compared to that of the observational network 2. Thus, more fluctuations of swTEC can be seen for the observational network 1. Whether such local swTEC variations are significantly different at different locations in Brazil requires a consequent study with more observational networks involved.

The MSTIDs have been observed over low altitude range (220-300 km) in the Brazilian territory. Features of the MSTIDs in Brazil have been summarized in research paper [43], from which a seasonal behavior of the MSTIDs shows a peak occurrence rate during the local Brazilian Winter time. In addition, most of the occurrences are observed during solar minimum years whereas no events are optically recorded during solar maximum. Therefore,

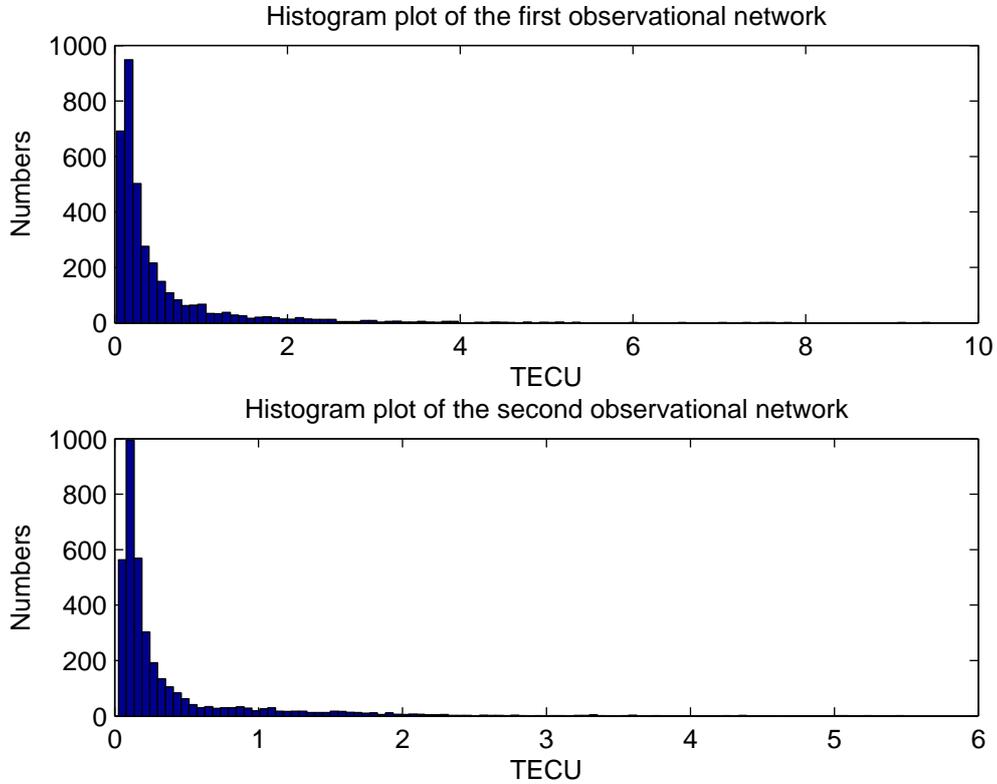


Figure 7.2: Histogram plot showing the occurrence frequency of the swTEC unit. Median value: 1.4392 (top) and 0.9677 (below).

the MSTIDs could possibly be one contribution source for the swTEC variation, especially during the local Winter period in solar minimum conditions. However, the swTEC contribution due to the MSTIDs as well as background noise can only result in small variation units which can be confirmed from small variations of the local Winter part in the Figure 7.1 for both observational networks.

Whether other ionospheric phenomenon such as the ionospheric storms or the SIDs have inconvenient influence on the swTEC still needs further verification. Because both the ionospheric storms and SIDs are closely related to the solar flare activities, correlation study can be made between the swTEC and records of the solar flare variation to further reveal how much contribution are coming from the solar flare activities.

## 7.2 Correlation Study

To study the relationship between two sets of the swTEC data from the two observational networks, correlation plots are generated in Figure 7.3 and Figure 7.4. Figure 7.3 refers to a whole year range correlation where dots do not show a close assemble behavior. Some dots even spread out towards the bottom right corner. Those dots imply low correlation and are assumed to be largely due to ionospheric inhomogeneities caused by the EPBs.

To validate this assumption, data influenced by the EPBs are removed. Reduced data are ranging from April to August because the EPBs active most from September to March in Brazil. Figure 7.4 shows an apparent correlation with relatively few divergent dots.

The close correlation can also be visible in Figure 7.1 between the 94<sup>th</sup> and the 241<sup>th</sup> day, corresponding to months from April to August. Within this period, the amount of the swTEC of both observational sites follow a quite similar variation behavior.

Note that in Figure 7.3 there are a lot of dots gathered together even when the EPBs persist. This can be explained by an illustration of plasma bubble observation scenario shown in Figure 7.5. Notation of Obs. Net. 1 and Obs. Net. 2 are just short for the observational network 1 and 2, respectively. In observation of the plasma bubble 3 case, both observational networks see the same bubble, thus, resulting in relative high correlation. Divergence dots of low correlation favored by either observational network are corresponding to either observation case of plasma bubble 1 or plasma bubble 2, respectively.

In conclusion, the active behavior of the EPBs in regions of Brazilian observational networks can be reflected from the correlation study. In 2013,

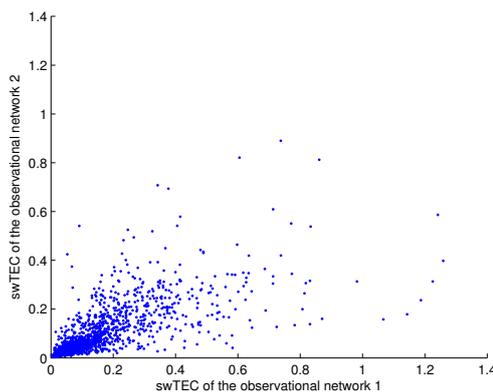


Figure 7.3: Whole year range correlation plot.

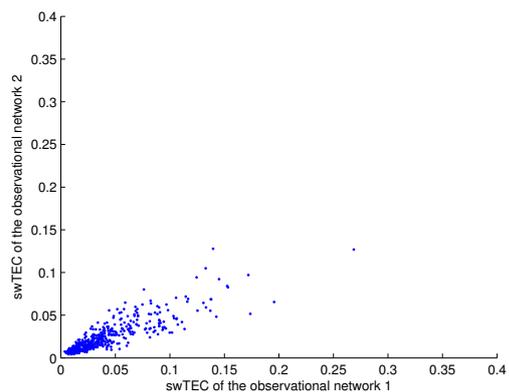


Figure 7.4: Correlation plot from April to August.

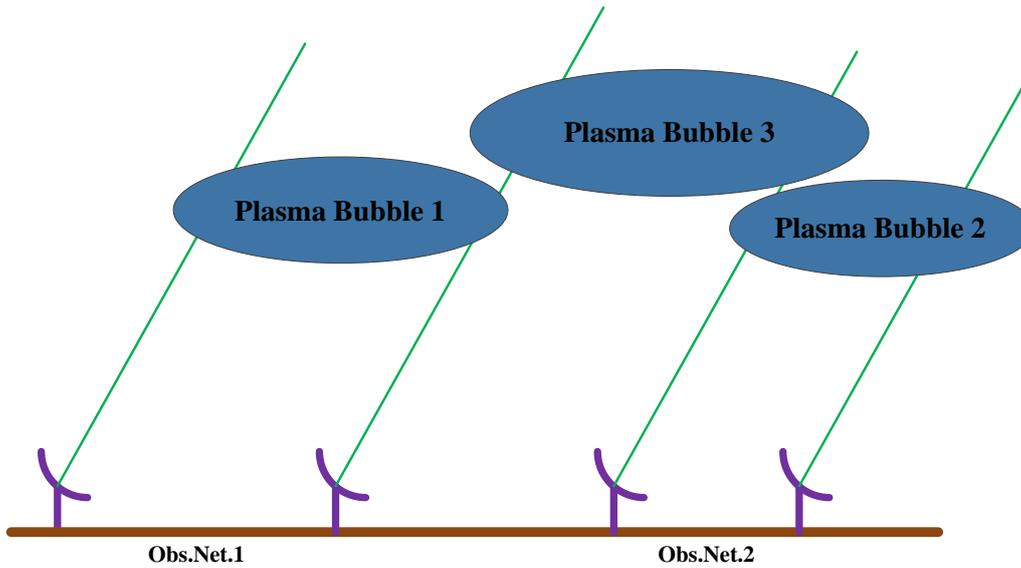


Figure 7.5: Illustration of plasma bubble observation.

non-active period of the EPBs roughly starts from the middle of April until the end of August.

### 7.3 Comparison of TEC and swTEC

Expressed within a 24-hours time interval, the whole year TEC and swTEC variation are compared and illustrated in Figure 7.6. The TEC amount estimated from data set of the IGS Ionosphere Working Group, plotting in green curve, can be read from the right Y-axis, while the swTEC amount estimated from data set of the two local observational networks can be read from the left Y-axis. Differences between these two estimations are quite obvious.

The evaluation of the TEC curve shown in Figure 7.6 is fairly straightforward. The amount of the TEC begins to increase right after the local sunrise each day until it reaches its maximum about two hours after the midday when the ionized electron density also reaches its daily maximum. Afterwards, the amount of the TEC begins to decay continually until to the dawn of the next day. This 24-hours TEC variation explicitly agree with the fact that the ionization increases the amount of electron during the day, while prevailing recombination declines the amount of electron during the night.

The behavior of the local swTEC variation for both observational networks, however, show a different tendency where the amount of the swTEC begins to quickly climb up after about 19 o'clock at local time. Then, peak values are reached for both observational networks at vicinity of the local midnight. Followed by a decay phase lasts until about two hours after the lo-

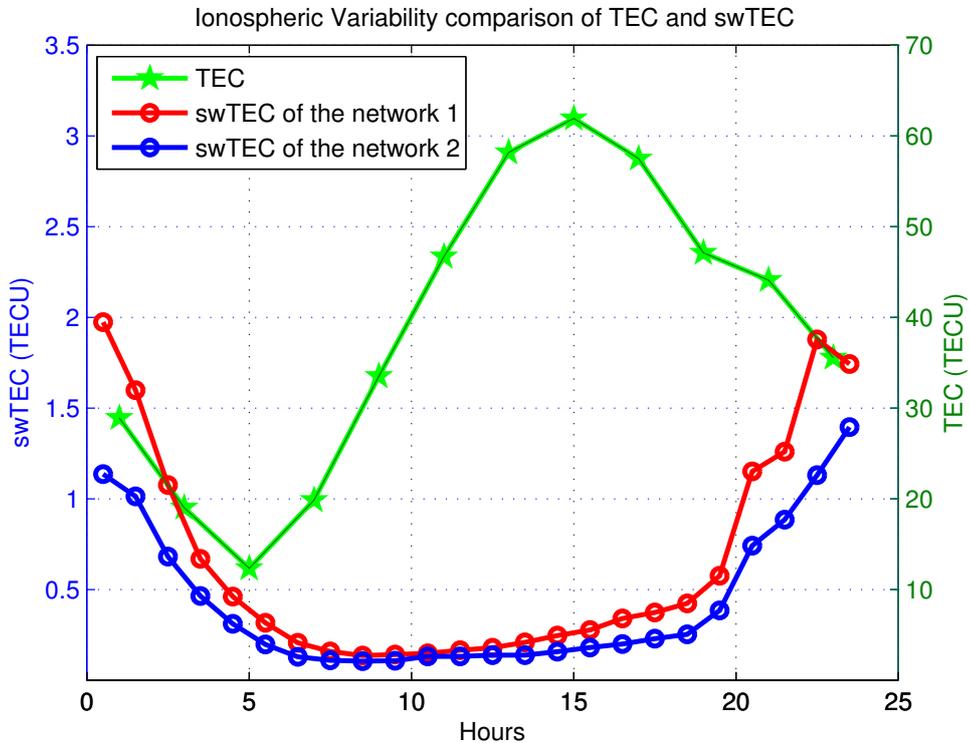


Figure 7.6: Whole year TEC and swTEC comparison.

cal sunrise. The amount of the local swTEC remains relatively steady within the whole morning. Small increases are visible during the afternoon for both observational networks until to the local sunset. Main reason for such kind of the swTEC behavior is thought to relate to the generation and vanishment of the EPBs during the night time.

The growth phase of the EPBs begins at approximate one to two hours after the local sunset at or near the geomagnetic equator. Then, the EPBs extend to low latitudes along the geomagnetic field lines due to pressure gradients. Around the local pre-midnight (one to two hours before the local midnight), the bubbles become gradually diluted and usually move to the east. The two observational networks are not located at the geomagnetic equator but beyond 20 degrees south of the geomagnetic latitude (see from Figure 3.2). As the plasma bubbles move along the magnetic field lines towards low latitude regions, the plasma density accumulation results in concentration gradient which also points to the low latitudes [27]. Such plasma diffusion takes some time to arrive the region where the two observational networks located. Therefore, it makes sense that the crest of the plasma

bubbles is observed at the time slightly before the local midnight.

From the Figure 7.6 both the swTEC curves fluctuate within two TECU. Another possible contribution may involve the MSTIDs which are usually observed with amplitudes of about one TECU. The spatial scale of the MSTIDs characterized by hundred of kilometers also matches the scale of the two triangular observational networks. However, the extent of the MSTIDs contributing to the swTEC in the observational regions still requires further analysis.

Similarly, the LSTIDs may contribute for the green colored TEC values. Research has reported that the amplitude associated with the LSTIDs can be observed greater than 10 TECU during strong magnetic storms [44]. In Brazilian ionospheric variability case, intense day time ionization combines with tropical ionospheric phenomena, for instance the daytime scintillations, as well as influence of the geomagnetic field distribution and the SAA. These phenomena can together pull up the TECU, far greater than those values calculated from the mid-latitude or high-latitude areas. Therefore, it may not surprise to see relative large TEC amount obtained in Brazilian area from data set of the IGS Ionosphere Working Group. However, to what extent could the LSTIDs contribute to the TEC amount still requires further investigation.

It should be pointed out that the local swTEC evaluation has superior temporal and spatial resolution. It has mentioned in previous data set chapter that the IGS VTEC maps are sampled with 2 hours in time by  $5^\circ$  in longitude and  $2.5^\circ$  in latitude, which maybe sufficient for a global TEC investigation. However, the IGS VTEC maps do not contain enough resolution for the local swTEC variation study. Therefore, less accuracy can be introduced regarding the local small-scale TEC variation evaluation done by the IGS Ionosphere Working Group's data set. In other words, the bilinear interpolation method used by both the triangular observational networks performs a better local small-scale TEC variability computation than those VTEC maps produced from data set of the IGS Ionosphere Working Group.

It is obvious that in Figure 7.6, the swTEC amount of the two observational networks are quite small compared to the TEC amount in a large dimensional scale. This is due to the spacial high-pass filtering effect on the true electron density variations from where the small-scaled highly fluctuation parts of the ionospheric variability can be captured by the two local observational networks. Illustration of spacial high-pass filtering effect is shown in Figure 7.7. The blue colored oscillation curve corresponds to spacial high frequency terms of the ionospheric variability, whereas the black colored varying curve corresponds to the spacial low frequency terms of the ionospheric variability in a relative large scale dimension. In other words, the

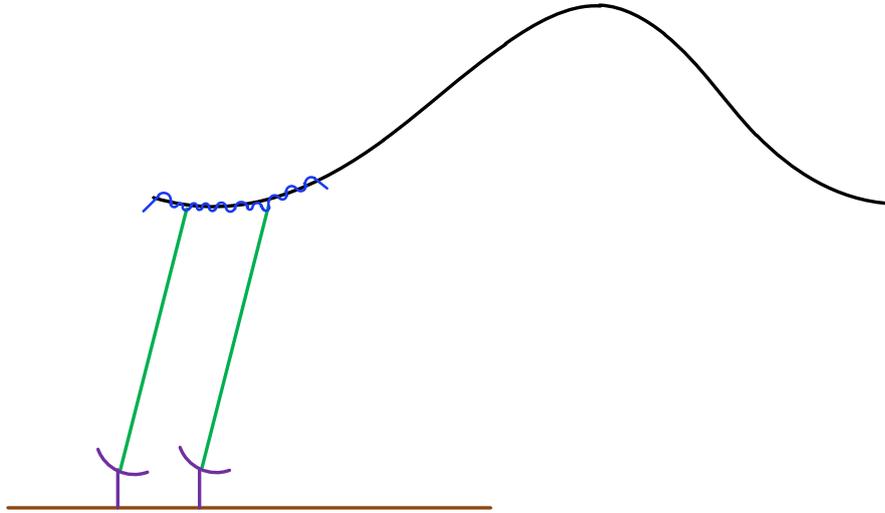


Figure 7.7: Spacial high-pass filtering effect can be resolved by the local observational network.

small-scaled ionospheric variability due to activities of the EPBs can be resolved and recorded by the local observational network, while the large-scaled variation can be reflected from the data of IGS VTEC map.

## 7.4 Conclusion

This project presents a whole year ionospheric variability study at local sites in Southern part of Brazil. The analysis has been performed for two nearby observational networks with spatial scale of hundred kilometers, matching with the spacial scale of the plasma bubbles which are normally of 100 kilometers in dimension. Results of the analysis show that the whole year ionospheric variation tendency is strongly affected by the annual active period of the EPBs from September to March. Daily behavior of the local swTEC variability has tight connection with the emergence and vanishment of the EPBs during the night time. Compared to the VTEC map data offered by the IGS Ionosphere Working Group, the methodology applied in this project yields fine local results for the small scale ionospheric variability. Because the local ionospheric variation details can be resolved by the spatial high-pass filtering effect performed by the observational networks.

## 7.5 Future Work

Future work may include several aspects of improvement:

1). Due to sudden bursts of the tropical ionospheric irregularities, the data sampled each Thursday may have lost some useful ionospheric variation information. Thus, perform a ‘truly’ whole year rms delay error analysis (then converted to TECU) including successive days may reveal some more features.

2). Select more triangular observational networks to track the the local variation of the ionospheric intensity. Further ground stations may come from the eastern regions of Brazil. Thus, traces of the plasma bubbles can be studied in detail.

3). Integrate other GNSS receivers with the Russian GLONASS (Globalnaya Navigazionnaya Sputnikovaya Sistema) signal enabled can provide supplement for those lost observational data records. GNSS signals from the European GALILEO system can offer a better resolution.

4). The IGS Ionosphere Working Group plan to reduce the temporal interval from 2 hours to 1 hour in the near future, thus chose data with fine temporal-spatial resolution provided by the IGS Ionosphere Working Group can produce a better global VTEC map in Brazil.

5). Try to find a method to separate the ionospheric variation made by the EPBs and the MSTIDs or other equatorial ionospheric phenomena and find out to what extent these phenomena can contribute to the ionosphere variation.

6). A correlation study can be performed between the swTEC and the solar flare variation to further reveal how much of TECU can be contributed from the solar flare activities.

# References

- [1] G. Seeber, *Mathematik : Satellite Geodesy : Foundations, Methods, and Applications*, ch. 2,7, pp. 44,54,212,220,235. Berlin: Walter de Gruyter, 2nd ed., October 2008.
- [2] P. Enge and F. van Diggelen, “GPS: An Introduction to Satellite Navigation, With an Interactive Worldwide Laboratory Using Smartphones.” <https://class.coursera.org/gpslab-001>, November 2014. Retrieved September 17, 2015.
- [3] A. El-Rabbany, *Introduction to GPS: The Global Positioning System*. Artech House, 2nd ed., August 2006.
- [4] National Geospatial-Intelligence Agency, “Maps of Magnetic Elements from the WMM2015.” <https://www.ngdc.noaa.gov/geomag/WMM/image.shtml>, May 2015. Retrieved October 8, 2015.
- [5] ESA, “June 2014 Magnetic Field.” [http://www.esa.int/spaceinimages/Images/2014/06/June\\_2014\\_magnetic\\_field](http://www.esa.int/spaceinimages/Images/2014/06/June_2014_magnetic_field), June 2014. Retrieved September 17, 2015.
- [6] Wikipedia, “GPS Signals.” [http://en.wikipedia.org/wiki/GPS\\_signals#Navigation\\_message](http://en.wikipedia.org/wiki/GPS_signals#Navigation_message), November 2014. Retrieved April 17, 2015.
- [7] E. D. Kaplan and C. Hegarty, *Understanding GPS: Principles and Applications*. Artech House Mobile Communications, Artech House, 2nd ed., November 2006.
- [8] U.S. Naval Observatory, “USNO NAVSTAR Global Positioning System.” <http://www.usno.navy.mil/USNO/time/gps/gps-info>. Retrieved April 16, 2015.
- [9] J. Johansson, “Lecture Notes of Satellite Positioning,” tech. rep., Chalmers University of Technology, November 2010.

- [10] M. S. Grewal, L. R. Weill, and A. P. Andrews, *Global Positioning Systems, Inertial Navigation, and Integration*. John Wiley & Sons.
- [11] G. Xu, *GPS Theory, Algorithms and Applications*, ch. 4,5, pp. 37,41,51. Springer, 2007.
- [12] A. Leick, *GPS Satellite Surveying*. New York: John Wiley, 1st ed., 1995.
- [13] R. B. Langley, “Dilution of Precision.” [http://www.nrem.iastate.edu/class/assets/nrem446\\_546/week3/Dilution\\_of\\_Precision.pdf](http://www.nrem.iastate.edu/class/assets/nrem446_546/week3/Dilution_of_Precision.pdf), May 1999. Retrieved September 23, 2015.
- [14] National Earth Science Teachers Association, “The Ionosphere.” <http://www.windows2universe.org/earth/Atmosphere/ionosphere.html>, 2012. Retrieved March 23, 2015.
- [15] Wikipedia, “Ionosphere.” <https://en.wikipedia.org/wiki/Ionosphere>, October 2015. Retrieved November 1, 2015.
- [16] J. K. Hargreaves, *The Upper Atmosphere and Solar-Terrestrial Relations*. Cambridge University Press, 3rd ed., May 1992.
- [17] LUXORION Project, “Ionospheric Perturbations.” <http://www.astrosurf.com/luxorion/qs1-perturbation.htm>. Retrieved October 10, 2015.
- [18] Integrated Publishing, “Variations In The Ionosphere.” <http://www.tpub.com/neets/book10/40h.htm>. Retrieved October 19, 2015.
- [19] S. Skone, “Lecture Notes of ENGO 633,” tech. rep., University of Calgary, 2007.
- [20] M. J. Buonsanto, “Ionospheric Storms – A Review,” *Space Science Reviews*, vol. 88, pp. 563–601, March 1999.
- [21] X. Pi, “Ionospheric Storms.” <http://iono.jpl.nasa.gov/storm.html>. Retrieved March 19, 2015.
- [22] A. Coster and T. Tsugawa, “The Impact of Traveling Ionospheric Disturbances on Global Navigation Satellite System Services .” <http://www.ursi.org/proceedings/procga08/papers/fgp1.pdf>. Retrieved October 13, 2015.
- [23] Z. T. Katamzi and L.-A. McKinnell, “Observations of Traveling Ionospheric Disturbances Associated with Geomagnetic Storms,” *IEEE 2011 XXXth URSI General Assembly and Scientific Symposium*, 2011.

- [24] Y. Memarzadeh, *Ionospheric Modeling for Precise GNSS Applications*. PhD thesis, Netherlands Geodetic Commission, 2009.
- [25] National Oceanic and Atmospheric Administration, “Sudden Ionospheric Disturbance (SID).” [http://www.ngdc.noaa.gov/stp/space-weather/ionospheric-data/sids/documentation/readme\\_sudden-ionospheric-disturbances.pdf](http://www.ngdc.noaa.gov/stp/space-weather/ionospheric-data/sids/documentation/readme_sudden-ionospheric-disturbances.pdf), March 2014. Retrieved October 13, 2015.
- [26] IPS Radio and Space Services, Bureau of Meteorology, Australian Government, “About Ionospheric Scintillation.” <http://www.sws.bom.gov.au/Satellite/6/3>. Retrieved March 21, 2015.
- [27] M. A. Hei, R. A. Heelis, and J. P. McClure, “Seasonal and Longitudinal Variation of Large-scale Topside Equatorial Plasma Depletions,” *Journal of Geophysical Research - Space Physics*, vol. 110, December 2005.
- [28] D. Shvarts, U. Alon<sup>1</sup>, D. Ofer, R. L. McCrory, and C. P. Verdon, “Non-linear Evolution of Multi-mode Rayleigh–Taylor Instability in Two and Three Dimensions,” in *Proceedings of the Fifth International Workshop on Compressible Turbulent Mixing*, World Scientific, 1996.
- [29] A. Taori, J. J. Makela, and M. Taylor, “Mesospheric Wave Signatures and Equatorial Plasma Bubbles: A Case Study,” *Journal of Geophysical Research: Space Physics (1978-2012)*, vol. 115, June 2010.
- [30] M. C. Kelley, J. J. Makela, O. de La Beaujardire, and J. Retterer, “Convective Ionospheric Storms: A Review,” *Reviews of Geophysics*, vol. 49, pp. 445–446, June 2011.
- [31] Mori, “The South Atlantic Anomaly.” [http://www.bibliotecapleyades.net/ciencia/ciencia\\_earthchanges49.htm](http://www.bibliotecapleyades.net/ciencia/ciencia_earthchanges49.htm), November 2007. Retrieved June 1, 2015.
- [32] E. Stassinopoulos, “Forty–Year Drift and Change of The SAA,” tech. rep., NASA Goddard Spaceflight Center, 2007.
- [33] A. Kapusta, A. J. Coster, and P. J. Erickson, “The South Atlantic Anomaly and Its Influence on The Total Electron Content Distribution,” *American Geophysical Union, Fall Meeting 2006*, December 2006.
- [34] M. Kruglanski, E. D. Donder, N. Messios, L. H. Emmanuel Gamby, and S. Calders, “SPENVIS Homepage.” [www.spennis.oma.be/](http://www.spennis.oma.be/), October 2013. Retrieved March 17, 2015.

- [35] M. Bauer, *Vermessung und Ortung mit Satelliten*. Wichmann, 3rd edn. ed., 1994.
- [36] B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins, *Global Positioning System: Theory and Practice*, ch. 6, pp. 97–98,102. Springer, 5th revised ed., 2001.
- [37] K. Davies, *Ionospheric Radio*. IEEE Electromagnetic Waves Series 31, London: Peter Peregrinus, 1990.
- [38] L. R. Hartmann, G., “Range Errors Due to Ionospheric and Tropospheric Effects for Signal Frequencies Above 100 MHz,” *Bull. Geod.*, vol. 58, pp. 109–136, 1984.
- [39] C. Rieck, P. Jarlemark, K. Jaldehag, and J. Johansson, “Thermal Influence on the Receiver Chain of GPS Carrier Phase Equipment for Time and Frequency Transfer,” *IEEE International Frequency Control Symposium and PDA Exhibition Jointly with the 17th European Frequency and Time Forum*, vol. 13, pp. 326–331, 2003.
- [40] T. R. Emardson, P. O. J. Jarlemark, J. M. Johansson, and S. Schfer, “Spatial Variability In The Ionosphere Measured With GNSS Networks,” *Radio Science*, vol. 48, pp. 646–652, 2013.
- [41] R. Emardson, P. Jarlemark, S. Bergstrand, T. Nilsson, and J. M. Johansson, “Measurement Accuracy In Network-RTK,” tech. rep., SP Boras, 2009.
- [42] L. F. C. de Rezende, E. R. de Paula, I. J. Kantor, and P. M. Kintner, “Mapping and Survey of Plasma Bubbles Over Brazilian Territory,” *Journal of Navigation*, vol. 60, pp. 69–81, 2007.
- [43] D. C. M. Amorim, A. A. Pimenta, J. A. Bittencourt, and P. R. Fagundes, “Longterm Study of Medium-scale Traveling Ionospheric Disturbances Using O I 630 nm Allsky Imaging and Ionosonde Over Brazilian Low Latitudes,” *Journal of Geophysical Research. A. Space Physics*, vol. 116, 2011.
- [44] J. Jakobsen, *Kinematic GNSS: Ionospheric Variations and Characteristics of Multipath*. PhD thesis, University of Copenhagen, 2010.
- [45] P. O. J. Jarlemark, T. R. Emardson, J. M. Johansson, and G. Elgered, “Ground-Based GPS for Validation of Climate Models: The Impact of Satellite Antenna Phase Center Variations,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, pp. 3847–3854, 2010.

# Appendix A

## Methodology for Estimating the Ionospheric Variability

The measurement of the small scale ionospheric spatial variability of swTEC, denoted as  $\sigma_{\Delta l}$ , is based on the deviation between interpolated and ‘true’ measured ionospheric electron content,  $\Delta l$ , at the center of a triangular observational network. The  $\Delta l$  can be obtained by forming the L4-combination via the carrier phase measurement model and eliminating the non ionospheric contributions.

According to the carrier phase measurement model, the phase observed by a single GPS receiver at L1 and L2 carrier can be modeled as:

$$\Phi^{L1} = \frac{\rho}{\lambda_{L1}} + N_{L1} + f^{L1}(\Delta t_{L1}^s + \Delta t_{L1}^r) + \frac{l_i^{L1}}{\lambda_{L1}} + \frac{l_t}{\lambda_{L1}} + \mu_{L1} \quad (\text{A.1})$$

$$\Phi^{L2} = \frac{\rho}{\lambda_{L2}} + N_{L2} + f^{L2}(\Delta t_{L2}^s + \Delta t_{L2}^r) + \frac{l_i^{L2}}{\lambda_{L2}} + \frac{l_t}{\lambda_{L2}} + \mu_{L2} \quad (\text{A.2})$$

where  $\Phi$  is the measured phase in fraction of cycles;  $\lambda$  is the wavelength of carrier signal;  $\rho$  is the true geometrical satellite-to-receiver distance;  $N$  is the phase ambiguity parameter which is the integer number of cycles;  $f$  is the carrier frequency;  $\Delta t^s$  and  $\Delta t^r$  refer to the satellite and receiver clock bias respectively;  $l_i$  and  $l_t$  refer to the signal delay in the ionosphere and troposphere respectively; and  $\mu$  contains the signal multi-path and receiver measurement error.

By combining the above equations, a new expression, which removes all frequency independent terms such as the geometrical distance ( $\rho$ ) as well as the signal delay in the troposphere ( $l_t$ ), is achieved.

$$\lambda_{L1}\Phi^{L1} - \lambda_{L2}\Phi^{L2} = l_i^{L1} - l_i^{L2} + \lambda_{L1}N_{L1} - \lambda_{L2}N_{L2} + C(\Delta t_{L1}^s - \Delta t_{L2}^s + \Delta t_{L1}^r - \Delta t_{L2}^r) + \lambda_{L1}\mu_{L1} - \lambda_{L2}\mu_{L2} \quad (\text{A.3})$$

where  $C$  refers to the speed of the light in vacuum. As can be seen in Equation (A.3), the difference regarding the ionospheric effect between the two frequencies is however produced. Thus this combination can be used to derive the TEC along the LOS propagation in the ionosphere.

Based on Equation (A.3), the slant ionospheric delay can be obtained. A reorganized form of equation shows

$$\Theta_p^s \equiv \kappa(\lambda_{L1}\phi^{L1} - \lambda_{L2}\phi^{L2}) = I^s + n + \tau_s + \tau_r + \mu \quad (\text{A.4})$$

where  $I^s$  is the slant TEC and  $\kappa$  is the scaling factor between L4 measurements and the TEC. This derived L4-combination,  $\Theta_p^s$ , contains information about the ionosphere but also potential non-ionospheric variations from the multi-path and measurement noise,  $\mu$ ; the receiver inter frequency L1-L2 bias,  $\tau_r$ ; the satellite inter frequency L1-L2 bias,  $\tau_s$ ; and the phase ambiguities,  $n$ .

Two assumptions should be noted here. One is that the GPS receivers have constant temperatures, thus the receiver L1-L2 bias variation is considered to be relatively small [39]. The satellite L1-L2 biases are generally considered to be constant over several hours. The phase ambiguity term is by definition constant over satellite passes. Therefore, by applying the time average of  $\Theta_p^s$  over a satellite passage we can form a new variable with the time-averaged value removed containing only contributions from the ionosphere,  $I$ , as well as the multi-path and measurement noise  $\mu$ .

$$\Theta^s = \Theta_p^s - a \{ \Theta_p^s \} \quad (\text{A.5})$$

where the time averaged term

$$a \{ \Theta_p^s \} = a \{ I^s + n + \tau_s + \tau_r + \mu \} = I_m^s + n + \tau_s + \tau_r + \mu_m \quad (\text{A.6})$$

The method during the analysis procedure is based on by using observations of the slant ionospheric delay from the three sites in a triangular observational network in order to interpolate the expected ionospheric delay for a site at the center. Therefore, for each site in a triangle the L4-combination for a specific satellite observation can be formed. By applying the L4-combination to each of the three surrounding sites at the vertices of the triangular network, an expected L4-combination at the inner site can be interpolated. The interpolation is performed through bilinear interpolation along both latitude and longitude components of plane coordinates. The difference between the interpolated and measured L4 values, at the inner site, is calculated as

$$\Delta\Theta^s = \Theta^s - \Theta_i^s \quad (\text{A.7})$$

where

$$\Theta_i^s = \sum_{n=1}^3 a_n \Theta_n^s, \quad \sum_{n=1}^3 a_n = 1 \quad (\text{A.8})$$

The weights of  $a_n$  can be uniquely determined during the bilinear interpolation as long as both the latitude and longitude components of each site are known in advance. And  $n$  just refers to the three surrounding sites.

The next step followed is to map the L4 deviations to the zenith direction,

$$\Theta = \frac{\Theta^s}{m_i(\varepsilon)} \quad (\text{A.9})$$

where  $m_i$  is the mapping function

$$m_i(\varepsilon) = \frac{R + h}{\sqrt{(R + h)^2 - R^2 \cos^2 \varepsilon}} \quad (\text{A.10})$$

and  $\varepsilon$  is the observational elevation angle;  $R$  is the radius of the earth and  $h$  is the height of the ionosphere, usually between ranges of 300 km and 400 km, represented as a thin shell. The interpolation operation and the mapping to the zenith result in a  $\Delta\Theta$  containing the seeking ionospheric electron content deviation  $\Delta l$ .

Now all available  $\Delta\Theta$  values from all visible satellites during one hour can be employed to construct the hourly rms values. To guarantee continuity of the integer number of cycles in the measured carrier phase, only cycle slip free satellites with passage time longer than 1 hour are utilized. In order to receive LOS GPS signals, satellite observation elevation angle above  $20^\circ$  are chosen. In fact, the duration of the satellite passes we use is typically a couple of hours in this project.

$$\sigma_{\Delta\Theta} = \sqrt{\sum \frac{\Delta\Theta^2}{N}} \quad (\text{A.11})$$

This rms value contains contribution from the ionosphere,  $I$ , and the multi-path and measurement noise,  $\mu$ . Since these contributions can be assumed to be uncorrelated, a squared term of swTEC can be written as

$$\sigma_{\Delta l}^2 = \sigma_{\Delta\Theta}^2 - \sigma_{\Delta\mu}^2 \quad (\text{A.12})$$

The expected contribution from the multi-path and measurement noise can be removed according to a model discussed in [41] and [45].

Last, the ionospheric variability  $\sigma_{\Delta l}$  can be obtained from Equation (A.12).

# Appendix B

## Observational Data Set

This part describes how the observational data were selected from the IBGE website. For the during of year 2013, 155 days' data are used in this investigation. Basically, data of each Thursday were used to produce the ionospheric variation. Specifics regarding the data can be found in Table B.1.

According to calendar of the year 2013, the first Thursday should be January 3<sup>rd</sup>, however, station of MGIN lost data on that day. The same station also lost data the following day. Therefore, the 5<sup>th</sup> day was chosen to represent the first week data.

Data sampled every 7 days go smoothly from 10<sup>th</sup> day (second Thursday) to 52<sup>th</sup> day. Then, instead of the 59<sup>th</sup> day, the 60<sup>th</sup> day was selected. This is because at the very beginning we downloaded the observational data from the NASA website instead of the IBGE website and there was no observational data record for the ILHA station at the 59<sup>th</sup> day in NASA website, thus the next day was chosen. When we realized the downloading error we

Table B.1: Specific date used from IBGE website each Thursday

Date							
5 <sup>th</sup>	10 <sup>th</sup>	17 <sup>th</sup>	24 <sup>th</sup>	31 <sup>th</sup>	38 <sup>th</sup>	45 <sup>th</sup>	52 <sup>th</sup>
60 <sup>th</sup>	66 <sup>th</sup>	73 <sup>th</sup>	80 <sup>th</sup>	87 <sup>th</sup>	94 <sup>th</sup>	101 <sup>th</sup>	108 <sup>th</sup>
115 <sup>th</sup>	122 <sup>th</sup>	129 <sup>th</sup>	136 <sup>th</sup>	143 <sup>th</sup>	150 <sup>th</sup>	157 <sup>th</sup>	164 <sup>th</sup>
171 <sup>th</sup>	178 <sup>th</sup>	185 <sup>th</sup>	192 <sup>th</sup>	199 <sup>th</sup>	206 <sup>th</sup>	213 <sup>th</sup>	220 <sup>th</sup>
227 <sup>th</sup>	234 <sup>th</sup>	241 <sup>th</sup>	248 <sup>th</sup>	255 <sup>th</sup>	262 <sup>th</sup>	269 <sup>th</sup>	274 <sup>th</sup>
283 <sup>th</sup>	290 <sup>th</sup>	297 <sup>th</sup>	304 <sup>th</sup>	311 <sup>th</sup>	318 <sup>th</sup>	325 <sup>th</sup>	331 <sup>th</sup>
339 <sup>th</sup>	346 <sup>th</sup>	353 <sup>th</sup>	360 <sup>th</sup>				

still decided to use the navigation file of the 60<sup>th</sup> day, but selected the observational data from the IBGE website. Data of the 60<sup>th</sup> day leads to an unexpected outcome because this specific day yielded large rms delay differences which imply strong local ionospheric variability occurred (see ‘Result’ chapter Figure 7.1).

Next, data sampled every 7 days go smoothly from 66<sup>th</sup> day to 360<sup>th</sup> day, the last Thursday (26 of December). Specific data lost are:

(1) Missing data for the SPJA station at 276<sup>th</sup> day. Same happened at 275<sup>th</sup> day for this station, therefore, data of 274<sup>th</sup> day was selected instead.

(2) Missing data for the ILHA station at 332<sup>th</sup> day, thus data of 331<sup>th</sup> day was used instead.

Besides data sampled every 7 days through the whole year, another two two-months-datasets were collected and processed for supplementary purpose. One dataset starts from the 30<sup>th</sup> day to the 89<sup>th</sup> day with a missing data at the 61<sup>th</sup> day (at this time, data set of the 59<sup>th</sup> day were included). Another starts from the 283<sup>th</sup> day to the 346<sup>th</sup> day with a missing data at the 332<sup>th</sup> day. Both missing data result in discontinuities (need to zoom in to see clearly) in the plot of Figure 7.1.