Precise Ionosphere Modeling Using Regional GPS Network Data

Y. Gao and Z.Z. Liu

Department of Geomatics Engineering The University of Calgary Calgary, Alberta, Canada T2N 1N4
e-mail: gao@geomatics.ucalgary.ca; Tel: 403-220-6174; Fax: 403-284-1980

Received: 2 July 2002 / Accepted: 18 July 2002

Abstract. The ionosphere affects the electromagnetic waves that pass through it by inducing an additional transmission time delay. The ionosphere influence has now become the largest error source in GPS positioning and navigation after the turn-off of the Selective Availability (SA). In this paper, methods of 2D grid-based and 3D tomography-based ionospheric modeling are developed based on regional GPS reference networks. Performance analysis was conducted using data from two different regional GPS reference networks. The modeling accuracy of the vertical TEC (VTEC) is at the level of several TECU for 2D ionospheric modeling and about one TECU for 3D tomographic modeling after a comparison to independent ionospheric map data or directly measured ionospheric TEC values. The data analysis has indicated that the modeling accuracy based on the 3D tomography method is much higher than the 2D grid-based approach.

Key words: GPS, Ionosphere, Ionosphere Grid, Tomography.

1 Introduction

The ionosphere affects the electromagnetic waves that pass through it by inducing an additional transmission time delay. The magnitude of this effect is determined by the amount of total electron content (TEC) and the frequency of electromagnetic waves. Under normal solar activity conditions, this influence on GPS signals is usually in the range from a few meters to tens of meters but it could reach more than 100 meters during severe ionosphere storms. After the turn-off of the Selective Availability (SA), the ionosphere effect has become the largest error source in GPS positioning and navigation. For high precision GPS positioning, the ionosphere effect must be estimated so that a correction can be made to eliminate it from the GPS observations. Precise ionosphere effect estimates are also important for space weather research and earth observation applications (Komjathy, 1997).

The currently available ionosphere correction models include the Klobuchar ionosphere parameters broadcast from the GPS satellites but the Klobuchar model could only correct about 50% of the total ionosphere effects (Klobuchar, 1987). More precise ionosphere model is therefore required. Ionosphere modeling methods using GPS data from ground GPS networks have been extensively investigated in the past several years (Komjathy, 1997; Skone, 1998; Jakowski et al., 1998; Liao, 2000; Fedrizzi et al., 2001). Komjathy (1997) established a polynomial ionospheric model based on data from reference stations of International GPS Service (IGS) and has compared it to the TOPEX/Poseidon-derived (T/P) TEC data. Agreements at the level of 5 TECU were reported under medium and low solar activity conditions (Komjathy, 1997). Skone (1998) employed a two-dimensional grid-based model to characterize the ionosphere activities over the auroral region and an accuracy of about 34 cm was obtained. In Gao et al. (2002), ionosphere parameters are estimated along with satellite and receiver biases using data from a regional area GPS network. To date, all proposed ionospheric models could be classified into two different categories: grid-based and function-based. The early modeling methods are mostly based on the function fitting techniques such as the broadcast ionosphere model from the GPS satellites (Klobuchar, 1987), the polynomial functions (Coster et al., 1992; Komjathy, 1997) and the spherical harmonics (Schaer, 1999; Walker, 1989). On the other hand, the grid-based method, first proposed by the MITRE Corporation and the Air Force Phillips Laboratory (El-Arini et al., 1994 and 1995), has demonstrated its capability for higher modeling accuracy when compared to the function-based algorithms. The grid-based modeling technique has since
been extensively used for both global and regional network-based ionosphere recovery (Gao et al., 1994; FAA, 1997; Skone, 1998; Liao, 2000).

No matter whether it is function or grid-based, current ionosphere modeling is two-dimensional in nature, which assumes that the ionosphere is condensed on a single shell at a fixed altitude above the earth surface. This assumption, however, is only an approximation to the reality and it is not physically true. In order to further improve the ionosphere modeling accuracy and the model’s sensitivity to temporal ionosphere variations, ionospheric tomography modeling method has started to receive more attentions in the recent years (Raymund et al., 1990 and 1994; Raymund, 1995; Howe, 1997; Liu et al., 2001a and 2001b). An ionosphere tomographic model can describe the ionosphere field in a three-dimensional frame instead of a two-dimensional frame as used by previous methods. As a result, the ionosphere tomography method would allow for more precise exploration of the ionospheric characteristics and subsequently for more precise modeling accuracy.

This paper describes the recent research results in the area of high precision ionosphere modeling using regional GPS reference network data and focuses on 2D grid-based and 3D tomography-based ionospheric modeling. The paper is organized as follows. In Section 2, ionospheric delays are derived from GPS dual-frequency observations including an algorithm for carrier phase leveling on code-derived ionospheric delay measurements. Section 3 discusses a 2D ionosphere modeling method while a 3D ionosphere tomographic model is presented in Sections 4. Numerical results and performance analysis are provided in Section 5. Conclusions are given in Section 6.

2 GPS IONOSPHERE MEASUREMENTS

A dual-frequency GPS receiver used at a reference station consists of both code and carrier phase observations on L1 (1575.42 MHz) and L2 (1227.60 MHz) frequencies, denoted as $P_i$ and $\Phi_i$ ($i = 1, 2$) in the following. Mathematically the corresponding observations can be described as

\[ P_1 = \rho + c(dt - dT) + d_{\text{amb}} + d_{\text{trop}} + k_1 I + b_{P_1} - B_{P_1} + d_{\text{mult}/P_1} + \varepsilon(P_1) \] (1)

\[ \Phi_1 = \rho + c(dt - dT) + \lambda_i N_i + d_{\text{amb}} + d_{\text{trop}} - k_1 I + b_{\Phi_1} - B_{\Phi_1} + d_{\text{mult}/\Phi_1} + \varepsilon(\Phi_1) \] (2)

\[ P_2 = \rho + c(dt - dT) + d_{\text{amb}} + d_{\text{trop}} + k_1 I + b_{P_2} - B_{P_2} + d_{\text{mult}/P_2} + \varepsilon(P_2) \] (3)

\[ \Phi_2 = \rho + c(dt - dT) + \lambda_2 N_2 + d_{\text{amb}} + d_{\text{trop}} - k_1 I + b_{\Phi_2} - B_{\Phi_2} + d_{\text{mult}/\Phi_2} + \varepsilon(\Phi_2) \] (4)

where

- $k_i = f_i^2/(f_1^2 - f_2^2)$, $i = 1, 2$
- $\rho$ is the true geometric range between receiver and satellite (m);
- $c$ is the speed of light (m/s);
- $dt$ is the satellite clock error with respect to GPS time (s);
- $dT$ is the receiver clock error with respect to GPS time (s);
- $\lambda_i$ is the wavelength of GPS signal on $L_i$ (m);
- $N_i$ is the carrier phase integer ambiguity (cycle);
- $d_{\text{trop}}$ is the tropospheric delay (m);
- $I$ is the ionospheric delay parameter (m);
- $d_{\text{amb}}$ is the satellite orbit error (m);
- $d_{\text{mult}}$ is the multipath effect (m);
- $b$ is the satellite hardware delay (m);
- $B$ is the receiver hardware delay (m); and
- $\varepsilon(\cdot)$ is the measurement noise (m).

Differencing the code observations from L1 and L2 results in the following ionosphere measurements:

\[ P_1 - P_2 = -I + b - B + \varepsilon(P_1 - P_2) \] (5)

where $b = b_{P_1} - b_{P_2}$, $B = B_{P_1} - B_{P_2}$. $b$ and $B$ represent the differential hardware delays between the L1 and L2 frequencies and they are often referred to as satellite and receiver L1/L2 inter-frequency biases. Although these biases are actually time dependent, in practice they are very stable over time on a scale of days to months so that they can be treated as constants during ionosphere modeling (Gao et al., 1994; Schauer, 1999). As to the bias magnitude, the satellite inter-frequency bias is usually in the range of several nanoseconds while the receiver inter-frequency bias could be as large as more than 10 nanoseconds (Gao et al., 1994).

Considering the much higher noise level of the ionosphere measurements derived from the code measurements, the carrier phase observations from L1 and L2 described in equations (2) and (4) can be used to smooth the code observation for a more precise vertical TEC estimate. Such carrier phase smoothing technique is also often referred as “carrier phase leveling”. Given below is a smoothing function that has been described in Gao et al. (2002).
\[(P_1 - P_2)_k = \frac{(w_1)_k}{(w_1)_k + (w_2)_k} (P_1 - P_2)_k + \frac{(w_2)_k}{(w_1)_k + (w_2)_k} [(P_1 - P_2)_{k-1} + \delta(\Phi_1 - \Phi_2)_{k,k-1}] \quad (6)\]

where \(k\) is the time epoch index; \((P_1 - P_2)\) is the smoothed ionosphere measurement and

\[(w_1)_k = \frac{1}{\sigma^2_{(P_1 - P_2)_k}} \quad (7)\]

\[(w_2)_k = \frac{1}{\sigma^2_{(P_1 - P_2)_k} + \sigma^2_{(\Phi_1 - \Phi_2)_k}} \quad (8)\]

\[\delta(\Phi_1 - \Phi_2)_{k,k-1} = (\Phi_1 - \Phi_2)_k - (\Phi_1 - \Phi_2)_{k-1} \quad (9)\]

Note that the smoothed ionosphere measurements in equation (6) are still corrupted by the inter-frequency biases \(b_p\) and \(B_p\) which therefore need to be estimated along with the ionospheric delay parameters.

### 3D Ionosphere Modeling

The ionosphere is a region of ionized plasma that extends from roughly 50km to 2000km above the surface of the earth. Generally, the ionosphere can be divided into several layers in altitude according to electron density, which reaches its peak value at about 350km in altitude. For 2D ionospheric modeling, the ionosphere is assumed to be concentrated on a spherical shell of infinitesimal thickness located at the altitude of about 350km above the earth’s surface (Gao et al., 1994). The implementation of the single-layer grid model requires computation of the intersection of the line-of-sight between the GPS receiver and the observed satellite on the ionosphere shell as illustrated in Figure 1. The intersection point of the GPS signal with the ionospheric shell is defined as pierce point at which the slant ionospheric delay has an elevation angle of \(E\).

An ionospheric grid model consists of grids distributed on the ionospheric shell in preset spacing usually 3° x 3°. Any pierce point therefore will fall within a specific grid defined by its surrounding four grid points. The slant ionospheric delay at a pierce point, which has elevation angle of \(E\), can be linked to the vertical ionospheric delay at the same location by a mapping function \(sf(E)\). Meanwhile the vertical ionospheric delay at the pierce point can be described by the vertical ionospheric delays at its surrounding grid points. The relationship between the slant ionospheric delay at the pierce point and the vertical ionospheric delays of its surrounding four grid points can be expressed by the following equation:

\[sf(E) = 1 - \frac{1}{\sqrt{1 - \cos(E)/(1 + h/R)}} \quad (11)\]

where
- \(E\) is the satellite elevation angle;
- \(R\) is the earth radius;
- \(h\) is the height of the ionosphere shell above the earth’s surface;
- \((I_v)_k\) is the vertical ionospheric delay parameter at the grid point \(k\);
- \(P_k\) is a weighting function which is used to project the vertical ionospheric delay at grid point \(k\) to the pierce point.

Taking into account the existence of L1/L2 inter-frequency biases \(b\) and \(B\), we can establish the following ionosphere observation equation for a satellite (j) and receiver (i) pair:

\[\overline{(P_1 - P_2)}_{ij} = sf(E_j) \sum_{k=1}^{4} \left( (I_v)_k \cdot p_k + b_{ij} - B_{ij} \right) \quad (12)\]

Equation (12) is the fundamental equation for the grid point vertical ionospheric delay estimation in grid-based 2D ionospheric modeling using carrier smoothed code-derived ionosphere measurements. As to parameter estimation, the standard least squares or Kalman filtering method can be used to facilitate the optimal estimation of the vertical ionospheric delay parameters and the satellite/receiver inter-frequency biases.

### 4D Tomographic modeling

A two-dimensional (2D) ionospheric model as described in the last section has difficulty to characterize the ionospheric field in the full spatial dimensions because it is unable to provide the vertical ionospheric profile. A three-dimensional (3D) ionospheric model therefore is expected. A 3D ionospheric modeling method using
Tomography technique will be described in the following which is able to characterize not only the ionospheric horizontal profile but also the vertical one.

Tomography-based modeling consists of two fundamental steps. First, integral measurements are made of the medium of interest, ideally along many paths from many different viewing angles. Second, these integral measurements are inverted to obtain an estimate of the field (Howe, 1997). In ionospheric tomography, the integral measurements are the lines of sight between the GPS receivers and satellites, which pass through the entire ionosphere. A 3D ionospheric model can be constructed using the tomography technique, horizontally by spherical harmonics functions (SHFs) and vertically by the empirical orthogonal functions (EOFs). The harmonics functions are primarily formed by the first several orders of harmonics. The high-order harmonics however improve the sharpness of the fronts. EOFs are derived from existing data set (observation data or model data) of the medium of our interest. The observation equation of ionospheric tomography can be given as follows (Liu et al., 2001a and 2001b):

\[ \text{TEC} = \int_{\text{sat}} N_{\text{s}}(\lambda, \phi, z) \text{ds} = \int_{\text{sat}} [N_{\text{s}}^0(\lambda, \phi, z) + \delta N_{\text{s}}(\lambda, \phi, z)] \text{ds} \]  

(13)

where TEC is the total electron contents along the line of sight from a GPS satellite to a ground GPS receiver and \( N_{\text{s}}(\lambda, \phi, z) \) is the ionospheric electron density at the geospatial position of \((\lambda, \phi, z)\). \( N_{\text{s}}^0(\lambda, \phi, z) \) is the a priori value of \( N_{\text{s}}(\lambda, \phi, z) \) which could be an output from an empirical model that reflects the deterministic portion of our a priori information. \( \delta N_{\text{s}}(\lambda, \phi, z) \) is the correction to the a priori value. \( \lambda, \phi, z \) are longitude, latitude and altitude, respectively, referenced to a solar-geomagnetic coordinate system. For the convenience purpose, we could let \( N_{\text{s}}^0(\lambda, \phi, z) \) be approximately equal to zero, namely \( N_{\text{s}}^0(\lambda, \phi, z) = 0 \). Thus we have,

\[ \text{TEC} = \int_{\text{sat}} \delta N_{\text{s}}(\lambda, \phi, z) \text{ds} \]  

(14)

The correction for the electron density function \( \delta N_{\text{s}}(\lambda, \phi, z) \) can be modeled by a series of space-location related functions. More specifically we can employ spherical harmonic expansions horizontally and empirical orthogonal functions (EOFs) vertically to model the ionospheric electron correction term. The integration of these two sets of functions enables us to depict the ionosphere field in a 3D mode as follows:

\[ \delta N_{\text{s}}(\lambda, \phi, z) = \sum_{k=1}^{K} \sum_{m=-M}^{M} \sum_{n=-[m]}^{[m]} \left[ a_{mk}^n \cos(m\lambda) + b_{mk}^n \sin(m\lambda) \right] \cdot P_{mk} \cos(\phi) Z_k(z) \]  

(15)

where \( P_{mk} \cos(\phi) \) is the associated Legendre polynomial of order \( m \) and degree \( n \); \( Z_k(z) \) is the empirical orthogonal function (EOF); \( a_{mk}^n \) and \( b_{mk}^n \) are the coefficients to be estimated. Combining Equations (14) and (15), the observation equation for ionospheric 3D modeling can then be established by the following equation:

\[ \text{TEC} = \sum_{k=1}^{K} \sum_{m=-M}^{M} \sum_{n=-[m]}^{[m]} \left[ \int_{\text{rx}} a_{mk}^n \cos(m\lambda) + b_{mk}^n \sin(m\lambda) \right] \cdot P_{mk} \cos(\phi) Z_k(z) \text{ds} \]

(16)

\[ + \sum_{k=1}^{K} \sum_{m=-M}^{M} \sum_{n=-[m]}^{[m]} \left[ \int_{\text{sat}} \cos(m\lambda) P_{mk} \cos(\phi) Z_k(z) \text{ds} \right] \]

Equation (16) is the fundamental observation equation for 3D ionospheric modeling using tomography inversion technique, through which the GPS derived total electron content (TEC) and the coefficients describing the ionosphere field are linked. The rest task of ionospheric tomography is to optimally estimate the model coefficients in equation (16) in which the number of the unknown model parameters is determined by the truncation limits of SHFs and EOFs. A so-called weighted, damped least squares technique has been developed via a combination of weighting and damping procedures described in Liu et al. (2001a and 2001b).

5 Results AND ANALYSIS

The data from two regional GPS reference networks, SWEPOS and SCIGN, were processed to generate regional ionospheric model using the developed 2D grid-based and 3D tomography techniques.

5.1 Results of 2D Grid-Based Modeling

The data from the Swedish GPS reference network (SWEPOS) has been used to evaluate the performance of the 2D grid-based ionosphere modeling method. The SWEPOS network (Figure 2) consists of 21 continuously operating GPS stations extending from latitude 55 to 69 degrees north with an average station separation of around 200km (Liao, 2000). Data used in the numerical analysis includes a total of five days of consecutively tracked GPS data collected during May 17-21, 1999. For the purpose of this research, only ten stations in the southern portion of the SWEPOS (the two stations in the block frame are not included due to an accident loss of...
data) were used in the numerical computation. The ten stations are located within a range of approximately 55°N to 60°N in geographic latitude and 11°E to 18°E in geographic longitude.

![Reference Stations in SWEPOS](image)

Shown in Figure 3 and 4 are the vertical TEC estimates at latitudes of 59°N and 56°N respectively with a grid size of 3° × 3° based on the five-day data. The results show clearly the diurnal behavior of the ionosphere and the variations from day to day. These vertical TEC estimates are consistent with diurnal ionosphere behavior considering the overall change pattern. For external comparison purpose, the vertical TEC estimates derived from the global ionosphere maps (GIM) provided by the Center of Orbit Determination in Europe (CODE) are also shown in Figure 3 and 4. CODE is one of the five International GPS Service (IGS) ionosphere analysis centres that have supplied global TEC maps since June 1, 1998 on a regular basis to the Crustal Dynamics Data Information Systems (CDDIS), a global data centre of IGS. The overall consistency in terms of RMS values between the global TEC maps derived from different analysis centres is reported to be at the level of 3–5 TECU although the inconsistency could reach up to 7–10 TECU between some centres. Provided in Table 1 are the daily RMS values of the TEC estimate differences between UofC and CODE estimates as well as the RMS value over the consecutive five days. The TEC estimates in the central region of the ground SWEPOS network (56°N and 59°N) are very consistent to the estimates derived from the CODE global TEC maps. The overall agreement was at the level 2.4–3.0 TECU. On the other hand, the consistency has been degraded for grids off the central latitude region (53°N and 62°N) to a level of up to 4.6 TECU. This is due to the fact that the ionosphere measurement density in the central region is higher than other boundary regions over a regional network.

![VTEC at Geomagnetic Latitude 59°N](image)

![VTEC at Geomagnetic Latitude 56°N](image)

Tab. 1 RMS Values of TEC Estimate Differences (TECU)

<table>
<thead>
<tr>
<th>Grid Latitude</th>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
<th>Day 4</th>
<th>Day 5</th>
<th>Day 1–5</th>
</tr>
</thead>
<tbody>
<tr>
<td>53°N</td>
<td>3.47</td>
<td>5.78</td>
<td>2.50</td>
<td>2.78</td>
<td>3.71</td>
<td>3.83</td>
</tr>
<tr>
<td>56°N</td>
<td>1.90</td>
<td>3.60</td>
<td>1.93</td>
<td>1.42</td>
<td>2.34</td>
<td>2.36</td>
</tr>
<tr>
<td>59°N</td>
<td>1.89</td>
<td>4.53</td>
<td>3.89</td>
<td>1.64</td>
<td>1.29</td>
<td>2.96</td>
</tr>
<tr>
<td>62°N</td>
<td>3.46</td>
<td>6.98</td>
<td>5.79</td>
<td>3.08</td>
<td>1.62</td>
<td>4.61</td>
</tr>
</tbody>
</table>
5.2 Results of 3D Tomographic Modeling

The performance analysis of 3D tomographic modeling was conducted using GPS data observed from six GPS reference stations within the Southern California Integrated GPS Network (SCIGN) on May 15, 2000. The SCIGN consists of 250 GPS stations and is primarily designed to monitor the crustal deformation and earthquake activities in southern California region (SCIGN, 2002). The geographical distribution of the selected six stations is shown in Figure 5. A total of 60 epochs of dual frequency GPS measurements from six stations at a data rate of 30 seconds were used for the construction of an ionospheric tomography model for the region. To assess the modeling accuracy, the 300 measurements, collected immediately after the 60 epochs, were used as prediction data points for comparison purpose.

Once the tomography model coefficients were obtained, the obtained model parameters are used to calculate the ionospheric delay for any line of sight between a GPS receiver and satellite. If the model derived ionospheric delays can be compared to its directly measured value from the dual frequency GPS observations, the 3D tomography ionospheric modeling accuracy can then be assessed. The comparison results are provided in Figure 6 where the differences indicated an agreement for the vertical ionospheric delays between the model-derived estimates and the direct measurements at the level of 1.2 TECU.

Shown in Figure 7 are the error estimates in percentage calculated using the following equation:

\[
RE = \frac{TEC_{direct} - TEC_{model}}{TEC_{direct}} \times 100\%
\]

The results indicated that the model-derived slant ionospheric delays have a mean error of 1.4\% with respect to the directly measured ionospheric delay values. Most relative errors varied between −4.0\% and +4.0\%. The 3D tomographic modeling therefore has produced much lower ionospheric modeling errors or better modeling accuracy than the 2D grid-based modeling method. It is also worth to mention that a “carrier phase leveling” procedure was not implemented during the 3D tomography ionospheric modeling computation. So it is expected that a further accuracy improvement can be obtained if the noise level of the code-derived ionospheric delay measurements is reduced via carrier phase smoothing.

6 Conclusions

Both two and three-dimensional ionospheric modeling methods have been described in this paper and their performance have been assessed based on data from two different regional GPS reference networks. The agreement for 2D grid-based modeling is at the level of several TECU when compared to an independent source while the agreement for 3D tomographic modeling is at the level of about one TECU compared to the directly measured ionosphere values from dual-frequency GPS.
observations. Based on the developed 3D tomographic ionospheric model, the slant ionospheric delays could be well recovered with a mean error of 1.4%, indicating a significant improvement of accuracy over the 2D grid-based modeling. More numerical analysis, however, is required to further investigate the performances of the 2D and 3D ionospheric models presented in this paper.

Acknowledgements

This research was supported by a strategic grant from the Natural Sciences and Engineering Research Council of Canada (NSERC). The Southern California Integrated GPS Network and its sponsors, the W.M. Keck Foundation, NASA, NSF, USGS, SCEC, are thanked for providing data used in this study.

References


Liao, X. (2000). Carrier Phase Based Ionospheric Recovery Over A Regional Area GPS Network, UCGE Reports, Number 20143, University of Calgary, Calgary, Alberta, Canada, pp120.


