Ionosphere Modelling using Spherical harmonics with in-equality constraints over IRNSS service area

Gurpreet Singh*(1), Megha Maheshwari(1) and Nirmala Srini(1)
(1) U R Rao Satellite Center, ISRO, Bangalore, India, e-mail: gurpreet.singhj29@gmail.com, megham@ursc.gov.in; nirmalas@ursc.gov.in

Abstract

The ionosphere is a part of the atmosphere composed of free electrons and ions. Due to its ionized character, it affects the propagation of the radio signals passing through it. Therefore, for a satellite navigation system, ionosphere error corrections are required to accurately determine the user's position. The Indian Regional Navigation Satellite System (IRNSS) is an independent, regional navigation system meant to provide user position for single frequency as well as dual frequency users. In order to reduce errors in the user position, IRNSS provides two types of ionospheric corrections: Grid based as well as ionosphere coefficient based on L5 frequency. Users are advised to use grid based correction as compared to the coefficient based correction since grid based corrections are more accurate than coefficient based. In IRNSS, kriging method is applied to generate an ionosphere grid at IRNSS service area. However, spherical harmonics based ionosphere grid data is also frequently used by researchers worldwide. These geo-statistical methods (kriging and spherical harmonics) are equally good for the reconstruction of the global ionosphere data. However, the superiority of any of the geo-statistical methods over the other geo-statistical methods for the reconstruction of regional ionosphere data has not been explored enough. Therefore, in this paper, we reconstructed the ionosphere map over the IRNSS service area using a spherical harmonics method. Since the IRNSS service area lies in the ionosphere equatorial anomaly region, we directly implement non-negative physical constraints of the ionosphere in our algorithm. Mathematically, an inequality-constrained least squares method is proposed by imposing non-negative inequality constraints in the areas to avoid negative vertical Total Electron Content (TEC) values. The proposed technique utilizes a priori information from the Global Ionosphere Maps (GIM) of the Center for Orbit Determination in Europe (CODE). The GIM map has been degraded randomly to compare our techniques for reconstruction of TEC maps.

1. Introduction

In the last decade, the modelling of the ionosphere using the NAVigation with Indian Constellation (NavIC)/IRNSS has been thoroughly explored, and the study is still ongoing. NavIC is a regional navigation system that operates at L5 (1176.45 MHz) and S-band (2492.028 MHz) with the aim of providing precise and accurate navigation and timing solutions within 1500 km of the Indian boundaries. However, various error sources limit the position. The ionospheric irregularities present in the Earth's ionosphere are considered a significant source of error, which degrades the receiver performance or even results in loss of lock[10]. This phenomenon turns out to be a significant concern for a variety of navigational communications technologies. It is difficult to predict the ionosphere due to its high temporal and spatial variability. However, navigation satellites and the geodesic distribution of reference stations on the ground permit the processing of signals along many pathways. Since the ionospheric delay is the main source of inaccuracy on the route along which the navigation signal is delivered, it lowers the precision of the user's location and the navigation data. The contemporary study of ionospheric models is for detailed research and analysis of features of the ionosphere through the combination of various observations measured by instruments such as ionosondes and incoherent scatter radar, as compared to previous models that were developed as a means to reduce the delay error of navigation signals and improve the accuracy of the user's position.

There are three primary classes of ionospheric models (a) physical, (b) empirical, and (c) mathematical. To capture the physical ionospheric variations over space and time, physical models such as the Global Assimilative Ionospheric Model [2] have been developed. The NeQuick [5] and International Reference Ionosphere [7] are empirical models based on the relationships between solar radiation and the spatial-temporal changes of ions and electrons in the ionosphere. Lastly, the mathematical models that are the subject of this work are generated by analyzing measurements (TEC values) collected from sources such as navigational satellites and receivers at reference stations, thus estimating the electron content analytically. A Total Electron Content Map (TEC MAP) is produced by solving an inversion to determine the electron content at each grid point. Single frequency navigation receivers on the ground need ionosphere TEC information to eliminate the ionosphere's impact in order to attain position, navigation, and timing (PNT) accuracy. However, from an ionospheric point of view, the NavIC is a unique regional system with a service area covering the ionospheric equatorial anomaly region(IEA). Compared to the rest of the world, the ionosphere in the IEA area is very active and unpredictable, which can severely degrade the navigation signals or even lead to loss of lock [9], thus hindering the process of reconstruction of the TEC map over the IRNSS service area.

There are several techniques for the reconstruction of TEC maps. The geo-statistical interpolation known as Kriging was first used for reconstruction[4]. The approach by Schaer in 1999 [3] models the global TEC map using an expansion of spherical harmonic (SH) functions whose coefficients are determined by fitting the observed TEC data using the least squares algorithm. Additional constraints may be introduced to the fitting process to enhance the resultant model, for instance by adding an inequality constraint to exclude negative TEC values[4]. The benefit of the spherical harmonic expansion approach is that large-capacity data may be represented by a small number of coefficients. The downside is that information that fluctuates dramatically over a relatively short period of time is discarded. Although the Kriging method, which is an interpolation technique in which distance-dependent weights are applied, is excessively sensitive to measurements, it has the benefit of being easily applicable to the rapid detection of significant change[9, 10]. The current objective of this work is to provide a spherical harmonic based reconstruction of ionospheric maps with missing values and, in addition, apply the inequality constraint to eliminate any negative TEC values.

2. Methodology

To recreate the global TEC map, we use a least-squares fit to the observed TEC data values to calculate the coefficients of spherical harmonic (SH) functions and expand those functions using the optimized coefficients. Due to the location of the IRNSS service area inside the ionosphere's equatorial anomaly zone, the ionosphere's non-negative physical restrictions have been directly included in our method. An inequality-constrained least-squares approach uses non-negative inequality restrictions in the affected regions, which leads to constrained least squares, which in turn is equivalent to the linear complementarity problem [1]. To carry out the analysis for our proposed method, we have randomly eliminated the TEC values for the given GIM data. Then, We compared our algorithm to the original TEC Map, which had no missing values.

2.1. Mathematical Background

The spherical harmonic functions arise in physical applications as the solutions of Laplace's equation in spherical coordinates. As the angular component of the solutions to Laplace's equation in spherical coordinates, spherical harmonic functions form a complete orthogonal basis and can therefore approximate a smooth surface function in the following form:

$$f(\theta,\ \phi) \ \approx \ \sum_{l=0}^{l_{max}} \sum_{m=-l}^{l} a_{lm} \ \ Y_{lm} \ (\theta,\ \phi) \ \ , \tag{1}$$

where θ and ϕ are elevation and azimuth angles in the spherical coordinates. Y_{lm} denotes a spherical harmonic function with degree m and order l ($|m| \leq l$) and a_{lm} is the corresponding coefficient. Just like the Fourier series, the expansion becomes a precise description of the function as the maximum order l_{max} approaches infinity.

By examining the global TEC distribution at a particular moment as a function of latitude and longitude, the spherical harmonic expansion has been used as an approximation of the whole TEC map[6]. Each single measurement of TEC at a location (θ_i, ϕ_i) will correspond to a linear equation :

$$f(\theta_i, \phi_i) \ = \ TEC_i \ = \ \sum_{l=0}^{l_{max}} \sum_{m=-l}^{l} a_{lm} \ \ Y_{lm} \ (\theta_i, \phi_i) \tag{2}$$

By solving the above system of linear equations formed using all the available measurements on a TEC map, a set of coefficients are obtained for a given order, l_{max}, and the resulting expansion over our chosen grid gives the least square approximation of the global TEC map (Figure 1). In addition, θ means the geocentric latitude and ϕ is the sun-fixed longitude of the point where the ionosphere penetrates.

2.2. Coefficient estimation using least square

The system in equation (2) can be represented in matrix notation by taking D measurements d_i at (θ_i, ϕ_i). If spherical harmonics are considered up to order l_{max} , then there will be P = (l_{max} +1)^2 coefficients (a_{lm}'s) to estimate. This can be ordered into one unknown vector $[x_i]_{i=1,.P.}$ and the spherical harmonic coefficients as $[s_i]_{i=1,.P.}$ .The equations of condition (3) below show the relationship of coefficients and harmonics with observed data. Usually, we have more measurements than the number of coefficients, i.e., D > P.

$$x_1 s_1(\theta_1, \phi_1) + \ldots x_i s_i(\theta_1, \phi_1) + \ldots.. + x_1 s_i(\theta_1, \phi_1) \ = d_1$$

$$x_1 s_1(\theta_D, \phi_D) + \ldots x_i s_i(\theta_D, \phi_D) + \ldots.. + x_P s_P(\theta_D, \phi_D) = d_D \tag{3}$$

In matrix form

$$G\mathbf{x} = \mathbf{d}, \quad (4)$$

where G is a DxP matrix, **x** is a vector of length P, and **d** is a vector of length D.

The least square solution of (4) by minimizing the squared error, i.e., $||G\mathbf{x} - d||^2$ is given as follows

$$\mathbf{x} = (G^TG)^{-1}G^T\mathbf{d} \quad (5)$$

These estimated coefficients are then used to reconstruct the TEC map on any choice of grid over the globe. The TEC values are computed by substituting the estimated coefficients in equation (1). This technique of reconstructing ionospheric TEC maps is called spherical harmonic fitting.

2.3. Inequality constrained least squares (ICLS) for negative TEC values.

In the previous section, the spherical fit to distributed TEC values was addressed. Due to our region of interest, i.e., the IRNSS service area, which is an ionospheric equatorial anomaly (IEA) region, there are more constraints that have to be applied while solving equation(4). For a particular physical measuring system, previous information about the physical quantities to be obtained via measurements and estimates is often available. In the case of an ionospheric inversion, estimated TEC values have a high probability of being negative in the IEA region. Such prior limitations should be explicitly included in the estimating procedure in order to ensure that predicted outcomes have meaningful physical significance. Inequality constraint adjustment of least squares is a well-known method in geodesy for addressing the negative value issue of negative TEC values.

The principle of ICLS begins with imposing an extra constraint on the least square problem in equation(4).

$$\mathbf{d} = G\mathbf{x} + \epsilon,\ \ E(d) = G\mathbf{x},\ Cov(\mathbf{d}) = Cov(\epsilon) = P^{-1}\sigma^2 \quad (6a)$$

subject to the following constraints: $A\mathbf{d} \geq \mathbf{c}$ (6b)

where **d** is the Dx1 vector of observations, G is the DxP design matrix, and A is the DxP coefficient matrix.

The constraint can be converted from inequality to equality by putting a non-negative slack variable $\mathbf{h} \geq 0$ in equation (6b), which becomes

$$A\mathbf{d} - \mathbf{h} = \mathbf{c} \quad (7)$$

The problem of least square transforms into weighted least square under the constraint (7), we formulate extended objective function for this new problem:

$$Min:\ E(\mathbf{d},\mathbf{x},\mathbf{q},\mathbf{h}) = \frac{1}{2}(\mathbf{d} - G\mathbf{x})^TP(\mathbf{d} - G\mathbf{x}) - \mathbf{q}^T(A\mathbf{d} - \mathbf{c} - \mathbf{h}) \quad (8)$$

Both h and q are non-zero vectors and with a dot product equal to zero, which plays a key role in separating the active inequality from the non-active one. [ref]. To minimize the new objective function $E(\mathbf{d},\mathbf{x},\mathbf{q},\mathbf{h})$, we differentiate equation (8) and we get $\underline{\mathbf{d}}_{ICLS}$ as an estimator of d as follows:

$$\begin{aligned}\underline{\mathbf{d}}_{ICLS} &= (G^TPG)^{-1}(G^TP\mathbf{d} + A^Tq)\\ &= N^{-1}G^TP\mathbf{d} + N^{-1}A^T\mathbf{q}\\ &= \underline{\mathbf{d}}_0 + N^{-1}A^T\mathbf{q} \quad (9)\end{aligned}$$

where $N = G^TPG$ and $\underline{\mathbf{d}}_0 = N^{-1}G^TP\mathbf{d}$ is the solution of the weighted least square problem, and now $\underline{\mathbf{d}}_{ICLS}$ solution will be substituted in equation(7), which leads to the following expressions:

$$\mathbf{h} = M\mathbf{q} + \mathbf{w} \quad (10)$$

where h , q are satisfying non-zero constraints along with $\mathbf{q}^T\mathbf{h} = 0$, $M = AN^{-1}A^T$ and $w = A\underline{\mathbf{d}}_0 - \mathbf{c}$. The equation(10) is very well known as the linear complementarity problem (LCP), and the solution to this problem exists and is unique[1,8].

The Implementation Details for ICLS procedure are as follows:

1. First step is to carry out the fitting of TEC values using least squares discussed in section 2.2., which provides $\underline{\mathbf{d}}_0 = N^{-1}G^TP\mathbf{d}$
2. Then forming the matrix A in equation (6b) by identifying the grid point where TEC values have negative values. Matrix A is also a spherical harmonic functions matrix like G but it contains only those elements of G where TEC values are negative using $\underline{\mathbf{d}}_0$ from step-1 for estimation.
3. After getting matrix A, the equation (10) is calculated which leads to LCP.
4. Solution of this LCP gives the final $\underline{\mathbf{d}}_{ICLS}$ which is an optimized solution.
5. Repeat Until there are no more negative TEC values.

3. Results and Discussions

The method developed in section 2 is applied using the GIM-CODE data for the day:02-JUN-2022. Figure 1 is a contour map showing the original TEC values across the globe. The missing values have been introduced randomly to verify the consistency and stability of our implemented method and Figure 2 shows the TEC map with 20% missing values which is an attempt to simulate the conditions similar to the IEA region. Figure 3 is a reconstructed TEC map using equation (5) and (2). To further demonstrate the elimination of negative TEC values, we apply equation (8) and (10) which resulted in Figure 4. The results here show over-smoothening the variations in original TEC values, however these techniques of reconstruction will greatly help in more difficult real scenarios such as over the IEA region. The average root mean square error in spherical harmonic fitting without inequality constraints 2.14 TECU and with inequality 2.18 TECU.

4. Figures

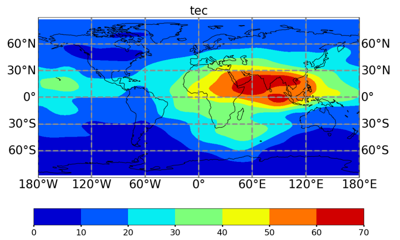

Figure 1. Original GIM-TEC map from CODE-GIM on 02-JUN-2022 at 11:00:00

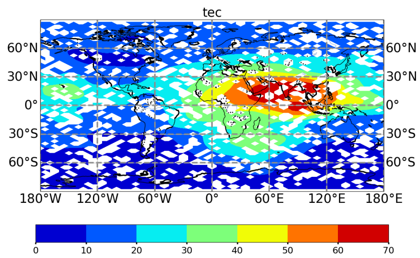

Figure 2. GIM-TEC map with nearly 20% missing values introduced randomly.

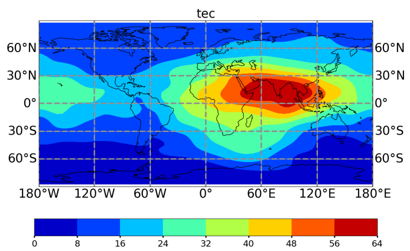

Figure 3. Reconstructed TEC map using spherical harmonic approximation.

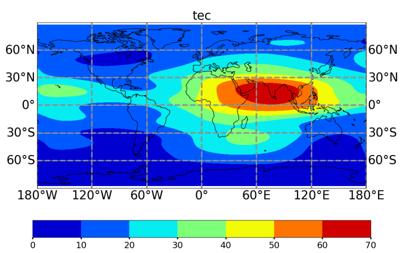

Figure 4. Reconstructed TEC map using spherical harmonic with inequality constrained least square.

5. Conclusions

Applying spherical harmonics functions with the degree and order to 13 using least squares(LS) fitting, we reconstructed the local ionospheric maps in this paper. The weighted LS method often results in negative VTEC values in GIM Map reconstruction derived with SH functions due to the uneven distribution in the IEA region. Since VTEC cannot be negative physically, we have directly implemented this physical non-negative constraint to reconstruct GIM models. Spherical Harmonic approximation with or without inequality constraint approximates the original TEC maps

6. Acknowledgements

The authors would like to thank the Director, U.R. Rao Satellite Center (URSC), ISRO, Bangalore; Deputy Director, Communication and Power Area, URSC, ISRO; and Group Director, Space Navigation Group, URSC, ISRO for their guidance, support, help, and encouragement. Valuable suggestions given by Space Navigation Group members are also acknowledged. Here, data from https://cddis.nasa.gov/ CODE-GIM is also acknowledged.

7. References

1.Liew,C.K., Inequality constrained least-squares estimation. Journal of the American Statistical Association 71 (355), 746–751, 1976 b.

2.Schunk R, A mathematical model of the middle and high latitude ionosphere. Pure Appl. Geophys,1988, 128, 255.

3.Stefan Schaer,"Mapping and predicting the earth's ionosphere using the Global Positioning System". Dissertation of Astronomical Institute, 1999page 205,

4. Hernández-Pajares,Oru´ s, R., M., Juan, J.M., Sanz, J. Improvement of global ionospheric VTEC maps by using Kriging interpolation technique. Journal of Atmospheric and Solar-Terrestrial Physics 67 (16), 2005,1598–1609,.

5.Radicella S . The NeQuick model genesis, uses and evolution. ANNALS OF GEOPHYSICS, 52, (2009)417 - 422.

6.Byung‑Kyu Choi, Woo Kyung Lee, Sung‑Ki Cho, Jong‑
Uk Park, Pil‑Ho Park, "Global GPS Ionospheric modeling using spherical harmonics expansion approach", Journal of Astronomy and Space Sciences, 2010, pp. 359‑366

7.Bilitza D, McKinnell L, Reinisch B Fuller-Rowell T . The international reference ionosphere today and in the future. (2011) J. Geod., 85, 909 - 920

8. Hongping Zhang, Peiliang Xu, Wenhui Han, Maorong Ge, Chuang Shi, "Eliminating negative VTEC in global ionosphere maps using inequality-constrained least squares,"Advances in Space Research,Volume 51, Issue 6, 2013,Pages 988-1000, ISSN 0273-1177

9. M. Maheshwari and S. Nirmala, "IRNSS ionosphere modelling for single frequency receiver using data ingestion technique",2018,IEEE 4th International Conferencefor Convergence in Technology (I2CT), 1-5,.

10.Ercha Aa, Shasha Zou, Aaron Ridley,et al "Merging of Storm Time Midlatitude Traveling Ionospheric Disturbances and Equatorial Plasma Bubbles," February 2019, Space Weather, 17(2):285–298