Using Neural Network to Construct Regional Reference Total Electron Content Model

Takashi Maruyama

National Institute of Information and Communications Technology
2-1 Nukuikita 4-chome, Koganei, Tokyo, 184-8795 Japan
(TEL: +81-42-327-7512, FAX: +81-42-327-6163, e-mail: tmaru@nict.go.jp)

Abstract

A regional reference model of total electron content (TEC) was constructed using data from the GPS Earth Observation Network (GEONET), which consists of more than 1000 Global Positioning System (GPS) satellite receivers distributed over Japan. The data covered almost one solar activity period from April 1997 to June 2007. First, TECs were determined for 32 grid points expanding from 27 to 45°N in latitude and 127 to 145°E in longitude at 15-minute intervals. Secondly, the time-latitude variation of TEC was determined by using the surface harmonic functional expansion. The coefficients of the expansion were then modeled by using a neural network technique with input parameters of the season (day of the year) and solar activity (F10.7 index and sunspot number). Thus, two-dimensional TEC maps (time vs. latitude) can be obtained for any given set of solar activity and day.

1. Introduction

One of the important effects of the ionosphere on radio waves is a propagation delay in the ionosphere. This delay depends on the frequency and the total electron content (TEC) along the propagation path. Several attempts have been made to specify the ionospheric electron density using theoretical and empirical approaches. Considerable effort has resulted in the continuous development and improvement of the International Reference Ionosphere (IRI) [1], which describes the density at various heights for any specified geophysical conditions, based on long-term observations. TEC can be derived by integrating the height profile of the electron density. The contribution of the plasmaspheric electron density to TEC cannot be neglected. However, observations of the electron density to construct a reliable model are lacking in the topside ionosphere and plasmasphere compared with the bottomside and the F-layer peak [2].

Direct measurements of TEC using radio waves transmitted from the Global Positioning System (GPS) satellites have been collected in this decade. Thus, GPS-based TEC data are now available to construct empirical models of TEC. Meanwhile, artificial neural network (NN) techniques have been applied to a variety of topics in the study of upper atmosphere. Multilayer feed-forward networks [3] are used to specify the ionosphere by approximating a relationship between geophysical conditions (seasons, solar activities, local times, longitude/ latitude etc.) and observed ionospheric parameters (foF2, h'F2, hmF2 etc.) [4-6]. Because of the input-output mapping features of NNs, they could be used to generate reference ionospheric models for possible incorporation into the IRI. For this purpose, a so-called training data set must cover a whole range of possible input parameter variations, say, a data period longer than one solar cycle.

In Japan, a dense GPS receiver network, GEONET (GPS Earth Observation Network) has been developed, and data from more than 1000 locations have been available since April 1997, close to a solar minimum. An algorithm that simultaneously determines satellite/receiver biases and vertical TEC using GEONET data has been developed [7]. By using this algorithm, we constructed a TEC database that nearly covered one solar cycle from 1997 to 2007. This paper describes an empirical model of TEC variations based on this database.

2. Data Set and Methodology

About 300 GEONET receivers were chosen for this study to ensure uniform coverage over Japan. The major issues in deriving TECs from GPS radio signal observations are the instrumental biases both in the satellites and receivers and the conversion process from the observed TECs along the slant path to the vertical ones. In this paper, the slant TECs were converted to vertical TECs (vTEC) at the piercing point where the ray path crossed a shell at a height of 400 km (thin shell model). We calculated the daily instrumental biases and quarter-hourly values of vTEC

at the 32 grid points, 2° increments in longitude and latitude, shown in Figure 1, using the least square fitting method for a data set that covered 24 hours. More details of the method are described elsewhere [7]. The vertical TEC obtained in this way is referred to as the grid TEC (gTEC).

The major factors that determine the TEC are the solar activity, season, local time, and geographic and geomagnetic coordinates. Our process to generate a model incorporating these factors consists of three steps: Step 1 is the procedure described in the previous paragraph in which the gTECs at the grid points are determined at 15-minute intervals. Step 2 is the procedure where variations in time and latitude are expressed as a two-dimensional distribution map for three consecutive days. Because the data grid is limited to a narrow longitudinal extent and geomagnetic conditions do not change greatly among the east-west aligned grids, the longitudinal dependence is assumed to be equivalent to the local mean time (LMT) in this step. Step 3 is the procedure where the solar activity and seasonal changes of the TEC map are modeled by using a neural network technique.

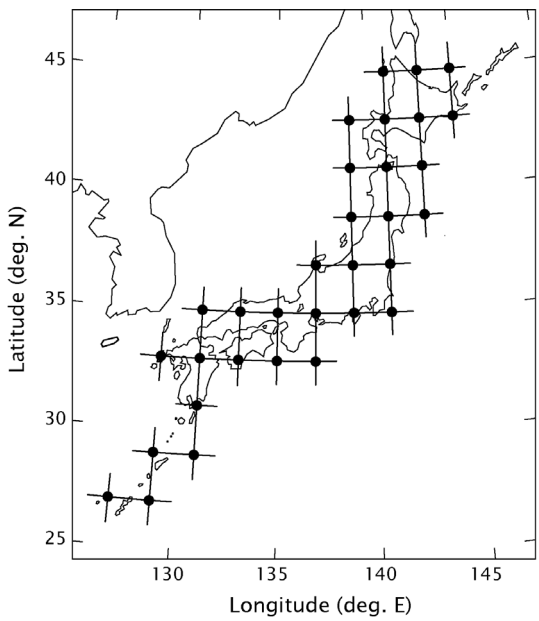

Figure 1. 32 grid points on which TEC values were determined

To generate TEC maps from the gTECs, we used the surface harmonic expansion method based on the associated Legendre's function as shown in (1) taking LMT (hour) at each grid point as the azimuth parameter, $\varphi = 2\pi(\mathrm{LMT}/24)$

$$\mathrm{TEC} = \sum_{m=0}^{M} \sum_{n=m}^{N} (A_{nm} \cos m\phi + B_{nm} \sin m\phi) P_n^m(\cos\theta) \tag{1}$$

Where θ is the colatitude, we took both degree and order up to 7 (N=M=7). As the grid point distributes at only northern mid-latitudes, dummy data were set, for mathematical convenience, in the southern hemisphere as a mirror image of the northern hemisphere with respect to the equator. The functional fitting was performed to determine coefficients A_{nm} and B_{nm} in (1). As dummy data were set in the southern hemisphere, the whole global distribution map is symmetrical with respect to the equator (resultant map data outside the latitude range from 29 to 45° N were disregarded). In other words, A_{nm} and B_{nm} with the odd number of n+m are equal to zero. Thus, a total of 36 target parameters needed to be determined.

The solar activity expressed by two proxies, the F10.7 solar flux and the sunspot number, R, for the whole data period is shown in Figure 2. The figure shows that both proxy parameters vary in a similar way but are not exactly the same. Thus, both parameters were included in the input parameters of the network.

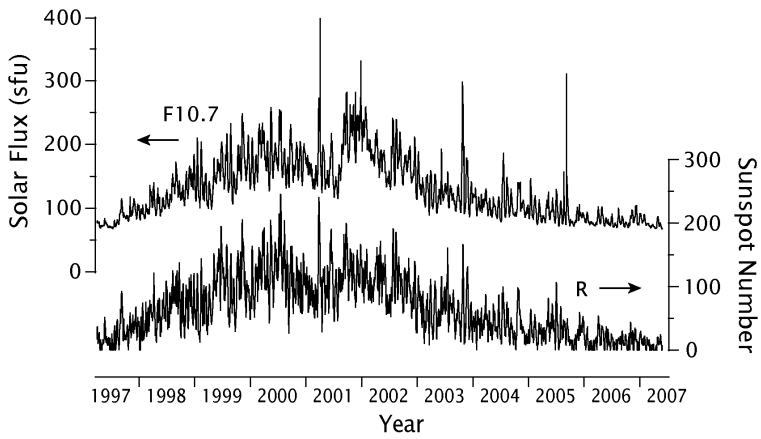

Figure 2. Daily solar flux and sunspot number

We adopted the multilayer feed-forward network [3] that consisted of the input layer, one hidden layer, and

the output layer. The schematic diagram of the network is shown in Figure 3. The input layer had 8 nodes for solar activity ($F10.7$ and R averaged over three days, a week, and three solar rotations (81 days), including the days in which the TEC was specified and prior to those days) and season (sin and cos components of day of the year). The number of nodes in the hidden layer was chosen to be 200. The output parameters were the 36 Legendre's coefficients. For the first several tenths of epochs in the back-propagation learning process, weight updating was performed by the pattern mode in which weights were updated after the presentation of each training example [3]. The order of the presentation of training examples was randomized from one epoch to the next. After the weights were coarsely determined, weight updating was continued by the batch mode in which weights were updated after the presentation of all the training examples that constituted an epoch [3].

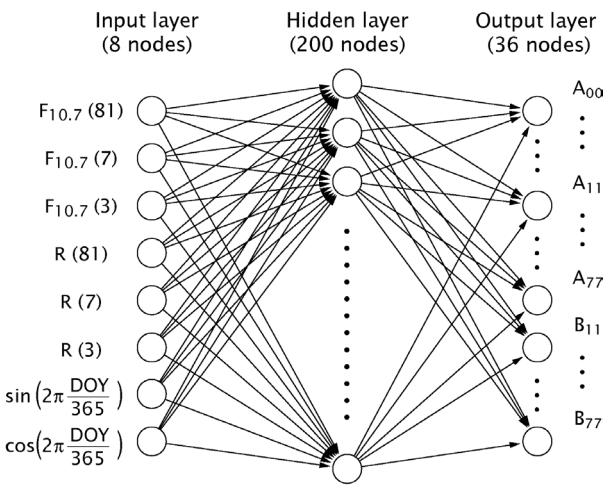

Figure 3. Schematic diagram of NN. For solar activity indices, [n] denotes an average over n days.

3. Results

To evaluate the performance of the neural network mapping, we ran the network learning for the data set, excluding a partial data set for 2003. After the learning was completed, the network outputs were compared with observations for 2003. Figure 4 demonstrates the performance of the network for the TECs at noon at grid point 14 (35°N, 137°E), near central Japan. The top panel is the solar activity inputs, $F10.7$ and R, averaged over three days. The middle panel is the grid TECs (circles connected with thin line) and network outputs (thick solid line). The bottom panel compares the grid TECs and IRI outputs with STORM option (thick solid line). Not only seasonal variations, but also solar activity dependences are reproduced by the network. Interestingly, at the end of October, the solar flux was quite large, but the grid TEC did not increase very much. This was well reproduced by the network. When the solar activity indices averaged over three solar rotations were not incorporated into the input parameter, this moderate increase in TEC against the extremely intense solar activity at the end of October was not successfully reproduced. On the other hand, the enhanced grid TEC found in the latter half of May was not reproduced by the network. In this period, the solar activity was not so high as compared with the solar activity in the solar rotation before and after this period. This suggests that the solar activity indices we used are not entirely accurate proxies of solar EUV flux.

4. Summary

A large amount of GPS-derived total electron content data have been collected in this decade covering almost one solar cycle, which now allows an empirical model of TEC to be constructed. We constructed a regional reference TEC model over Japan based on the dense GPS receiver network, GEONET. The process consisted of three steps: (1) determining vertical TECs at grid points separated by 2° in latitude and longitude (gTECs), (2) approximating by using surface harmonic functional fitting (time-latitude maps), and (3) using neural network mapping to relate the solar activity and season with the pattern of the time-latitude map. In the first step,

instrumental biases were simultaneously determined and vTECs were averaged in 2×2° longitude/latitude cells. Averaging and smoothing were also performed in step 2 by using limited degree and order of surface harmonic function to approximate gTEC over three days. Step 3 successfully worked to separate the solar activity and seasonal dependences of the TEC distribution pattern with respect to time and latitude.

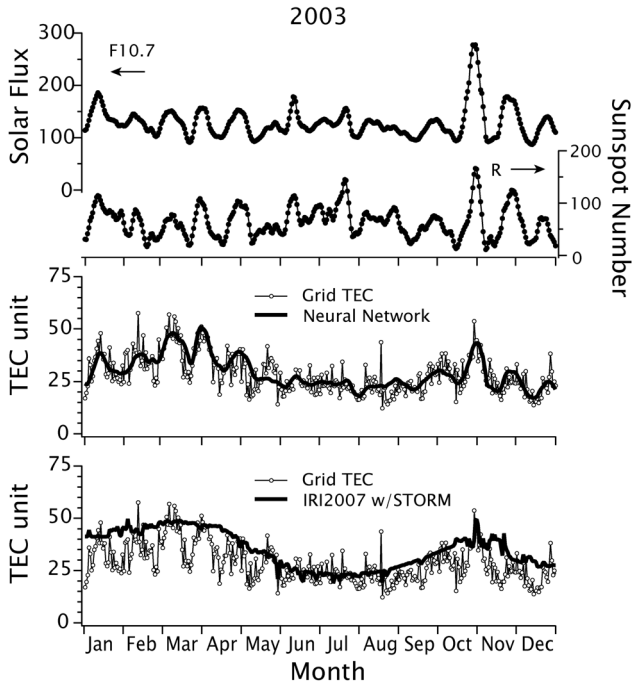

Figure 4. Solar activity averaged over three days (top panel), comparison between gTEC from GEONET observations and NN outputs (middle panel), and comparison between gTEC from GEONET observations and IRI outputs (bottom panel)

5. References

1. D. Bilitza, "International Reference Ionosphere 2000," *Radio Sci.*, **36**, 2001, pp. 261-275.

2. D. Bilitza and R. Williamson, "Towards a better representation of the IRI topside based on ISIS and Alouette data," *Adv. Space Res.*, **25(1)**, 2000, pp. 149-152.

3. S. Haykin, *Neural networks - A comprehensive foundation*, Macmillan College Publishing Company, Inc., 1994.

4. L. A. Williscroft and A. W. V. Poole, "Neural networks, foF2, sunspot number and magnetic activity," *Geophys. Res. Lett.*, **23**, 1996, pp. 3659-3662.

5. L. A. McKinnell and A. W. V. Poole, "Predicting the ionospheric F layer using neural networks," *J. Geophys. Res.*, **109**, 2004, A08308, doi:10.1029/2004JA010445.

6. E. O. Oyeyemi, A. W. V. Poole, and L. A. McKinnell, "On the global model for foF2 using neural networks," *Radio Sci.*, **40**, 2005, RS6011, doi:10.1029/2004RS003223.

7. G. Ma and T. Maruyama, "Derivation of TEC and estimation of instrumental biases from GEONET in Japan," *Ann. Geophys.*, **21**, 2003, pp. 2083-2093.