Pertanika J. Sci. & Technol. 25 (S): 19 - 28 (2017)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Total Electron Content Forecasting using Artificial Neural Network

**Rohaida Mat Akir[1,3]*, Kalaivani Chellappan[1,2] and Mardina Abdullah[1,2]**

[1]*Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia*
[2]*Space Science Centre (ANGKASA), Institute of Climate Change, Universiti Kebangsaan Malaysia (UKM), 43600 UKM Bangi, Selangor, Malaysia*
[3]*Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia (UTHM), 86400 Parit Raja, Batu Pahat, Johor, Malaysia*

## ABSTRACT

Space weather forecasting and its importance for the power and communication industry have inspired research related to TEC forecasting lately. Research has attempted to establish an empirical model approach for TEC prediction. In this paper, artificial neural networks (ANNs) have been applied in total electron content using GPS Ionospheric Scintillation and TEC Monitor (GISTM) data from UKM Station. The TEC prediction will be useful in improving the quality of current GNSS applications, such as in automobiles, road mapping, location-based advertising, personal navigation or logistics. Hence, a neural network model was designed with relevant features and customised parameters. Various types of input data and data representations from the ionospheric activity were used for the chosen network structure, which was a three-layer perceptron trained by feed forward back propagation method and tested on the chosen test data. We found that the optimum RMSE occurred with 10 nodes as the best NN for GISTM UKM station for the studied period with RMSE 1.3457 TECU. An analysis was made to compare the TEC from the measured TEC with neural network prediction and from IRI-corr model. The results showed that the NN model forecast the TEC values close to the measured TEC values with 9.96% of relative error. Thus, the forecasting of total electron content has the potential to be implemented successfully with larger data set from multi-centred environment.

*Keywords:* Forecasting, GPS, IRI-corr model, Neural network, Total electron content

ARTICLE INFO

*Article history:*
Received: 15 August 2016
Accepted: 18 May 2017

*E-mail addresses:*
rohaida261@gmail.com (Rohaida Mat Akir),
kckalai@ukm.edu.my (Kalaivani Chellappan),
mardina@ukm.edu.my (Mardina Abdullah),
*Corresponding Author

ISSN: 0128-7680 

## INTRODUCTION

Ionosphere is part of the upper atmosphere that can affect the propagation of radio waves. Total Electron Content (TEC), defined as the total number of electrons along the signal path between GPS satellite and the receiver, describes the ionospheric ionisation content. A unit of TEC $=10^{16}$ electron/$m^2$. (Habarulema, McKinnell, Cilliers, & Opperman, 2009), and it is one of the physical parameters that can be derived from GPS data and provides an indication of ionospheric variability against climate, weather and meteorology.

The TEC data derived from global positioning system (GPS) have been collected to construct empirical models. GPS signal is mostly affected by the ionosphere during travel from the GPS satellite to the receiver on the ground. The IRI is a widely used global empirical ionospheric model for approximation or prediction of ionospheric parameters (Wichaipanich & Supnithi, 2014). It provides physical parameters such as the monthly average TEC as a function of height, location, local time and sunspot number for magnetically quiet conditions. However, due to the general scarcity of data from the Southern hemisphere, the IRI model does not give accurate predictions for Southeast Asia (Watthanasangmechai et al., 2012; Wichaipanich & Supnithi, 2014).

The neural network (NN) is an artificial intelligence method that can be used as a tool for solving problems in various fields such as control system, medicine, weather and meteorology, finance and ionospheric prediction. The ANN techniques have been used in the study of the upper atmosphere. Typically, NN is organised in three layers: input layer, one or several hidden layers, and an output layer. The ANN has been employed previously in modelling of TEC using GPS data over several locations and different periods of observation (Bagiya et al., 2009). As stated in (Wichaipanich & Supnithi, 2014), the feed forward with BP algorithm is widely used for modelling of ionospheric behaviour. A feed forward NN has a layer or subgroup of processing elements. This layer of processing elements can make independent computations on the data that it receives and passes the results to the next layer (Abhishek, Kumar, Ranjan, & Kumar, 2012).

Several studies have been carried out to forecast and predict the TEC and critical frequency (foF2) using neural network. In South Africa, the researcher employed the NN method to predict the parameter on foF2 and TEC and to determine the optimum parameter for modelling (Habarulema, McKinnell, & Cilliers, 2007; Habarulema et al., 2009). Nakamura, Maruyama and Shidama (2009) studied the ionospheric variation value of foF2 at mid-latitude over Kokubunji in Japan and China, presenting and forecasting foF2 up to five hours in advance (Chen, Sun, & Ban, 2010). Thai researchers (Watthanasangmechai et al., 2012; Wichaipanich & Supnithi, 2014) used NN to predict the foF2 and TEC values at the magnetic equator and compared the foF2 and TEC prediction using neural network model with IRI model in Thailand while Malaysian researchers proposed the prediction of hourly TEC value in ionospheric modelling with NN (Homam, 2014).

This study focuses on the application of ANN in forecasting the GPS TEC at Universiti Kebangsaan Malaysia (2°55' N, 101°46'E) using data from the GPS Ionospheric Scintillation and TEC Monitor (GISTM) receiver installed there. We also compared the neural network technique for GPS TEC with the IRI model and the measured TEC.

## MATERIALS AND METHODS

The study was divided into two phases. Phase 1 focused on data acquisition and preparation while predictive model development and testing against an established model were done in Phase 2.

### Data Collection

Data was obtained from the GPS Ionospheric Scintillation and TEC Monitor (GISTM) system installed in Universiti Kebangsaan Malaysia (UKM) as shown in Figure. 1. It consists of an antenna, a NovAtel OEM4 dual frequency Global Positioning System (GPS) receiver and computer for data storage. The GPS receiver system can track up to 11 GPS satellites at the L1 frequency (1575.42 MHz) and the L2 frequency (1227.6 MHz). The GPS data from 2011 was recorded as Universal Time (UT) system and TEC rates were logged every 60 seconds from all available satellites.

All the raw data from the GISTM receiver was analysed and sorted using MATLAB software. Missing data were replaced by not a number (NaN) in the MATLAB program. GPS TEC values at each hour were used in this paper, depending on the availability of data from the UKM station. The input parameters of the neural network were derived from the parameters known to influence TEC variability such as seasonal variation, diurnal variation, solar activity and magnetic activity. Solar activity and magnetic activity variation were represented by sunspot number (SSN) or solar flux F10.7 and magnetic A index values. These data can be freely downloaded from the NOAA website (ftp.ngdc.noaa.gov). Seasonal variation and diurnal variations were represented by a day number (DN) and hour (HR). Both DN and HR were divide into two cyclic components to allow numerical continuous trend of the data (Habarulema et al., 2009)

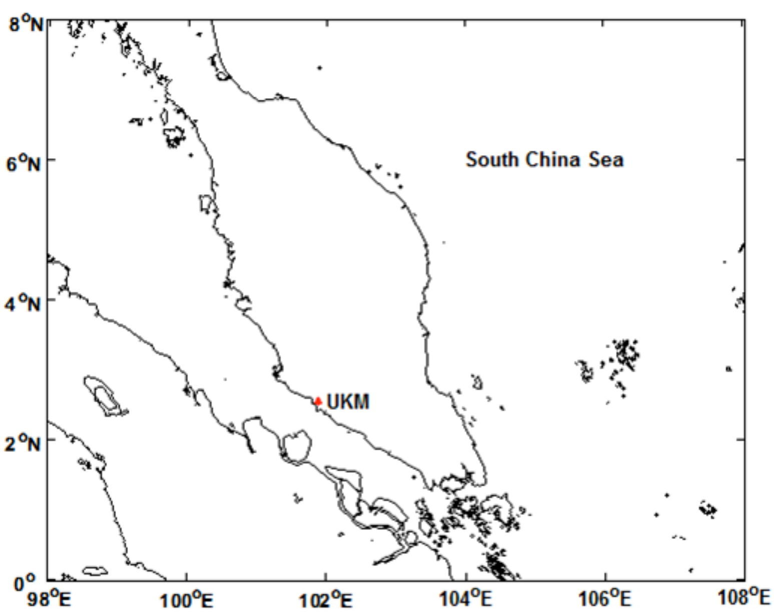


*Figure 1.* The GISTM station at UKM

$$DNS = sin\left(\frac{2\pi \times DN}{365}\right) \quad DNC = cos\left(\frac{2\pi \times DN}{365}\right) \tag{1}$$

$$HRS = sin\left(\frac{2\pi \times HR}{24}\right) \quad HRC = cos\left(\frac{2\pi \times HR}{24}\right) \tag{2}$$

Where, the sine and cosine component of the DN and HR are shown in both equations (1) and (2). In the NN model, the number of nodes in the hidden layer is important. If the number of nodes is too small, the result values might not be reliable, while if we were to use too large number of nodes, the computation time will be too long (Huang, Li, & Yuan, 2014). The default maximum values of iteration is 1000 iterations to stop the training process.

### Artificial Neural Network, ANN

The neural network is a mathematical model or computational model based on biological neural processing. The neural network has many different types, simple or complex structures.

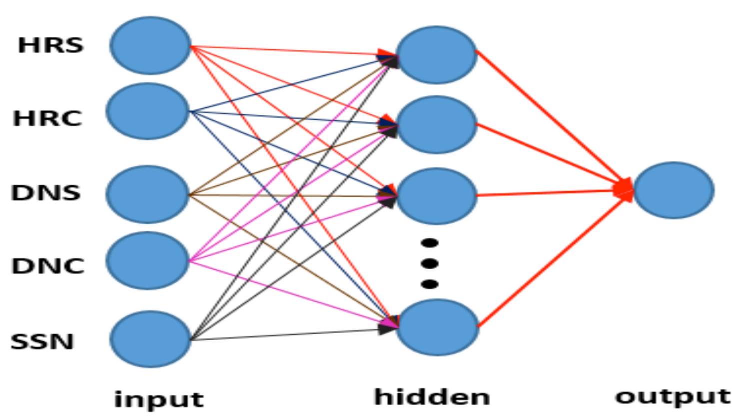


*Figure 2.* A schematic diagram of the inputs and output ANN for TEC prediction

From Figure 2, the NN consists of three layers. These layers are the input layer, the hidden layer and lastly the output layer. The hidden layer can be one or several layers. Information from the input layer will be transferred to and processed by the hidden layer, and then it continues until reaches the output layer. Every layer is connected to each other by weight. A basic structure known as a feed forward network with a back propagation was used in this model. The number of units in the hidden layer was determined on a trial and error basis. The learning procedure was the same as that described by (Haykin, 1999).

### Training and Testing Process

Each of the networks must be trained and tested to calculate its performance configuration before deciding on the best network. First, data for the input and output had to be normalised between -1 and 1. Then, the input data had to be trained using a MATLAB BP algorithm. The default algorithm in MATLAB takes only 70% of the input data for training and another 30%

for testing and validation. Every time the data is trained, the training data will be selected randomly from the whole data set.

The testing process evaluates the performance of the trained model. A validation set of data was used to evaluate the NN model by estimating its root mean square error (RMSE) as described in (Otsuka et al., 2002).

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(TEC_{pred} - TEC_{meas}\right)^2} \quad (3)$$

Where N is the number of data, $TEC_{pred}$ is the value of TEC from the NN prediction and $TEC_{meas}$ is the value from measured the TEC GISTM. In addition, the absolute error, |α| was computed as follows:

$$|\alpha| = |TEC_{pred} - TEC_{meas}| \quad (4)$$

The relative error, ε also can be determined as follows (Leandro & Santos, 2007). This process was repeated for each NN model.

$$\varepsilon = \frac{|\alpha|}{TEC} \times 100 \quad (5)$$

## RESULTS

The results presented here are for the GISTM UKM station. During the training of the data, only 70% of the input data was used. From the total of 576 data set, only 404 were used for training and these data were selected randomly from all the data set. The other 86 were used for validation while the rest were kept for the testing process. Based on previous studies, parameters that influence TEC variability are seasonal variation, diurnal variation and solar activity. These parameters are used as input parameters while the TEC data were used as the output parameter.

The number of nodes in the hidden layer was determined by choosing the lowest RMSE; 0.2 and 0.95 were chosen as the learning rate and momentum constant respectively. These values were kept constant for each model design at this stage in order to observe the effect of different numbers of neurons in the hidden layer. NN must be run 10 times in order to obtain the best results.

In Figure.3, the best values for RMSE of the NN with 7 to 13 nodes in the hidden layer are shown. The number of nodes in a single hidden layer was chosen from the lowest RMSE of NN. It was discovered that the optimum RMSE occurred with 10 nodes as the best NN for the GISTM UKM station for the studied period, with RMSE of 1.3457 TECU. Thus, we referred to 10 hidden nodes as the proposed NN.

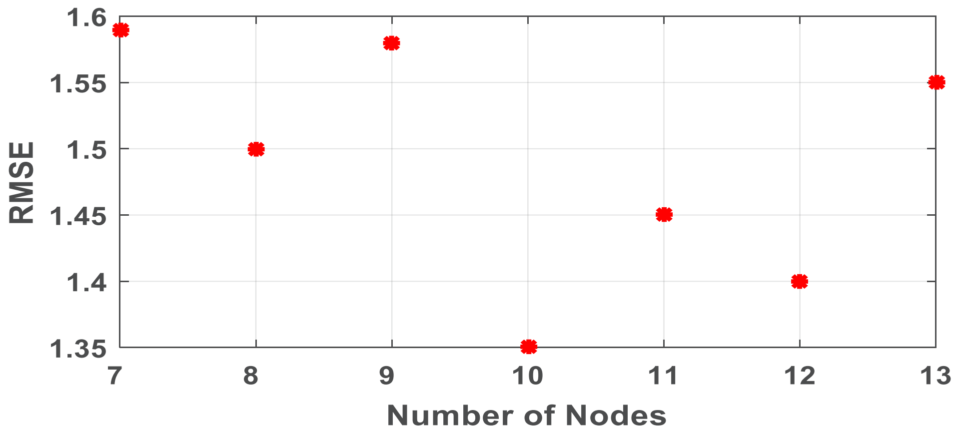


*Figure 3.* RMSE value computed for NN model

After choosing the best number of neurons in the hidden layer, the NN model was trained using both the training and evaluation set as the training data and then assessed by using testing data. In order to improve the performance of the NN model, learning rate (lr) and momentum constant (mc) were varied between 0 and 1. Table 1 and Table 2 show the RMSE, absolute error and relative error values of the changing value of the learning rate and momentum are constant. From the table, 0.2 was chosen as learning rate and 0.95 for momentum constant, as the best performance configuration for forecasting the GPS TEC. The optimum RMSE value obtained from this model was 1.3251 TECU. Table 3 shows the best properties for the NN model in order to forecast the TEC value. In the network configuration, the input layer consisted of five nodes corresponding with 5 input parameters, each with 576 data points, 10 nodes in the hidden layer and 1 node as the output layer which refers to the TEC value. In addition, the maximum iteration was 1000 iterations to stop the training process.

Table 1
*Performance of the momentum constant (mc)*

| Momentum constant (mc) | Learning rate (lr) | RMSE (TECU) | Absolute error (rmse) | Relative error (%) |
|---|---|---|---|---|
| 0.2 | 0.1 | 1.9292 | 1.5362 | 17.73 |
| | 0.2 | 1.5623 | 1.2127 | 15.28 |
| | 0.3 | 1.5355 | 1.1045 | 11.27 |
| | 0.4 | 1.9348 | 1.5163 | 14.79 |
| | 0.5 | 1.6371 | 1.2850 | 17.20 |
| | 0.6 | 1.8653 | 1.5681 | 17.58 |
| | 0.7 | 1.7413 | 1.3845 | 15.18 |
| | 0.8 | 1.5139 | 1.1671 | 11.29 |
| | 0.9 | 1.5456 | 1.1992 | 12.97 |
| | 1.0 | 1.4628 | 1.1368 | 9.96 |

Table 2
*Performance of the learning rate (lr)*

| Momentum constant (mc) | Learning rate (lr) | RMSE (TECU) | Absolute error (rmse) | Relative error (%) |
|---|---|---|---|---|
| 0.95 | 0.1 | 1.8754 | 1.4196 | 18.87 |
| | 0.2 | 1.3251 | 0.9735 | 9.96 |
| | 0.3 | 1.8819 | 1.4690 | 13.84 |
| | 0.4 | 1.8292 | 1.5005 | 18.71 |
| | 0.5 | 1.4665 | 1.2636 | 13.37 |
| | 0.6 | 1.8489 | 1.5342 | 16.51 |
| | 0.7 | 1.5531 | 1.2048 | 11.49 |
| | 0.8 | 1.7373 | 1.3732 | 15.59 |
| | 0.9 | 1.8855 | 1.4892 | 14.29 |
| | 1.0 | 1.5483 | 1.2244 | 15.78 |

Table 3
*Properties to develop the network for TEC forecasting*

| ANN Properties | Properties |
|---|---|
| Network Configuration | [5,10,1] |
| Transfer Function | Tansig, purelin |
| Training Function | trainlm |
| Learning rate (lr) | 0.2 |
| Momentum constant (mc) | 0.95 |
| Epochs | 1000 |

After the testing process was carried out, a graph was plotted between the measured VTEC and predicted VTEC so that a comparison can be made. Figure 4 shows both the measured TEC and predicted TEC hourly for the network configuration using [5, 10, 1] setting. From the graph, comparison between the measured TEC and predicted values show a high degree of similarity. While, the IRI-corr model overestimate the measured TEC for all hours of the day except during sunset to sunrise hours. This figure also shows that the predicted TEC model is close to the measured TEC value. The RMSE value for ANN is 1.3251 TECU and the relative error is 9.96%. While the RMSE value for IRI-corr model is 5.2301 TECU and the relative error is 40.24%. Therefore, it can be proven that neural network model is accurate compared with IRI-corr model.

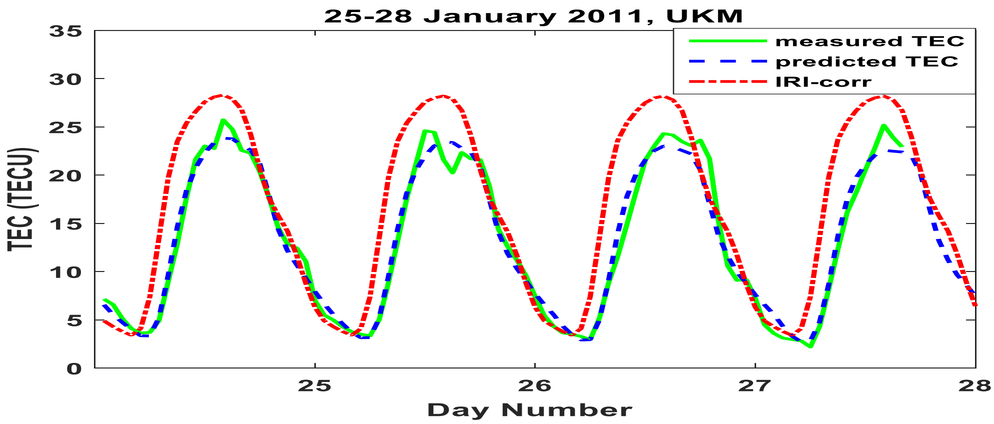

*Figure 4.* Measured and predicted TEC for 25-28 January 2011

## CONCLUSION

This paper presents an approach to forecast the TEC four days ahead based on GISTM data. In this study, feed forward back propagation network was developed. The values 0.2 and 0.95 were chosen as the optimum values of the momentum constant and learning rate. This result was obtained by testing all possible values of the momentum constant and learning rate within the range of 0 to 1. The TEC predicted by both ANN and IRI model generally shows the same results as the measured TEC. However, the line figure between the measured and predicted neural network values show a high degree of similarity between them. n future, it is recommended that larger data size is used to improve accuracy and reliability of the TEC forecasting.

## ACKNOWLEDGEMENT

The authors are grateful to Universiti Tun Hussein Onn Malaysia (UTHM) for enabling them to conduct this research. Appreciation is also due to Universiti Kebangsaan Malaysia (UKM) for the installation and maintenance of the GISTM in UKM. This research was funded by the Malaysian Government through National Space Agency of Malaysia (ANGKASA) under the SA1213001.

## REFERENCES

Abhishek, K., Kumar, A., Ranjan, R., & Kumar, S. (2012, July). A rainfall prediction model using artificial neural network. In C*ontrol and System Graduate Research Colloquium (ICSGRC), 2012 IEEE* (pp. 82-87). IEEE.

Bagiya, M. S., Joshi, H. P., Iyer, K. N., Aggarwal, M., Ravindran, S., & Pathan, B. M. (2009). TEC variations during low solar activity period (2005–2007) near the equatorial ionospheric anomaly crest region in India. *Ann. Geophys*, *27*, 1047-1057

Chen, C., Sun, S., & Ban, P. (2010, November). Forecasting the ionospheric f o F 2 in Chinese region by neural network technique. *9th International Symposium in Antennas Propagation and EM Theory (ISAPE)* (pp. 399-402). IEEE.

Habarulema, J. B., McKinnell, L. A., & Cilliers, P. J. (2007). Prediction of global positioning system total electron content using neural networks over South Africa. *Journal of Atmospheric and Solar-terrestrial Physics, 69*(15), 1842-1850.

Habarulema, J. B., McKinnell, L. A., Cilliers, P. J., & Opperman, B. D. (2009). Application of neural networks to South African GPS TEC modelling. *Advances in Space Research, 43*(11), 1711-1720.

Haykin, S. (1999). Neural Networks: A Comprehensive Foundation. Prentice Hall International. *The Knowledge Engineering Reviews* (2nd ed.). doi:org/10.1017/S0269888998214044

Homam, M. J. (2014, December). Initial prediction of Total Electron Content (TEC) at a low latitude station using neural network. *Asia-Pacific Conference in Applied Electromagnetics (APACE)* (pp. 111-114). IEEE.

Huang, Z., Li, Q. B., & Yuan, H. (2015). Forecasting of ionospheric vertical TEC 1-h ahead using a genetic algorithm and neural network. *Advances in Space Research, 55*(7), 1775-1783.

Leandro, R. F., & Santos, M. C. (2007). A neural network approach for regional vertical total electron content modelling. *Studia Geophysica et Geodaetica*, 51(2), 279-292.

Nakamura, M., Maruyama, T., & Shidama, Y. (2009). Using a neural network to make operational forecasts of ionospheric variations and storms at Kokubunji, Japan. *Journal of the National Institute of Information and Communications Technology, 56*(1–4), 391–406.

Otsuka, Y., Ogawa, T., Saito, A., Tsugawa, T., Fukao, S., & Miyazaki, S. (2002). A new technique for mapping of total electron content using GPS network in Japan. *Earth, Planets and Space, 54*(1), 63-70.

Watthanasangmechai, K., Supnithi, P., Lerkvaranyu, S., Tsugawa, T., Nagatsuma, T., & Maruyama, T. (2012). TEC prediction with neural network for equatorial latitude station in Thailand. *Earth, Planets and Space, 64*(6), 473-483.

Wichaipanich, N., & Supnithi, P. (2014, May). Prediction of foF2 using Neural Network at Thailand equatorial latitude station, Chumphon. *11th International Conference in Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)* (pp. 1-4). IEEE.