

Prediction of ionospheric total electron content data using NARX neural network model

Nayana Shenvi, Hassanali Gulamali Virani

Department of Electronics and Telecommunication Engineering, Goa College of Engineering, Goa University, Farmagudi, India

Article Info

Article history:

Received Apr 27, 2023

Revised Jul 12, 2023

Accepted Aug 2, 2023

Keywords:

Geomagnetic index

Ionosphere

Nonlinear autoregressive with exogenous inputs

Prediction

Solar index

Total electron content

ABSTRACT

Successful prediction of ionospheric total electron content (TEC) data will help in correction of positioning errors in global navigation satellite systems (GNSS) caused by the ionosphere. This research paper proposes a prediction model for ionospheric TEC using a nonlinear autoregressive with exogenous inputs (NARX) neural network that utilizes past TEC data along with solar and geomagnetic indices namely F10.7, disturbed storm (Dst), Kp, Ap, and time of the day. We assess the prediction capability of our model at different latitudes during different solar activity years. We compare our results with another NARX model which uses previous TEC data along with time of the day, day of the year and season as exogenous parameters. The results show that for the solar minimum year the TEC prediction accuracy improves by 35.71% and for the solar maximum year it improves by 31.20%. The results using root mean square error (RMSE), mean absolute error (MAE), correlation coefficient, and symmetric mean absolute percentage error (sMAPE) clearly indicate that solar and geomagnetic indices along with time of the day help in enhancing prediction accuracy of TEC across different latitudinal regions during both solar minimum and maximum years.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Nayana Shenvi

Department of Electronics and Telecommunication Engineering, Goa College of Engineering

Goa University

Farmagudi, Ponda, Goa 403401, India

Email: nayana@gec.ac.in

1. INTRODUCTION

The signals from the global positioning system (GPS) satellites orbiting round the earth propagate through the ionosphere. Ionosphere is the ionized region of the earth's upper atmosphere ranging from about 60-1,000 km which contains significant amount of free electrons. The ionosphere is known for its dispersive nature and thus delays the electromagnetic signals and affects precise measurements in satellite navigation and communication. This influence is directly proportional to the density of electrons between the satellite transmitter and the receiver which quantifies the total electron content (TEC). Ionospheric TEC exhibits complex spatial and temporal fluctuations along with certain solar and geomagnetic related ionospheric disturbances [1]. The daily diurnal variations, seasonal variations, 27-day variations, 11-year solar cycle variations, geographic latitude, and some events like geomagnetic storms which generate sudden disturbances in the ionosphere are some obvious variations of ionospheric TEC [2]. Prediction of ionospheric TEC can thus help in improving the performance of GPS navigational systems.

There are a number of empirical models like international reference ionosphere (IRI) [3], NeQuick [4], Klobuchar [5] which are available for reconstructing the state of the ionosphere. These models effectively predict the state of the ionosphere during quiet times but are not very accurate during periods of

high ionospheric activity. Several time-series models developed for prediction of TEC include fourier series [6], discrete cosine transform [7], and autoregressive distributed lag [8]. However, these models are not effective in predicting the non-stationary fluctuations present in the ionosphere. During last few decades, neural network techniques have been applied to predict the ionospheric TEC as well as model the various parameters in the ionospheric study [9]–[11]. These models consider the solar and geomagnetic factors in improving the prediction accuracy. Neural network models are found to outperform statistical methods. However, these models do not consider the temporal feature of TEC [12]. The TEC data exhibits non-stationarity which implies that the behavior of the ionosphere is extremely complex and nonlinear [13]. nonlinear autoregressive with exogenous inputs (NARX) neural network has been widely used for modeling non-linear dynamic systems [14]. Hence, in this paper we are proposing a NARX neural network model fed with time of the day, solar and geomagnetic indices as exogenous parameters to predict the TEC. Literature suggests number of applications modeled by NARX neural network. NARX model was tested on chaotic laser time series and traffic time series and it outperformed standard neural network architectures like time delay neural network (TDNN) and Elman architectures [14]. NARX gave better results over autoregressive integrated moving average (ARIMA) for wind speed prediction [15]. NARX was successfully tested for prediction of rainfall and showed improved results compared to linear algorithms like support vector machines [16]. NARX was used for traffic prediction and had superior prediction accuracy to RNN models [17]. The ability to predict symmetric horizontal (SYM-H) storm time index using NARX showed better accuracy compared to back propagation and recurrent Elman networks [18]. NARX has also been used for TEC prediction using time, day of the year and season as exogenous inputs [1].

In this study, we discuss prediction of TEC using NARX neural network model using exogenous parameters such as solar index F10.7 and geomagnetic indices such as disturbed storm (Dst), Kp, and Ap along with local time. We have studied TEC prediction at different latitudinal regions namely equatorial (0° - 10° N), low-latitude (10° - 30° N), mid-latitude (30° - 60° N), and high-latitude (60° - 80° N), during both the minimum solar activity year (2008) and maximum solar activity year (2014). We compare the results of our NARX model (henceforth called NARX1) with the NARX model developed for TEC prediction by Guoyan *et al.* [1] (henceforth called NARX2).

2. METHOD

2.1. Data and Input parameters

The ionospheric TEC data is downloaded from IONOLAB (www.ionolab.org) [19]. To explore latitudinal influence on the prediction of ionospheric TEC, we selected four locations in the northern region, each situated at different latitudes. The selected stations are qaq1 (Greenland and Denmark), baie (Baie-Comeau and Canada), mas1 (Maspalomas and Spain), and bogt (Bogota and Columbia). The latitude and longitude co-ordinates of these stations is shown in Table 1.

Table 1. Selected stations in the northern hemisphere

Region	Stations	Latitude and longitude
High latitude region (60° to 80°)	qaq1 (Greenland and Denmark)	60.7° N and 46.04° W
Mid latitude region (30° to 60°)	baie (Comeau and Canada)	49.18° N and 68.26° W
Low latitude region (10° to 30°)	mas1 (Maspalomas and Spain)	27.76° N and 15.63° W
Equatorial region (0° to 10°)	bogt (Columbia)	4.64° N and 74.08° W

The exogenous parameters for the NARX model were selected based on the factors influencing TEC. We calculated the Pearson correlation coefficient for various exogenous parameters and selected the ones that correlate well with TEC data. Table 2 shows the values of Pearson correlation coefficient for the various exogenous parameters with TEC data. From the table it is clear that Ap and Kp show high positive correlation, Dst shows high negative correlation and F10.7 shows moderate correlation with TEC. Time of the day was selected since TEC exhibits diurnal characteristics. Thus, we selected solar index F10.7 and geomagnetic indices Dst, Kp and Ap along with time of the day.

Table 2. Pearson correlation coefficient

Exogenous parameters	Pearson correlation coefficient
Ap	0.67
Kp	0.65
Dst	-0.63
F10.7	0.37

The exogenous parameters were obtained from the space physics data facility at NASA's Goddard Space Flight Center, accessible through their website at <https://omniweb.gsfc.nasa.gov/form/dx1.html> [20]. The TEC data along with the solar and geomagnetic indices were sampled every one hour. We used the data from 1st July, 2008 to 26th July, 2008 (total samples=26 days*24= 624) for training the model and then tested the prediction capability of the model for unseen data from 27th July, 2008 to 31st July, 2008 (total samples= 5 days*24=120). The same period was selected for training and testing the model for solar maximum year 2014. The various exogenous parameters used are as follows:

- Local time: TEC exhibits diurnal variation because the electron density progressively rises to a maximum value at noon before falling to a minimum value at midnight. We compute $\sin\left(\frac{2\pi}{24}\right)h$ and $\cos\left(\frac{2\pi}{24}\right)h$ where local time, h ranges from 0-23. The sin and cos functions help in normalizing the time input in the range (-1,1) [1].
- Solar activity: solar activity significantly influences TEC values, with lower TEC during solar minimum years and higher TEC during solar maximum years [21]. F10.7, which correlates directly with the number of sunspots and solar radiation in the ultraviolet and visible spectrum, serves as a valuable predictor of solar activity [22].
- Geomagnetic activity (Dst, Kp, and Ap index): the Dst, Kp, and Ap represent geomagnetic activity.
- Dst index: the Dst is a metric utilized to gauge the intensity of geomagnetic storms. It quantifies the deviation of the Earth's horizontal magnetic field component from its typical variation during calm days [23]. Under tranquil conditions, the Dst ranges from 0 to -50 nT, but during the occurrence of a powerful storm, it can surpass -200 nT.
- Kp index: the planetary index Kp is the average standardized K-index obtained from 13 geomagnetic observatories between 44^o and 60^o in either hemisphere [21]. This parameter serves as a measure of geomagnetic activity. The higher the Kp index, more severe the geomagnetic activity, and greater the impact on the TEC.
- Ap index: the Ap index, also known as the planetary index, is a numerical scale that ranges from 0 to 400, with higher values signifying more intense geomagnetic activity. This index serves as a valuable tool for monitoring the effects of space weather on the earth's magnetic field and assessing potential impacts on critical systems such as power grids, satellites, and radio communications [21].

2.2. Architecture of NARX1 model

A NARX neural network is a dynamic model that incorporates recurrent feedback connections from the output layer to the input layer with time-delayed units. This unique architecture makes NARX networks an efficient tool for modeling and validation purposes, as they demonstrate better generalization and faster convergence compared to other conventional neural network models [24]. This can model nonlinear time-series systems and can be mathematically represented by:

$$\hat{y}(t+1) = f(y(t), y(t-1), \dots, y(t-n_y), x(t), x(t-1), \dots, x(t-n_x)) + e(t) \quad E_v - E = \frac{h}{z.m} (k_x^2 + k_y^2) \quad (1)$$

where $\hat{y}(t+1)$ is the predicted TEC value at time $(t+1)$. This depends on historical TEC values $y(t), y(t-1), \dots, y(t-n_y)$ and historical values of the independent exogenous inputs $x(t), x(t-1), \dots, x(t-n_x)$ which in our case is time of the day, F10.7, Dst, Kp and Ap. n_x and n_y are the time delays for $x(t)$ and $y(t)$ respectively as shown Figure 1, $e(t)$ is residual value at a specific point in time, t [14].

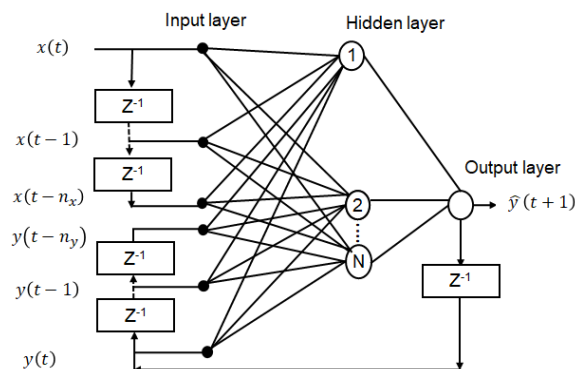


Figure 1. NARX model with tapped delay line at the input and output, z^{-1} represents unit time delay

The NARX neural network, like a multilayer perceptron (MLP), comprises an input layer, a hidden layer, and an output layer in its architecture. However, the key difference is that the NARX network includes recurrent feedback connections from the output layer to the input layer with time-delayed units, enabling it to model dynamic systems and time-dependent data more effectively. The past exogenous inputs along with the past predicted TEC values are fed at the input of the model. The feedback connections allow the network to take into account the previous behavior of the system and to capture any patterns that may exist between the input and output variables over time. Figure 2 depicts the framework of NARX neural network. The input layer takes in the exogenous inputs along with past predicted TEC values. The output of the intermediate hidden layer $H_i(t)$ is given by:

$$H_i(t) = F_1 \left[\sum_{p=0}^{n_x} w_{ip} x(t-p) + \sum_{q=0}^{n_y} w_{iq} y(t-q) + a_i \right] \tag{2}$$

where, each hidden layer node is activated by the function F_1 . The input neuron $x(t-p)$ is connected to the i^{th} hidden layer node with a connection weight of w_{ip} and the output feedback neuron $y(t-q)$ is also connected to the i^{th} hidden layer node with a connection weight of w_{iq} . Additionally, each hidden layer node is associated with a bias term a_i . These connection weights and bias terms are learned during the training process and they determine the strength and direction of the connections between the neurons in the network [25]. The predicted TEC value at the output $\hat{y}(t+1)$ is given by:

$$\hat{y}(t+1) = F_2 \left[\sum_{i=1}^N w_i H_i(t) + b_i \right] \tag{3}$$

where w_i is the connection weight between the i^{th} hidden layer node and output layer, b_i is the bias of the network, N is the number of hidden layer nodes. Therefore, the predicted TEC value at time $t+1$ is a function of the weighted sum of the inputs passed through a nonlinear activation function and is determined by the weights and biases of the neural network. NARX neural network can be used in open-loop mode and closed-loop mode. The open-loop architecture and the closed-loop architecture of NARX neural network model is shown in Figures 3(a) and (b) respectively.

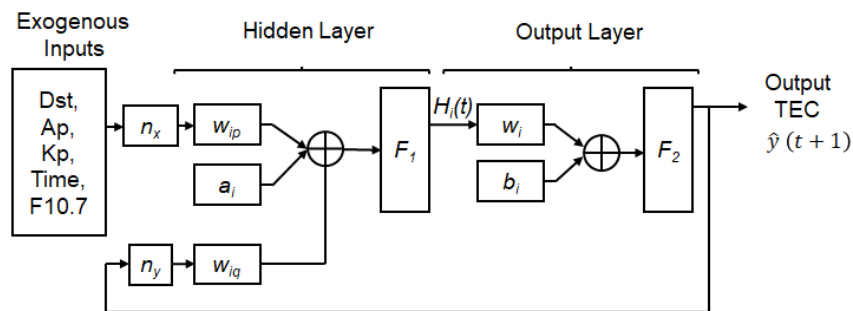


Figure 2. NARX neural network structure

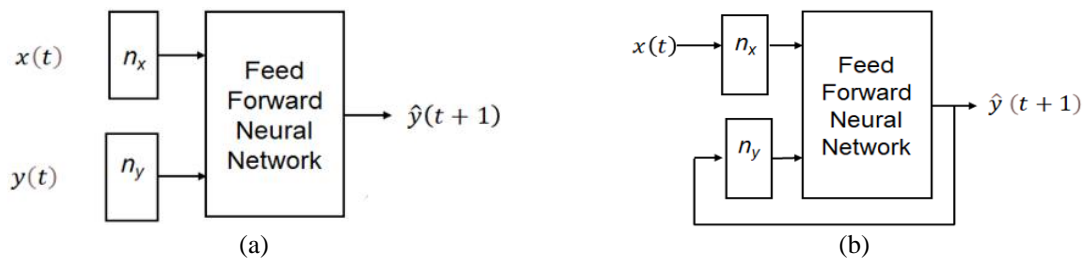


Figure 3. NARX; (a) open-loop architecture and (b) closed-loop architecture

In the open-loop architecture, as depicted in Figure 3(a), the NARX neural network predicts the output TEC $\hat{y}(t+1)$ using historical values of both the exogenous data and the actual past TEC values. This configuration enables the model to utilize accurate information from the actual TEC time series, contributing to more precise inputs [26]. In the closed-loop configuration, as depicted in Figure 3(b), the NARX model feeds back the predicted TEC output to the input of the feedforward neural network. Consequently, the

prediction of $\hat{y}(t+1)$ is based on the previous values of the exogenous data and the previously predicted TEC values. This closed-loop architecture allows the model to exploit its own predictions to make further estimations, potentially capturing dynamic dependencies and time-dependent patterns effectively. The mathematical expression for the NARX open-loop architecture is given by:

$$\hat{y}(t+1) = f(y(t), y(t-1), \dots, y(t-n_y), x(t), x(t-1), \dots, x(t-n_x)) \tag{4}$$

and for the closed-loop architecture;

$$\hat{y}(t+1) = f(\hat{y}(t), \hat{y}(t-1), \dots, \hat{y}(t-n_y), x(t), x(t-1), \dots, x(t-n_x)) \tag{5}$$

where $f(\cdot)$ is the mapping function of the neural network, $\hat{y}(t+1)$ is the predicted TEC output of the NARX using either the past true TEC values $y(t), y(t-1), \dots, y(t-n_y)$ or the past predicted TEC values $\hat{y}(t), \hat{y}(t-1), \dots, \hat{y}(t-n_y)$ and the past values of the exogenous time series $x(t), x(t-1), \dots, x(t-n_x)$ and n_x and n_y are the input and output delays.

2.3. Implementation of NARX1 model in MATLAB

We have simulated the NARX model using the neural net toolbox in MATLAB. The simulated model is shown in Figures 4 and 5. The number of hidden layer neurons was set to 20. The delays n_x and n_y was set to two. These hyperparameters were finalized after multiple simulations. The model was run with different number of network feedback delays and then finalized to two where the prediction model performed well. The neurons in the hidden layer determine the accuracy of the model. Less neurons in the hidden layer fail to establish relationship of the data but increasing the number of neurons leads to more computational time [27]. Initially we trained the model in open loop configuration as shown Figure 4. The advantage of doing this is that the NARX model uses the previous values of the actual TEC time series and hence it receives more precise input during the training phase. Then we converted the model into closed loop for making predictions into the future as shown Figure 5. We perform recursive multi-step forecasting in which we predict one-step TEC value and use this predicted value in the next step to make next prediction. This process is continued till we forecast TEC values for the next 5 days.

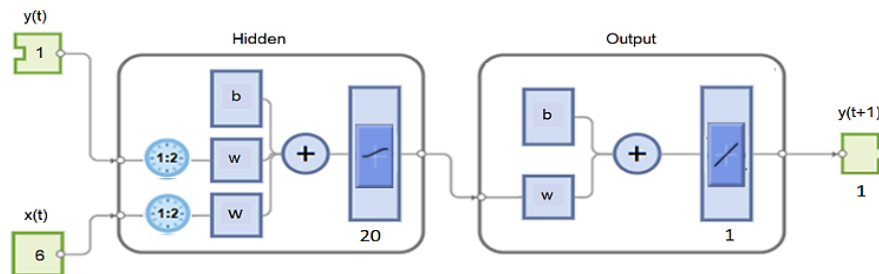


Figure 4. NARX open loop architecture simulated in MATLAB during the training phase

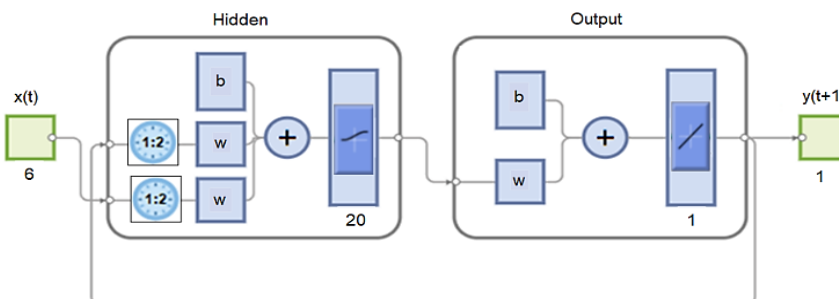


Figure 5. NARX closed loop architecture simulated in MATLAB for step ahead predictions

Before training the network, the data was divided into three sets. 70% of the data was used for the training, 14% for validation, and 16% for testing. The network weights and biases were updated and the

gradient was calculated using the training data. The validation samples aid in fine-tuning the hyperparameters and the test samples help in computing the generalization performance of the model. The training algorithm used was Levenberg Marquardt. This algorithm updates weights and biases using less memory. Training stops when error starts increasing indicating that there is no further improvement in generalization. The output provides the predicted TEC values. The NARX model was developed using a hyperbolic tangent function in the hidden layer and a linear transfer function in the output layer. We calculate the root mean square error (RMSE), mean absolute error (MAE), correlation coefficient (r), and symmetric mean absolute percentage error (sMAPE) using the following formulas to evaluate the performance of the model.

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N [y_i - \hat{y}_i]^2} \quad (6)$$

$$\text{MAE} = \frac{\sum_{i=1}^N |y_i - \hat{y}_i|}{N} \quad (7)$$

$$r = \frac{\sum_{i=1}^N (y_i - \bar{y}_i)(\hat{y}_i - \bar{\hat{y}}_i)}{\sqrt{\sum_{i=1}^N (y_i - \bar{y}_i)^2 \sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}}_i)^2}} \quad (8)$$

$$\text{sMAPE} = \frac{100}{N} \sum_{i=1}^N \frac{|\hat{y}_i - y_i|}{(|y_i| + |\hat{y}_i|)/2} \quad (9)$$

where \hat{y}_i is the predicted TEC by the model and y_i is the actual TEC value, \bar{y}_i is the mean of the actual TEC values, $\bar{\hat{y}}_i$ is the mean of the predicted TEC values and N is the frequency of observations.

3. RESULTS AND DISCUSSION

3.1. Stability of the model

After training the model we check the stability of the model by plotting the error autocorrelation function, error histogram, and regression coefficient. Figure 6 shows the error autocorrelation plot for qaq1 station during the solar maximum year 2014. The error autocorrelation function is a valuable tool for validating the performance of a model. It helps describe the relationship between prediction errors over time, revealing how errors are correlated at different time lags.

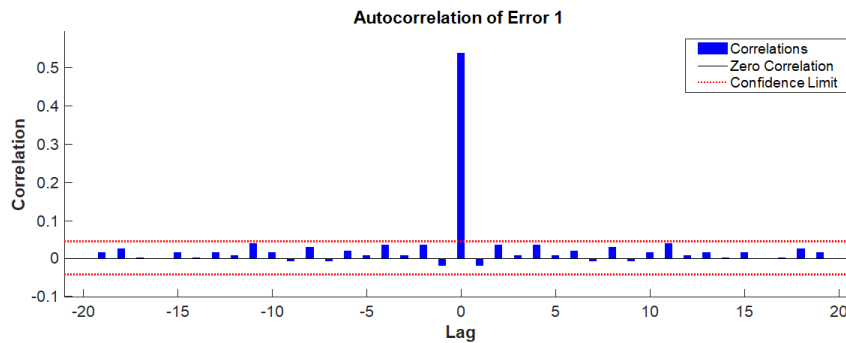


Figure 6. Error autocorrelation plot for qaq1 station during solar active year 2014

An ideal prediction model is indicated by only one non-zero value in the autocorrelation function, which occurs at zero lag, signifying a perfect match between the model's predictions and the actual data [28]. The predicted errors are below 95% confidence limit (red dotted line) and thus are uncorrelated with each other. This also shows that the network has been trained satisfactorily and can be used for prediction. The error histogram plot in MATLAB is a graphical representation of the distribution of errors in a dataset. The plot shows the frequency of occurrence of errors, binned into intervals, with the x-axis representing the difference between predicted and actual TEC values and the y-axis representing the instances of those errors. Figure 7 shows a well-behaved error histogram plot which has a roughly symmetric distribution around the mean error value, with very few or no outliers. The plot clearly indicates that maximum samples from the training, validation and test dataset have a very small error equal to 0.01736.

The regression plot denotes an equation between the predicted TEC (output) and true (target) value of TEC. Figure 8 shows the regression line plot for baie station during training, validation and test phases during the solar maximum year 2014. Table 3 shows the slope and intercept values for the regression plot shown in Figure 8. The regression equation relating the predicted and the true TEC is shown on the y-axis. The coefficient of the equation shows proportionality between predicted TEC and the true TEC and is very close to 1. The constant term represents an error which should be added to the scaled target to make it as close as possible to the predicted output and is very close to 0 (see Figure 8). Overall, high values of R were achieved, specifically 99.75% for the training set, 99.04% for the validation set, 99.16% for the test data set and overall 99.54%. R values greater than 95% suggest excellent fit between true and predicted values [27]. Figures 6-8 clearly indicate that NARX network has been trained efficiently and can be effectively used for prediction of TEC data.

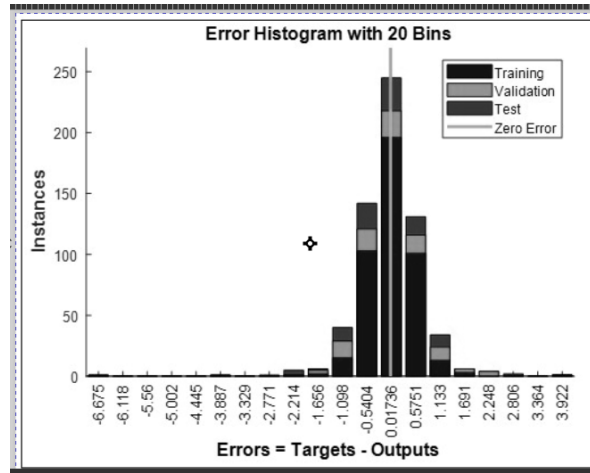


Figure 7. Error histogram plot for qa1 station during solar active year 2014

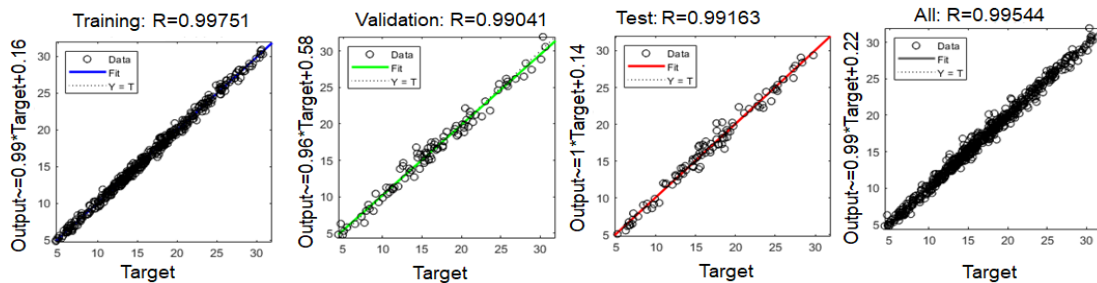


Figure 8. Regression line plot for baie station during solar active year 2014

Table 3. Slope and intercept values for the regression plot in Figure 8

Training set		Validation set		Test set	
Slope	Intercept	Slope	Intercept	Slope	Intercept
0.99	0.16	0.96	0.58	1.00	0.14

The validity of our model's results could not be meaningfully compared to existing literature due to variations in the methodologies used in each study. These differences include the use of different sets of stations, varying amounts of training data, and use of different solar minimum and active years. To accurately compare our results and validate our model, we simulated the NARX2 model using the identical architecture and parameters as described by Guoyan *et al.* [1] using time of the day, day of the year and season as exogenous inputs.

3.2. For the solar quiet year 2008

We check the prediction capability of our model by testing it with unseen data from 26th July to 31st July, 2008. The plots in Figure 9 show the observed and the predicted TEC at Figure 9(a) qa1, Figure 9(b) baie, Figure 9(c) mas1, and Figure 9(d) bogt stations respectively during the solar minimum year 2008. The

graph illustrates how well the predictions align with the actual TEC values for each station during this specific time frame.

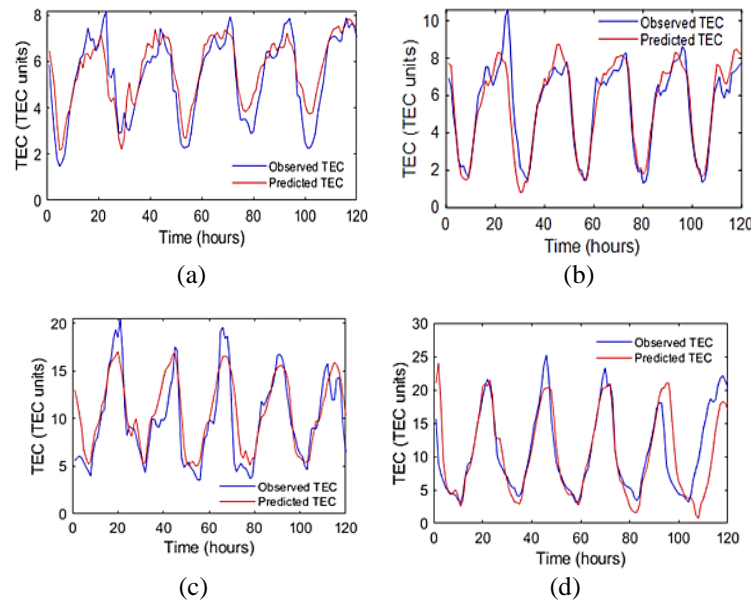


Figure 9. Observed and predicted TEC using NARX model for the period from 26th July 2008 to 31st July 2008 for; (a) qaql, (b) baie, (c) mas1, and (d) bogt stations during the solar minimum year 2008

3.3. Solar maximum year 2014

We also check the prediction capability of our model by testing it with unseen data from 26th July to 31st July, 2014 for the solar maximum year 2014 (see Figure 10). The plots in Figure 10 show the observed and the predicted TEC at Figure 10(a) qaql, Figure 10(b) baie, Figure 10(c) mas1, and Figure 10(d) bogt stations respectively during the solar maximum year 2014. Compared to solar minimum year 2008, there is a more deviation between the original and predicted TEC. This is specifically due to increased levels of ionization and increase solar activity and is consistent with the literature.

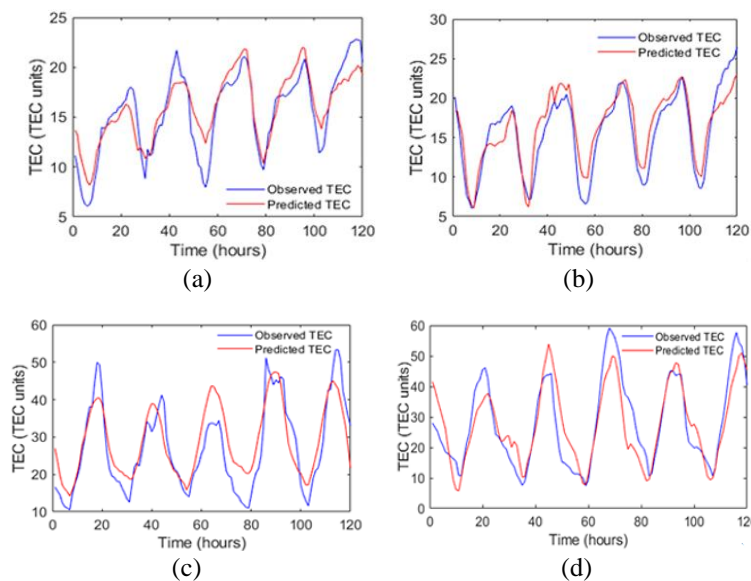


Figure 10. Observed and predicted TEC using NARX model for the period from 26th July 2014 to 31st July 2014 for; (a) qaql, (b) baie, (c) mas1, and (d) bogt stations during the solar maximum year 2014

3.4. Comparison with NARX2 model

To validate our model, we compared our results with NARX2. NARX1 model was implemented using time, solar index (F10.7), and geomagnetic indices (Dst, Kp, and Ap). The literature used NARX2 model for TEC prediction using time of the day, day of the year and season as exogenous inputs. This model was implemented for China region for low and mid latitude regions for the solar maximum year 2011 and quiet year 2017. To ensure a fair comparison, we implemented NARX2 for our regions and also for the same solar minimum (2008) and active years (2014). Using the same model architecture (NARX), we wanted to see the improvement in prediction accuracy using solar and geomagnetic indices as exogenous parameters at all latitudinal regions and for both the solar minimum and active years. The RMSE, MAE, correlation coefficient and sMAPE values obtained for each of the station for both the solar minimum and active years using both NARX1 and NARX2 model are plotted in Figures 11(a)-(d) respectively.

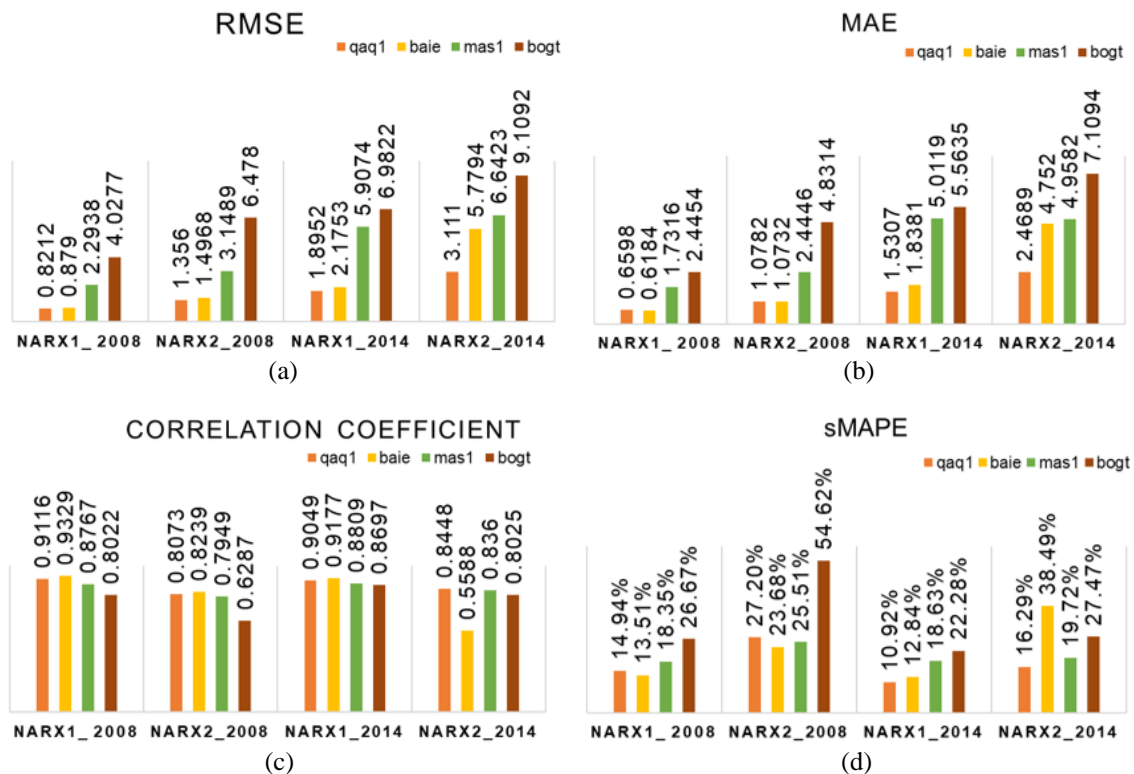


Figure 11. Comparison of; (a) RMSE, (b) MAE, (c) correlation coefficient, and (d) sMAPE values for NARX1 and NARX2 model for qaq1, baie, mas1 and bogt stations during the year 2008 and 2014

Both the RMSE and MAE metrics show that our NARX1 model can predict TEC with better accuracy during both the solar minimum and active years. The prediction accuracy for the solar minimum year 2008 shows an improvement of 35.71% (RMSE) and 42.14% (MAE) over NARX2. For the solar maximum year 2014, the prediction accuracy shows an improvement of 31.20% (RMSE) and 27.7% (MAE) over NARX2. A high correlation coefficient between true and predicted values using NARX1 model indicates that the predicted values are closely aligned with the true values. Further, low values of sMAPE for NARX1 model indicate a higher level of accuracy and better performance of the forecasting model compared with NARX2. A lower sMAPE suggests that the forecasted values are closer to the actual values, with smaller percentage errors. As anticipated, the RMSE increases as latitude decreases, implying better prediction performance in higher and mid-latitude regions compared to lower and equatorial regions. The challenging aspect of modeling and predicting ionospheric TEC in low-latitude regions is attributed to the presence of the equatorial ionization anomaly [29]. This anomaly introduces complex and variable ionospheric behavior, making accurate modeling and prediction more difficult in these specific areas. The RMSE and MAE values for solar maximum year are higher compared to solar minimum year. This is to increased solar radiation, ionization, and disturbances in the ionosphere during solar maximum year and is consistent with the literature. The RMSE, MAE, correlation coefficient and sMAPE metrics show that

NARX1 model can predict TEC with better accuracy even for solar maximum year. According to our study findings, the NARX1 model consistently demonstrated better predictive performance for TEC values compared to the NARX2 model.

4. CONCLUSION

This study examined the use of NARX neural network for prediction of TEC data during the solar minimum and active years using time of the day, solar, and geomagnetic indices as exogenous parameters. The NARX model's ability to capture nonlinear dynamics, incorporate autoregressive terms, and include exogenous inputs make it a powerful tool for predicting TEC data. The error analysis based on RMSE, MAE, correlation coefficient and sMAPE indicate that use of solar index (F10.7) and geomagnetic indices (Dst, Kp, and Ap) as exogenous data helps in effective prediction of TEC. These parameters deviate from their normal values with changes in the ionospheric TEC. Any fluctuations in TEC due to disturbance in the ionosphere is quickly captured by these exogenous parameters and are effective in TEC prediction. But, NARX2 uses time of the day, day of the year and season as exogenous parameters. These parameters cannot capture the sudden ionospheric disturbances and are only useful in TEC prediction during quiet stable ionospheric conditions. Further, the TEC data, the solar and the geomagnetic data are all used in raw form without normalizing indicating the suitability of this technique for real time TEC prediction. The prediction of TEC at different latitudinal regions is also consistent with the literature and shows increasing RMSE with decreasing latitudes. The error autocorrelation function, error histogram and regression line plots also indicate the best fit of the model for TEC prediction. The RMSE between 0.8212 TECU to 4.0277 TECU for the solar minimum year and between 1.8952 TECU to 6.9822 TECU for the solar maximum year indicate the suitability of this model and correct selection of the exogenous parameters for ionospheric TEC prediction. All the evaluation metrics clearly showed that the NARX1 model outperformed the NARX2 model in accurately predicting TEC values.




REFERENCES

- [1] L. Guoyan, G. Wang, Z. Zhengxie, and Z. Qing, "Prediction of Ionospheric TEC Based on the NARX Neural Network," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–10, Oct. 2021, doi: 10.1155/2021/7188771.
- [2] C.-C. Wu, K. Liou, S.-J. Shan, and C.-L. Tseng, "Variation of ionospheric total electron content in Taiwan region of the equatorial anomaly from 1994 to 2003," *Advances in Space Research*, vol. 41, no. 4, pp. 611–616, Jan. 2008, doi: 10.1016/j.asr.2007.06.013.
- [3] D. Bilitza *et al.*, "International Reference Ionosphere 2016: From ionospheric climate to real-time weather predictions," *Space Weather*, vol. 15, no. 2, pp. 418–429, Feb. 2017, doi: 10.1002/2016SW001593.
- [4] B. Nava, P. Coisson, and S. M. Radicella, "A new version of the NeQuick ionosphere electron density model," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 70, no. 15, pp. 1856–1862, Dec. 2008, doi: 10.1016/j.jastp.2008.01.015.
- [5] J. Klobuchar, "Ionospheric Time-Delay Algorithm for Single-Frequency GPS Users," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-23, no. 3, pp. 325–331, May 1987, doi: 10.1109/TAES.1987.310829.
- [6] S. Stanimir, K. Ivan, J. Norbert, and W. Andreas, "A New Method for Total Electron Content Forecasting Using Global Positioning System Measurements," *ESA space weather workshop*, pp. 169–172, 2001.
- [7] A. García-Rigo *et al.*, "Global prediction of the vertical total electron content of the ionosphere based on GPS data," *Radio Science*, vol. 46, no. 6, Dec. 2011, doi: 10.1029/2010RS004643.
- [8] N. Shenvi, H. Virani, and E. Chandrasekhar, "Forecasting of Ionospheric Total Electron Content Data using Autoregressive Distributed Lag Model for Mid-Latitude Region During Solar Minimum and Maximum," in *2022 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI)*, Dec. 2022, pp. 1–6, doi: 10.1109/ICDSAAI55433.2022.10028905.
- [9] A. M. El-naggar, "Artificial neural network as a model for ionospheric TEC map to serve the single frequency receiver," *Alexandria Engineering Journal*, vol. 52, no. 3, pp. 425–432, Sep. 2013, doi: 10.1016/j.aej.2013.05.007.
- [10] R. F. Leandro and M. C. Santos, "A neural network approach for regional vertical total electron content modelling," *Studia Geophysica et Geodaetica*, vol. 51, no. 2, pp. 279–292, Apr. 2007, doi: 10.1007/s11200-007-0015-6.
- [11] A. Yilmaz, K. E. Akdogan, and M. Gurun, "Regional TEC mapping using neural networks," *Radio Science*, vol. 44, no. 3, Jun. 2009, doi: 10.1029/2008RS004049.
- [12] W. Sun *et al.*, "Forecasting of ionospheric vertical total electron content (TEC) using LSTM networks," in *2017 International Conference on Machine Learning and Cybernetics (ICMLC)*, Jul. 2017, pp. 340–344, doi: 10.1109/ICMLC.2017.8108945.
- [13] S. Bhardwaj, E. Chandrasekhar, G. K. Seemala, and V. M. Gadre, "Characterization of ionospheric total electron content data using wavelet-based multifractal formalism," *Chaos, Solitons & Fractals*, vol. 134, p. 109653, May 2020, doi: 10.1016/j.chaos.2020.109653.
- [14] J. M. P. Menezes and G. A. Barreto, "Long-term time series prediction with the NARX network: An empirical evaluation," *Neurocomputing*, vol. 71, no. 16–18, pp. 3335–3343, Oct. 2008, doi: 10.1016/j.neucom.2008.01.030.
- [15] E. Cadenas, W. Rivera, R. Campos-Amezcuca, and C. Heard, "Wind Speed Prediction Using a Univariate ARIMA Model and a Multivariate NARX Model," *Energies*, vol. 9, no. 2, p. 109, Feb. 2016, doi: 10.3390/en9020109.
- [16] M. Peña, A. Vázquez-Patiño, D. Zhiña, M. Montenegro, and A. Avilés, "Improved Rainfall Prediction through Nonlinear Autoregressive Network with Exogenous Variables: A Case Study in Andes High Mountain Region," *Advances in Meteorology*, vol. 2020, pp. 1–17, Sep. 2020, doi: 10.1155/2020/1828319.
- [17] Z. Song, F. Sun, R. Zhang, Y. Du, and C. Li, "Prediction of Road Network Traffic State Using the NARX Neural Network," *Journal of Advanced Transportation*, vol. 2021, pp. 1–17, Nov. 2021, doi: 10.1155/2021/2564211.
- [18] L. Cai, S. Ma, H. Cai, Y. Zhou, and R. Liu, "Prediction of SYM-H index by NARX neural network from IMF and solar wind data," *Science in China Series E: Technological Sciences*, vol. 52, no. 10, pp. 2877–2885, Oct. 2009, doi: 10.1007/s11431-009-0296-9.




- [19] U. Sezen, F. Arikan, O. Arikan, O. Ugurlu, and A. Sadeghimorad, "Online, automatic, near-real time estimation of GPS-TEC: IONOLAB-TEC," *Space Weather*, vol. 11, no. 5, pp. 297–305, May 2013, doi: 10.1002/swe.20054.
- [20] J. H. King and N. E. Papitashvili, "Solar wind spatial scales in and comparisons of hourly Wind and ACE plasma and magnetic field data," *Journal of Geophysical Research: Space Physics*, vol. 110, no. A2, Feb. 2005, doi: 10.1029/2004JA010649.
- [21] R. Mukesh, V. Karthikeyan, P. Soma, and P. Sindhu, "Forecasting of ionospheric TEC for different latitudes, seasons and solar activity conditions based on OKSM," *Astrophysics and Space Science*, vol. 365, no. 1, p. 13, Jan. 2020, doi: 10.1007/s10509-020-3730-x.
- [22] W. Zhang *et al.*, "Predicting the Daily 10.7-cm Solar Radio Flux Using the Long Short-Term Memory Method," *Universe*, vol. 8, no. 1, p. 30, Jan. 2022, doi: 10.3390/universe8010030.
- [23] S. B. Xu, S. Y. Huang, Z. G. Yuan, X. H. Deng, and K. Jiang, "Prediction of the Dst Index with Bagging Ensemble-learning Algorithm," *The Astrophysical Journal Supplement Series*, vol. 248, no. 1, p. 14, May 2020, doi: 10.3847/1538-4365/ab880e.
- [24] Tsung-Nan Lin, C. L. Giles, B. G. Horne, and Sun-Yuan Kung, "A delay damage model selection algorithm for NARX neural networks," *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2719–2730, 1997, doi: 10.1109/78.650098.
- [25] R. Magallanes-Quintanar, C. E. Galván-Tejada, J. I. Galván-Tejada, S. de J. Méndez-Gallegos, A. García-Domínguez, and H. Gamboa-Rosales, "Narx Neural Networks Models for Prediction of Standardized Precipitation Index in Central Mexico," *Atmosphere*, vol. 13, no. 8, p. 1254, Aug. 2022, doi: 10.3390/atmos13081254.
- [26] Z. Boussaada, O. Curea, A. Remaci, H. Camblong, and N. Mrabet Bellaaj, "A Nonlinear Autoregressive Exogenous (NARX) Neural Network Model for the Prediction of the Daily Direct Solar Radiation," *Energies*, vol. 11, no. 3, p. 620, Mar. 2018, doi: 10.3390/en11030620.
- [27] G. Sivavaraprasad, V. S. Deepika, D. SreenivasaRao, M. Ravi Kumar, and M. Sridhar, "Performance evaluation of neural network TEC forecasting models over equatorial low-latitude Indian GNSS station," *Geodesy and Geodynamics*, vol. 11, no. 3, pp. 192–201, May 2020, doi: 10.1016/j.geog.2019.11.002.
- [28] Y. Raptodimos and I. Lazakis, "Application of NARX neural network for predicting marine engine performance parameters," *Ships and Offshore Structures*, vol. 15, no. 4, pp. 443–452, Apr. 2020, doi: 10.1080/17445302.2019.1661619.
- [29] I. Srivani, G. Siva Vara Prasad, and D. Venkata Ratnam, "A Deep Learning-Based Approach to Forecast Ionospheric Delays for GPS Signals," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 8, pp. 1180–1184, Aug. 2019, doi: 10.1109/LGRS.2019.2895112.

BIOGRAPHIES OF AUTHORS



Nayana Shenvi    completed her B.E. and M.E. (Electronics Communication and Instrumentation) degree from the Electronics and Telecommunication Department of Goa College of Engineering. She is presently pursuing her Ph.D. degree in the domain of application of machine learning techniques in Ionospheric TEC. She is currently working as Associate Professor in the Department of Electronics and Telecommunication Engineering of Goa College of Engineering, Goa, India. She has teaching experience of more than 23 years. She has more than 15 publications in journal and conference proceedings at the national and international levels. Her areas of interest include neural networks, machine learning, embedded systems, microprocessors, and microcontrollers. She is a life member of IETE and ISTE. She can be contacted at email: nayana@gec.ac.in.



Hassanali Gulamali Virani    is currently working as Professor and Head of the Electronics and Telecommunication Department of Goa College of Engineering, Goa, India. He completed his Bachelor's Degree in Electronics and Telecommunication Engineering from Goa College of Engineering, Goa University. He completed his Master's Degree in Microelectronics from National Institute of Technology, Surathkal, India. He has Ph.D. degree from Indian Institute of Technology, Bombay. He has more than 27 years of teaching experience. His areas of research include devices, nanoelectronics, control systems, and machine learning. He has paper publications in IEEE Transactions on Electron Devices and several other reputed journals and conferences. He is a fellow member of IETE and ISTE. He can be contacted at email: virani@gec.ac.in.