International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-2 Issue-5, April 2013

Predictive Models for Vertical Total Electron Content in Ionosphere

S.Priya, A.Parameswari

Abstract -*The ionosphere is defined as a region of the earth's upper atmosphere where sufficient ionisation can exist to affect the propagation of radio waves. Prediction of ionosphere vertical total electron content (TEC) are crucial and remain as a challenge for GPS positioning and navigation system , space weather forecast, as well as many other Earth Observation System. TEC is an important descriptive quantity for the ionosphere of the Earth. TEC is strongly affected by solar activity. This ionospheric characteristic constitutes an important parameter in trans ionospheric links since it issued to derive the signal delay imposed by the ionosphere. This paper gives an overview of the various predictive models that can be used to predict Total electron content in ionosphere.*

Keywords— K Nearest neighbor, Linear Predictive coding, Vertical Total Electron Content.

I. INTRODUCTION

The concept of "space weather" is now widely used in quantitative descriptions of the physical changes in the near-Earth space environment in response to variations in solar radiation, solar plasma ejection and the electromagnetic status of the interplanetary medium. In the last decade, the utility of radio wave transmissions from the Global Positioning System (GPS) satellites in obtaining information about the Earth's ionosphere simultaneously from a global network of stations has been demonstrated. Irregularities in the ionosphere due to space weather events caused by solar flares and coronal mass ejection can scatter trans-ionospheric radio signals producing fluctuations in both amplitude and phase and GPS cycle slips disrupting satellite communications and navigation.

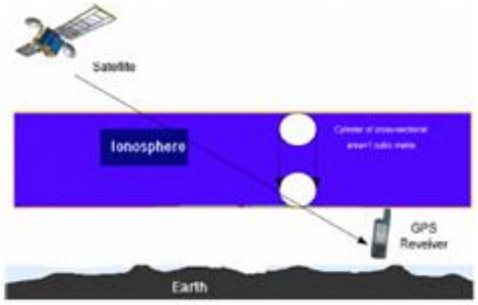

Figure 1. TEC representation and typical electron density profile

Manuscript published on 30 April 2013.
***Correspondence Author(s)**
S.Priya, Assistant Professor,Department of Computer Applications & ISM, Theivanai Ammal College for Women,Villupura, India.
A.Parameswari, Assistant Professor,Department of Computer Applications & ISM, Theivanai Ammal College for Women,Villupuram, India.

Over the past decade, several numerical and empirical models for the regular variations of TEC have been developed at regional and global scales. Numerical ionospheric models use different techniques, such as polynomial expansions, grid-based techniques and spherical harmonics in latitude and longitude, to model VTEC on one or more ionospheric layers as function of time. The parameters for these models are estimated from slant TEC (STEC) measurements. STEC is converted into VTEC using a mapping function, and vise versa. In these applications, the estimation of VTEC as function of time is a vital part. However, there are varieties of error sources which affect the estimation of VTEC. Rideout and Coster (2006) extensively discussed mapping function errors, Ciraolo et al. (2007) have studied the effects of the multipath on carrier-phase smoothed code observations using co-located receivers and intra-daily DCB variations for a zero-baseline experiment using single difference observations. Thus, it is important to predict VTEC accurately and efficiently in order to avoid errors. In this research work, based on time series analysis theory and technology, ionospheric VTEC prediction techniques are proposed. In particular, the applicability of using neural network (NN), K Nearest Neighbor (KNN) and Linear Predictive Coding model (LPC) to perform short-term regional ionospheric VTEC prediction is analyzed

II. REVIEW OF LITERATURE

At the start of the 1980s, state space models were only beginning to be used by statisticians for forecasting time series, although the ideas had been present in the engineering literature since Kalman's (1960) ground-breaking work. State space models provide a unifying framework in which any linear time series model can be written. The key forecasting contribution of Kalman (1960) was to give a recursive algorithm (known as the Kalman filter) for computing forecasts. Statisticians became interested in state space models when Schweppe (1965) showed that the Kalman filter provides an efficient algorithm for computing the one-step-ahead prediction errors and associated variances needed to produce the likelihood function. Shumway and Stoffer (1982) combined the EM algorithm with the Kalman filter to give a general approach to forecasting time series using state space models, including allowing for missing observations. In 2010, Xiuhai LI1, Dazhi GUO1 [3], in their research on modeling and prediction of Ionospheric total electron content by time series analysis. Precise modeling and accurate prediction for the ionospheric total electron content(TEC) are crucial and remain as a challenge for GPS positioning and navigation , space weather forecast, as well as many other Earth Observation System(EOS).

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

This research develops and analyzes a new prediction technique for the regional ionospheric TEC, based on time series analysis theory using autoregressive model (AR) to perform short-term ionospheric TEC prediction. The predicted TEC were then compared with the TEC measured by IGS, and with TEC from the International Reference Ionosphere(IRI) to assess the performance of the model. Preliminary results show that AR model could well describe the variation trend of the regional ionospheric TEC and has a good short-term performance of the ionospheric TEC prediction. The forecasting methodology based on the time series for the regional ionospheric TEC prediction is feasible.

In 2010, LI Shuhui and PENG Junhuan [4] presented Ionospheric TEC Prediction and Analysis Based on Phase Space Reconstruction. The time series analysis and prediction methods based on the chaotic theory don't need the subjective model having been constructed for the system beforehand. Actually, they are predicted according to a group of new phase point series in the phase space constructed by the existing time series and the characters of the phase space series.

III. PROPOSED METHODOLOGY

Vertical Total Electron Content is a quantity that concern for predicting space weather effects on telecommunications, improving the accuracy of satellite navigation, fly control vehicles and other systems that use ionospheric signals, because the ionospheric layer affects the mentioned signals. Prediction of VTEC values helps one to be prepared of future calamities and the research work focus on this point. This chapter presents the various techniques used for predicting VTEC values over time and compares their efficiency in terms of accuracy and speed.

General Approach

The general methodology used in the present study in shown in Figure 2.

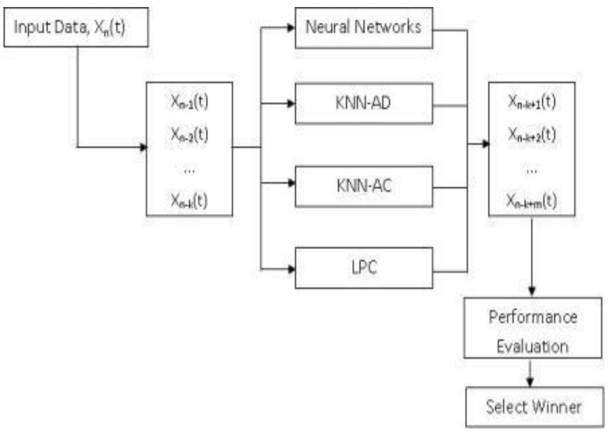

Figure 2: Proposed Methodology

The research methodology consists of three stages.

Stage 1: Preparation of input data into training and testing set
Stage 2: Prediction of future values
Stage 3: Comparison of the results with respect to efficiency in prediction

In stage 1, given an input dataset (X) with 'n' VTEC values obtained over a period of time't', the proposed methodology first groups X into two sets, namely, training dataset and testing dataset. To make meaningful forecasts, the predictors has to be trained on an appropriate data series. Data in the form of <input, output> pairs are extracted from X, where input and output are vectors equal in size to the number of network inputs and outputs, respectively. In the present research work, a 70% and 30% division was adopted, that is, 70% of the records in X is taken as training data and the rest of the 30% is taken as testing data.

Throughout the literature, many techniques have been implemented to perform time series forecasting. This study focus on three techniques: back propagation neural networks, K Nearest-Neighbor and linear predictive coding. In stage 2, the selected three predictors are used to predict of next 'm' VTEC values. Initially the predictors are trained using the training set, which are then verified using the testing set.

In stage 3, ten performance metrics to evaluate the efficiency of the three predictors. They are, Mean Squared Error (MSE), Normalised Mean Squared Error (NMSE), Root Mean Squared Error (RMSE), Normalised Root Mean Squared Error (NRMSE), Mean Absolute Error (MAE), Mean Absolute Relative Error (MARE), Coefficient of Correlation (R), Coefficient of Determination (D), Coefficient of Efficiency (E), Maximum Absolute Error (A), Maximum Absolute Relative Error (AE) and execution speed. From the results, a performance comparison is conducted to select a predictor among the three that produces accurate results, that is, to pick out a winner that produces least error in a time saving manner.

Proposed Predictive Models

The main purpose of predictive model is the discovery of valuable information about the probability distribution that generated the data by making predictions about new incoming data. For example, a regression model will be used for predicting the value of the response variable y for any new value of the predictor x. Even descriptive models make predictions. A histogram predicts the value of the density of a distribution where a new observation happens to fall.

K Nearest Neighbor (KNN) Method

K-nearest neighbors algorithm (k-NN) is a method for classifying objects based on closest training examples in the feature space. K-NN is a type of instance-based learning, or lazy learning where the function is only approximated locally and all computation is deferred until classification. The k-nearest neighbor algorithm is amongst the simplest of all machine learning algorithms: an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its k nearest neighbors (k is a positive integer, typically small). If k = 1, then the object is simply assigned to the class of its nearest neighbor.

In the KNN Predictor, the final data points of the data series are called the 'reference' and the length of the reference is called the 'window size'. The data series without the last data point is the 'shortened data series'. To forecast the data series' next data point, the reference is compared to the first group of data points in the shortened data series, called a 'candidate', and an error is computed.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Then the reference is moved one data point forward to the next candidate and another error is computed and the process is repeated for all data points. All errors thus calculated are stored and sorted. The smallest k errors correspond to the k candidates that closest match the reference. Finally, the forecast will be the average of the k data points that follow these candidates. Then, to forecast the next data point, the process is repeated with the previously forecasted data point appended to the end of the data series. This can be iteratively repeated to forecast n data points. This process is explained in Figure 3.

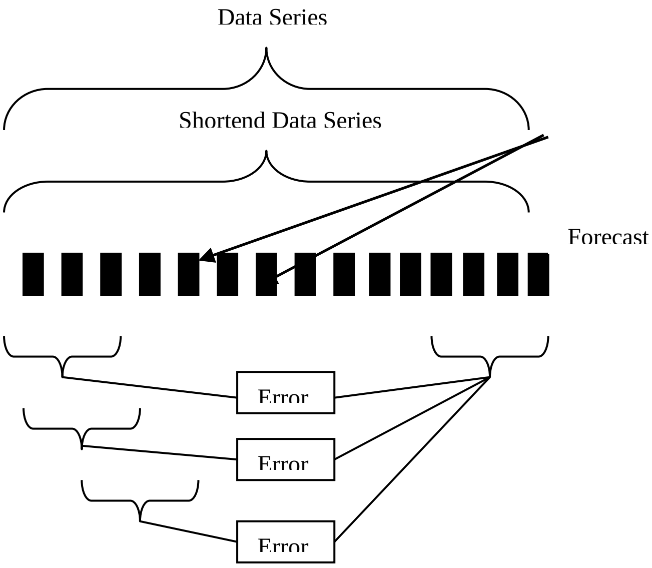

Figure 3: K Nearest Neighbor Forecasting (Window Size = 4, K=2)

Linear Predictive Coding (LPC)

Linear prediction is a technique of time series analysis that emerges from the examination of linear systems. Using linear prediction, the parameters of a future system can be determined by analyzing the systems inputs and outputs.
In proposed LPC method the input data series VTEC represent as a column vector or a matrix with series organized in columns. We have to defines the number of predictor coefficients to use (>=2).In this method we have to specify the number of data values to return in output. It is necessary to specify the position that is the string ‘pre’ or ‘post’. This determines whether extrapolation occurs before or after the observed series x.
Given x(n-1), x(n-2), , x(n-M), the problem here is to predict the value of VTEC denoted as x(n). In LPC, this predicted VTEC value can be expressed as a linear function of the given M past samples (Equation 1).

$$X(n \mid n-1, n-2, \ldots n-M) = \psi(x(n-1), x(n-2), \ldots x(n-m)) \quad (1)$$

When a value is predicted using the above equation, then it is said to be predicted linearly.

Neural Networks Model

A neural network is a computational model that is loosely based on the neuron cell structure of the biological nervous system. Given a training set of data, the neural network can learn the data with a learning algorithm; in this research, the most common algorithm, backpropagation, is used. Through backpropagation, the neural network forms a mapping between inputs and desired outputs from the training set by altering weighted connections within the network. A brief history of neural networks follows.

Backpropagation Training For Vtec Prediction

The backpropagation training (Gonzalez and Woods, 2008) stage uses the training dataset from Stage 1 and train the network in of three steps,

1. Present an input vector VTEC to the network for training. Compute activation functions sequentially forward from the first hidden layer to the output layer (Figure 4.2, from layer A to layer C).
2. Compute the difference between the desired output, and the actual network output (output of unit(s) in the output layer). Propagate the error sequentially backward from the output layer to the first hidden layer (Figure 4.2, from layer C to layer A).
3. For every connection, change the weight modifying that connection in proportion to the error.

When these three steps have been performed for every input from the training data set, one epoch occurs. Training usually lasts thousands of epochs, possibly until a predetermined maximum number of epochs (epochs limit) is reached or the network output error (error limit) falls below an acceptable threshold. Training can be time-consuming, depending on the network size, number of examples, epochs limit, and error limit.

IV. RESULTS AND DISCUSSION

The present research work proposes the use of neural networks, K Nearest Neighbor and linear predictive coding for predicting VTEC values using time series analysis. This chapter obtained while testing the proposed 3 predictors for efficiency is presented in this Chapter.

Dataset Used

The TEC data is obtained by using the dual frequency of GPS receivers installed for the satellite navigation project of ISRO called GAGAN. These TEC data recorded by the Agatti satellite on January 2008 with the elevation angle 60^0. TEC dataset used in this research work contains 1065 records with the columns IST time, latitude, longitude, VTEC (Vertical Total Electron Content) and STEC (Slant Total Electron Content).VTEC values are taken for predictions. From the dataset 70% are taken for training and 30% are taken for testing.

PERFORMANCE METRICS

Several experiments were conducted to evaluate the proposed predictors to analyze their efficiency in terms of accuracy in prediction and time taken to predict. Importance is given to accuracy parameter which is analyzed using several error metrics which gives the error difference between the predicted and original values. The performance metrics used during experimentation are explained in this section.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Mean Squared Error (MSE)

The mean squared error (MSE) of a predictor is one of many ways to quantify the difference between values implied by an predictor and the true values of the values being estimated. MSE is a risk function, corresponding to the expected value of the squared error loss or quadratic loss. MSE measures the average of the squares of the "errors." The error is the amount by which the value implied by the estimator differs from the quantity to be estimated. It is calculated using Equation (1).

$$\text{MSE(h)} = \frac{1}{N-h-m+1}\sum_{i=m}^{N-h}(x_{i+h} - x_i(h))^2 \qquad (1)$$

where x_{i+h} is the actual sample and $x_i(h)$ is the h-step ahead prediction at current time i, for i=m,...,N-h, and N is the length of the time series.

Normalised Mean Squared Error (NMSE)

The normalized mean square error (NMSE) measure of goodness of fit is the MSE divided by the variance of the samples included in the sum of MSE (Equation 2).

$$\text{NMSE}(h) = \frac{\sum_{i=m}^{N-h}(x_{i+h} - x_i(h))^2}{\sum_{i=m}^{N-h}(x_{i\_h} - \bar{x})^2} \qquad (2)$$

Where $\bar{x}$ is the mean of the samples in the sum. Where x_{i+h} is the actual sample and $x_i(h)$ is the h-step ahead prediction at current time i, for i=m,...,N-h, and N is the length of the time series

V. EXPERIMENTAL RESULTS

This section presents the results obtained for each of the performance metrics used.

Mean Squared Error (MSE)

Table 1 shows the MSE values obtained by the three predicting algorithms.

TABLE 1 COMPARISON OF MEAN SQUARED ERROR

Predictor	MSE
BPNN	0.0022
LPC	9.4391
KNN	1.6114

From the results, it could be seen that the LPC predictor performance is very poor when compared with other three predictors. The best performance in terms of prediction accuracy using MSE metrics is given by BPNN. Among the KNN algorithms, the performance of KNN-AC is better than KNN-AD. Thus, from the table, it is clear that BPNN is the best followed by KNN-AC, KNN-AD and LPC.

Normalised Mean Squared Error (NMSE)

The second performance metric used for analysis is NMSE and the results obtained are projected in Figure 5.1.From figure, it can be seen that the BPNN based predictor produces the best performance (0.0001 NMSE) followed by KNN-AC (0.1031) and KNN-AD (0.1037). The worst performance was given by the LPC predictor with a high NMSE value of 0.6071. Even though the difference between KNN-AC and KNN-AD seems to be very small, while considering the efficiency gain obtained, the KNN-AC shows a 0.52% increase, which is very important in time critical prediction environments like ionosphere.

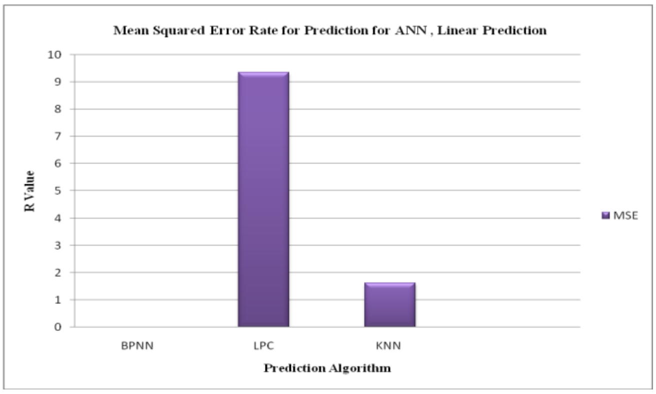

Figure 5.1: Comparison of Normalized Mean Squared Error

VI. CONCLUSION

This research introduced four prediction algorithm based on Neural Networks, Linear Predictive Coding and K Nearest Neighbor for predicting VTEC values over time. Various experiments were conducted with the primary aim of identifying an algorithm, among the four, that produces accurate VTEC values. From the results, it could be seen the BPNN algorithm shows maximum improvement in terms of accuracy and lower error rate and is the best choice for predicting VTEC values.

REFERENCES

1. Acharya, R., Roy, B., Sivaraman, M.R. and Dasgupta, A. (2009) Kalman Filter Approach for Prediction of Ionospheric Total Electron Content, International Conference on Computers and Devices for Communication, 1-4.
2. Adya, M. and Collopy, F. (1998) How effective are neural networks at forecasting and prediction. A review and evaluation. Journal of Forecasting, 17, 481–495.
3. Balkin, S. D. and Ord, J. K. (2000) Automatic neural network modeling for univariate time series. International Journal of Forecasting, 16, 509–515.
4. Bates, J. M. and Granger, C.W.J. (1969) Combination of forecasts. Operations Research Quarterly, 20, 451–468.
5. Bhansali, R.J. (1996) Asymptotically efficient autoregressive model selection for multistep prediction, Annals of the Institute of Statistical Mathematics, Vol 48, 577–602.
6. Chatfield, C. (1988) What is the 'best' method of forecasting? Journal of Applied Statistics, 15, 19–38.
7. Coomes, P. A. (1992) A Kalman filter formulation for noisy regional job data. International Journal of Forecasting, 7, 473–481.
8. Coster, A. and Komjathy, A. (2008) Space weather and the global positioning system, Space Weather, Vol, 6, Article No. S10D07.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

9. Darbellay, G. A. and Slama, M. (2000) Forecasting the short-term demand for electricity: Do neural networks stand a better chance? International Journal of Forecasting, 16, 71– 83.
10. Geriner, P. T. and Ord, J. K. (1991) Automatic forecasting using explanatory variables: A comparative study. International Journal of Forecasting, 7, 127– 140.
11. Gomez, V. and Maravall, A. (2001) Seasonal adjustment and signal extraction in economic time series. In D. Pen˜a, G. C. Tiao and R. S. Tsay (Eds.), Chapter 8 in a course in time series analysis. New York, John Wiley and Sons.
12. Gonzalez, R.C. and Woods, R.E. (2008) Digital Image Processing, 3rd ed., Prentice Hall, Upper Saddle River, NJ.
13. Habarulema, J.B., McKinnell, L.-A., Cilliers, P.J., Opperman, B.D.L. Application of neural networks to South African GPS TEC modelling. Adv. Space Res., 43(11), 1711–1720. doi:10.1016/j.asr.2008.08.020, 2009.
14. Hafer, R. W. and Sheehan, R. G. (1989) The sensitivity of VAR forecasts to alternative lag structures. International Journal of Forecasting, 5, 399– 408.
15. Hernández-Pajares, M., Juan, J.M., Sanz, J. and Colombo, O.L. (2000) Application of ionospheric tomography to real-time GPS carrier-phase ambiguities resolution, at scales of 400–1000 km and with high geomagnetic activity, Geophys. Res. Lett., 27(13), 2009–2012.
16. Ladiray, D. and Quenneville, B. (2004) Implementation issues on shrinkage estimators for seasonal factors within the X-11 seasonal adjustment method. International Journal of Forecasting, 20, 557– 560.
17. Leandro, R.F., Santos, M.C. (2007) A neural network approach for regional vertical total electron content modelling. Stud. Geophys. Geod. 51 (2), 279–292.
18. Novales, A. and de Fruto, R. F. (1997) Forecasting with time periodic models: A comparison with time invariant coefficient models. International Journal of Forecasting, 13, 393– 405.
19. Olson, D. and Mossman, C. (2003) Neural network forecasts of Canadian stock returns using accounting ratios. International Journal of Forecasting, 19, 453– 465.
20. Thompson, P. A. (1991) Evaluation of the M-competition forecasts via log mean squared error ratio. International Journal of Forecasting, 7, 331– 334.
21. Tkacz, G. (2001) Neural network forecasting of Canadian GDP growth. International Journal of Forecasting, 17, 57– 69.
22. Tulunay, E., Senalp, E., Radicella, S., Tulunay, Y. Forecasting total electron content maps by neural network technique. Radio Sci. 41, doi:10.1029/2005RS003285, 2006.
23. Ullah, T. A. (1993) Forecasting of multivariate periodic autoregressive moving-average process. Journal of Time Series Analysis, 14, 645– 657.
24. Vinod, H. D. and Basu, P. (1995) Forecasting consumption, income and real interest rates from alternative state space models.International Journal of Forecasting, 11, 217– 231.
25. Wang, Z. and Bessler, D. A. (2004) Forecasting performance of multivariate time series models with a full and reduced rank: An empirical examination. International Journal of Forecasting, 20, 683–695.
26. Wells, J. M. (1997) Modelling seasonal patterns and long-run trends in U.S. time series. International Journal of Forecasting, 13, 407– 420.
27. West, K.D. (1996) Asymptotic inference about predictive ability, Econometrica, 68, 1084–1097.
28. West, M., Harrison, P. J. and Migon, H. S. (1985) Dynamic generalized linear models and Bayesian forecasting (with discussion) Journal of the American Statistical Association, 80, 73– 83.
29. Wieringa, J. E. and Horva´th, C. (2005) Computing level-impulse responses of log-specified VAR systems. International Journal of Forecasting, 21, 279– 289.
30. Winkler, R. L. and Makridakis, S. (1983) The combination of forecasts. Journal of the Royal Statistical Society (A), 146, 150– 157.

AUTHOR PROFILE

S.Priya , Assistant Professor in Theivanai Ammal College for Women ,Villupuram, recieved Master of Computer Applications fromTheivanai Ammal College for Women and completed Master of Philosophy in Avinashilingam university for women, Comibatore. Area of interests are Data mining and Data structures.

A.Parameswari, Assistant Professor in Theivanai Ammal College for Women ,Villupuram, recieved Master of Computer Applications from Theivanai Ammal College for Women, Villupuram and completed Master of Philosophy in Avinashilingam university for women,Comibatore. Area of interest are Data mining and software Engineering.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)