Journal of Statistics & Management Systems
ISSN 0972-0510 (Print), ISSN 2169-0014 (Online)
DOI : 10.47974/JSMS-1553

Prediction of Ionospheric TEC based on EEMD-Prophet-BiLSTM Model

Anjan Kumar Das [†]
Santanu Kalita *
Surajit Nath [§]
Department of Computer Application
Mahapurusha Srimanta Sankaradeva Viswavidyalaya
Nagaon
Assam
India

Bornali Chetia [‡]
Department of Physics
Royal Global University
Guwahati
Assam
India

Abstract

Predicting ionospheric GPS Total Electron Content (TEC) data is quite challenging due to the need for processing large datasets and handling variations across time and space. Machine learning techniques have increasingly been adopted in recent years for predicting complex time series characterized by nonlinear and non-stationary behavior. This work emphases on using machine learning techniques to predict GPS-derived ionospheric TEC. A unique hybrid version combining EEMD (Ensemble Empirical Mode Decomposition), facebook's Prophet, and BiLSTM is proposed to forecast TEC values for distinctive GNSS stations in India. The model demonstrated reliable performance during periods of both

[†] *ORCID-ID:* **0009-0009-9425-1719**
* *ORCID-ID:* **0000-0003-2715-9225**
[§] *ORCID-ID:* **0000-0002-7355-6013**
[‡] *ORCID-ID:* **0009-0002-7565-8547**

[†] *E-mail:* **anjandasmssv@gmail.com**
* *E-mail:* **santanu.kalita@rediffmail.com (Corresponding Author)**
[§] *E-mail:* **surajitnath2@gmail.com**
[‡] *E-mail:* **bornalichetia.physics@gmail.com**

© Taru Publications

minimal and heightened geomagnetic activity. Its effectiveness is evaluated using various performance metrics and compared with existing models.

Subject Classification: *37M10, 68T07.*

Keywords: *GPS, TEC, EMD, Fbprophet, BiLSTM.*

1. Introduction

For the effective planning, operation, and maintenance of critical technologies such as RADAR, communication systems, and navigation services, accurate forecasting of the Earth's ionosphere is essential. These systems specially those relying on radio frequency (RF) transmissions and global navigation satellite systems (GNSS), are highly sensitive to the effects of space weather. The transionospheric signals in the L-band from RF and GNSS sources pass via the ionosphere and experience signal delays caused by adjustments inside the ionosphere's refractive index. Mannucci et al. [1] reported that these delays are primarily influenced by the ionospheric Total Electron Content (TEC), which are influenced by factors such as time of day, seasonal and annual cycles, and geomagnetic conditions. Global Navigation Satellite System (GNSS) signals have proven to be effective in monitoring the dynamic behavior of the ionosphere through both terrestrial and satellite-based receivers. These signals when passing through the ionosphere experience delays that are influenced by fluctuations in electron density. Such delays, referred to as ionospheric delays are particularly significant in the L-band frequency range and can affect various communication and navigation systems.

The availability of historical ionospheric dataset acquired under diverse space weather environment have facilitated the practice of deep learning techniques for understanding and forecasting ionospheric dynamics. Two key indicators of ionospheric activity — solar flux (F10.7 cm) and geomagnetic disturbances (Dst index) are commonly used to model and forecast these delays. However, due to the stochastic and highly variable nature of ionospheric behavior, especially in equatorial and low-latitude regions developing robust prediction models remains a challenge.

Recent studies have explored a variety of deep learning techniques to predict Total Electron Content (TEC) and other related parameters. These include traditional neural networks (NN) (see Sivavaraprasad et al. [2], Tulunay et al. [3], Mallika et al. [4], convolutional neural networks (CNN) Boulch, Cherrier, and Castaings [5]), and long short-term memory

networks (LSTM) (see Sun et al. [6], Tan et al. [7], Pala and Atici [8]). In recent years, incorporation of recurrent neural networks (RNNs) with CNNs has been developed to effectively model both temporal and spatial variations in ionospheric TEC data (see Boulch, Cherrier, and Castaings [9]). While recurrent neural networks (RNNs) are commonly used for sequential data analysis, however their performance is often limited due to the vanishing gradient issue, which hinders the capturing of long-term dependencies. These limitations were overcome by using Long Short-Term Memory (LSTM) networks which introduce memory cells and gating mechanism including the constant error carousels which enable the model to retain relevant information over extended time periods.

In addition to machine learning signal processing techniques like Empirical Mode Decomposition (EMD) have been utilized to analyze non-linear and non-stationary TEC time series. Yadav, Pal, and Tripathi [10] applied EMD with a back-propagation neural network for short-term photovoltaic power forecasting. Bairagi and Harpale [11] utilized singular spectrum EMD combined with machine learning techniques for accurate epileptic seizure detection. The EMD algorithm is often affected by mode mixing problem due to noise, which limits its effectiveness. To overcome this challenge, Ensemble Empirical Mode Decomposition (EEMD) has emerged as an enhanced version of EMD by incorporating white noise and performing repeated decompositions whose averaged outcome significantly suppresses mode mixing.

The previous research has highlighted the influence of seasonal variation on TEC patterns (see Song et al. [12], Tebabal et al. [13, 14], Mukesh et al. [15], Ruwali et al. [16]). We consider the limitations of using a single predictive model in this current study and adopts a hybrid methodology. It combines EEMD for signal decomposition, Facebook Prophet for trend forecasting, and a Bidirectional LSTM (BiLSTM) model for learning both forward and backward temporal dependencies. This integrated framework aims to improve the accuracy of TEC forecasts using GNSS data from multiple monitoring stations across India.

2. Data and Analysis

In this present work, Total Electron Content (TEC) data from various GNSS stations across India were utilized. The raw data collected from these stations were first corrected for instrumental bias and then converted into .csv format for further analysis. To transform slant TEC into vertical

TEC, matrix inversion and smoothing techniques were applied. The input features used for training the hybrid model included:

(i) TEC measurements obtained from dual-frequency GNSS receivers operating at 1.575 GHz and 1.227 GHz
(ii) Geomagnetic indices (Kp, Ap, Dst)
(iii) Solar flux (F10.7 index)

2.1 *EEMD Algorithm*

Empirical Mode Decomposition (EMD) is a data-driven signal processing method that decompose time series data into characteristic components, such as intrinsic mode functions (IMFs) and residuals, based on distinct characteristic scales. However, despite its effectiveness in reducing continuous noise such as Gaussian noise, EMD exhibits limited capability in dealing with intermittent and mixed noise components. To overcome these shortcomings, an enhanced approach known as Ensemble Empirical Mode Decomposition (EEMD) was proposed by Wu and Huang [17]. The technique adds different realizations of random noise to the original signal and performs multiple decomposition trials, which aids in reducing mode mixing and enhancing scale separation. By averaging the resulting IMFs from multiple trials, the additional noise can be eliminated, assuming a sufficient number of trials.

The steps for constructing the Ensemble Empirical Mode Decomposition (EEMD) model for Total Electron Content (TEC) data are as follows:

The EEMD approach starts through introducing white noise to the original TEC signal, developing a new version of the data The extrema of the resulting signal are then identified, and cubic spline interpolation is used to form upper and lower envelopes. This average is then subtracted from the modified signal to extract a component, which is evaluated against predefined standards to decide whether it qualifies as an Intrinsic Mode function (IMF). If it does, the extracted mode is saved as an IMF, and the residual signal is updated; if not, the modified signal is reprocessed. These steps are repeated until no further IMFs can be extracted. The entire process is then performed multiple times with different sets of random noise added to the original signal in each iteration. On the final step all the IMFs obtained throughout different runs are averaged to produce a more stable and correct decomposition of the TEC signal.

2.2 *Prophet Algorithm*

Taylor and Letham [18] proposed Prophet, a time series forecasting tool which is designed to effectively captures nonlinear trends and multiple seasonal patterns in temporal datasets. It employs an additive modeling approach, where the overall prediction is expressed as the sum of numerous parameters: trend, seasonality, and effects from holidays or special events. Prophet is capable of managing missing values and outliers effectively, making it suitable for real-world time series data. It also accounts for trend shifts, commonly known as change points. The forecasting process generally involves four stages: model construction, evaluation of predictions, issue resolution, and manual review. It utilizes a Bayesian modeling framework to estimate parameters, enabling it to capture uncertainty in forecasts. As a result, the output includes confidence intervals, providing both upper and lower bounds for the predicted values. Specifically, the time series is decomposed as: $X_t=g_t + s_t+ h_t+ \varepsilon_t$ where g_t captures the long-term trend, s_t accounts for periodic seasonal effects (daily, weekly, or yearly), h_t models the impact of holidays or special events, and ε_t is the error. By modeling trend, seasonality, and holiday effects as separate components, Prophet provides a flexible framework for time series forecasting. This allows it to detect and adjust for irregularities such as sudden changes and recurring patterns in the data. Additionally, it incorporates both seasonal and event-based influences, providing a reliable and effective solution for handling complex forecasting tasks across various domains.

2.3 *Bidirectional Long Short-Term Memory (BiLSTM)*

Although Long Short-Term Memory (LSTM) networks are effectively used in capturing temporal dependencies in sequential data, they generally analyze information only in the forward direction, which may lead to missing important context from the reverse temporal sequence. This limitation becomes mainly important in the analysis of ionospheric total Electron content (TEC), in which both past and future contextual information influence current values. To overcome this constraint, the Bidirectional long short-term memory (BiLSTM) architecture is employed, presenting a more holistic method to sequence modeling.

Bidirectional LSTM (BiLSTM) enhances the traditional LSTM architecture by incorporating two parallel layers: a forward LSTM, which processes the input sequence in its natural chronological order to capture past dependencies, and a backward LSTM, which reads the sequence in

reverse to extract future dependencies. This bidirectional architecture facilitates the learning of temporal structure from the TEC data, thereby increasing the reliability of predictions

For each time step *t*, the BiLSTM model generates forward and backward hidden states using the following relations:

$$\overrightarrow{h_t} = \sigma(W_{F1xt} + W_{F2}\overrightarrow{h_{t-1}} + \vec{b})$$

$$\overleftarrow{h_t} = \sigma(W_{B1xt} + W_{B2}\overleftarrow{h_{t-1}} + \overleftarrow{b})$$

Here, $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ signify the hidden states obtained from the forward and backward LSTM layers, respectively. The input at the t-th time step is denoted by *xt*, while W and b represent the trainable weight parameters and bias vectors for each direction. The final output y_t at time step t is a weighted combination of these two directional outputs:

$$y_t = W_{O1}\overrightarrow{h_t} + W_{O2}\overleftarrow{h_t} + b_y$$

This architecture is especially effective in TEC forecasting, in which capturing each historical and future trends enhances the model's capability to predict complex ionospheric behavior inspired via solar and geomagnetic situations.

2.4 *Model Construction*

In this study, a hybrid prediction framework has been developed to forecast Total Electron Content (TEC) using time series data with a 10-minute average, that were filtered using commonly employed geomagnetic activity indicators namely the Kp, Ap and Dst index. The proposed model integrates Ensemble Empirical Mode Decomposition (EEMD), Facebook Prophet (Fbprophet), and Bidirectional Long Short-Term Memory (BiLSTM) networks to improve prediction accuracy. The methodology involves the folowing steps:

1. **Preprocessing:** The raw TEC data are first pre-processed to fit to the input structure required by the EEMD algorithm. This stage prepares the data for subsequent decomposition and modeling.
2. **Decomposition:** Using EEMD technique, the pre-processed time series is separated into several intrinsic mode functions (IMFs) and a residual component R_n. The resulting components consisted of subseries: IMF1, IMF2, ..., IMF_{n-1} and R_n.

3. **Fbprophet Prediction:** Each decomposed component, including the residual, is individually forecasted using the Fbprophet model, which accounts for seasonal trends, holidays, and changepoints. The outputs are then aggregated to produce a combined TEC prediction.
4. **BiLSTM Prediction:** In parallel, the decomposed components are also modelled using the BiLSTM network, which captures bidirectional temporal dependencies within the data. The resulting forecasts from each BiLSTM submodel are then integrated.
5. **Fusion:** To generate the final forecast, the predictions from the Fbprophet and BiLSTM models are merged using a weighted combination technique. This fusion strategy leverages the strengths of both statistical and deep learning approaches to produce a more robust forecast.
6. **Evaluation:** The performance of the hybrid forecasting framework is assessed by comparing the predicted TEC values against observed data. Key evaluation metrics include Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and the coefficient of determination (R^2). These indicators provide comprehensive assessment of the model's reliability and predictive capability during both testing and validation phases. The overall workflow of the proposed framework is illustrated in Figure 1.

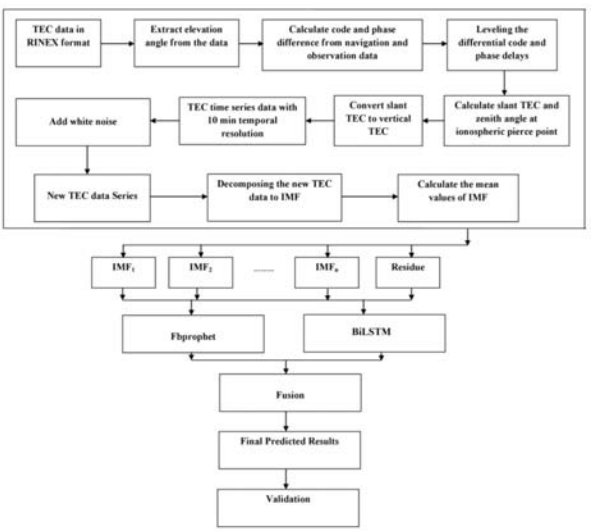


Figure 1

The architecture of Fbprophet-BiLSTM-EEMD prediction model

3. Experimental Results

After preprocessing, the TEC dataset is subsequently analyzed using the Ensemble Empirical Mode Decomposition (EEMD) method into nine intrinsic mode functions (IMFs).

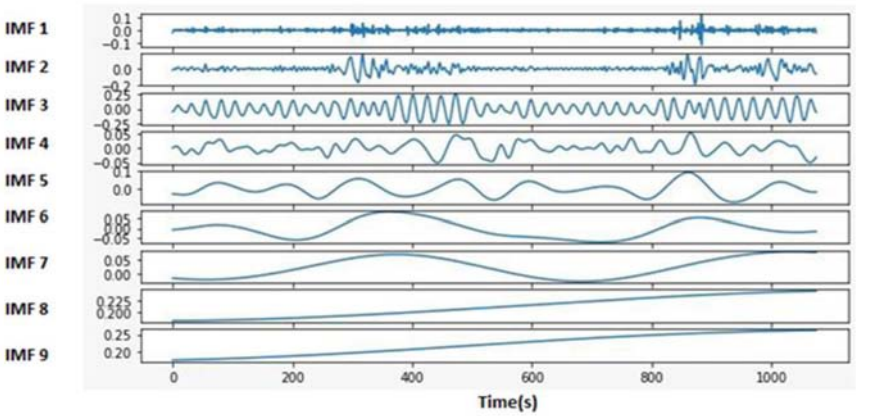


Figure 2

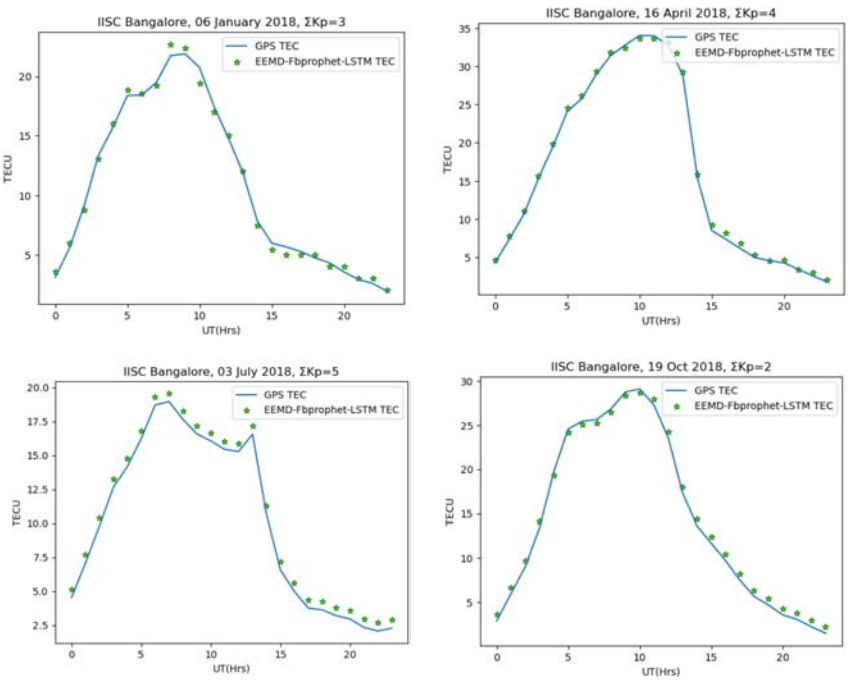
EEMD decomposition results

Figure 2 represents nine Intrinsic Mode functions (IMFs), ordered by their frequencies which capture the periodic behavior of various influencing elements in the TEC time series. The higher-frequency components—namely IMF1, IMF2, and IMF3—reflect short-term fluctuations within the original series, while IMF4 reveals more evident periodic patterns. The lower-frequency IMFs (IMF5 through IMF9) represent long-time period trends and slower variations. these decomposed components collectively highlight the presence of multi-scale affects affecting the TEC signal.

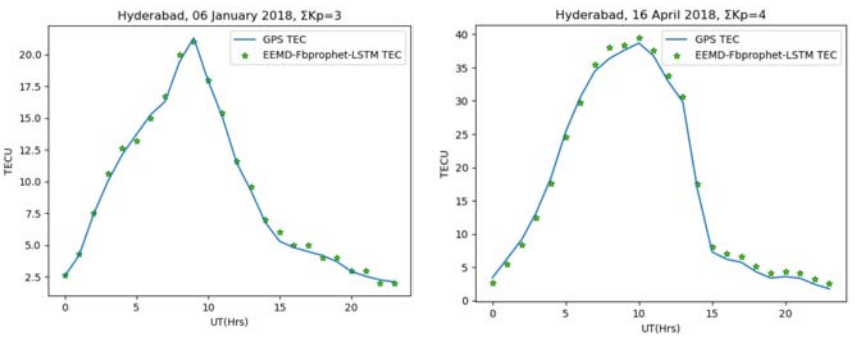
Considering the spatial variability of **Total Electron Content (TEC)**, data from three geographically distinct GNSS stations in India **IISc Bangalore**, **Lucknow**, and **Hyderabad** - have been selected for analysis. The dataset spans a decade (2008–2018) and is divided based on low and high solar activity period for model training, validation and performance evaluation of the forecasting model.

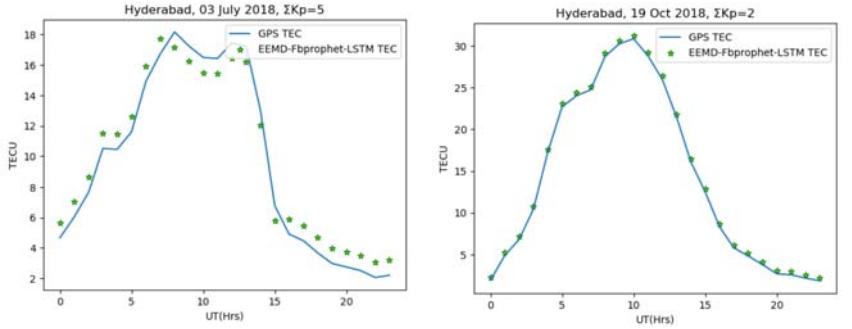
Previous studies have emphasized the significance of seasonal variations on TEC dynamics. To account for this, the current study evaluates the model's overall performance throughout four seasonal

categories: **winter season** (December to February), **march equinox** (March to May), **summer** (June to August), and **September equinox** (September to November), for periods of both weak and strong solar activity conditions. Figure 3 (a - c) compares the forecasted TEC values with GPS-derived observations for the year 2018 at the three selected stations across the identified seasons.

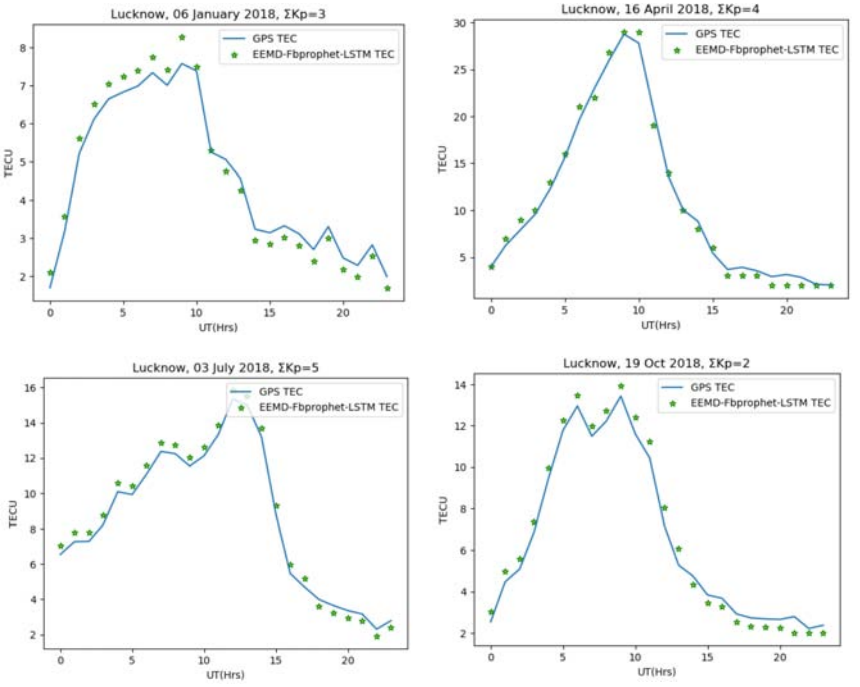


(a)





(b)

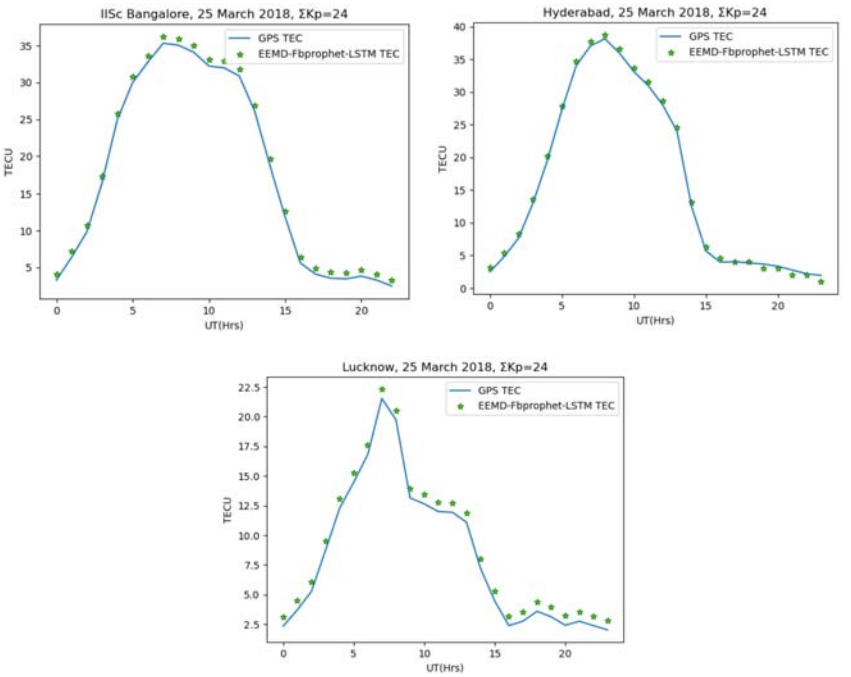


(c)

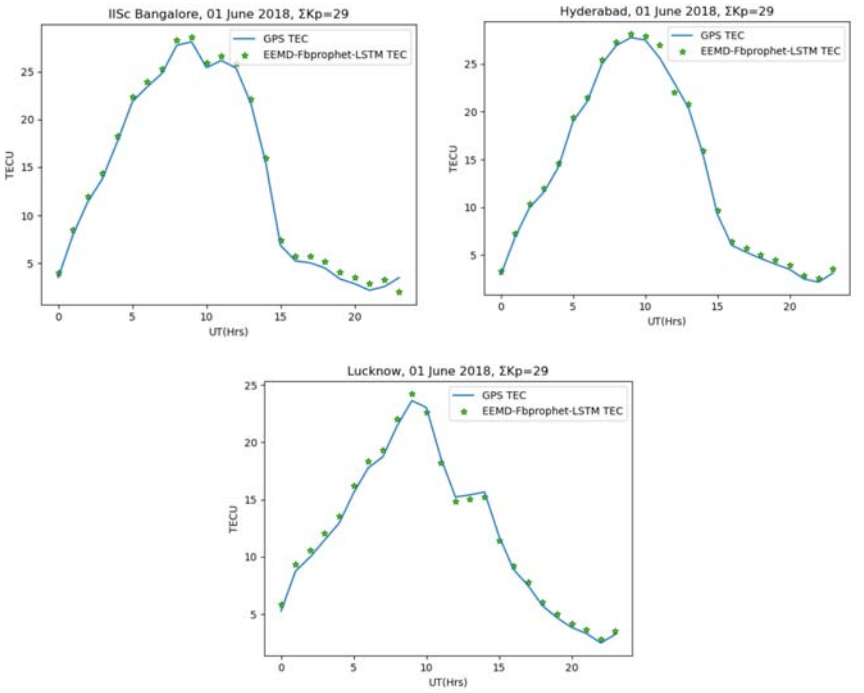
Figure 3

(a,b,c): Comparison of model-predicted and observed quiet day TEC profiles during different periods of 2018 at: (a) IISc Bangalore, (b) Hyderabad and (c) Lucknow.

As a standard measure of geomagnetic disturbances, the Kp index is derived from 3-hourly magnetic field variations recorded by magnetometers located at different sites around the world. A day is classified as geomagnetically quiet when the cumulative Kp value is below 20. Based on experimental results, the proposed hybrid EEMD–Fbprophet–BiLSTM model demonstrates robust performance across all seasonal conditions—winter, vernal, summer, and autumnal equinox—in predicting quiet day TEC variations at all three observation sites: IISc Bangalore, Hyderabad, and Lucknow. A comparison of the model-predicted and observed TEC profiles for disturbed days at these stations in 2018 is shown in Figure 4. The analysis show that the hybrid model can capture TEC variations during geomagnetic disturbances, like storm-induced ones, at all chosen locations, in addition to being effective during quiet geomagnetic conditions.



(a)



(b)

Figure 4

Comparison of model-predicted and observed disturbed day TEC profiles during on: (a) 25 March 2018 and (b) 01 June 2018, across three GNSS stations: IISc Bangalore, Hyderabad and Lucknow.

To evaluate the effectiveness of the proposed EEMD-Fbprophet-BiLSTM framework, its prediction results are benchmarked against those of EEMD-LSTM, EMD-LSTM, standalone Prophet and standalone LSTM models. Here the model performance is evaluated using RMSE, MAPE and R^2 as evaluation criteria. Lower values of RMSE and MAPE, along with higher R^2 values, indicate better predictive accuracy. It is essential to recognize that no single metric can fully represent model performance. Hence, a comprehensive evaluation framework incorporating multiple statistical indicators is adopted. In this study, we employ RMSE, MAPE, and R^2 to provide a robust assessment of the forecasting models. The mathematical formulations of these metrics are presented below:

$$MAPE = \frac{\sum_{i=1}^{N} \frac{\left|TECP_i - TECO_i\right|}{TECO_i}}{N} \times 100$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(TECP_i - TECO_i)^2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(TECP_i - TECO_i)^2}{\sum_{i=1}^{N}(TECP_i - \overline{TEC})^2}$$

Here, N denotes the total number of observations, $TECP_i$ represent the forecasted TEC value, and $TECO_i$ correspond to observed TEC value and $\overline{TEC}$ is the mean TEC value. Table 1 represents the comparison of RMSE, MAPE, and R^2 values using different models.

Table 1

Performance comparison of proposed model with other existing models based on evaluation metrics

Models	MAPE	RMSE	R^2
EEMD- Fbprophet-BiLSTM	1.0032	0.6020	0.9989
EEMD LSTM	1.3112	0.7011	0.9969
EMD LSTM	1.5912	0.9621	0.9188
Prophet	2.0183	1.1871	0.8161
LSTM	2.1626	1.2155	0.8466

A comparative analysis of MAPE, RMSE, and R^2 values across various prediction models indicates that the proposed EEMD-Fbprophet-BiLSTM model demonstrates superior performance and higher predictive accuracy compared to the other models.

4. Conclusion

This study presents a new hybrid **EEMD-Fbprophet-BiLSTM** framework for forecasting GPS TEC data across multiple GNSS stations in India. The experimental analysis signify that the model effectively captures TEC dynamics under both geomagnetically quiet and disturbed conditions. Its predictive capability is assessed using performance metrics such as

MAPE, RMSE, and **R^2**, and comparisons reveal that it consistently outperforms existing models including EEMD-LSTM, EMD-LSTM, Prophet, and standalone LSTM approaches. Furthermore, this hybrid framework shows potential for broader applications in areas like weather prediction, stock market analysis, and sales forecasting.

Acknowledgements

The authors gratefully acknowledge the availability of high-quality data from organizations such as the International GNSS Services (IGS), UNAVCO, NOAA, and the World Data Center for Geomagnetism, Kyoto, which significantly contributed to this study.

References

[1] A. J. Mannucci, B. D. Wilson, D. N. Yuan, C. M. Ho, U. J. Lindqwister, and T. F. Runge, "A global mapping technique for GPS-derived ionospheric total electron content measurements," *Radio Science*, vol. 33, no. 3, pp. 565–582 (1998), doi: 10.1029/97RS02707.

[2] G. Sivavaraprasad, V. S. Deepika, D. Sreenivasa Rao, 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. 2, pp. 127–136 (2020), doi: 10.1016/j.geog.2019.11.002.

[3] Y. Tulunay, E. T. Senalp, E. Tulunay, and S. M. Radicella, "Forecasting total electron content maps by using neural network," *Radio Science*, vol. 39, no. 1, Art. no. RS1010 (2004), doi: 10.1029/2002RS002854.

[4] I. L. Mallika, D. V. Ratnam, Y. Otsuka, G. Sivavaraprasad, and S. Raman, "Implementation of hybrid ionospheric TEC forecasting algorithm using PCA-NN method," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 1, pp. 371–381 (2019), doi: 10.1109/JSTARS.2018.2881791.

[5] A. Boulch, N. Cherrier, and T. Castaings, "Ionospheric activity prediction using convolutional recurrent neural networks," *arXiv* preprint arXiv:1810.13273 (2018). [Online]. Available: https://arxiv.org/abs/1810.13273.

[6] W. Sun, Y. Wang, B. Xu, Z. Wang, and Y. Yuan, "Forecasting of ionospheric vertical total electron content (TEC) using LSTM networks,"

in *Proc. Int. Conf. Machine Learning and Cybernetics (ICMLC)*, pp. 340–344 (2017).

[7] Y. Tan, Q. Hu, Z. Wang, and Q. Zhong, "Geomagnetic index Kp forecasting with LSTM," *Space Weather*, vol. 16, no. 4, pp. 406–416 (2018), doi: 10.1029/2017SW001764.

[8] Z. Pala and R. Atici, "Forecasting sunspot time series using deep learning methods," *Solar Physics*, vol. 294, no. 5, Art. no. 50 (2019), doi: 10.1007/s11207-019-1449-8.

[9] A. Boulch, N. Cherrier, and T. Castaings, "Ionospheric activity prediction using convolutional recurrent neural networks," *arXiv* preprint arXiv:1810.13273 (2018). [Online]. Available: https://arxiv.org/abs/1810.13273.

[10] H. K. Yadav, Y. Pal, and M. M. Tripathi, "Short-term PV power forecasting using empirical mode decomposition in integration with back-propagation neural network," *Journal of Information and Optimization Sciences*, vol. 41, no. 1, pp. 25–37 (2020), doi: 10.1080/02522667.2020.1714181.

[11] V. K. Bairagi and V. K. Harpale, "Improved epileptic seizure detection using singular spectrum empirical mode decomposition and machine learning approach," *Journal of Statistics and Management Systems*, vol. 25, no. 1, pp. 103–123 (2022), doi: 10.1080/09720510.2020.1862958.

[12] R. Song, X. Zhang, C. Zhou, J. Liu, and J. He, "Predicting TEC in China based on the neural networks optimized by genetic algorithm," *Advances in Space Research*, vol. 62, no. 4, pp. 745–759 (2018), doi: 10.1016/j.asr.2018.03.043.

[13] A. Tebabal, S. M. Radicella, M. Nigussie, B. Damtie, B. Nava, and E. Yizengaw, "Local TEC modeling and forecasting using neural networks," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 172, pp. 143–151 (2018), doi: 10.1016/j.jastp.2018.03.004.

[14] A. Tebabal, S. M. Radicella, B. Damtie, Y. Migoya-Orúe, M. Nigussie, and B. Nava, "Feed forward neural network based ionospheric model for the East African region," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 191, Art. no. 105051 (2019), doi: 10.1016/j.jastp.2019.05.016.

[15] 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, Art. no. 9 (2020), doi: 10.1007/s10509-020-3730-x.

[16] A. Ruwali, A. J. S. Kumar, K. B. Prakash, G. Sivavaraprasad, and D. V. Ratnam, "Implementation of hybrid deep learning model (LSTM-CNN) for ionospheric TEC forecasting using GPS data," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 9, pp. 1605–1609 (2021), doi: 10.1109/LGRS.2020.3004378.

[17] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 1, pp. 1–41 (2009), doi: 10.1142/S1793536909000047.

[18] S. J. Taylor and B. Letham, "Forecasting at scale," *The American Statistician*, vol. 72, no. 1, pp. 37–45 (2018), doi: 10.1080/00031305.2017.1380080.

Received April, 2025