Prediction of Total Electron Content by using Machine Learning Approach over Low Latitude Regions during Low Solar Activity Period

Adrika Kakoty[1], Lipika Soni[2], Barsha Dutta[3], S. Aruna[4]
[1,2,4] *Computer Science and Engineering (AIML), SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu 603203, Chennai, India*
[3]*Department of Basic Science and Humanities, Dibrugarh University Institute of Engineering and Technology, Dibrugarh University, Dibrugarh 786004, Assam, India*

***Abstract*—Total Electron Content (TEC) present in the ionosphere plays a crucial role in the transmission of signals through the ionosphere. So, the study of ionospheric TEC, its variations (diurnal, monthly, seasonal, yearly, etc.), and the various effects like solar activity, geomagnetic activity, etc. on TEC becomes an important topic for navigation. The machine learning methods are very useful to solve non-linear prediction challenges, especially for short-term GPS TEC forecasting due to the composite spatio-temporal variations. This study analyzes the performances of seven different machine learning models viz, RandomForest, XGBoost, GradientBoosting, LightGBM, LSTM, GBDT, GRU, and compares them to predict the ionospheric TEC over two low-latitude stations Pathum Wan (13.74° N, 100.54° E) and Bangalore (12.97° N, 77.59° E) separated by 23° longitude in the year 2009, low solar activity period (65.8 < F10.7 < 86.9). The model outputs have been validated by the real time data over the two locations. Solar quiet period is selected for the better analysis of day-to-day changes of the ionosphere which is important to improve the accuracy of physics-based models of the ionosphere. The monthly average TEC values predicted by all the models show strong agreement with the observed TEC values at both stations, with correlation coefficients 0.88 < R^2 < 0.93 over Pathum Wan and 0.85 < R^2 < 0.92 over Bangalore. The Root Mean Square Errors (RMSE) in the correlation plots are found between 2.13 to 2.87 over Pathum Wan and 2.29 to 3.25 over Bangalore. The values of Mean Absolute Errors (MAE) over these two locations are in between 1.48 to 2 and 1.51 to 2.27 respectively. Among all seven models LightGBM (R^2 = 0.926 over Pathum Wan and 0.918 over Bangalore) and XGBoost (R^2 = 0.923 over Pathum Wan and 0.914 over Bangalore) show the best correlations with lowest RMSE (2.1 over Pathum Wan and 2.3 over Bangalore) and MAE (1.5 over both Pathum Wan and over Bangalore).**

***Index Terms*—TEC, VTEC, Ionosphere, Machine Learning, RMSE, MAE, Correlation**

1. INTRODUCTION

The total electron content (TEC) refers to the integrated number of electrons along the signal's transmission path. The delay experienced in any transionospheric satellite radio communication is directly proportional to the TEC along the signal's path. TEC is essential to provide an overall explanation of the ionosphere with high temporal resolution. Hence, measurement of total electron content under various geophysical conditions has become crucial because of the high demand and growing reliance of aircraft and ground-based navigation on transionospheric communication systems. Equatorial Ionization Anomaly (EIA) regions, influenced by equatorial electrodynamics, exhibit distinct characteristics compared to other regions. Because of the huge variability of the equatorial ionization anomaly, TEC shows considerable daily variations, especially in equatorial and low-latitude regions. The formation and variation of EIA has been investigated by using ionosonde data and TEC obtained from various sources like satellite beacons, in situ satellite measurements etc. [1], [2], [3], [4], [5]

The thorough study of TEC in different solar-terrestrial events provides a complete understanding of the ionosphere and helps in modeling. TEC can be estimated by physical models [6], [7] mathematical models [8], [9] and empirical models [10], [11] in addition to the scientific analysis. The international reference ionosphere (IRI) is one of the most popular empirical models which is basically used to validate other models, [12], [13]. It is considered as a recommended international standard. The most recent version of the IRI model is IRI-2016 [14] where electron density is calculated by extending the anchor points of peak electron densities of E and F-layer. NeQuick model is used to estimate the maximum electron density [15].

Machine learning (ML) and deep learning (DL) have made significant contributions in ionospheric modeling with better accuracy [10]. Wavelet neural networks [16] are used for ionospheric tomography whereas PCA (Principal Component Analysis) and Neural Networks for ionospheric forecasting [17]. Machine learning techniques have been increasingly utilized to enhance the understanding and prediction of ionospheric Total Electron Content (TEC), presenting promising avenues for global and regional ionospheric studies.

The use of Neural Networks (NNs) for modeling of Vertical Total Electron Content (VTEC) for both quiet and disturbed geomagnetic conditions is increasingly attracting the attention of researchers because of the availability of advanced Graphics Processing Units (GPUs) and the capability of production and handling of big data [18], [19], [20], [21] used the neural network approach trained using the Levenberg-Marquardt algorithm to predict ionospheric TEC over various regions of Europe, capturing effectively the nonlinear behavior associated with space weather events. The results of Neural Networks (NN) has been evaluated against the International Reference Ionosphere (IRI) model over Chumphon, South Africa, and low-latitude regions of India (from equator to EIA region) [22], [23], [17] . To predict total electron content (TEC), recurrent neural network (RNN) based upon deep learning technique was suggested by [24]. LSTM is used to develop some ionospheric models to predict TEC [12], [25]. On the otherhand, some ionospheric TEC maps were built by using bidirectional LSTM algorithms [26], [27]. Generative adversarial networks (GAN) are also used in some TEC forecasting methods [28], [29] to improve the IRI [30]. [31] used four different ML approaches for ionospheric TEC, such as LSTM, adaptive neuro-fuzzy inference system, ANNs, and GBDT for the development of the model of ionospheric TEC. Some other ML methods like the nearest neighbor method, the prophet model, decision trees were also introduced for ionospheric TEC modeling [32], [33], [34]. The scientific data analysis methods and modeling approaches are developed continuously to enhance the accuracy of Total Electron Content (TEC) prediction for both military and civilian applications.

This study offers a comparative evaluation of seven different machine learning models viz. Random Forest Model, XGBoost Model, Gradient Boosting Model, LightGBM Model, LSTM, GBDT, GRU for the ionospheric TEC prediction over low latitude stations Bangalore (India) and Pathum Wan (Thailand) during low solar activity period. Different parameters indicating their performance are chosen and listed in detail.

2. METHODOLOGY

2.1. Dataset:

Daily fluctuations in VTEC over the two locations, Pathum Wan (13.74° N, 100.54° E) and Bangalore (12.97° N, 77.59° E), were acquired from International GPS Service (IGS) through the use of the GPS-TEC program [35] at a 5-min interval. Fig. 1 illustrates the latitudes and the longitudes of stations on the map. The hourly Dst data, Ap index and Kp index were derived from the online source https://wdc.kugi.kyoto-u.ac.jp/dstdir/. Solar flux data are acquired from the online source https://psl.noaa.gov/data/timeseries/month/SOLAR/ (NOAA Physical Sciences Laboratory)

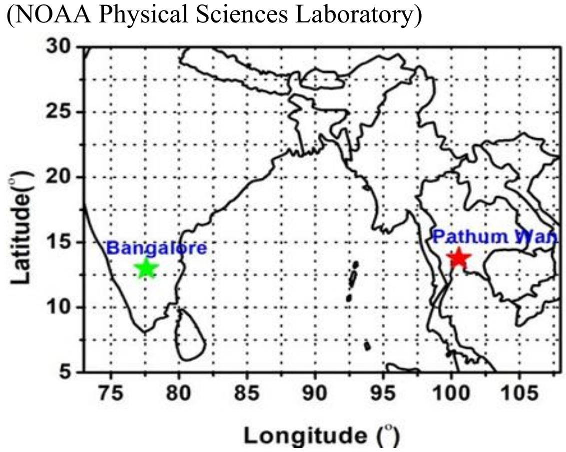

Fig.1 Map showing the locations

2.2. Model approach:

This work aims to estimate the TEC in the regions of Pathum Wan and Bangalore with the help of a comparative analysis of 7 machine learning models: LSTM, GBDT, GBM, XGBoost, LightGBM, GRU, and Random Forest Regressor for the year 2009. Fig. 2 shows the flow chart of the model. Over 8000 data points for each station were used for the hourly average of VTEC. The dataset was split into two subsets: 75% was utilized for model training, whereas the remaining 25% was reserved for testing to evaluate model performance.

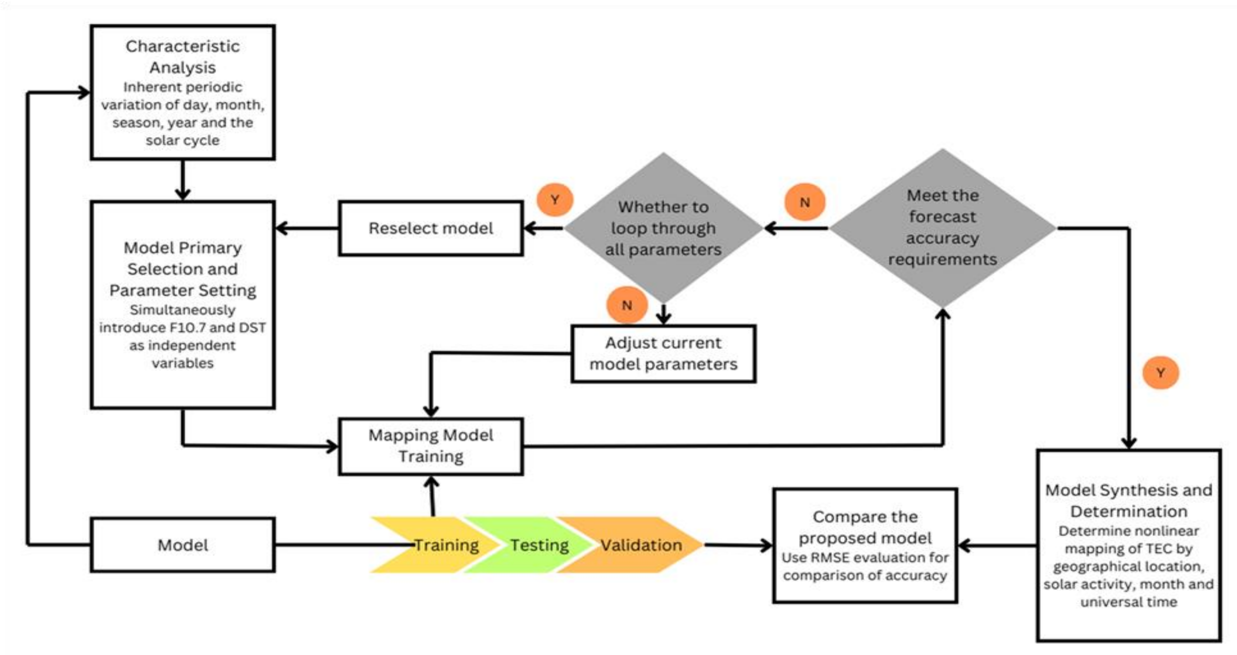

Fig. 2 Flowchart of the models

2.2.1. LSTM (Long Short-Term Memory)

LSTM networks are a variant of RNN created to resolve the difficulty of learning long-term dependencies in sequential data [36]. As illustrated in Fig. 3, the core feature that enables LSTMs to retain information over long periods is the memory block. It is a recurrent sub-network consisting of a memory cell and three gates. The memory cell acts as an internal storage unit, preserving temporal information across time steps. The system consists of three types of gates: input, forget, and output. The input gate controls the flow of new information into the memory cell. Forget gates decide the amount of previously stored information that should be retained or discarded. Output gates manage the amount of information from the memory cell that is used to generate the output and passed to the next layer. This gate-based system enables LSTMs to selectively store, update, and utilize information, making them highly effective for tasks like time series forecasting.

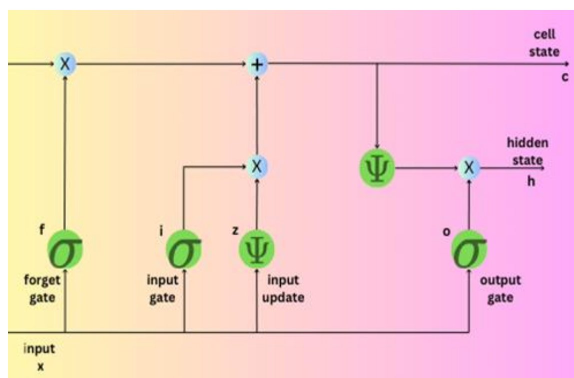

Fig. 3 LSTM memory block

A. GBM (Gradient Boosting Machines)

GBM is a technique that utilizes ensemble learning to build a robust predictive model by progressively merging several weak learners, usually decision trees. The model is optimized iteratively by reducing the residual errors from earlier predictions. [37] . At each step m, the model is represented as:

f_m(x) = f_{m-1}(x) + ρ_m h_m(x) where f_m(x) denotes the model at iteration m, f_{m-

$1\}(x)$ is the model from the prior iteration, $h\_m(x)$ is the newly added regression tree trained on the residuals, and ρ_m is the learning rate controlling the contribution of the new tree. The objective function minimized by GBM is typically the loss function $L(y, f(x))$, which guides the model updates. The residuals are computed as the negative gradient of the loss function:

$$r_{mi} = - \partial L(y_i, f_{m-1}(x_i)) / \partial f_{m-1}(x_i)$$

By fitting each new regression tree to the residuals, the model gradually improves its predictive accuracy. This approach enables GBM to model complex non-linear relationships and handle various data types, making it suitable for regression and classification tasks. GBM also demonstrates robustness against outliers and can effectively capture intricate patterns in the data [38].

B. XGBoost

XGBoost is a widely used decision tree-based machine learning algorithm in data science due to its high accuracy and efficiency. It is an advanced implementation of Gradient Boosting which combines the predictions of multiple weak classifiers, such as Classification and Regression Trees (CART) or linear classifiers, into a stronger classifier by iteratively adding weak trees with different weights. XGBoost enhances model performance by performing second-order Taylor expansion on the cost function, which provides richer information during optimization [39]. The objective function for this model can be expressed as:

$$Obj = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$$

where n denotes the number of training instances, y_i denotes the true label, $\hat{y}_i$ is the predicted value, K is the total number of regression trees, f_k represents the k-th regression tree, and $\Omega(f_k)$ is the regularization term that penalizes the complexity of the tree to avoid overfitting. XGBoost improves computational speed through parallel and distributed calculations and can effectively handle missing values. This comprehensive approach enables XGBoost to deliver highly accurate predictions while maintaining generalization ability [40].

C. GBDT (Gradient Boosting Decision Trees)

GBDT is a machine learning technique based on ensemble methods that merges several weak learners and creates a robust predictive model. It is widely used for the problems of regression and classification. The GBDT model builds decision trees sequentially, with each tree correcting the errors of its predecessor by minimizing the specified loss function. The model can be expressed as a summation of decision trees:

$$f(x) = \sum_{m=1}^{M} h_m(x, \Theta_m)$$

where $h_m(x, \Theta_m)$ denotes a decision tree defined by its set of parameters Θ_m, and M is the total number of trees. The model is developed incrementally using a forward stage-wise approach, where the m-th step updates the model by adding a new regression tree:

$$f_m(x) = f_{m-1}(x) + \rho_m h_m(x, \Theta_m)$$

The parameters Θ_m of each tree are determined by minimizing the loss function:

$$\Theta_m = \mathrm{argmin}_\Theta \sum_{i=1}^{N} L(y_i, f_{m-1}(x_i) + h_m(x_i, \Theta))$$

The loss function, typically the mean squared error (MSE), is given by:

$$L(y, f(x)) = 1/2 (y - f(x))^2$$

At each iteration, the GBDT algorithm computes the pseudo-residuals:

$$r_{mi} = -\partial L(y_i, f_{m-1}(x_i)) / \partial f_{m-1}(x_i)$$

The new regression tree is fitted to these pseudo-residuals, improving the model's performance. GBDT efficiently captures nonlinear interactions and mitigates overfitting through the use of learning rates and regularization terms, making it highly effective in modeling GNSS time series data and other complex patterns [41].

D. LightGBM

LightGBM is a gradient boosting framework that leverages decision trees to achieve high performance in machine learning tasks. It uses a histogram-based algorithm to discretize continuous feature values, significantly reducing memory consumption and improving training speed. Unlike traditional level-wise tree growth, LightGBM employs a leaf-wise growth approach, splitting the node with the highest split gain, which enhances accuracy but requires regularization techniques to prevent overfitting. The model's hyperparameters play a crucial role in performance optimization, including max_depth, which controls tree depth and mitigates overfitting, n_estimator to specify the number of boosting iterations, and learning_rate to regulate the step size of each iteration. Grid Search Cross-Validation (GSCV) is used to optimize hyperparameters by systematically searching combinations of parameter values and evaluating model performance using cross-validation. LightGBM's loss function uses the negative gradient to approximate residuals through Taylor expansion, with complexity which is managed via regularization terms such as L1 regularization. This approach makes LightGBM highly efficient, capable of handling large datasets, and robust against overfitting while maintaining high accuracy [42], [43].

E. GRU (Gated Recurrent Units)

GRUs, proposed by Kyunghyun Cho in 2014, are a simplified variant of LSTM networks. They were developed to mitigate issues like vanishing and exploding gradient problems in RNNs. GRUs utilize two gating mechanisms: the update gate and reset gate, which control the flow of information within the network

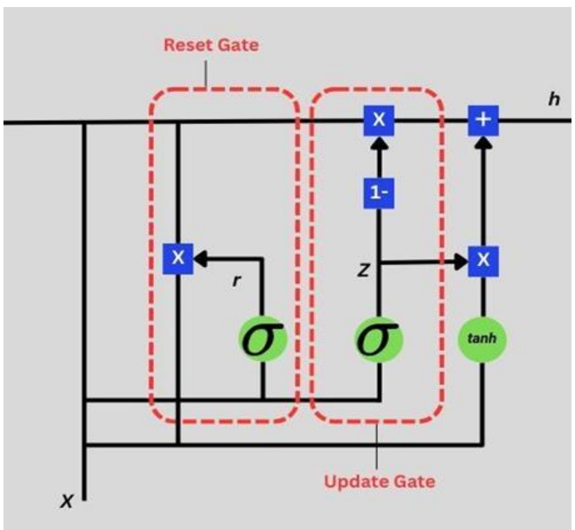

Fig. 4 GRU Architecture

The update gate zt controls the extent to which past information is preserved, calculated as $zt=(w(t)xt+u(t)ht-1+b)$

Where, w and u represent weights, xt is the input, $ht-1$ is the hidden state, and b is the bias.. The reset gate rt defines the degree to which past information is forgotten, expressed as $rt=(w(r)xt+u(r)ht-1+b)$.

The candidate hidden state ht' is computed as $ht'=\tanh(wxt+rt\odot uht-1)$, where $\odot$ represents the Hadamard product.

Finally, the current hidden state ht is derived as $ht=zt\odot ht-1+(1-zt)\odot ht'$.

GRUs are computationally efficient due to their simpler architecture, which requires fewer parameters and less training time compared to LSTMs, making them ideal for sequential data tasks [44].

F. Random Forest Regressor

Random Forest (RF) is an ensemble-based technique that constructs several decision trees to perform regression tasks, reducing prediction variance and enhancing accuracy. The RF algorithm operates by creating M regression trees, each tree built from bootstrap samples of training data. The predicted value at input u for the j-th tree is denoted by m_s(u; θ_j, S_s), where θ_j represents random variables, and S_s is the fitting sample. The final prediction is the average of all tree predictions:

m(u) = E[v|u = u]

The training data consists of:

S_s = ((u_1, v_1), …, (u_s, v_s))

For each tree, a random subset of predictor variables is selected for node splitting, reducing correlation between trees. The number of trees (M), number of variables (m_try), number of sampled data points (bs), and minimum node size (node size) are critical hyper-parameters. Cross-validation methods like bootstrap resampling are used to optimize m_try, often minimizing the Root Mean Squared Error (RMSE). PCA is then applied to the ensemble of regression tree predictors, decomposing them into principal component scores and eigen functions:

T_i = ∑_(j=1)^B α_ij φ_j

where α_ij represents principal component scores, and φ_j represents eigenfunctions. The Bayes-Linear-Gauss method is used to find new coefficients α ensuring perfect fit at sampled locations. This method is defined by:

α ~ N(μ, C)

with mean μ and covariance matrix C computed from PCA scores. Conditional coefficients are obtained through Monte Carlo sampling, reconstructing regression tree predictors:

T̂_i = ∑_(j=1)^B α̂_ij φ_j

The final prediction is the average of the reconstructed regression tree predictors:

m̂(u) = 1/T ∑_(i=1)^T T̂_i(u)

This method ensures accurate predictions that match observed values while maintaining computational efficiency [45].

3. RESULTS AND DISCUSSION

3.1. Monthly variations of hourly averaged vertical TEC over the stations

Fig. 5 and Fig. 6 represent the diurnal variations of VTEC over Pathum Wan and Bangalore, the two low latitude stations. The solid line represents the result obtained from the real time data analysis whereas the dotted lines show the outputs of the seven models viz., RandomForest, XGBoost, Gradient Boosting, Light GBM, Gradient Boosting Decision Trees, GRU, and LSTM. Similar TEC variations are observed in both real-time measurements and model-generated outputs. The observed TEC is lower than the model results during January, June, July, August. In contrast, real-time TEC exceeds model predictions in March, April, and October. In February, May, September, November, and December, the real-time TEC variations closely match the predictions of the models. This similarity suggests that the models are effectively capturing the general trends and variations of TEC, which could be influenced by factors like geomagnetic activity, solar radiation, and ionospheric dynamics. However, slight deviations might exist due to dynamic space weather conditions, localized anomalies, or model limitations.

The low latitude stations have been selected in our study because the E×B drift, particularly near the equatorial region, causes significant vertical plasma transport, leading to higher charge densities and more pronounced TEC variations in the low latitude stations. This makes low-latitude regions ideal for studying ionospheric dynamics, such as equatorial ionization anomaly (EIA) effects and space weather influences. The longitude difference between the selected low-latitude stations allows for an assessment of how well the model captures ionospheric variations across different longitudes. Since ionospheric conditions can vary with longitude due to factors like geomagnetic field asymmetry, zonal winds, and electrodynamic processes, this approach helps evaluate the model's accuracy and efficiency in different longitudinal sectors.

This study investigates and predicts Total Electron Content (TEC) during a period of low solar activity and geomagnetically quiet conditions, aiming to better understand the day-to-day variability of ionospheric behavior with minimal influence from space weather disturbances.

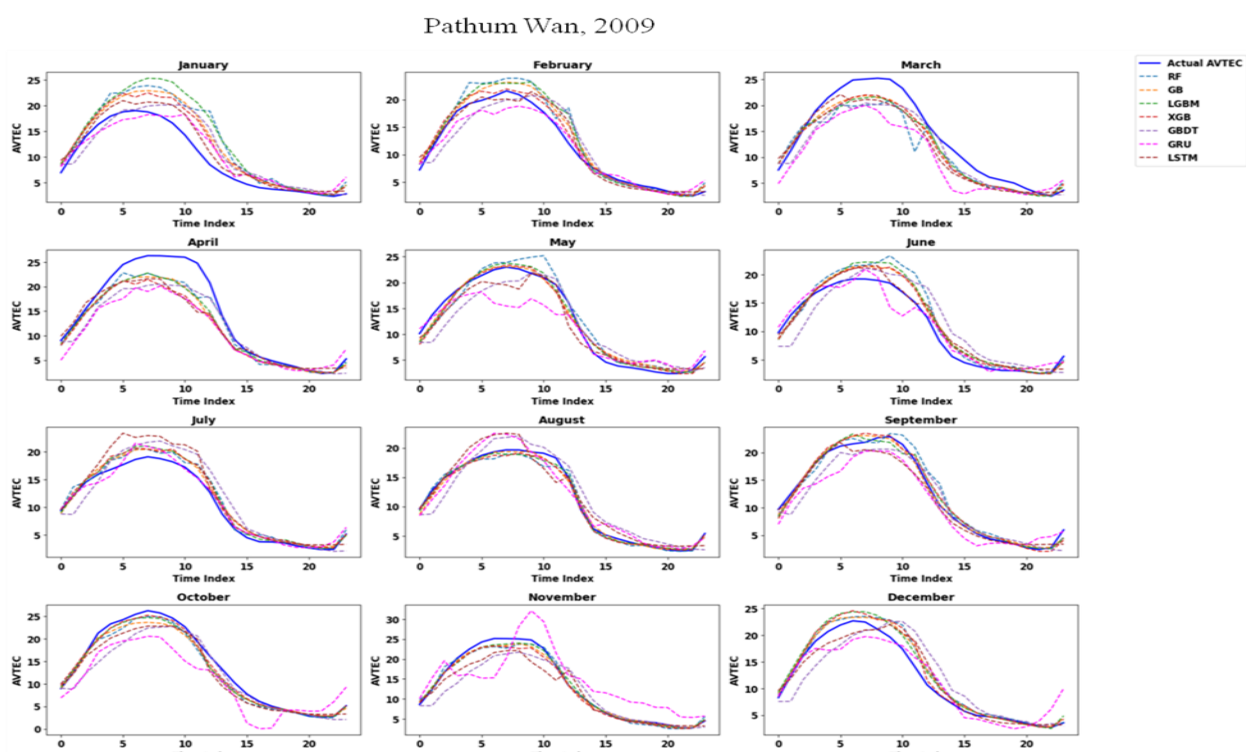

Fig. 5 Diurnal variations of real time VTEC along with model outputs over Pathum Wan, 2009

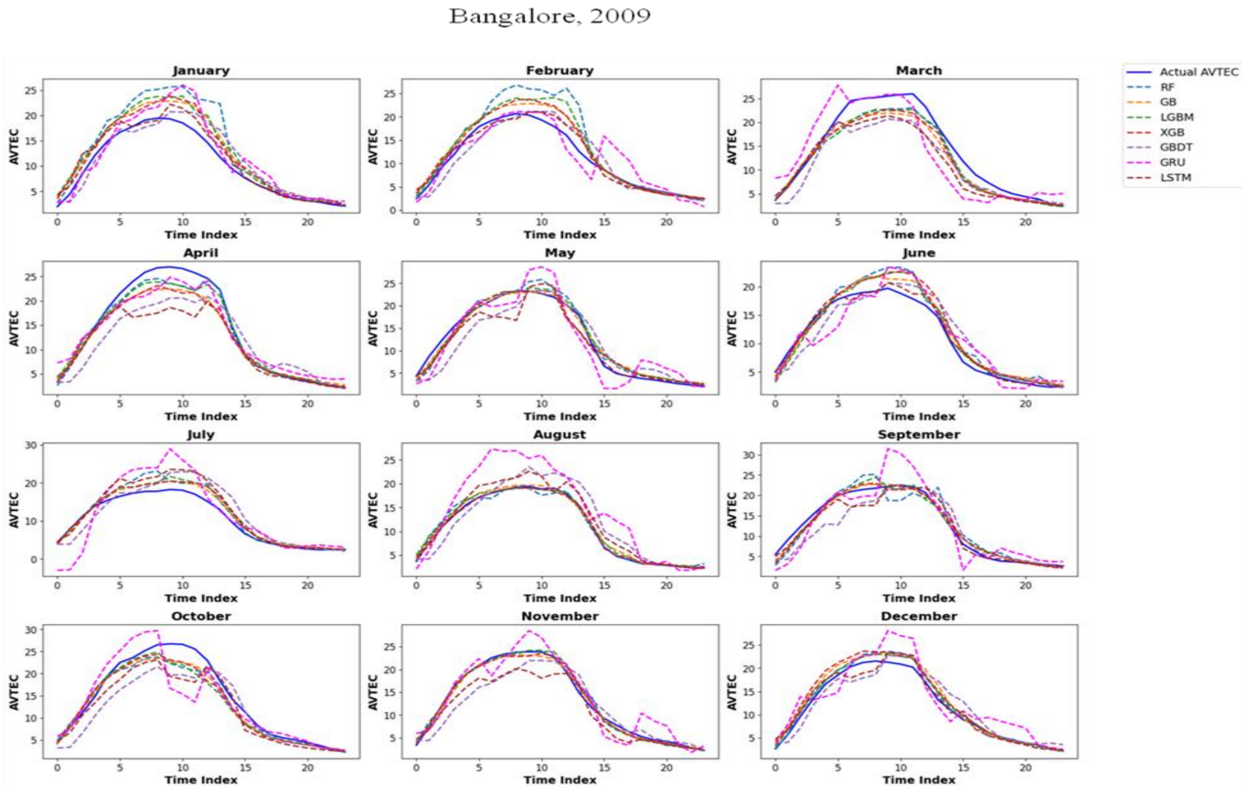

Fig. 6 Diurnal variations of real time VTEC along with model outputs over Bangalore, 2009

3.2. Performance comparison of various model outputs with the observed TEC

To evaluate the effectiveness of the machine learning models, the predicted TEC values are compared against the observed TEC measurements over the stations Pathum Wan and Bangalore. Fig. 7 and Fig. 8 represent the correlation between all seven model outputs with the observed TEC values, with the RMSE and MAE. From the figures it is seen that the predictions of TEC by all models are highly consistent with the observed value. The results over Pathum Wan and Bangalore are shown in Table 1 and Table 2 respectively. R^2 values lie in between 0.88 to 0.93 over Pathum Wan and 0.85 to 0.92 over Bangalore with RMSE in between 2.125 to 2.872 and 2.290 to 3.248 over the locations respectively. MAE is in between 1.481 to 1.998 over Pathum Wan and 1.508 to 2.268 over Bangalore.

Among all seven models LightGBM and XGBoost show the best correlations. The R^2 value between the LightGBM output and the observed TEC is found to be 0.926 over Pathum Wan and 0.918 over Bangalore and that between the XGBoost output and the observed TEC was 0.923 over Pathum Wan and 0.914 over Bangalore. Moreover, these two models give the lowest error values with RMSE = 2.1 over Pathum Wan and 2.3 over Bangalore and MAE = 1.5 over both Pathum Wan and over Bangalore.

[30] applied four distinct machine learning models, i.e., Artificial Neural Network (ANN), LSTM, Adaptive Neuro-Fuzzy Inference System using Subtractive Clustering (ANFIS-SC) and Gradient Boosting Decision Tree (GBDT) to forecast ionospheric TEC over three IGS stations of low-latitude region (16°S to 10°S) during solar and geomagnetic extreme conditions and improvised the prediction accuracy 37.93% to 49.28% during high solar activity period and from 28.16% to 67.39% during the magnetic storm period in compare to Global Ionospheric Map (GIM) prediction model. [46] proposed machine learning based model and used that model as an empirical model of TEC over some regions of Europe. Their prediction was highly compatible with the observed value of TEC with RRMSE (relative RMSE) value 12.76%.

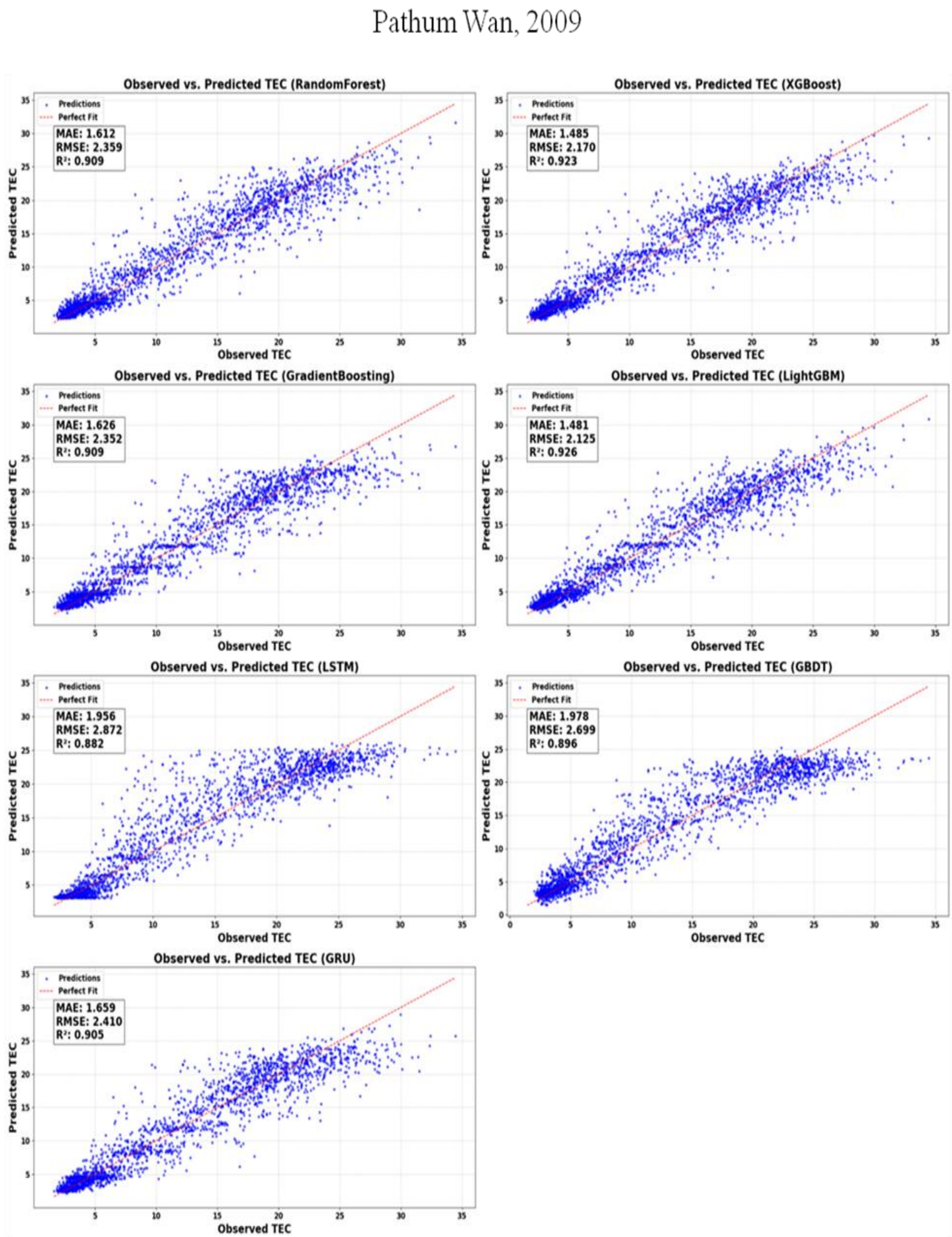

Fig. 7 Correlations of VTEC obtained from all seven model outputs with the observed value over Pathum Wan

Bangalore, 2009

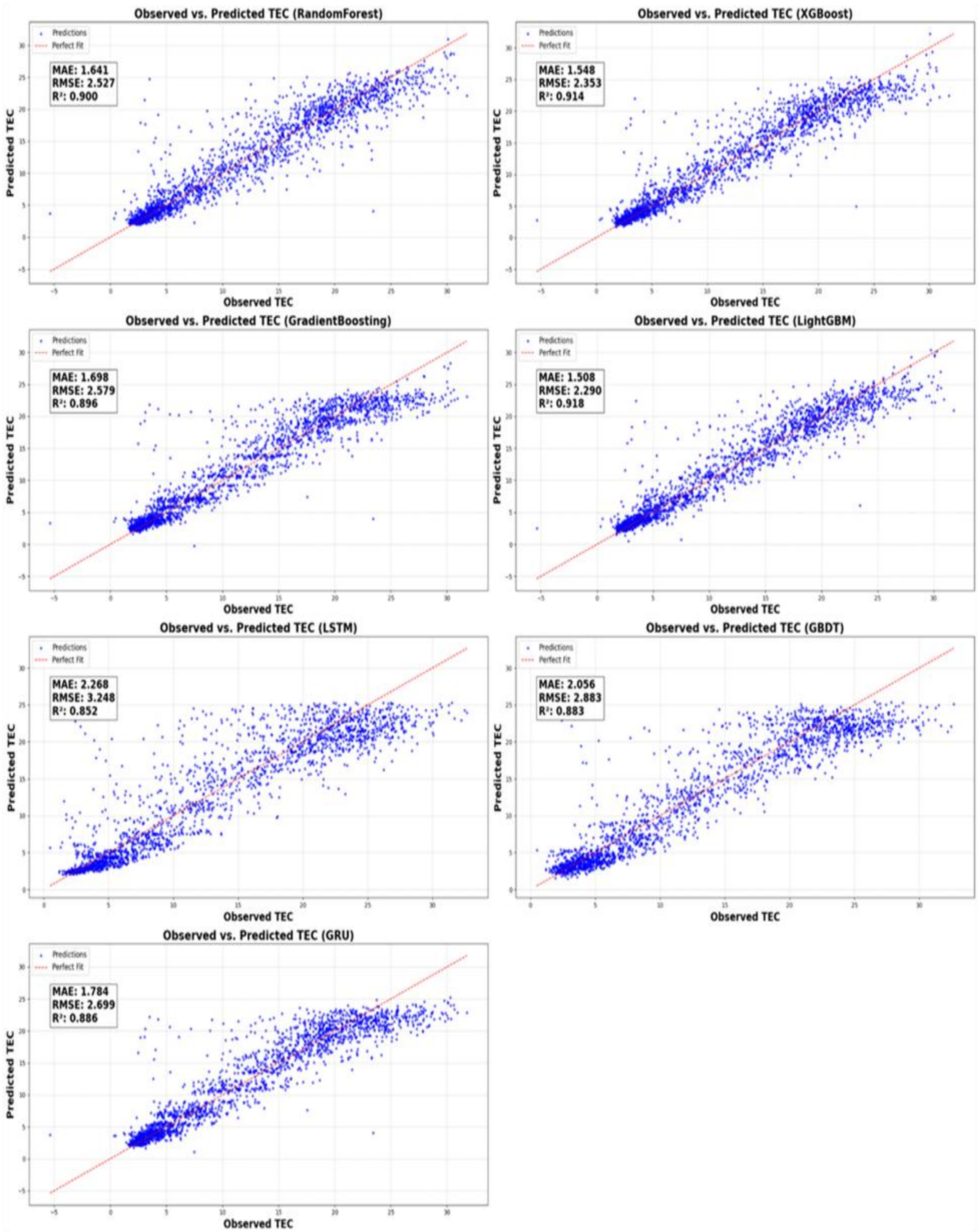

Fig. 8 Correlations of VTEC obtained from all seven model outputs with the observed value over Bangalore

Table. 1. Test data predictions by all seven models over Pathum Wan

Pathum Wan			
Model	Root Mean Square Error (RMSE)	Mean Absolute Error (MAE)	Linear correlation (R^2)
RandomForest Model	2.359	1.612	0.909
XGBoost Model	2.170	1.485	0.923
Gradient Boosting Model	2.352	1.626	0.909
Light GBM Model	2.125	1.481	0.926
LSTM	2.872	1.956	0.882
GBDT	2.699	1.998	0.896
GRU Model	2.403	1.665	0.905

Table. 2. Test data predictions by all seven models over Bangalore

Bangalore			
Model	Root Mean Square Error	Mean Absolute Error	Linear correlation (R^2)
Random Forest Model	2.527	1.641	0.900
XGBoost Model	2.353	1.548	0.914
Gradient Boosting Model	2.579	1.698	0.896
Light GBM Model	2.290	1.508	0.918
LSTM	3.248	2.268	0.852
GBDT	2.883	2.056	0.883
GRU Model	2.699	1.784	0.886

4. CONCLUSIONS

Seven Machine Learning approaches have been employed to predict the variations of TEC over two low latitude stations, Pathum Wan and Bangalore. A comparative analysis of the performance of these seven models against observed TEC has been conducted in this study.

The following conclusions have been drawn:

i) Diurnal Variation Capture: The diurnal variations of TEC predicted by all models closely match the observed TEC for all months. This indicates that the models effectively capture the general trends and temporal variations in TEC.

ii) Monthly Mean Accuracy: The monthly mean TEC values predicted by the models show strong agreement with the observed values at both stations. The correlation coefficients range from 0.88 to 0.93 over Pathum Wan and from 0.85 to 0.92 over Bangalore, demonstrating the models' reliability in predicting long-term TEC trends.

iii) Error Metrics:

- The RMSE values in the correlation plots range from 2.125 to 2.872 over Pathum Wan and 2.290 to 3.248 over Bangalore.
- The Mean Absolute Error (MAE) values range from 1.481 to 1.998 over Pathum Wan and 1.508 to 2.268 over Bangalore.

These results collectively indicate that the machine learning models used in this study are capable of accurately predicting TEC variations over both stations.

Declaration:

Funding: No funding has been received for this work

Authors Contribution:

Adrika Kakoty and Lipika Soni were responsible for Model Development, Data Analysis, Writing; Barsha Dutta was responsible for Conceptualization, Supervising, Methodology, Validation, Investigation, Writing Original Draft and S Aruna was responsible for Conceptualization, Supervising.

REFERENCES

[1] G.O. Walker, and A.E. Strickland, “A comparison of the ionospheric equatorial anomaly in the East Asian and the American regions at Sunspot minimum,” Journal of Atmospheric and Terrestrial Physics, vol. 43, no. 5–6, pp. 589-595, 1981. https://doi.org/10.1016/0021-9169(81)90121-5.

[2] D. N. Anderson, “A theoretical study of the ionospheric F region equatorial anomaly. I. Theory,” Planetary and Space Science, vol. 21, pp. 409–419, 1973. https://doi.org/10.1016/0032-0633(73)90040-8.

[3] G.O. Walker, “Longitudinal structure of the F-region equatorial anomaly—a review,” Journal of Atmospheric and Terrestrial Physics, vol. 43, no. 8, pp. 763-774, 1981. https://doi.org/10.1016/0021-9169(81)90052-0.

[4] Y. Z. Su, S. Fukao, and G. J. Bailey, “Modeling studies of the middle and upper atmosphere radar observations of the ionospheric F layer,” Journal of Geophysical Research, vol. 102, pp. 319-327, 1997. https://doi.org/10.1029/96JA02994.

[5] C. C. Wu, K. Liou, S. J., Shan, and 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, 2008. https://doi.org/10.1016/j.asr.2007.06.013.

[6] J. He, X. Yue, H. Le, Z. Ren, and W. Wan, “Evaluation on the Quasi-Realistic Ionospheric Prediction Using an Ensemble Kalman Filter Data Assimilation Algorithm,” Space Weather 2020, 18, e2019SW002410, 2020. https://doi.org/10.1029/2019SW002410.

[7] J. Tang, S. Zhang, X. Huo, and X. Wu, “Ionospheric Assimilation of GNSS TEC into IRI Model Using a Local Ensemble Kalman Filter,” Remote Sensing, vol. 14, pp. 3267, 2022. https://doi.org/10.3390/rs14143267.

[8] S. Farzaneh, and E. Forootan, “Reconstructing regional ionospheric electron density: A combined spherical Slepian function and empirical orthogonal function approach,” Surveys in Geophysics, vol. 39, pp. 289–309, 2018. https://doi.org/10.1007/s10712-017-9446-y.

[9] M. A. Sharifi, and S. Farzaneh, “The ionosphere electron density spatio-temporal modeling based on the Slepian basis functions,” Acta Geodaetica et Geophysica, vol. 52, pp. 5–18, 2017. https://doi.org/10.1007/s40328-016-0165-5.

[10] L. Wang, E. Wei, S. Xiong, T. Zhang, and Z. Shen, “Evaluation of NeQuick2 Model over Mid-Latitudes of Northern Hemisphere,” Remote Sensing, vol. 14, pp. 4124, 2022. https://doi.org/10.3390/rs14164124.

[11] M. Pietrella, M. Pezzopane, B. Zolesi, L.R. Cander, & A. Pignalberi, “The Simplified Ionospheric Regional Model (SIRM) for HF Prediction: Basic Theory, Its Evolution and Applications,” Surveys in Geophysics, 41, 1143–1178, 2020. https://doi.org/10.1007/s10712-020-09600-w

[12] W. Zhang, X. Huo, Y. Yuan, Z. Li, & N. Wang, “Algorithm Research Using GNSS-TEC Data to Calibrate TEC Calculated by the IRI-2016 Model over China,” Remote Sensing, 13, 4002, 2021. https://doi.org/10.3390/rs13194002

[13] Y. Yao, X. Chen, J. Kong, C. Zhou, L. Liu, L. Shan, & Z. Guo, “An Updated Experimental Model of IG12 Indices Over the Antarctic Region via the Assimilation of IRI2016 With GNSS TEC,” IEEE Transactions on Geoscience and Remote Sensing, 59, 1700–1717, 2021. https://doi.org/10.1109/TGRS.2020.2999132

[14] D. Bilitza, “IRI the International Standard for the Ionosphere,” Advances in Radio Science, 16, 1–11, 2018. https://doi.org/10.5194/ars-16-1-2018

[15] A. Pignalberi, M. Pezzopane, B. Nava, “Optimizing the NeQuick Topside Scale Height Parameters Through COSMIC/FORMOSAT3 Radio Occultation Data,” IEEE Geoscience and Remote Sensing Letters, 19, 1–5, 2021.

[16] M. R. G. Razin and B. Voosoghi, “Modeling of ionosphere time series using wavelet neural networks (case study: NW of Iran),” Advances in Space Research, vol. 58, pp. 74–83, 2016. [Online]. Available: https://doi.org/10.1016/j.asr.2016.04.006

[17] I. L. Mallika, D. V. Ratnam, Y. Ostuka, 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, pp. 371–381, 2018.

[Online]. Available: https://doi.org/10.1109/JSTARS.2018.2877445
[18] R. Stamper, A. Belehaki, D. Buresová, L. R. Cander, I. Kutiev, M. Pietrella, I. Stanislawska, S. Stankov, I. Tsagouri, and Y. K. Tulunay, "Nowcasting, forecasting and warning for ionospheric propagation: tools and methods," Annals of Geophysics, vol. 47, 2004. [Online]. Available: https://doi.org/10.4401/ag-3281
[19] S. M. Radicella and E. Tulunay, "Space plasma effects on Earth-space and satellite-to-satellite communications: Working Group 4 overview," Annals of Geophysics, vol. 47, 2004. [Online]. Available: https://doi.org/10.4401/ag-3300
[20] L. A. McKinnell and A. W. Poole, "Predicting the ionospheric F layer using neural networks," Journal of Geophysical Research: Space Physics, vol. 109, 2004. [Online]. Available: https://doi.org/10.1029/2004JA010445
[21] E. Tulunay, E. T. Senalp, S. M. Radicella, and Y. Tulunay, "Forecasting total electron content maps by neural network technique," Radio Science, vol. 41, 2006. [Online]. Available: https://doi.org/10.1029/2005RS003285
[22] J. B. Habarulema, L.-A. McKinnell, P. J. Cilliers, and B. D. Opperman, "Application of neural networks to South African GPS TEC modelling," Advances in Space Research, vol. 43, pp. 1711–1720, 2009. [Online]. Available: https://doi.org/10.1016/j.asr.2008.08.020
[23] K. Watthanasangmechai, P. Supnithi, S. Lerkvaranyu, T. Tsugawa, T. Nagatsuma, and T. Maruyama, "TEC prediction with neural network for equatorial latitude station in Thailand," Earth, Planets and Space, vol. 64, pp. 473–483, 2012. [Online]. Available: https://doi.org/10.5047/eps.2011.05.025
[24] M. Kaselimi, A. Voulodimos, N. Doulamis, A. Doulamis, and D. Delikaraoglou, "Deep recurrent neural networks for ionospheric variations estimation using GNSS measurements," IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1–15, 2021. [Online]. Available: https://doi.org/10.1109/TGRS.2021.3090856
[25] X. Ren, P. Yang, H. Liu, J. Chen, and W. Liu, "Deep learning for global ionospheric TEC forecasting: Different approaches and validation," Space Weather, vol. 20, p. e2021SW003011, 2022. [Online]. Available: https://doi.org/10.1029/2021SW003011
[26] K. Sivakrishna, D. Venkata Ratnam, and G. Sivavaraprasad, "A bidirectional deep-learning algorithm to forecast regional ionospheric TEC maps," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, pp. 4531–4543, 2022. [Online]. Available: https://doi.org/10.1109/JSTARS.2022.3180940
[27] S. Shi, K. Zhang, S. Wu, J. Shi, A. Hu, H. Wu, and Y. Li, "An investigation of ionospheric TEC prediction maps over China using bidirectional long short-term memory method," Space Weather, vol. 20, p. e2022SW003103, 2022. [Online]. Available: https://doi.org/10.1029/2022SW003103
[28] Z. Chen, W. Liao, H. Li, J. Wang, X. Deng, and S. Hong, "Prediction of global ionospheric TEC based on deep learning," Space Weather, vol. 20, p. e2021SW002854, 2022. [Online]. Available: https://doi.org/10.1029/2021SW002854
[29] Y. Pan, M. Jin, S. Zhang, and Y. Deng, "TEC map completion through a deep learning model: SNP-GAN," Space Weather, p. e2021SW002810, 2011. [Online]. Available: https://doi.org/10.1029/2021SW002810
[30] E. Ji, Y. Moon, and E. Park, "Improvement of IRI global TEC maps by deep learning based on conditional generative adversarial networks," Space Weather, vol. 18, p. e2019SW002411, 2020. [Online]. Available: https://doi.org/10.1029/2019SW002411
[31] Y. Han, L. Wang, W. Fu, H. Zhou, T. Li, and R. Chen, "Machine learning-based short-term GPS TEC forecasting during high solar activity and magnetic storm periods," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, pp. 115–126, 2022. [Online]. Available: https://doi.org/10.1109/JSTARS.2021.3132049
[32] E. Monte-Moreno, H. Yang, and M. Hernández-Pajares, "Forecast of the global TEC by nearest neighbour technique," Remote Sensing, vol. 14, p. 1361, 2022. [Online]. Available: https://doi.org/10.3390/rs14061361
[33] J. Tang, Y. Li, D. Yang, and M. Ding, "An approach for predicting global ionospheric TEC using machine learning," Remote Sensing, vol.

14, p. 1585, 2022. [Online]. Available: https://doi.org/10.3390/rs14071585

[34] R. Natras, B. Soja, and M. Schmidt, "Ensemble machine learning of Random Forest, AdaBoost and XGBoost for vertical total electron content forecasting," Remote Sensing, vol. 14, p. 3547, 2022. [Online]. Available: https://doi.org/10.3390/rs14153547

[35] G. K. Seemala and C. E. Valladares, "Statistics of total electron content depletions observed over the South American continent for the year 2008," Radio Science, vol. 46, p. RS5019, 2011. [Online]. Available: https://doi.org/10.1029/2011RS004722

[36] Y. Hua, Z. Zhao, R. Li, X. Chen, Z. Liu, and H. Zhang, "Deep learning with long short-term memory for time series prediction," IEEE Communications Magazine, vol. 57, no. 6, pp. 114–119, 2019. [Online]. Available: https://doi.org/10.1109/MCOM.2019.1800155

[37] A. Natekin and A. Knoll, "Gradient boosting machines, a tutorial," Frontiers in Neurorobotics, vol. 7, p. 21, 2013. [Online]. Available: https://doi.org/10.3389/fnbot.2013.00021

[38] K. Nakagawa and K. Yoshida, "Time-series gradient boosting tree for stock price prediction," International Journal of Data Mining, Modelling and Management, vol. 14, no. 2, p. 110, 2022. [Online]. Available: https://doi.org/10.1504/IJDMMM.2022.123357

[39] Z. Wang, X. Wu, and Y. Wu, "A spatiotemporal XGBoost model for PM2.5 concentration prediction and its application in Shanghai," Heliyon, vol. 9, no. 12, p. e22569, 2023. [Online]. Available: https://doi.org/10.1016/j.heliyon.2023.e22569

[40] Z. Fang, S. Yang, C. Lv, S. An, and W. Wu, "Application of a data-driven XGBoost model for the prediction of COVID-19 in the USA: A time-series study," BMJ Open, vol. 12, no. 7, p. e056685, 2022. [Online]. Available: https://doi.org/10.1136/bmjopen-2021-056685

[41] S. Si, H. Zhang, S. S. Keerthi, D. Mahajan, I. S. Dhillon, and C.-J. Hsieh, "Gradient boosted decision trees for high dimensional sparse output," in Proceedings of the 34th International Conference on Machine Learning (ICML), Proceedings of Machine Learning Research, vol. 70, pp. 3182–3190, 2017. [Online]. Available: https://proceedings.mlr.press/v70/si17a.html

[42] Y. Guo, Y. Li, and Y. Xu, "Study on the application of LSTM-LightGBM model in stock rise and fall prediction," MATEC Web of Conferences, vol. 336, p. 05011, 2021. [Online]. Available: https://doi.org/10.1051/matecconf/202133605011

[43] A. D. Hartanto, Y. N. Kholik, and Y. Pristyanto, "Stock price time series data forecasting using the Light Gradient Boosting Machine (LightGBM) model," JOIV International Journal on Informatics Visualization, vol. 7, no. 4, p. 2270, 2023. [Online]. Available: https://doi.org/10.30630/joiv.7.4.01740

[44] M. Ghadimpour and S. B. Ebrahimi, "Forecasting financial time series using deep learning networks: Evidence from Long-Short Term Memory and Gated Recurrent Unit," Iranian Journal of Finance, vol. 6, no. 4, pp. 81–94, 2022. [Online]. Available: https://doi.org/10.30699/ijf.2022.313164.1286

[45] R. Davis and M. Nielsen, "Modeling of time series using random forests: Theoretical developments," Electronic Journal of Statistics, vol. 14, pp. 3644–3671, 2020. [Online]. Available: https://doi.org/10.1214/20-EJS1758

[46] Y. Liu, J. Wang, C. Yang, Y. Zheng, and H. Fu, "A machine learning-based method for modeling TEC regional temporal-spatial map," Remote Sensing, vol. 14, p. 5579, 2022. [Online]. Available: https://doi.org/10.3390/rs14215579