IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Prediction Of Ionospheric Total Electron Content Using Lstm Networks During The Ascending Phase Of Solar Cycle 25

G. V. Avanija Mahalakshmi and Bharati Bidikar

Department of Computer Science and Systems Engineering,
Andhra University College of Engineering (A), Andhra University
Visakhapatnam, India

Abstract: Accurate prediction of ionospheric Total Electron Content (TEC) is vital for the reliability of Global Navigation Satellite Systems (GNSS), especially in equatorial regions such as India, where ionospheric fluctuations are more pronounced. This paper presents a deep learning approach using GNSS data from the IISc Bangalore station (13.0170° N, 77.5659° E) over the ascending phase of Solar Cycle 25 (December 2019 – December 2024). A multivariate Long Short-Term Memory (LSTM) architecture is employed to forecast TEC variations, integrating historical TEC records and the planetary K-index (Kp) to account for geomagnetic disturbances driven by space weather events. The model has been improved with dense layers to fine-tune hidden state transformations, dropout regularization to avoid overfitting, and stacking LSTM layers for increased depth. Model performance is assessed through statistical measures: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Squared Error (MSE), and R-squared (R^2), evaluated across multiple training epochs. Results confirm that incorporating geomagnetic indices alongside temporal deep learning substantially increases prediction accuracy for TEC. The study underscores the importance of both model architecture and relevant feature selection in space weather forecasting, paving the way for more robust GNSS operations under fluctuating ionospheric conditions.

***Index Terms* -** Ionospheric TEC, LSTM, GNSS, Solar Cycle 25, Geomagnetic Index, Kp, Space Weather, Time Series Forecasting.

I. INTRODUCTION

Accurate prediction of Total Electron Content (TEC) is particularly critical for the Indian subcontinent, which lies in the equatorial and low-latitude region where ionospheric variability is most pronounced. TEC, defined as the total number of free electrons along the satellite–receiver path, directly impacts the reliability of Global Navigation Satellite Systems (GNSS). Fluctuations in TEC introduce signal delays, phase distortions, and amplitude fading, thereby degrading GNSS positioning accuracy [1], [2]. Figure 1 presents the month-wise average TEC values during the study period, highlighting the strong temporal and seasonal variations in the region.

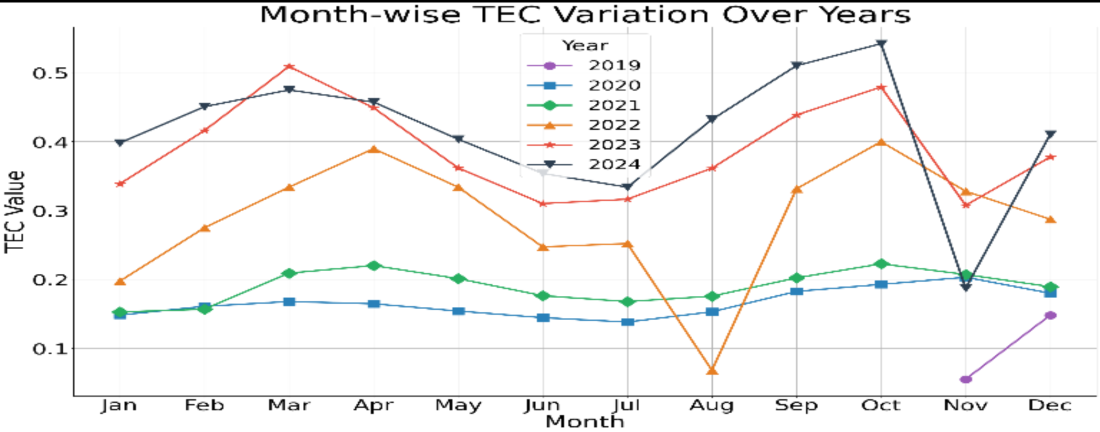
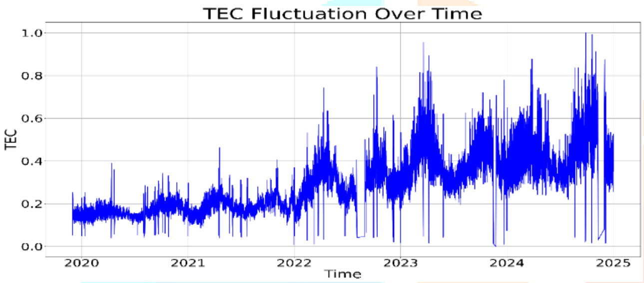


Figure 1: Month-wise Average TEC Values

The ascending phase of Solar Cycle 25 is characterized by intensified solar ultraviolet radiation, which amplifies ionospheric disturbances and exacerbates TEC anomalies over India [3]. In addition to solar forcing, the planetary geomagnetic K-index (Kp) is a widely used measure of space-weather-driven disturbances such as storms and substorms [4], [5]. Figure 2 illustrates TEC variations at the IISc Bangalore GNSS station (13.0170° N, 77.5659° E) from 2019 to 2024, while Figure 3 shows the corresponding Kp fluctuations, together reflecting the complex spatiotemporal dynamics governing ionospheric behavior.



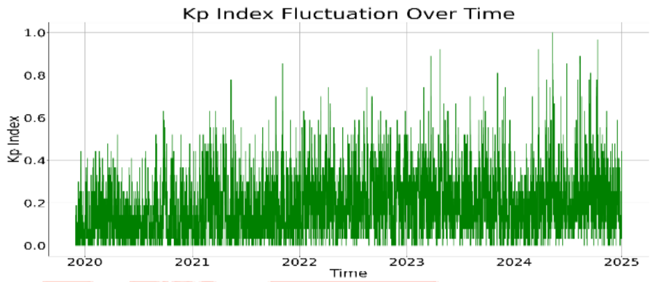


Figure 2: TEC Variations Over the Study Period (2019–2024) *Figure 3: Kp Index Trend Over the Study Period (2019–2024)*

Reliable TEC forecasting is crucial for critical applications, including civil aviation, military operations, navigation, and precision agriculture, where GNSS performance has a direct impact on safety and accuracy [2], [6]. Traditional empirical and statistical models are often insufficient for TEC modeling, as they fail to capture the nonlinear, non-stationary, and multi-scale features of ionospheric variability [7].

Recent advances in deep learning (DL) have enabled more robust approaches to modeling sequential data. Long Short-Term Memory (LSTM) networks, in particular, are effective for ionospheric studies because they capture long-term temporal dependencies and nonlinear dynamics [8], [9]. Prior research has demonstrated the potential of LSTM-based TEC prediction across different regions, including India and China [10], [11].

Building on these efforts, this work develops a multivariate stacked LSTM framework that integrates the Kp index with GNSS-derived TEC data from the IISc Bangalore station during the solar ascending phase (December 2019 – December 2024). Unlike earlier studies that emphasized solar flux, this approach prioritizes geomagnetic forcing by explicitly incorporating the Kp index. The framework evolves from a simple LSTM to a more complex stacked architecture with dense and dropout layers, thereby enhancing learning capacity and generalization.

The contributions of this study are threefold:

(i) demonstrating the predictive capability of multivariate LSTMs for TEC modeling in the Indian equatorial region during Solar Cycle 25,
(ii) emphasizing the role of geomagnetic indices in improving forecasting accuracy, and
(iii) analyzing the impact of model depth and architecture on predictive performance. Furthermore, this work establishes a foundation for future hybrid TEC prediction systems that can integrate multiple space weather indicators and advanced DL models to strengthen GNSS reliability under disturbed ionospheric conditions.

II. GEOPHYSICAL BACKGROUND: SOLAR CYCLE AND KP INDEX

The Sun's activity follows an approximately **11-year solar cycle**, characterized by variations in solar radiation, sunspot numbers, coronal mass ejections (CMEs), and solar wind streams. These changes have a significant impact on the Earth's upper atmosphere and ionosphere, resulting in variations in Total Electron Content (TEC) and other space weather phenomena. The current study focuses on the **ascending phase of Solar Cycle 25 (December 2019–December 2024)**, a period known for increasing solar and geomagnetic disturbances [1], [2].

One of the critical manifestations of solar activity on Earth's space environment is through **geomagnetic storms**, which result from enhanced interactions between solar wind and Earth's magnetosphere. These disturbances cause large-scale ionospheric fluctuations that severely impact GNSS signal propagation. To quantify the level of geomagnetic disturbance, the **planetary K-index (Kp)** is widely used. The Kp index is a quasi-logarithmic measure of geomagnetic activity derived from magnetometer observations at subaerial stations worldwide [3].

The Kp index ranges from 0 (quiet conditions) to 9 (extreme geomagnetic storm). Values above 4 typically indicate significant ionospheric variability. Unlike solar flux indices like **F10.7**, which reflect continuous solar output, the Kp index captures **short-term disturbances** such as CMEs and high-speed solar wind streams, making it highly relevant for modeling the dynamic effects of space weather on ionospheric TEC [4].

Incorporating the **Kp index** as an input feature in TEC prediction models enables a more accurate representation of externally driven disturbances, particularly during geomagnetic storm periods, which are challenging to model using TEC history alone. Prior studies have shown that models augmented with geomagnetic parameters such as Kp or Dst improve TEC forecasting accuracy in both equatorial and mid-latitude regions [5], [6].

This study utilizes the Kp index within a multivariate LSTM architecture to capture both the internal temporal dependencies in TEC and its external modulation due to geomagnetic activity. This combination is crucial during periods of solar cycle disturbance.

III. METHODOLOGY

Since the behavior of TEC is highly nonlinear and dynamic, particularly during geomagnetic disturbances, we must consider the complex manner in which TEC can change. To address this issue, we developed a deep learning framework based on Long Short-Term Memory (LSTM) networks to characterize TEC observations using GNSS and Kp observations during disturbed geomagnetic conditions. This section describes the data sources, preprocessing steps, sequence generation process, LSTM fundamentals, and the two prediction models evaluated in this study.

A. Data Collection

Data on ionospheric and geomagnetic phenomena are included in this dataset:

- **TEC Data**: Vertical TEC values were downloaded from a site located at the Indian Institute of Science (IISc), Bengaluru (13.0170° N, 77.5659° E), from the International GNSS Service (IGS), via the Scripps Orbit and Permanent Array Center (SOPAC) [1].
- **Geomagnetic Data:** Planetary Kp index values, a global measure of geomagnetic activity, were downloaded from NASA's OMNIWeb [2].

The time period for the data set spans from December 2019 to December 2024, during the ascending phase of Solar Cycle 25 and periods of increased variability in the ionosphere.

B. Data Preprocessing

To prepare the raw data for model training:

- **Outlier removal and interpolation:** Spurious data points were corrected, and missing values were filled to maintain continuity.
- **Normalization:** Both TEC and Kp values were scaled using Min–Max normalization to the range [0,1], improving learning stability.

- **Temporal alignment:** The corresponding TEC and Kp indices were aligned at a day level.
- **Train-test split:** The data was separated chronologically into a training (80%) and a test (20%) set, provided that it was not shuffled and temporal dependencies were preserved.

C. Sequence Generation

The LSTMs operate on sequential data; hence, a **sliding window approach** was adopted.

- For each prediction, the model takes the **previous 30 days** of TEC and Kp values as input and predicts the TEC value of the next day.
- Thus, the input tensor is (samples, 30, 2 features).
- This step takes unformatted time series and formats them into supervised sequences, allowing the model to learn short

fluctuations and medium-term dependencies.

E. Fundamentals of the LSTM Cell

The Long Short-Term Memory (LSTM) network is a variant of the Recurrent Neural Network (RNN) designed to overcome vanishing and exploding gradient problems. Thus, LSTM is particularly effective in modeling ionospheric TEC, which has time dependencies and has longer cycle dependencies linked to solar and geomagnetic cycles.

An LSTM cell consists of three main gates**: the forget gate, the input gate, and the output gate, which** control the flow of information and facilitate the selective remembering or forgetting of historical context.

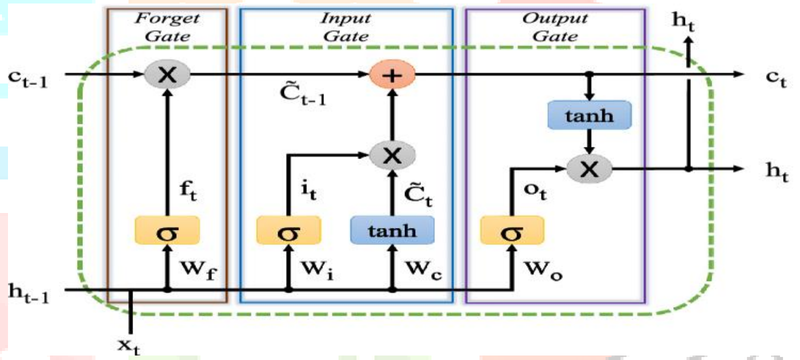


Figure 4: LSTM Cell Architecture

- **Forget Gate**

  The forget gate determines which information from the previous cell state, Ct−1, should be discarded. It takes the last hidden state, h_{t-1}, and the current input x_t and passes them through a sigmoid function.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

  A value close to 0 indicates forgetting, while a value close to 1 indicates retaining the information.

- **Input Gate and Candidate State**

The input gate decides what new information will be added to the cell state. It comprises two components:

  - **Input gate activation**:

$$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i \qquad (2)$$

  - **Candidate values** (to be added to state):

$$\tilde{C}_t = tanh(W_C.[h_{t-1}, x_t] + b_C \qquad (3)$$

These are combined to update the cell state:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \qquad (4)$$

- **Output Gate and Hidden State**

$$ot = \sigma(Wo \cdot [ht - 1, xt] + bo) \quad (5)$$

The output gate decides what part of the cell state becomes the output (hidden state) for the current time step:

$$h_t = o_t \cdot tanh\,(C_t) \quad (6)$$

The internal architecture of the LSTM cell is illustrated in Figure 4. This gating mechanism enables LSTM to capture short-term variations (e.g., daily TEC changes) and dependence on long-term information (e.g., solar and geomagnetic cycles).

Need for LSTM in TEC Prediction

The multiscale nature of ionospheric Total Electron Content (TEC) fluctuations stems from long-term and short-term processes, including solar cycle activity, geomagnetic disturbances, seasonal cycles, and diurnal cycles. Most conventional forecasting techniques (e.g., ARIMA or simple RNNs) struggle to capture these multi-scale characteristics because they suffer from nonlinearities and lengthy time dependencies.

Long Short-Term Memory (LSTM) networks can overcome these limitations due to their gated memory mechanism that can selectively maintain or ignore historical information. The ability of LSTMs to learn long-term trends and quickly adapt to sudden and unexpected changes in TEC makes it possible to use LSTMs for forecasting Tropopause conditions in both quiet and disturbed space weather conditions.

F. Model Development

To predict the ionospheric Total Electron Content (TEC) during the solar ascending phase (December 2019 to December 2024) over the Indian region, two Long Short-Term Memory (LSTM) based deep learning models were developed and evaluated. Both models utilize multivariate time series input, incorporating TEC and the planetary K-index (Kp), a critical geomagnetic indicator of ionospheric variability. The modeling framework was designed progressively, beginning with a basic LSTM architecture and advancing to an improved stacked LSTM structure to assess the impact of model depth and regularization on prediction performance.

1) Basic LSTM Model

The fundamental baseline model is a shallow LSTM architecture that employs a single LSTM layer followed by a dense output unit. This model is intentionally kept simple to serve as a benchmark and to assess the core learning capability of LSTM for capturing sequential dependencies in TEC variation.

- **Input Configuration**:
  The model processes a 30-day input window of two features, TEC and Kp index, represented as a 3D tensor of shape *($n_{samples}$, 30, 2)*.
- **LSTM Layer**:
  A single LSTM layer with 50 units is employed to learn temporal patterns across the input sequence. This layer captures medium-range dependencies and basic fluctuations relevant to ionospheric behavior.
- **Output Layer**:
  A dense layer with a single neuron and linear activation is used to predict the TEC value for the next day.
- **Compilation**:
  The model is trained using the Mean Squared Error (MSE) loss function and the Adam optimizer for efficient convergence.

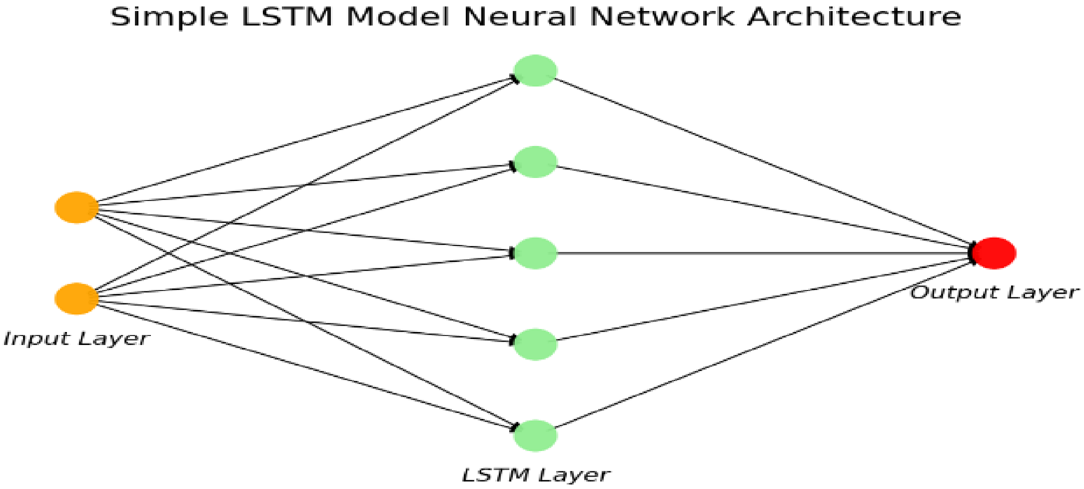


Figure 5: Simple LSTM Model Neural Network Architecture

This basic architecture provides a reference performance baseline and helps in quantifying improvements offered by deeper, regularized models.

2) Improved LSTM Model

To enhance the model's capacity for learning complex, nonlinear, and long-range temporal relationships, a deeper LSTM architecture was developed. This improved model consists of multiple stacked LSTM layers, followed by dense and dropout layers, which enable both temporal abstraction and regularization.

- **Input Configuration**:
  Similar to the basic model, this architecture also utilizes a 30-day sequence of TEC and Kp indices with a shape of *(samples, 30, 2)*.
- **First LSTM Layer (128 units)**:
  Captures broad temporal dependencies and overall trends in the input sequence.
- **Second LSTM Layer (64 units)**:
  Learns intermediate-level sequential features, refining the representation from the first layer.
- **Third LSTM Layer (32 units)**:
  Compresses the temporal data into a compact context vector, extracting salient features relevant to TEC dynamics.
- **Dense Layer 1 (128 neurons, ReLU)**:
  Transforms the learned temporal features into a high-dimensional space, allowing the network to capture intricate nonlinearities.
- **Dropout Layer 1 (rate = 0.3)**:
  Introduces regularization by randomly deactivating 30% of the neurons during training, reducing overfitting and enhancing generalization.
- **Dense Layer 2 (64 neurons, ReLU)**:
  Further processes the feature representation, enabling the model to learn higher-level abstractions.
- **Dropout Layer 2 (rate = 0.3)**:
  Provides additional regularization to prevent over-reliance on specific neurons.
- **Dense Layer 3 (32 neurons, ReLU)**:
  Acts as a compression layer to reduce dimensionality and eliminate noisy activations before prediction.
- **Output Layer**:
  A final dense layer with one linear activation unit outputs the predicted TEC value.

The enhanced architecture is designed to improve prediction accuracy during periods of heightened ionospheric disturbance. The use of stacked LSTM layers facilitates hierarchical temporal learning, while dropout and dense transformations aid in capturing nonlinear and robust patterns in the TEC–Kp relationship.

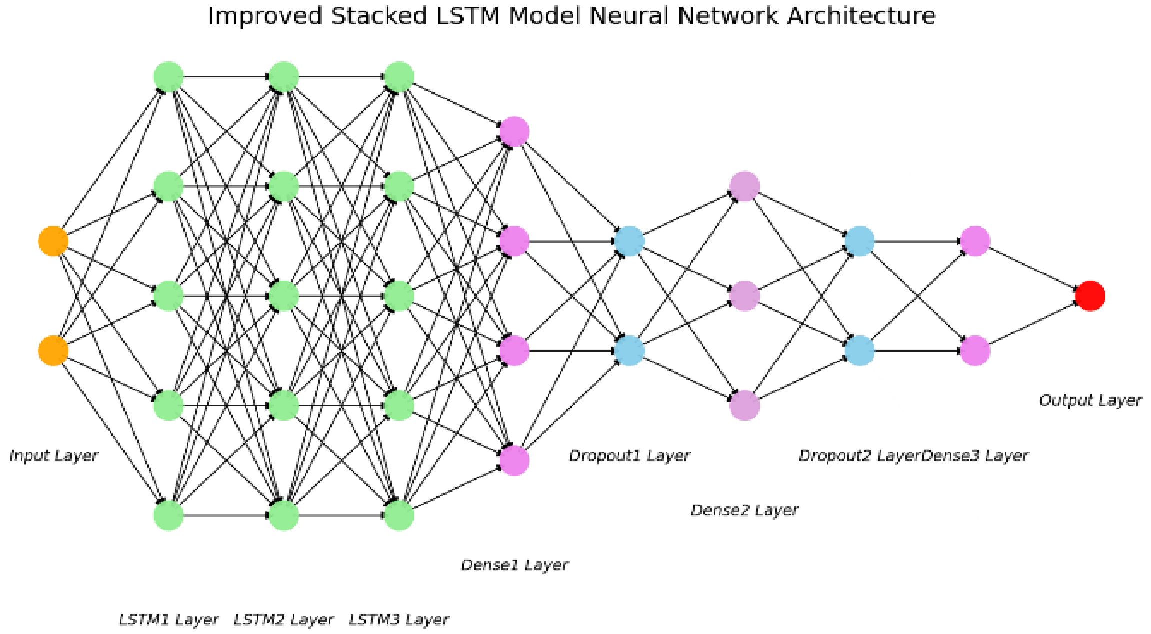


Figure 6: Improved LSTM Model Neural Network Architecture

IV. OUTCOMES AND DISCUSSION

This section presents a comparative evaluation of different Long Short-Term Memory (LSTM) models in forecasting ionospheric Total Electron Content (TEC) over the Indian region using data from the **IISc Bangalore GNSS station** (13.0170° N, 77.5659° E) during the solar ascending phase from **December 2019 to December 2024**.

The primary objective is to evaluate how architectural depth and the incorporation of geomagnetic features, such as the planetary Kp index, impact prediction accuracy. The models were assessed using standard statistical metrics: **Mean Absolute Error (MAE)**, **Mean Squared Error (MSE)**, **Root Mean Squared Error (RMSE)**, **Mean Absolute Percentage Error (MAPE)**, and the **Coefficient of Determination (R^2 Score)** [13].

Two model architectures were developed and compared:

- **Baseline LSTM**: A minimal model comprising a single LSTM layer with 50 units and a final Dense output layer, trained solely on input sequences of TEC and Kp index.
- **Improved LSTM**: A deeper architecture that stacks three LSTM layers (128, 64, and 32 units), followed by Dense layers with ReLU activations and Dropout (rate = 0.3) to prevent overfitting. This design aims to capture complex temporal dependencies and nonlinear interactions more effectively.

To evaluate the impact of training duration, both models were trained for **50, 100, and 150 epochs** using the Adam optimizer with a batch size of 32. The dataset was normalized using MinMax scaling, and an 80:20 chronological split was used for training and testing.

Model performances are tabulated and compared based on their error metrics. The findings are discussed in relation to the suitability of each model for real-world GNSS applications under varying ionospheric conditions.

A. Evaluation Criteria

The effectiveness of the developed LSTM model was evaluated using a variety of assessment metrics, each offering a unique view of the model's forecasting abilities:

- **Mean Absolute Error (MAE):**

  Represents the average of the absolute differences between predicted and actual TEC values, reflecting overall prediction error magnitude. A reduced MAE indicates improved model accuracy.

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

- **Mean Squared Error (MSE):**
  Emphasizes larger errors by squaring the differences, which is useful for penalizing large deviations in TEC predictions.

$$MSE = \frac{1}{N}\sum_{i=0}^{n}(y_i - \hat{y}_i)^2 \qquad (8)$$

- **Root Mean Squared Error (RMSE):**
  The square root of MSE provides an interpretable measure of average prediction error in the same unit as TEC.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (9)$$

- **R² Score (Coefficient of Determination):**

  Indicates how well the model explains the variance in the observed data. A higher R^2 denotes a better model fit.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \qquad (10)$$

- **Mean Absolute Percentage Error (MAPE):**
  Expresses prediction accuracy as a percentage, helpful for understanding relative errors across varying TEC values.

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_y}\right| \quad (11)$$

B. Performance Comparison

The table below summarizes the performance of the two LSTM architectures under different training epoch conditions.

Table 1: *Performance metrics for the Baseline LSTM Model*

Epochs	MAE	MSE	RMSE	R²	MAPE
50	0.0224	0.0012	0.0358	0.9404	5.94%
100	0.0232	0.0013	0.0365	0.9380	5.79%
150	0.0230	0.0013	0.0362	0.9388	5.87%

Table 2: *Performance metrics for the Improved LSTM Model*

Epochs	MAE	MSE	RMSE	R²	MAPE
50	0.0191	0.0010	0.0323	0.9513	6.25%
100	0.0187	0.0009	0.0306	0.9562	4.80%
150	0.0187	0.0009	0.0304	0.9570	5.18%

C. Analysis and Observations

- The **Simple LSTM model** shows consistent learning with increasing epochs, but its performance plateaus after 100 epochs. This suggests it captures the basic temporal dependencies in TEC data but lacks the architectural capacity to model more complex patterns.
- The **Improved LSTM architecture**, which incorporates **stacked LSTM layers**, **Dropout regularization**, and a **Dense output layer**, significantly outperforms the baseline across all metrics. This highlights the effectiveness of deeper architectures in capturing the nonlinear and dynamic nature of ionospheric TEC variations influenced by geomagnetic activity.
- The **lowest MAE and RMSE** are observed at **150 epochs** in the improved model, suggesting better generalization and retention of temporal features.
- The model achieves an **R^2 score of 0.95700**, demonstrating a strong fit to the target TEC values and high explanatory power. This validates its robustness in accurately and reliably forecasting TEC for real-world GNSS error mitigation.

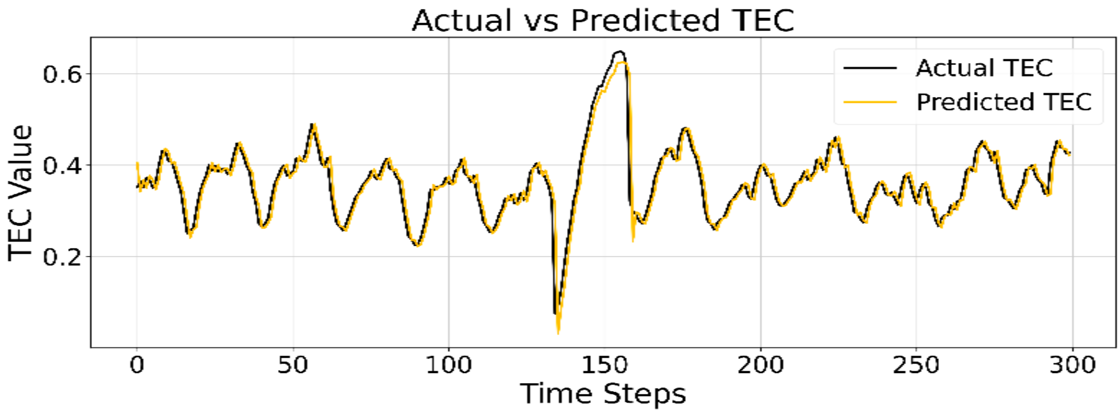


Figure 7: Actual vs Predicted TEC — for Simple LSTM at 100 epochs.

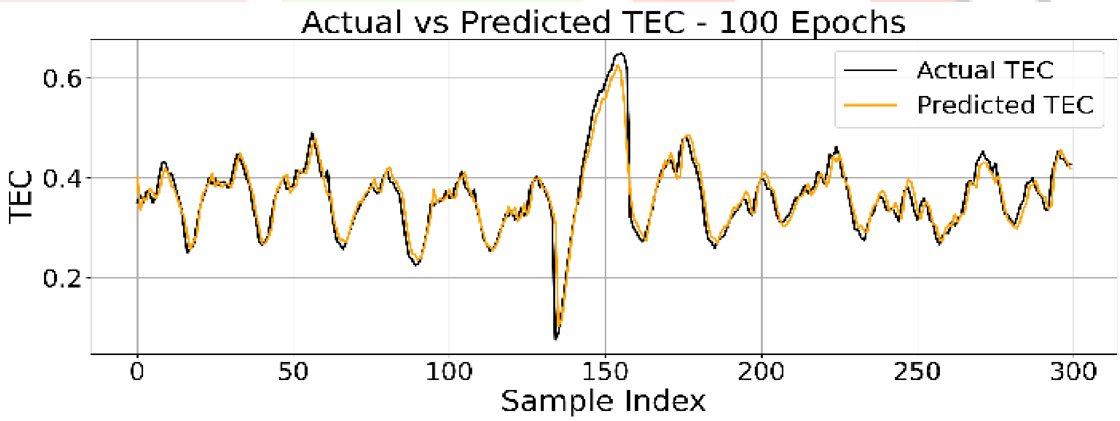


Figure 8: Actual vs Predicted TEC — for Improved LSTM at 100 epochs.

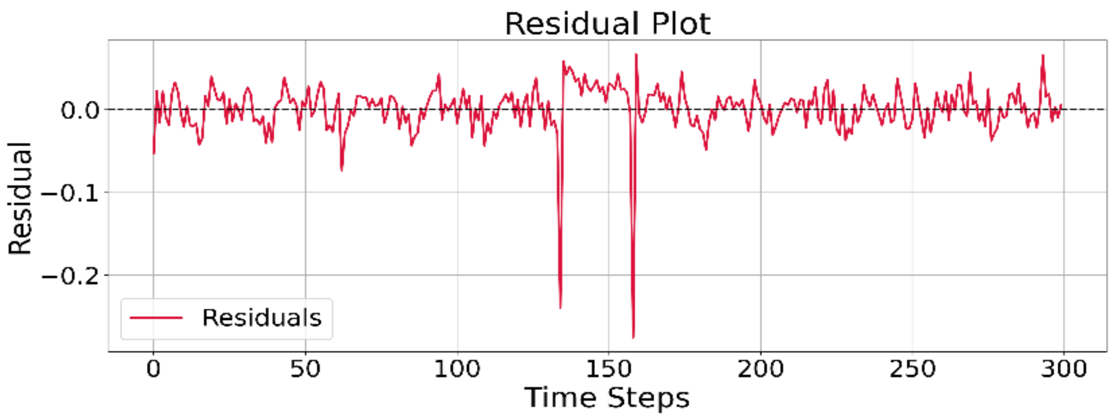


Figure 9: Residuals plot – simple LSTM at 100 epochs

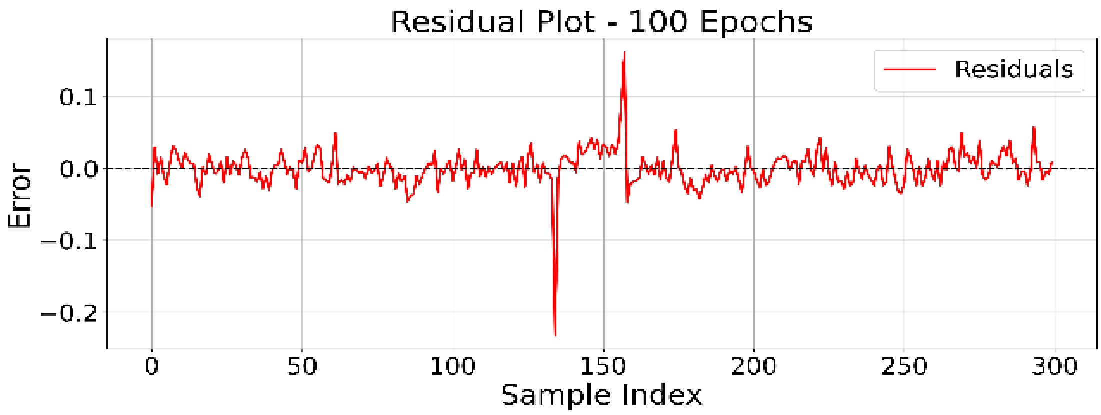


Figure 10: Residuals plot – improved LSTM at 100 epochs

V. CONCLUSION

In this study, a deep learning-based predictive framework leveraging Long Short-Term Memory (LSTM) networks was developed to forecast ionospheric Total Electron Content (TEC) over the Indian region. The analysis focused on the ascending phase of **Solar Cycle 25** (December 2019 to December 2024), a period marked by heightened ionospheric variability due to increased solar and geomagnetic activity. The TEC data was sourced from the **IISc Bangalore GNSS station**, and supplemented with **Kp index** values to incorporate geomagnetic conditions as an additional feature. A progressive modeling approach was adopted—starting with a simple, single-layer LSTM model and evolving into a deeper, **stacked LSTM architecture** with **Dropout layers** and **dense output connections**. The improved model demonstrated **superior performance** across all standard evaluation metrics, including **Mean Squared Error (MSE)**, **Root Mean Squared Error (RMSE)**, **Mean Absolute Error (MAE)**, **Mean Absolute Percentage Error (MAPE)**, and **R^2 Score**. Notably, the **best results were observed at 150 training epochs**, with the improved LSTM achieving an **R^2 score of 0.95700**, reflecting a strong correlation between predicted and actual TEC values. These results underscore the **effectiveness of incorporating domain-relevant features** (like the Kp index) and **architectural depth** in neural network design for spatiotemporal forecasting tasks. The improved forecasting accuracy directly contributes to mitigating ionospheric delay errors in **GNSS-based positioning**, particularly in equatorial and low-latitude regions that are highly susceptible to space weather anomalies.

7. REFERENCES

[1] G. S. Rao, *Global Navigation Satellite Systems with Essentials of Satellite Communications*, McGraw-Hill Education, New Delhi, India, 2010.

[2] B. Bidikar, G. S. Rao, L. Ganesh, and M. N. V. S. Santosh Kumar, "Orthogonal decomposition algorithm for ionospheric delay estimation for precise GPS applications," *Procedia Computer Science*, vol. 85, pp. 303–310, 2016. doi: 10.1016/j.procs.2016.05.232

[3] N. Shiva Kumar, G. S. Rao, L. Ganesh, and B. Bidikar, "Machine learning approach for prediction of total electron content and classification of ionospheric scintillations over Visakhapatnam region," *AIP Conference Proceedings*, vol. 2889, 2023. doi: 10.1063/5.0176196

[4] L. Ciraolo, "Ionosphere total electron content from GPS data: How stable are the differential satellite and receiver biases?" *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 60, no. 4, pp. 410–413, 1998.

[5] P. Themens, P. Jayachandran, A. McCaffrey, et al., "Using machine learning to improve GNSS ionospheric correction models," *Radio Science*, vol. 54, no. 4, pp. 345–359, 2019.

[6] J. C. Hernández-Pajares, J. M. Juan, and M. Sanz, "GNSS Ionospheric Mapping: Methods and Applications," *Springer Handbook of Global Navigation Satellite Systems*, pp. 763–795, 2017.

[7] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[8] Y. He, Z. Wu, and Y. Chen, "A novel deep learning-based model for regional ionospheric TEC prediction," *Advances in Space Research*, vol. 63, no. 1, pp. 291–301, 2019.

[9] M. Jiao, W. Wan, and L. Liu, "A deep LSTM network for predicting ionospheric total electron content," *Journal of Geophysical Research: Space Physics*, vol. 124, no. 9, pp. 7892–7905, 2019.

[10] D. Yadav, M. Shrivastava, A. Dabas, et al., "Ionospheric TEC modeling using LSTM neural network for low-latitude stations," *GPS Solutions*, vol. 26, 2022.
[11] NASA OMNIWeb, Kp Index Data Source, https://omniweb.gsfc.nasa.gov/form/dx1.html
[12] Scripps Orbit and Permanent Array Center (SOPAC), UC San Diego, https://sopac-csrc.ucsd.edu/