Spatiotemporal prediction of total electron content using CONVLSTM, Patch- CONVLSTM and 3D-U-Net models

Arezoo Mokhtari[1*], Jamal Asgari[1], Alireza Amiri-Simkooei[2], Jamshid Maleki[1]

[1] Department of Geomatics Engineering, Faculty of Civil Engineering and Transportation, University of Isfahan, Isfahan, Iran
[2] Department of Control and Operations, Delft University of Technology, Delft, the Netherlands

Keywords: Total Electron Content, CONVLSTM, 3-D U-Net, Patch-CONVLSTM.

Abstract

Given the importance of predicting the total ionospheric electron content (TEC), many studies have attempted to predict its spatiotemporal nature. In this study, a patch-based convolutional neural network with long short-term memory (CONVLSTM) (with patch sizes of 5 and 15), simple CONVLSTM, and 3D-U-Net models were used to predict the spatiotemporal nature of the next day's TEC data (next 12 samples). The proposed models use the spatiotemporal nature of the previous day's TEC data (previous 12 samples) along with temporal data such as AP, KP, DST, SN, and F10.7 to predict the next day's TEC data. The results showed that the 3D-U-Net model and then the model with patch size 5 had a higher generalization ability than the classical CONVLSTM architectures, while reducing the RMSE and MAE. The execution time of the program in the 3D-U-Net model has been significantly reduced compared to other models and it has also been able to better extract the microstructural features of TEC maps.

1. Introduction

The ionosphere is located approximately 60 to 1200 km above the Earth's surface and is an important component of the Earth's atmosphere and a major area of space activity that has direct and indirect effects on radio communications, navigation, and positioning (Zhao et al., 2018). TEC is one of the important parameters in the study of the ionosphere. In order to evaluate the behavior of the ionosphere worldwide, TEC is considered as an important tool and its unit of measurement is TECU (1TECU=10^{16}electrons/m2) (Zhao et al., 2018). TEC in the ionosphere significantly affects the propagation of radio waves through the ionosphere (Yeh & CH, 1972; Zhao et al., 2018). When the frequency of a signal is known, the ionospheric delay can be calculated by simply determining the TEC in the signal transmission path. Therefore, this parameter can be used to describe the ionospheric delay of satellite signals (Zhao et al., 2018). Obtaining information about these structures and changes in electron content allows for a better understanding of space weather, tectonic activity such as earthquakes and tsunamis, and the identification of human-made events such as large explosions (Afraimovich et al., 2013; Afraimovich et al., 2000; Artru et al., 2005; Calais & Bernard Minster, 1996; Calais et al., 1998; Calais & Minster, 1995; Dickinson & Tamarkin, 1965; Fitzgerald, 1997; Kazimirovsky et al., 2003; Liu et al., 2004; Tsugawa et al., 2011). Given the importance of calculating the TEC parameter, various methods have been used for this purpose. Multi-frequency GNSS receivers, by utilizing multiple ionospheric delays at different frequencies, enable accurate TEC calculation (Ciraolo et al., 2007; Kim et al., 2014; Sezen et al., 2013; Yang et al., 2018), while single-frequency users and ground stations have access to only their local data. For global coverage, IGS has been integrating and providing TEC maps for free since 1998 (Benoit & Petry, 2021). The lack of continuous global measurements has underscored the need to develop empirical ionospheric models; these models use spatiotemporal analyses of long-term datasets to simulate electron density perturbations (Feng, Jiang, et al., 2017; Feng et al., 2016; Feng, Wang, et al., 2017; Hoque & Jakowski, 2012; Jakowski et al., 2011; Mukhtarov et al., 2013) and are capable of providing a comprehensive depiction of TEC variations (Feng et al., 2016). Some of these empirical models may not be accurate enough in all conditions. In addition, data interpolation also plays a key role in predicting instantaneous electron density changes throughout the day (Benoit & Petry, 2021). In recent years, various deep learning methods such as Long short-term memory (LSTM) neural networks, convolutional neural network (CNN) and CONVLSTM neural network have been used for TEC prediction (Kaselimi et al., 2021; Muhammad & Külahcı, 2022; Sun et al., 2017). Considering that global ionospheric TEC data have temporal and spatial features, in 2021, a CONVLSTM-based network was developed to extract features and reduce the dimensionality of historical ionospheric TEC data to train the model and finally achieve the prediction of spatiotemporal series of global ionospheric TEC maps (Fubin et al., 2021). Many studies have used this model for spatiotemporal prediction of TEC data (Jeong et al., 2024; Lin et al., 2020; Xia et al., 2022). The CONVLSTM model overcomes the shortcomings of the LSTM network in capturing spatial features with spatial dimension convolution operations, so it is often used in precipitation forecasting or object motion prediction (Shi et al., 2015). CONVLSTM method is able to use the spatial and temporal distribution of data in its prediction. Therefore, it is concluded that CONVLSTM method is a reliable and efficient method for accurate prediction of ionospheric conditions (Jeong et al., 2024).

also, several studies have investigated the role of image patching in improving the performance of deep learning models (Jalali et al., 2024; Wang et al., 2025; Zhang et al., 2024). Dividing the image into smaller patches allows the model to extract local features and spatial details with higher accuracy. In patch-based methods, the original image is divided into regular or overlapping sections, and each patch is used as an independent input to the neural network. This makes the model more sensitive to subtle changes in small areas in addition to understanding the overall structure of the image. On the other hand, combining this approach with spatiotemporal models such as CONVLSTM has been able to provide acceptable results in applications such as video prediction, climate change prediction, and satellite image analysis. Overall, using the patching strategy alongside CONVLSTM-based models is an efficient approach to more effectively extract local features and improve the accuracy of predictions based on image data (Nie, 2022).

Recent studies have applied U-Net variants (including 3D-U-Net) to TEC forecasting, frequently integrating

---

[*] Corresponding author

solar/geomagnetic indices and temporal fusion modules (Wang et al., 2023). Such studies justify the inclusion of a 3D-U-Net baseline in our experiments.

This paper presents a systematic empirical comparison of flattened patching strategy and multiple backbones for TEC forecasting. In this study, we have examined the prediction of TEC maps for the next 24 hours using the CONVLSTM method in grid and patched in dimensions 15*15 and 5*5, and a simple 3D-U-Net model. The results showed that 3D-U-Net model and then CONVLSTM model based on Patching have a higher generalization ability in predicting TEC values.

2. The Dataset

In this study, we employ a high-fidelity dataset that captures both the temporal and spatial dimensions of the ionosphere with great precision. Temporally, the data are collected at two-hour intervals (12 samples per day). The rapid combined TEC maps (named IGRG) are used as reference data to train and evaluate the model, which is obtained from NASA's Crustal Dynamics Data Information System (CDDIS) website (https://cddis.nasa.gov/archive/gnss/products/ionex/).

Spatially, the TEC maps represented as 71 × 73 matrices. These matrices span longitudes from –180° to +180° in 5° increments and latitudes from –87.5° to +87.5° in 2.5° increments.

In addition to the TEC maps, the dataset includes the following indices: F10.7 (solar radio flux), Ap (daily average geomagnetic activity), DST (magnetic storm index), SN (sunspot number), and Kp (global geomagnetic disturbance index). Each of these indices has been temporally interpolated (linear interpolation) to match the two-hour sampling rate and is incorporated as a complementary feature in the model. At each time step, a sequence of 12 consecutive samples (equivalent to one full day) of TEC values together with the synchronized index values and time information (day-of-year and hour) is fed into the model. The model then uses this input array to predict the next 12 samples (the following day).

3. Methodology

In this study, we forecast TEC maps using a CONVLSTM, Patched-CONVLSTM (patch size 5 ,15) and 3D-U-Net models.

3.1 Data Preprocessing

This study employs data from 2014 to 2022. The temporal resolution is set to two-hour intervals (12 samples per day). For each sample, a 71×73 spatiotemporal TEC map is extracted together with five indices: F10.7, Ap, DST, SN, and Kp.

At each time step i, an input sequence of T = 12time steps and a target sequence of T_{pe} = 12time steps are defined according to equations (1) and (2).

$$\mathbf{X_i} - [x_i, x_{i+1}, \ldots, x_{i+T-1}] \in R^{T\times F} \tag{1}$$

$$\mathbf{Y_i} = \left[y_{i+T}, y_{i+T+1}, \ldots, y_{i+T+T_{pred}-1}\right] \in R^{T_{pe}\times H\times W} \tag{2}$$

In equations (1) and (2), i is the sample index; T is the length of the input sequence; T_{pe} is the length of the target sequence; F is the number of input features per time step; H and W are the spatial dimensions of the TEC maps. Moreover, $x_k \in \mathbb{R}^f$ denotes the five-dimensional environmental feature vector at time step k, and $y_k \in \mathbb{R}^{H\times W}$ denotes the TEC map at time step k. For sample i, the input matrix $X_i \in \mathbb{R}^{T\times f}$ comprises the T time steps, and the target tensor $Y_i \in R^{T_{pe}\times H\times W}$ comprises the T_{pe} future TEC maps.

3.2 Padding and Patch Division

In the proposed model, to capture fine-scale spatial features of the 71 × 73 TEC maps, each map is first padded—according to equation (3)—to expand its dimensions to 75 × 75. It is then divided into 25 patches of size 15 × 15 and 225 patches of size 5×5 using equation (4). This partitioning not only reduces computational load but also allows each patch to be treated as an independent unit with its own localized features.

Each 15 × 15 and 5×5 patches is then flattened into a 225-dimensional and 25-dimentional vector and arranged along the temporal dimension (12time steps), forming an ordered sequence of spatial tokens for each input sample, as described in equation (5).

$$\widetilde{y_k} = \mathrm{Pad}(y_k, (75-H, 75-W)) \in R^{75\times 75} \tag{3}$$

$$\mathrm{Patch}_{p,q}^{(k)} = \widetilde{y_k}\lfloor pP:(p+1)P,\ qP:(q+1)P\rfloor \tag{4}$$

$$\mathbf{z_{p,q}^{(k)}} = \mathrm{vec}\left(\mathrm{Patch}_{p,q}^{(k)}\right) \in R^{P^2} \tag{5}$$

In equations (3) to (5), $\widetilde{y_k} \in R^{75\times 75}$ represents the zero-padded TEC maps at time step k; $y_{k\in R^{H\times W}}$ denotes the original TEC maps at time step k, H is the number of original rows of the TEC maps; W is the number of original columns; P is the size of each patch; n is the number of patches in each dimension; $\mathrm{Pad}(.)$ is the zero-padding function; vec is the function for converting a matrix into a vector; $\mathrm{Patch}_{p,q}^{(k)}$ is the patch at position (p,q) from map k and $\mathbf{z_{p,q}^{(k)}} \in R^{P^2}$ is the vector corresponding to patch (p,q), where p and q are the row and column indices of the patch, respectively.

Afterward, for each sample ii and each time step t, the combined token vector consisting of spatial and temporal features is constructed as shown in equation (6).

$$\mathbf{Z_i}[t,k] = \left[\mathbf{z_{p,q}^{(k)}};\ x_{i+t};\ (d_{i+t}, h_{i+t})\right]\mathbf{Z_i} \subset R^{T\times n_{pth}\times(P^2+F+2)} \tag{6}$$

In equation (6), $\mathrm{Z_i} \in R^{T\times n_{pth}\times(P^2+F+2)}$ represents the combined token vector for each sample, $\mathbf{z_{p,q}^{(k)}} \in R^{P^2}$ denotes the spatial patch data, $x_{i+t} \in R^F$ is the five-dimensional vector of environmental features, d_{i+t} is the day of the year at time step $i+t$, h_{i+t} is the hour of the day at time step $i+t$, P^2 is the spatial patch dimension, F is the number of input feature dimensions, n_{pth} is the total number of patches, and the number 2 refers to the inclusion of the day-of-year and hour-of-day indices.

3.3 Build Patch CONVLSTM Model

The model architecture consists of two parallel subnetworks:

3.3.1 Spatial Subnetwork

The input tokens first pass through a Time Distributed (Dense) layer and are embedded as patch-based representations with an embedding dimension of 32, as defined in equation (7). Finally, they are reshaped according to equation (8).

$$\mathbf{E} = \mathrm{Dense}_{\mathrm{emb}}(\mathbf{Z}) \in R^{T\times n_{\mathrm{pth}}\times D} \tag{7}$$

$$\mathbf{E}' = \mathrm{Reshape}(\mathbf{E}) \in R^{T\times n\times n\times D} \tag{8}$$

In equations (7) and (8), $\mathrm{Dense}_{\mathrm{emb}}$ refers to the fully connected (dense) layer, D is the embedding dimension, $\mathbf{E} \in R^{T\times n_{\mathrm{pth}}\times D}$ is the embedding output, Reshape(.) denotes the reshaping operation, *n* is the number of patches along each dimension, and $\mathbf{E}' \in R^{T\times n\times n\times D}$ is the output after reshaping. Then, two consecutive CONVLSTM layers with 32 and 16 filters, along with batch normalization, simultaneously modeled spatial and temporal structures as described in equations (9) to (12). Afterward, UpSampling3D and cropping operations (equation (13)) were used to restore the original dimensions of the TEC maps.

$$H_1 = ConvLSTM2D(32, (3,3), return\_sequences = True)(E') \tag{9}$$

$$\mathbf{H_1}' = \mathrm{BatchNorm}(\mathbf{H_1}) \tag{10}$$

$$H_2 = ConvLSTM2D(16, (3,3), return\_sequences = True)(H_1') \tag{11}$$

$$\mathbf{H_2}' = \mathrm{BatchNorm}(\mathbf{H_2}) \tag{12}$$

$$\mathbf{H_3} = \mathrm{Crop}(\mathrm{Upsample3D}(\mathbf{H_2}')) \in R^{T\times H\times W\times C_1} \tag{13}$$

In equations (9) to (13):

$\mathrm{ConvLSTM2D}(f, (k, \mathrm{k}), \mathrm{return\_sequences{=}True})$ represents the 2D CONVLSTM layer with f filters of size *(k × k)*, BatchNorm(.) is the batch normalization layer, H_1 and H_2 are the intermediate outputs before normalization, H_1' and H_2' are the outputs after normalization, Upsample3D(.) denotes the 3D upsampling operation, Crop(.) refers to the cropping operation used to precisely restore the original dimensions, C_1 is the number of output channels, and $\mathbf{H_3} \in R^{T\times H\times W\times C_1}$ represents the final output.

3.3.2 Temporal Subnetwork

The temporal features (indices and time vector) were processed through an LSTM layer with 32 units as shown in Equation (14), and then, using Equations (15) and (16), they were mapped into a shape of **1×71×73** through a Dense layer followed by a Reshape operation.

$$T_1 = LSTM(32, return\_sequences = True)([x_i \cdots x_{i+T-1}]) \tag{14}$$

$$\mathbf{T_2} = \mathrm{Dense}(H\times W)(\mathbf{T_1}) \tag{15}$$

$$\mathbf{T_3} = \mathrm{Reshape}(\mathbf{T_2}) \in R^{T\times H\times W\times 1} \tag{16}$$

In Equations (14) to (16), $\mathrm{LSTM}(32, \mathrm{return\_sequences{=}True})$ is an LSTM layer with 32 hidden units, $\mathrm{Dense}(H\times W)$ is a fully connected layer that maps the LSTM hidden vector to a length of $(H\times W)$, and $\mathbf{T_3} \in R^{T\times H\times W\times 1}$ is the output of the temporal branch.

The outputs of the two subnetworks were combined using a Concatenate operation along the channel dimension according to Equation (17), and then passed through two Conv3D layers as shown in Equations (18) and (19)—one with 8 filters for feature fusion and the other with 1 filter for the final prediction—to generate the predicted TEC maps.

$$\mathbf{M} = \mathrm{Concatenate}([\mathbf{H_3}, \mathbf{T_3}]) \in R^{T\times H\times W\times (C_1+1)} \tag{17}$$

$$\widehat{\mathbf{Y}} = \mathrm{Conv3D}(8, (1,3,3), \mathrm{ReLU})(\mathbf{M}) \tag{18}$$

$$\widehat{\mathbf{Y}} = \mathrm{Conv3D}(1, (1,3,3), \mathrm{linear})(\widehat{\mathbf{Y}}) \tag{19}$$

In Equations (17) to (19), $\mathrm{Concatenate}([.,.])$ denotes the merging of spatial and temporal outputs along the channel dimension, $C_1 + 1$ is the number of channels after concatenation, $\mathrm{Conv3D}(8, (1,3,3), \mathrm{ReLU})$ is a 3D convolution layer with 8 filters, a kernel size of (1,3,3), and a ReLU activation function, $\mathrm{Conv3D}(1, (1,3,3), \mathrm{linear})$ is the final 3D convolution layer with one filter and a linear activation function and $\widehat{\mathbf{Y}} \in R^{T\times H\times W\times 1}$ represents the final predicted TEC maps.

3.4 3D-U-Net Architecture

We implemented a 3D-U-Net as a spatio-temporal encoder–decoder that operates on short time sequences of global TEC maps. The network processes an input tensor of shape $\mathbf{X}_{\mathrm{in}} \in \mathbb{R}^{S\times H\times W\times C}$ where S is the temporal depth (sequence length, in our experiments S=12), H=71and W=73are the spatial dimensions, and C=1is the number of channels (TEC). The network produces a prediction tensor of the same temporal length and spatial resolution ($\widehat{\mathbf{Y}} \in \mathbb{R}^{S\times H\times W\times 1}$). The network design follows the canonical U-Net pattern but uses 3D convolutions so that temporal and spatial features are learned jointly. Key design choices and motivation:

Encoder blocks: Each encoder level contains two Conv3D layers with kernel size 3×3×3(time × height × width), each followed by Batch Normalization and ReLU. The number of filters increases with depth to capture progressively higher-level spatio-temporal features.

Spatial downsampling only: Downsampling is performed with MaxPooling3D (pool_size = (1,2,2)). This preserves the temporal resolution (the sequence length Sremains unchanged) while reducing spatial resolution (height and width halved). Preserving the time axis avoids temporal aliasing and makes it straightforward to output a prediction per input time step.

Bottleneck: The deepest layer (bottleneck) uses the highest filter count and the same two-Conv3D block structure, providing a compact spatio-temporal representation.

Decoder blocks: Each decoder level upsamples with UpSampling3D (size = (1,2,2)) (again preserving the time dimension), concatenates the corresponding encoder feature map (skip connection), and applies two Conv3D + BatchNorm + ReLU layers to refine features and restore spatial detail.

Output layer: A final Conv3D with kernel 1×3×3 and linear activation produces the predicted TEC maps per time step. Output shape matches the input spatial grid and temporal depth.

3.5 Train and Evaluate

To enhance training stability, all spatial and temporal data were normalized using Equations (20) and (21) prior to model training.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-4-W8-2025-549-2026 | © Author(s) 2026. CC BY 4.0 License.

$$X_{\mathrm{norm}} = \frac{X - X_{\min}}{X_{max} - X_{\min}} \tag{20}$$

$$Y_{\mathrm{norm}} = \frac{Y - Y_{min}}{Y_{max} - Y_{min}} \tag{21}$$

in Equations (20) and (21), X and Y denote the spatial and temporal input data before normalization; $X_{\min}$ and X_{max} are the minimum and maximum values within X. Y_{min} and Y_{max} are the minimum and maximum values within Y; and X_{norm} and Y_{norm} are the resulting normalized datasets.

Subsequently, the model was trained using the Mean Squared Error loss function as defined in Equation (22).

$$\mathcal{L}_{MSE} = \frac{1}{N}\sum_{i=1}^{N} |\mathbf{Y_i} - \widehat{\mathbf{Y}}_i|^2 \tag{22}$$

In Equation (22), N is the total number of samples, $\mathbf{Y_i}$ and $\widehat{\mathbf{Y}}_i$ is the predicted value for sample i, and $\mathcal{L}_{MSE}$ denotes the mean squared error loss.

Finally, to evaluate the model's performance, the mean RMSE, MAE, pearson's sample correlation and skill related to persistence metrics defined in Equations (23), (24), (25) and (26) were computed.

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N} |\mathbf{Y_i} - \widehat{\mathbf{Y}}_i| \tag{23}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \left(\mathrm{Y_i} - \widehat{\mathrm{Y}}_i\right)^2} \tag{24}$$

$$r = \frac{\sum_{i=1}^{N}(\hat{y}_i - \bar{\hat{y}})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(\hat{y}_i - \bar{\hat{y}})^2}\sqrt{\sum_{i=1}^{N}(y_i - \bar{y})^2}} \tag{25}$$

$$kill = 1 - \frac{MSE_{model}}{MSE_{persist}} \tag{26}$$

In Equations (23), (24), (25) and (26) N is the total number of samples; $\mathbf{Y_i}$ is real velue and $\widehat{\mathbf{Y}}_i$ is the predicted value for sample i, $\bar{y}$ and $\bar{\hat{y}}$ is sample means and $\mathrm{MSE}_{\mathrm{model}}$ be the model MSE and $\mathrm{MSE}_{\mathrm{persist}}$ the MSE of a persistence baseline. MAE denotes the mean absolute error; RMSE denotes the root mean squared error; r Compute Pearson's sample correlation coefficient between flattened prediction target arrays; and skill is skill related to persistence.

4. Results and Discussion

In this study, the datasets from 2014 to 2022 are used. The second half of the 2022 data is considered as validation data and the data from 2014 (high solar activity) and 2018 (low solar activity) are considered as testing data. All data were used with a temporal resolution of two hours. In addition, temporal indices such as F10.7, DST, SN, AP and KP were extracted and included as supplementary features in the model.

Main hyperparameters used in the experiments shown in table 1.

Parameter	Description	Value / Setting
Batch size	Number of samples per update	2
Epochs	Max training epochs (with early stopping)	20
Pooling type	Spatial downsampling	MaxPooling3D (1, 2, 2)
Upsampling	Spatial upsampling	UpSampling3D (1, 2, 2)
Encoder filters	Filters per encoder block	16 → 32 → 64
Decoder filters	Filters per decoder block	64 → 32 → 16
Dropout	Regularization in bottleneck (optional)	0.2
Batch Normalization	After each Conv layer	Yes

Table 1. hyperparameters used in the experiments

All models were trained under the same configuration using Adam optimizer (learning rate 1×10⁻⁴) and mean squared error loss. Batch normalization was applied after each convolution layer, and early stopping (patience = 5) prevented overfitting.

Table 1 shows the average RMSE and MAE values obtained from all models.

Model	MAE_{2014}	$RMSE_{2014}$	$RMSE_{2010}$	MAE_{2010}
CONVLSTM	6.16	4.12	1.84	1.20
Patch-CONVLSTM (15×15)	5.83	3.88	1.80	1.22
Patch-CONVLSTM (5×5)	5.81	3.90	1.83	1.25
3D-U-Net	5.05	3.36	1.73	1.10

Table 2. average RMSE and MAE

Model	r_{2014}	r_{2018}	$Skill_{2014}$	$Skill_{2018}$
CONVLSTM	0.96	0.90	0.96	0.93
Patch-CONVLSTM (15×15)	0.96	0.91	0.96	0.93
Patch-CONVLSTM (5×5)	0.96	0.91	0.96	0.93
3D-U-Net	0.97	0.93	0.97	0.94

https://doi.org/10.5194/isprs-annals-X-4-W8-2025-549-2026 | © Author(s) 2026. CC BY 4.0 License.

Table 3. Average Pearson Correlation (r) and Skill relative to Persistence.

According to the results presented in Table 2 and Table 3, overall, the 3D-UNet model performed best in both test years (2018 and 2014). The RMSE and MAE values in this model show the lowest values and the correlation coefficient (Pearson) and the highest Skill values. Trainable patching on CONVLSTM (patch-sized CONVLSTM) provides limited improvements, but this improvement is small and depends on the year/conditions. The results showed that for 2018 there is only a RMSE reduction of about 2% (patch=5), but in 2014 there were moderate improvements of ≈5–6% for patches.

The average monthly RMSE values over the 12-month test years period for the all of models are presented in Figure 1.

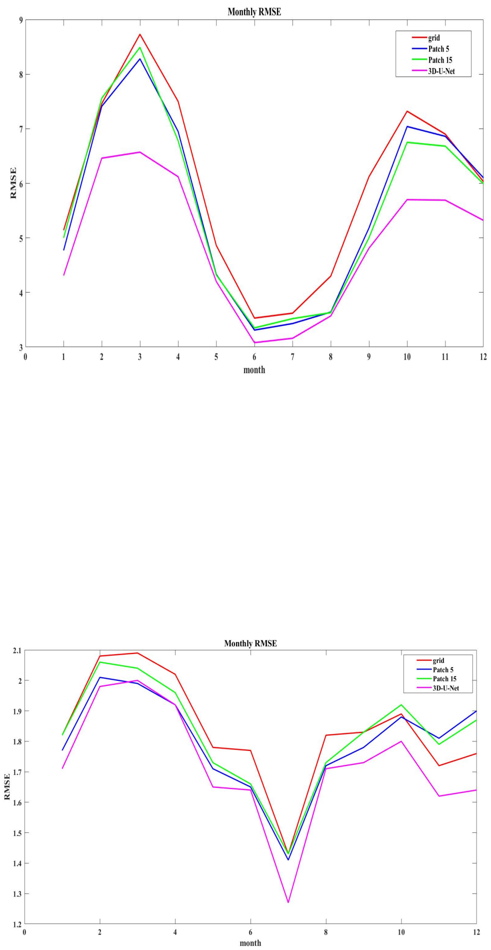

Figure 1. Monthly RMSE of CONVLSTM, Patch-CONVLSTM and 3D-U-Net models.

Over the course of a full year, our model errors behaved like the changing seasons of nature: each month was a snapshot of the interaction of sunlight, magnetic winds, and ion distributions that the models had to account for. If we take a quick look at the monthly table, a consistent but fluctuating pattern emerges—errors are higher in months aligned with peak solar activity and equinoxes, and decrease in the summer months, or the less active half of the relative solar cycle. A numerical comparison between the models shows that 3D-U-Net recorded lower RMSE and MAE in most months, telling us that this architecture has a more stable ability to represent spatiotemporal structures and better track seasonal fluctuations. Patchy models (especially patch5) are competitive in some months and show improvement in months where strong local spatial patterns develop (e.g. mid-summer/winter months depending on the year).

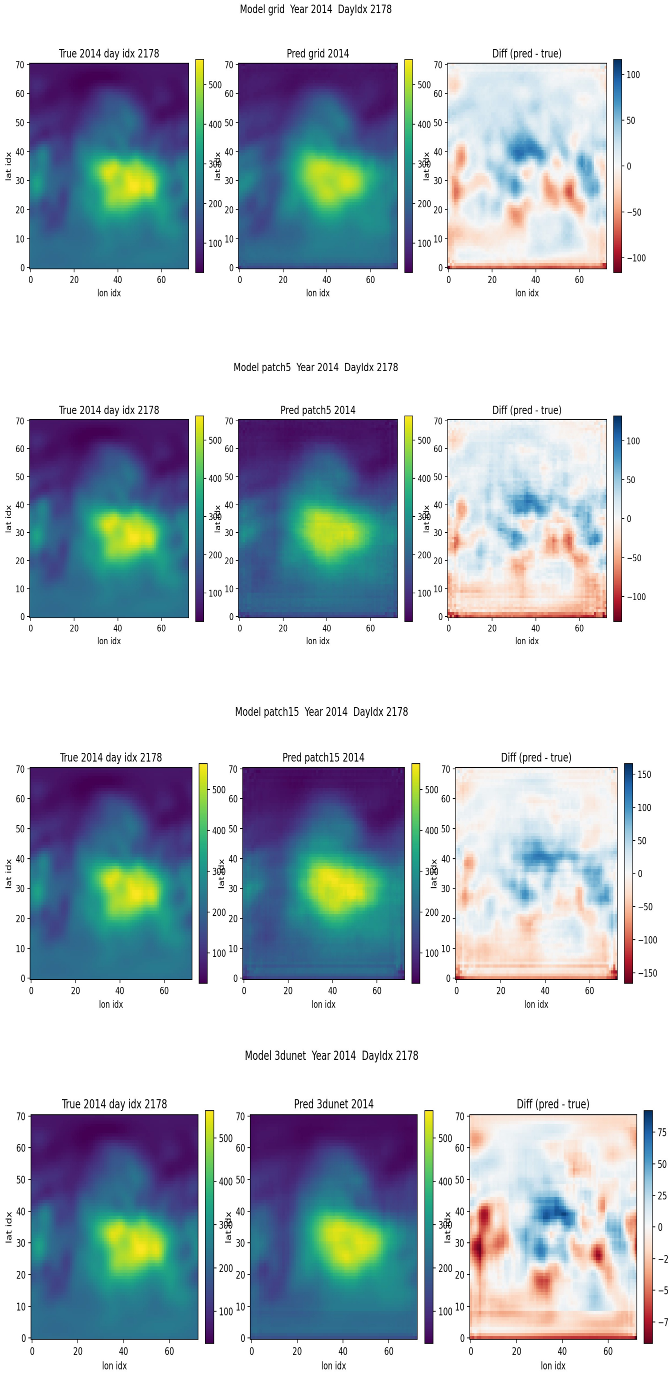

Figure 2. Predicted TEC map, real TEC map and differences between real and predicted Maps (year=20014)

https://doi.org/10.5194/isprs-annals-X-4-W8-2025-549-2026 |

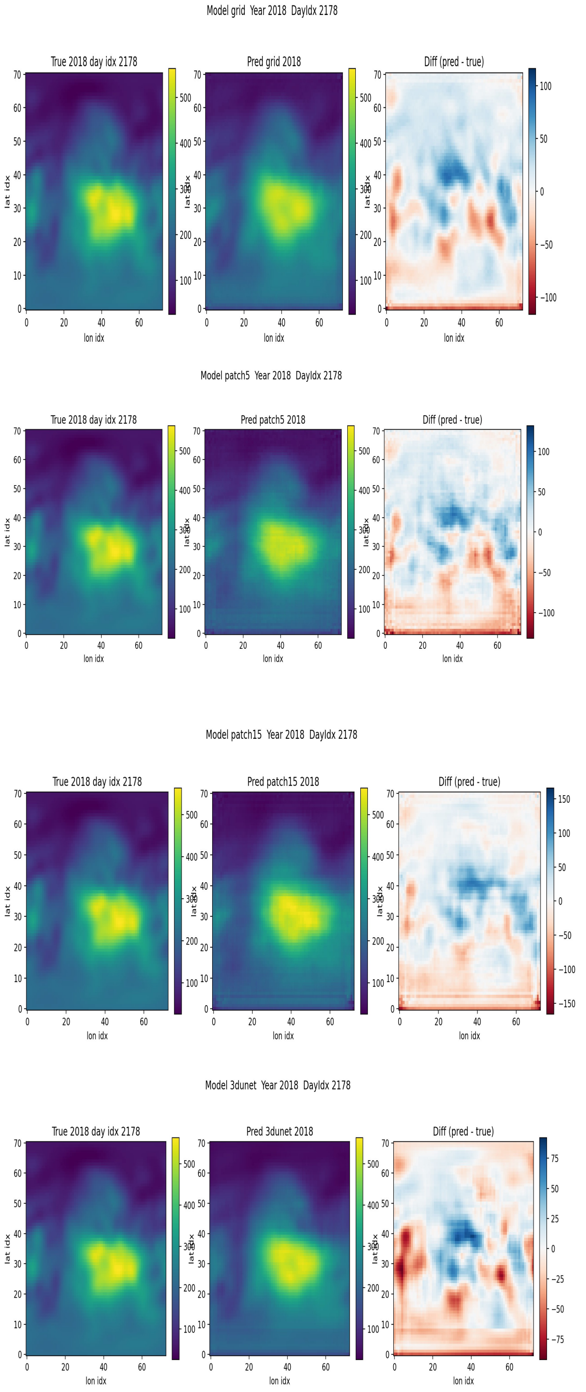

Figure 3. Predicted TEC map, real TEC map and differences between real and predicted Maps (year=2018)

The differences between real and predicted maps are shown in Figures 3 and 4, where, as expected, the errors are clearly higher in the mid-latitude regions. In these regions, local ion accumulation, magnetic excitations, and small-scale atmospheric changes cause rapid and local fluctuations. The results presented show that architectures such as 3D-U-Net, which are capable of capturing large spatiotemporal structures, often perform more stable; and local patching (e.g., patch=5) can provide improvements at time scales and latitudes where strong and local spatial features dominate.

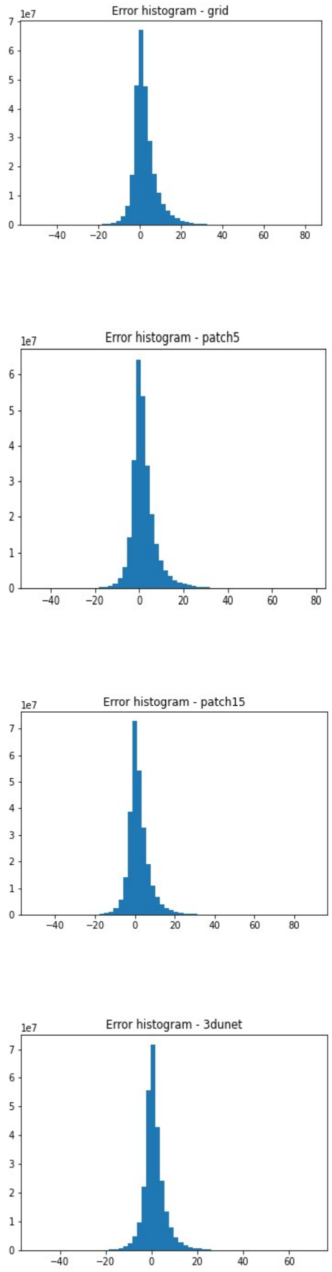

Figure 4. Histogram of model errors (year=2014)

https://doi.org/10.5194/isprs-annals-X-4-W8-2025-549-2026 | © Author(s) 2026. CC BY 4.0 License.

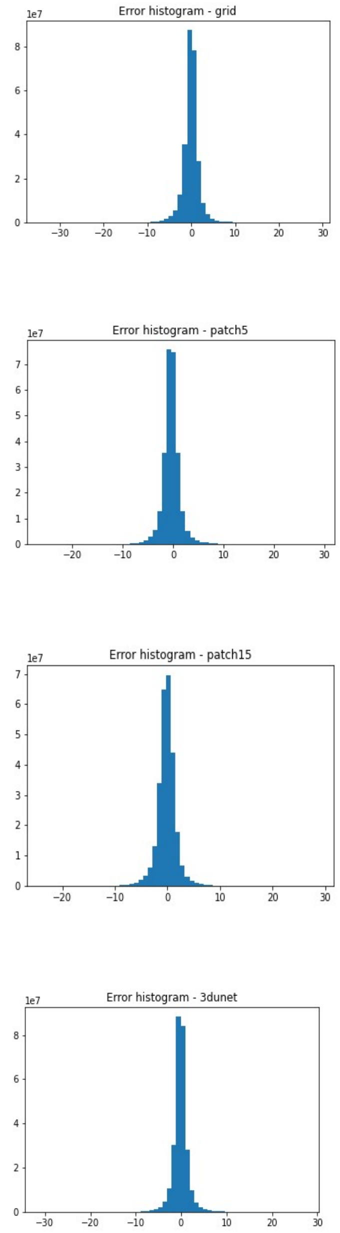

Figure 5. Histogram of mode errors (year=2018)

The error histogram (Figure 4,5) acts as a condensed picture of the error behavior: the width of the distribution indicates the variance of the error, and the position of the center of the distribution indicates the bias of the model. Here, we see that the 3D-UNet model has a narrower distribution and is more concentrated around zero—that is, it produces smaller and more stable errors—while the continuous (unpatched) CONVLSTM has longer tails, indicating outliers. The patched mode (especially patch=5) shows a distribution with shorter tails than the unpatched version, indicating a reduction in local extreme errors; but sometimes the center of the distribution is slightly shifted, which may be due to local downscaling or the loss of some pervasive features. In other words, 3D-UNet spreads the errors more evenly (less affected by strong local fluctuations) and patching helps reduce the number of very inaccurate forecasts, but may come at the cost of spatial detail at some latitudes.

5. Conclusion

In this study, a simple CONVLSTM, CONVLSTM based patch (patch-size=5 and 15) and simple 3D-U-Net architecture were employed to predict the next twelve samples of Total Electron Content (TEC) using the previous twelve samples of TEC data along with the geophysical indices F10.7 (solar radio flux), Ap (daily average geomagnetic index), DST (geomagnetic storm index), SN (sunspot number), and KP (geomagnetic disturbance index).

The results obtained from applying the models to the selected dataset demonstrate that the 3D-U-Net architecture not only improves the execution speed but also significantly enhances prediction accuracy by reducing RMSE and MAE values.

References

Afraimovich, E. L., Astafyeva, E. I., Demyanov, V. V., Edemskiy, I. K., Gavrilyuk, N. S., Ishin, A. B., Kosogorov, E. A., Leonovich, L. A., Lesyuta, O. S., Palamartchouk, K. S, 2013: A review of GPS/GLONASS studies of the ionospheric response to natural and anthropogenic processes and phenomena. *Journal of Space Weather and Space Climate*, *3*, A27.

Afraimovich, E. L., Kosogorov, E. A., Palamarchouk, K. S., Perevalova, N. P., Plotnikov, A. V, 2000: The use of GPS arrays in detecting the ionospheric response during rocket launchings. *Earth, planets and space*, *52*, 1061-1066.

Artru, J., Ducic, V., Kanamori, H., Lognonné, P., Murakami, M, 2005: Ionospheric detection of gravity waves induced by tsunamis. *Geophysical Journal International*, *160*(3), 840-848.

Benoit, A. G. M. d. S., Petry, A, 2021: Evaluation of F10. 7, sunspot number and photon flux data for ionosphere TEC modeling and prediction using machine learning techniques. *Atmosphere*, *12*(9), 1202.

Calais, E., Bernard Minster, J, 1996: GPS detection of ionospheric perturbations following a Space Shuttle ascent. *Geophysical Research Letters*, *23*(15), 1897-1900.

Calais, E., Bernard Minster, J., Hofton, M., Hedlin, M, 1998: Ionospheric signature of surface mine blasts from Global Positioning System measurements. *Geophysical Journal International*, *132*(1), 191-202.

Calais, E., Minster, J. B, 1995: GPS detection of ionospheric perturbations following the January 17, 1994, Northridge earthquake. *Geophysical Research Letters*, *22*(9), 1045-1048.

Ciraolo, L., Azpilicueta, F., Brunini, C., Meza, A., Radicella, S. M, 2007: Calibration errors on experimental slant total electron content (TEC) determined with GPS. *Journal of Geodesy*, *81*, 111-120.

Dickinson, H., Tamarkin, P, 1965: Systems for the Detection and Identification of Nuclear Explosions in the Atmosphere and in Space. *Proceedings of the IEEE*, *53*(12), 1921-1934.

Feng, J., Jiang, W., Wang, Z., Zhao, Z., Nie, L, 2017: Regional TEC model under quiet geomagnetic conditions and low-to-moderate solar activity based on CODE GIMs. *Journal of Atmospheric and Solar-Terrestrial Physics*, *161*, 88-97.

Feng, J., Wang, Z., Jiang, W., Zhao, Z., Zhang, B, 2016: A new regional total electron content empirical model in northeast China. *Advances in Space Research*, *58*(7), 1155-1167.

Feng, J., Wang, Z., Jiang, W., Zhao, Z., Zhang, B, 2017: A single-station empirical model for TEC over the Antarctic Peninsula using GPS-TEC data. *Radio Science*, *52*(2), 196-214.

Fitzgerald, T. J, 1997: Observations of total electron content perturbations on GPS signals caused by a ground level explosion. *Journal of Atmospheric and Solar-Terrestrial Physics*, *59*(7), 829-834.

Fubin, Z., Chen, Z., Cheng, W., Jiaqi, Z., Yi, L., Guozhen, X., Zhengyu, Z, 2021: Global ionospheric TEC prediction based on deep learning. *Chinese Journal of Radio Science*, *36*(4), 553-561.

Hoque, M. M., Jakowski, N, 2012: A new global model for the ionospheric F2 peak height for radio wave propagation. Annales Geophysicae,

Jakowski, N., Hoque, M., Mayer, C, 2011: A new global TEC model for estimating transionospheric radio wave propagation errors. *Journal of Geodesy*, *85*, 965-974.

Jalali, Y., Fateh, M., Rezvani, M, 2024: DABT-U-Net: Dual attentive BConvL-STM U-Net with transformers and collaborative patch-based approach for accurate retinal vessel segmentation. *Int. J. Eng*.

Jeong, S. H., Lee, W. K., Kil, H., Jang, S., Kim, J. H., Kwak, Y. S, 2024: Deep Learning-Based Regional Ionospheric Total Electron Content Prediction—Long Short-Term Memory (LSTM) and Convolutional LSTM Approach. *Space Weather*, *22*(1), e2023SW003763.

Kaselimi, M., Voulodimos, A., Doulamis, N., Doulamis, A., Delikaraoglou, D, 2021: Deep recurrent neural networks for ionospheric variations estimation using GNSS measurements. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1-15.

Kazimirovsky, E., Herraiz, M., De la Morena, B, 2003: Effects on the ionosphere due to phenomena occurring below it. *Surveys in Geophysics*, *24*, 139-184.

Kim, M., Seo, J., Lee, J, 2014: A comprehensive method for GNSS data quality determination to improve ionospheric data analysis. *Sensors*, *14*(8), 14971-14993.

Lin, Z., Li, M., Zheng, Z., Cheng, Y., Yuan, C, 2020: Self-attention convlstm for spatiotemporal prediction. Proceedings of the AAAI conference on artificial intelligence,

Liu, J. Y., Chuo, Y., Shan, S., Tsai, Y., Chen, Y., Pulinets, S., Yu, S, 2004: Pre-earthquake ionospheric anomalies registered by continuous GPS TEC measurements. Annales geophysicae,

Muhammad, A., Külahcı, F, 2022: A semi-supervised total electron content anomaly detection method using LSTM-auto-encoder. *Journal of Atmospheric and Solar-Terrestrial Physics*, *241*, 105979.

Mukhtarov, P., Pancheva, D., Andonov, B., Pashova, L, 2013: Global TEC maps based on GNSS data: 1. Empirical background TEC model. *Journal of Geophysical Research: Space Physics*, *118*(7), 4594-4608.

Nie, Y, 2022: A Time Series is Worth 64Words: Long-term Forecasting with Transformers. *arXiv preprint arXiv:2211.14730*.

Sezen, U., Arikan, F., Arikan, O., Ugurlu, O., Sadeghimorad, A, 2013: Online, automatic, near-real time estimation of GPS-TEC: IONOLAB-TEC. *Space Weather*, *11*(5), 297-305.

Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., Woo, W.-c, 2015: Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*, *28*.

Sun, W., Xu, L., Huang, X., Zhang, W., Yuan, T., Chen, Z., Yan, Y, 2017: Forecasting of ionospheric vertical total electron content (TEC) using LSTM networks. 2017 international conference on machine learning and cybernetics (ICMLC),

Tsugawa, T., Saito, A., Otsuka, Y., Nishioka, M., Maruyama, T., Kato, H., Nagatsuma, T., Murata, K, 2011: Ionospheric disturbances detected by GPS total electron content observation after the 2011 off the Pacific coast of Tohoku Earthquake. *Earth, planets and space*, *63*, 875-879.

Wang, Y., Jiang, H., Liu, T., Yao, L., Zhou, C, 2025: A Patch-wise Mechanism for Enhancing Sparse Radar Echo Extrapolation in Precipitation Nowcasting. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.

Wang, Z., Zou, S., Sun, H., Chen, Y, 2023: Forecast global ionospheric TEC: Apply modified U-net on VISTA TEC data set. *Space Weather*, *21*(8), e2023SW003494.

Xia, G., Zhang, F., Wang, C., Zhou, C, 2022: ED-ConvLSTM: A novel global ionospheric total electron content medium-term forecast model. *Space Weather*, *20*(8), e2021SW002959.

Yang, H., Yang, X., Zhang, Z., Zhao, K, 2018: High-precision ionosphere monitoring using continuous measurements from BDS GEO satellites. *Sensors*, *18*(3), 714.

Yeh, K. C., CH, L, 1972: Theory of ionospheric waves.

Zhang, J., Guo, L., Song, L., Gao, S., Hao, C., Li, X, 2024: PatchTCN: Patch-Based Transformer Convolutional Network for Times Series Analysis. Proceedings of the 2024 3rd International Symposium on Computing and Artificial Intelligence,

Zhao, Z., Feng, J., Han, B., Wang, Z, 2018: A single-station empirical TEC model based on long-time recorded GPS data for estimating ionospheric delay. *Journal of Space Weather and Space Climate*, *8*, A59.