Liu Lei (Orcid ID: 0000-0002-1040-6026)
Zou Shasha (Orcid ID: 0000-0001-7726-2349)
Wang Zihan (Orcid ID: 0000-0001-6601-2520)

Forecasting global ionospheric total electron content (TEC) using deep learning approach

Lei Liu[1, 2, 3], Shasha Zou[2], Yibin Yao[1], Zihan Wang[2]
[1]School of Geodesy and Geomatics, Wuhan University, Wuhan, Hubei, China.
[2]Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI, USA.
[3]College of Engineering's Space Weather Technology, Research & Education Center, University of Colorado Boulder, Boulder, CO, USA.
Corresponding author: Lei Liu (leiliu@whu.edu.cn)

Abstract Global ionospheric total electron content (TEC) maps are widely utilized in research regarding ionospheric physics and the associated space weather impacts, so there is a great interest in the community in short-term ionosphere TEC forecasting. In this study, the long short-term memory (LSTM) neural network (NN) is applied to forecast the 256 spherical harmonic (SH) coefficients that are traditionally used to construct global ionospheric maps (GIM). Multiple input data, including historical time series of the SH coefficients, solar extreme ultraviolet (EUV) flux, disturbance storm time (Dst) index, and hour of the day, are used in the developed LSTM NN model. Different combinations of the above parameters have been used in constructing the LSTM NN model, and it is found that the model using all four parameters performs the best. Then the best performing LSTM model is used to forecast the SH coefficients, and the global hourly TEC maps are reproduced using the 256 predicted SH coefficients. A comprehensive evaluation is carried out with respect to the CODE GIM TEC. Results show that the 1st/2nd hour TEC root mean square error (RMSE) is 1.27/2.20 TECU during storm time, and 0.86/1.51 TECU during quiet time, so the developed model performs well during both quiet and storm times. Moreover, typical ionospheric structures, such as equatorial ionization anomaly (EIA) and storm-enhanced density (SED), are well reproduced in the predicted TEC maps during storm time. The developed model also shows competitive performance in predicting global TEC when compared to the persistence model and two empirical models (IRI-2016 and NeQuick-2).

Key points:

- The LSTM neural network is adopted to predict the global ionosphere TEC.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2020SW002501

- The use of the external solar EUV flux and Dst index is able to improve the prediction performance of the spherical harmonic (SH) coefficients.
- The developed LSTM model performs well during both quiet and storm conditions.

1 Introduction

The ionosphere is the ionized part of the earth's upper atmosphere where sufficient plasma exists to affect the propagation of radio waves. It is arguably the most important region in terms of space weather impacts of modern technologic society, because irregularities in the ionosphere frequently degrade and disrupt satellite navigation and communication. The total electron content (TEC) refers to the total number of electrons integrated along a tube of one square meter cross section (unit: TECU, $1\ TECU = 10^{16} electrons/m^2$). TEC is a significant ionospheric parameter, which is widely applied to the satellite ionospheric delay correction and scientific research area about space weather impacts on the ionosphere and thermosphere (Liu et al., 2020). As a consequence, there is a great interest in the community in the TEC forecasting.

The ionospheric TEC along the ray path between the satellite and receiver can be derived by using dual frequency global navigation satellite system (GNSS) measurements, because of the dispersive characteristics of the ionosphere. For instance, the International GNSS Service (IGS) ionospheric working group, consist of seven Ionospheric Associate Analysis Centers (IAACs), has started generating reliable global TEC maps based on worldwide GNSS data since 1998, which are widely used in the ionosphere monitoring (Schaer, 1999; Hernández-Pajares at al., 2009; Li at al., 2015; Yao at al., 2018).

In terms of the short-term GNSS-based TEC forecasting, many approaches have been developed in recent decades, such as least-square collocation, discrete cosine transform (DCT), linear regression, adaptive autoregressive model, and neural networks, etc. Based on the least-square collocation method, Schaer (1999) predicted the global TEC parameters in the next two days by extrapolating the spherical harmonic (SH) coefficients used to fit the TEC map in the previous 30 days. Tulunay et al. (2006) presented a neural network technique to forecast the TEC maps over Europe, which is trained based on the Levenberg-Marquardt algorithm, and results show that the developed model learned the shape of the inherent nonlinearities during space weather events. García-Rigo et al. (2011) developed a global TEC prediction model by applying the Discrete Cosine Transform (DCT) to the TEC maps, and then used a linear regression module to forecast the time evolution of the DCT coefficients,

but physical information, including the solar and geomagnetic activity, was not included in the prediction model. Habarulema et al. (2011) developed an ionospheric TEC prediction model in South Africa using artificial neural network and evaluated its capability by using independent ionosonde data and observed TEC values. They found that the developed model performed well during quiet periods while the accuracy decreased under storm conditions. Wang et al. (2018) proposed an adaptive autoregressive model to predict the SH coefficients used in TEC map fitting, and then the global TEC maps can be generated based on the predicted SH coefficients. Results show that the predicted TEC has better performance during the low solar activity periods than the mid-to-high solar activity periods.

Nowadays, deep learning techniques, such as the convolutional neural network (CNN) and recurrent neural network (RNN), are very promising and have been increasingly applied to the short-term ionosphere forecasting. For instance, Sun et al. (2017) developed a single-station TEC forecasting model at Beijing station in China based on the long short-term memory (LSTM) neural network using five-day's sequence of TEC, AP and F10.7 as inputs. Results showed that their model outperformed the conventional multilayer perceptron network, though the prediction performance under disturbed conditions still needs further improvement. Similarly, Srivan et al. (2019) also developed a LSTM forecasting model over the Bengaluru station in India using similar inputs as Sun et al. (2017), but for different network configuration parameters (i.e., the input time span is 24 hours). They found the model is able to capture the TEC variations even during the active solar and geomagnetic activities. Boulch et al. (2018) presented three convolutional RNN architectures for global ionospheric TEC prediction, and their results showed that the prediction performance was comparable with state-of-the-art methods, such as the auto regressive (AR), auto regressive moving average (ARMA) model, and the radial basis function (RBF) neural network.

However, most of the above-mentioned deep learning approaches for the TEC forecast are either for a single-station forecasting model or have not considered the physical parameters representing the solar and geomagnetic disturbances. In this paper, we aim at developing a novel deep learning model to forecast the SH coefficients used in constructing the global TEC map by using time series of the SH coefficients and multiple physical parameters representing solar and geomagnetic activity levels. The model performance is then evaluated by comparing the prediction with independent TEC data under both geomagnetic quiet and storm conditions.

2 Data and methodology

2.1 Global ionosphere maps reconstruction

Based on deferential delay measurements from globally distributed GNSS tracking sites, the Centre for Orbit Determination in Europe (CODE) develops hourly global ionosphere maps (GIM) in the solar-geomagnetic reference frame using spherical harmonic (SH) expansion (Schaer 1999),

$$TEC(\beta,s)=\sum_{n=0}^{n_{\max}}\sum_{m=0}^{n}\tilde{P}_{nm}(\sin\beta)\cdot(\tilde{A}_{nm}\cos(ms)+\tilde{B}_{nm}\sin(ms)) \quad (1)$$

where β and s are the geomagnetic latitude and sun-fixed longitude of the ionospheric pierce point (IPP), respectively. The IPP is the intersection point between the satellite-receiver ray path and the single-layer height. The sun-fixed longitude is defined as $s=15(t-12)+\lambda$, where λ is the geographic longitude of the IPP, and t is the universal time. $TEC(\beta,s)$ is the vertical TEC (VTEC, hereafter referred as TEC for simplification) at the location of the IPP. $\tilde{P}_{nm}$ denotes the normalized associated Legendre function with degree n and order m. $\tilde{A}_{nm}$ and $\tilde{B}_{nm}$ are the hourly SH coefficients to be determined, which are then used for constructing the global TEC map. $n_{\max}$ is the maximum degree of the SH expansion ($n_{\max}=15$), which means there are 256 ($(n_{\max}+1)^2=256$) SH coefficients for one-hour interval.

In general, we are not interested in the individual SH coefficients but rather in the TEC value, which is represented as a function of geographic coordinates (β,λ) and time t, namely $TEC(\beta,\lambda,t)$. For global TEC map reconstruction, once obtaining the 256 SH coefficients ($\tilde{A}_{nm}$ and $\tilde{B}_{nm}$) at time t, we are able to generate GIM using formula (1) with a spatial resolution of 2.5° in latitude and 5° in longitude. These resolutions are chosen to be consistent with the TEC maps developed at CODE for model validation and performance evaluation purpose. These 256 SH coefficients are estimated routinely with one-hour interval and are available from the CODE analysis center (http://ftp.aiub.unibe.ch/CODE/) on a daily basis since Oct, 2014. We predict the time series of these SH coefficients in the next two hours using the CODE SH coefficients in the past and then construct the GIM.

2.2 SH coefficients forecast using the LSTM neural network

One challenge of this study is to design an appropriate neural network (NN) architecture to extract temporal trend and features from previous states, and to predict the SH coefficients from historical sequences and related external drivers, such as solar and geomagnetic activities.

2.2.1 Data preparation

In this study, the most important dataset are the time series of the SH coefficients, which is freely available from the CODE analysis center. Moreover, considering that the solar and geomagnetic activities play significant roles in changing the ionospheric TEC, realistic solar irradiance and geomagnetic activity index are also needed. Here, the solar extreme ultraviolet (EUV) flux from Flare Irradiance Spectral Model (FISM), and the Disturbance Storm Time (Dst) index representing the ring current strength and storm phases are used in this study. One of the reasons for choosing Dst instead of other geomagnetic indices, such as the 3-hour Kp index, is because the time resolution of Dst index is one hour, which is consistent with the SH coefficients provided by the CODE center. All of the above datasets are collected from October 19, 2014 to December 31, 2016 (see Figure S1). It should be noted that FISM estimates the solar irradiance at wavelengths from 0.1 to 190 nm at 1 nm resolution every minute, meaning that there are 190 wave bands for the solar EUV flux. Figure 1 (a) shows the Pearson correlation between the 190 FISM solar EUV fluxes. We recognize that these 190 features are highly correlated with each other, and not all of these parameters are needed as the input data of the network, because too many redundant features might result in over-fitting problems in the machine learning algorithm. The principal component analysis (PCA) is a very powerful procedure to transform many correlated parameters into a smaller number of uncorrelated parameters via dimension reduction, so here we extracted several principal components from the 190 FISM EUV fluxes using the PCA method. Figure 1(b) shows the first three principal components from the 190 original features. We found that the first principal component could account for more than 99.6% of the variability of the 190 features, so in this study only the first principal component extracted from 190 FISM features is used to represent the solar EUV flux.

In addition, it is well known that ionosphere has a diurnal variation, so we introduce time (hour of day, HD) as an input to enable the network learn the diurnal variations of the SH coefficient sequences. To allow continuity across the midnight boundary, the HD is defined as two quadrature components (Poole and McKinnell, 2000),

$$\left.\begin{aligned} HD_s &= \sin(\frac{2\pi \times HD}{24}) \\ HD_c &= \cos(\frac{2\pi \times HD}{24}) \end{aligned}\right\}, 0 \le HD \le 23 \qquad (2)$$

where HD_s and HD_c are the sine and cosine components of the HD, respectively.

2.2.2 Neural network architecture

The LSTM neural network (NN), a special type of the recurrent neural network (RNN), is capable of learning time dependence in sequential prediction through the forget gate weights, and is developed to avoid gradient exploding and vanishing problems existing in conventional RNN (Hochreiter and Schmidhuber, 1997; Gers et al., 1999; Chen et al., 2019). In this study, we try to forecast the SH coefficients separately two hours ahead by using LSTM NN first, and then use the predicted SH coefficients to construct the GIM by using the SH expansion function.

The overall architecture for a single SH coefficient prediction is shown Figure 2, which includes the input data in the input layer, two LSTM layers and one dense layer in the hidden layer, as well as the predicted SH coefficients in the output layer.

In the input layer, the input time span is set to be 24 hours, as the autocorrelation for each SH coefficient at 24 hours is the highest due to the ionospheric diurnal variation. The final input candidates after the data preparation procedure are the SH coefficients, HDc, HDs, the first component of the FISM solar irradiance and the Dst index of the previous 24 hours.

In the hidden layer, two stacked LSTM layers are used to extract representative features from historical input data, and then are connected to the fully connected (FC) hidden layer. The neuron node number in each hidden layer is chosen to be 48 (n=48) throughout the paper, which is determined by training the NN from node 1 to 60 (see details in in sector 3.1 and Figure 3).

In the output layer, the output parameters are the individual SH coefficients for the next two hours, which could be extended further in the future.

The entire dataset is split into two groups, namely, group 1 (Training set: Jan 1, 2015 to May 26, 2016) and group 2 (Testing set: Oct 19-Dec 31, 2014, and May 27-Dec 31, 2016) (see Figure S1 in the supporting information). To find the optimum NN parameters, 5-fold cross-validation strategy is utilized in the Training set of group 1. More specifically, group 1 is divided into five folds, where four folds are used for training the NN, and the remaining one fold is used for validation. This process is repeated for five times and the final performance is the averaged results of these five

times. Finally, the NN parameters showing best performance for both the four fold training and one fold validation will be used as the final NN parameters. Once the optimum NN parameters are determined by the 5-fold cross-validation, we test their performance in TEC prediction using group 2 (Testing set), which is not involved in training the NN model.

2.2.3 Evaluation metrics for the deep learning-based model

One of the core tasks in building deep learning model is to evaluate its effectiveness and reliability. In this paper, several important evaluation metrics, such as Mean Absolute Error (MAE), Mean Square Error (MSE) and Root Mean Square Error (RMSE), as well as Mean Error (ME), are used to qualify how well the proposed deep learning algorithm in predicting the SH coefficients and the resulting TEC results.

$$\left.\begin{aligned} MAE &= \frac{1}{N}\sum_{j=1}^{N}|\, y_j - \hat{y}_j \,| \\ MSE &= \frac{1}{N}\sum_{j=1}^{N}(y_j - \hat{y}_j)^2 \\ RMSE &= \sqrt{MSE} \\ ME &= \frac{1}{N}\sum_{j=1}^{N}(y_j - \hat{y}_j) \end{aligned}\right\} \qquad (3)$$

where y_j and $\hat{y}_j$ are observations and predictions, respectively. $j = 1,2,3,...,N$, and N is the total number of data samples. Considering the various ranges of the input and output features, each feature in this study is normalized to [0,1] before training the NN model, which is a commonly used method in machine learning. The outputs or predictions can be recovered to the actual dynamic range using an inverse normalization process. In our study, if the skill scores are calculated based on the normalized observations or predictions, we will call them as normalized skill scores.

2.3 Construction of global ionosphere maps using the predicted SH coefficients

As mentioned in sector 2.1, global ionospheric TEC maps can be represented as a function of geomagnetic coordinates and time when the 256 SH coefficients are plugged into formula (1). Thus, the hourly GIM can be produced when the 256 SH coefficients are predicted successfully from the proposed NN model in sector 2.2. The spatial resolution of the resulting GIM is 2.5° in latitude and 5° in longitude, which is consistent with the global TEC products released by the CODE ionosphere analysis center.

3 Results and evaluation

3.1 Determination of the NN parameters

Neural nodes and inputs are two types of key parameters in the NN architecture, so in this part we will try to determine them appropriately.

Here, we take the 1st SH coefficient prediction results as an example to identify the neural nodes, and similar procedures are performed for the rest of the 255 SH coefficients. We train the proposed NN by changing the node number from 1 to 60 in the step of 1, and evaluate the Training/Testing performance in terms of MAE and MSE. Figure 3 (a-d) shows the MAE/RMSE error distributions for various neural node numbers on Train/Test data sets, and Figure 3 (e) shows the learning curve variations with different training epochs. It can be seen that the SH prediction achieved the smallest error when the neural node number is 48, except in Figure 3(d) where that optimal number is 49. Moreover, the learning curves in Figure 3(e) show overall very good fitting, because the normalized MAE errors of both Train and Test sets decrease gradually with the training epochs, and are overlapped with each other. Therefore, the ideal neural node number in hidden layer is set to 48 throughout this paper.

Another important task is to determine the most appropriate input parameters for the input layer. As mentioned in sector 2.2, the candidates in the input layer are the SH coefficient, HDc, HDs, the 1st principal component of the FISM solar irradiance, and the Dst index. In this study, we try different input combinations and evaluate the corresponding prediction performance in terms of the normalized MAE/RMSE by using 5-fold cross-validation strategy. Table 1 presents the normalized MAE/RMSE for different input parameter combinations, including I (SH coefficient + 1st component of FISM solar irradiance + Dst + HDc + HDs), II (SH coefficient +1st component of FISM solar irradiance+ HDc + HDs), III (SH coefficient + Dst+ HDc + HDs), and IV (SH coefficient + HDc + HDs). It is found that the MAE/RMSE obtained from both the training and validation sets of group 1 reaches minimum when the input combination I is used in the input layer. This suggests that the combination of the solar and geomagnetic activity information is able to improve the SH coefficient prediction results to some degree. Moreover, similar prediction performances are achieved when the input combination I is applied to next four SH coefficients (see Tables S1-S4 in the supporting information), which are also important ones for TEC determination (see Figure S2 in the supporting information). Therefore, the input combination I is used in the input layer of the proposed NN architecture, although the improvement in predicting the SH coefficient is not significant compared to the other input combinations.

3.2 The predicted SH coefficient performance

Figure 4(a-f) shows examples of the 1st SH coefficient prediction results from our proposed NN, and there are also similar prediction performances for the rest 255 SH coefficients. Figure 4(a-d) is the scatter plot of the 1st SH coefficient predictions versus corresponding observations for the Train and Test dataset. The Pearson correlation values are generally higher than 0.99, meaning that there is awesome linear correlation between the observed SH coefficient and predicted ones. In terms of RMSE/MAE/ME, the prediction errors with respect to both Train and Test sets are really small when compared to the actual SH coefficient magnitude, so our proposed NN architecture is proper for predicting the SH coefficients for next two hours. Moreover, the predicted SH coefficient performance for the 1st and 2nd hour are somewhat different, and it appears that the first hour predictions have smaller errors and higher correlations with observed values as compared to the second hour results. The RMSE/MAE/ME on Train increases from 0.151/0.112/0.028 to 0.388/0.285/0.071, while from 0.182/0.132/0.043 to 0.296/0.226/0.059 on Test. Therefore, our trained NN model has slightly better performance in predicting the SH coefficient for the first hour than the second hour. Figure 4(e, f) gives the histogram statistics of predicted errors evaluated by both Train and Test set. Blue/orange refers to the difference between the SH coefficient prediction and the Train/Test observations. It is clear that the predicted residual distributions for both Train and Test sets are concentrated around zero with almost no dispersion values, which are well consistent with the evaluation results shown in Figure 4(a-d).

3.3 The performance of the predicted global TEC maps

After predicting the 256 SH coefficients separately, we are able to produce global TEC maps according to formula (1). To test the accuracy of the proposed algorithm, the resulting TEC maps obtained from the predicted SH coefficients are evaluated by comparing with the CODE GIM products. It should be noted that the 256 SH coefficients used for generating the CODE GIM are the Test data mentioned in sector 2, and they are not involved in training the LSTM NN. Figure 5 shows the resulting predicted TEC maps and the corresponding CODE GIM for the first hour, along with their difference, for randomly selected geomagnetic quiet and storm days. Compared to the CODE TEC maps, the typical equatorial ionization anomaly (EIA) crest regions can also be clearly observed from the predicted TEC maps, and the global residuals between them are very small, mostly within 1 TECU. These indicate that the resulting TEC prediction is satisfactory. Residuals may reach to 5 TECU near the EIA region.

Moreover, the storm time TEC maps also agree well with CODE TEC products, since complicated ionosphere features with large TEC magnitudes, i.e., storm-enhanced density (SED) structure (e.g., Foster 1993; Foster et al., 2005; Zou et al., 2013, 2014), are well reproduced in the predicted TEC map (see Figure 5(b)). This result suggests that the developed deep learning approach is also reliable even for the storm-time TEC forecasting.

Figure 6 shows the forecasted TEC residual histogram during both quiet (blue) and storm (red) days for the Testing set. It is clear that the 1st/2nd hour TEC residual distributions for both quiet and storm time are concentrated around zero with very small RMSE errors, though the 1st/2nd hour RMSE for the storm time (1.27/2.20 TECU) is slightly larger compared to those during quiet time (0.86/1.51 TECU). Therefore, the comparison demonstrates that the proposed deep learning algorithm performs well in predicting the global TEC for both storm and quiet times. It is also evident that the TEC residual distribution for the first hour is more concentrated around zero than that for the second hour, and the averaged RMSE increases slightly from 1.27/0.86 to 2.20/1.51 TECU. This suggests that our proposed deep learning approach performs better in predicting the first hour TEC than the second hour, but the degradation with time is minor.

We also compared the proposed method's performance with respect to the persistence model. The persistence implies that the predicted values remain unchanged between current time t and future time $t+\delta t$ (Kleissl, 2013), namely,

$$TEC(t+\delta t) = TEC(t) \tag{4}$$

where δt=1,2 $hour$. $TEC(t)$ represents the TEC observation at current time t, and the $TEC(t+\delta t)$ is the predicted TEC of the 1st/2nd hour based on $TEC(t)$.

Table 2 shows the comparison of the proposed NN and the persistence model in terms of RMSE for the entire Testing set (October 19 to December 31, 2014, and May 27 to Dec 31, 2016), and percentages listed in in Table 2 refer to the RMSE improvement relative to the persistence model ($\frac{RMSE_{NN} - RMSE_{Persistence}}{RMSE_{Persistence}} \times 100\%$). It is found that the TEC forecasting errors of both the 1st and 2nd hours from the proposed NN model are much smaller than that from the persistence model, and the proposed NN model shows consistent improvements of above 60% when compared to the persistence model. This suggests that the proposed NN model significantly outperforms the

persistence model.

To better demonstrate the merit of the proposed NN model, NeQuick-2 and IRI-2016 models are also involved in comparison. The NeQuick-2 is the latest version of the NeQuick ionosphere electron density model, and its outputs (e.g., electron density, TEC) depend on solar activity (monthly-mean sunspot number R12 or 10.7 cm solar radio flux F10.7), location and time (Nava et al., 2008). The IRI-2016 model, as the most widely used empirical climatological model of the ionosphere, can also provide electron density and TEC parameters in the ionospheric altitude range for a given location, time and date, and it is driven by solar (F10.7, R12), ionospheric and magnetic indices (Bilitza et al., 2017; Liu et al., 2019). In our study, the daily F10.7 indices are used for driving both IRI-2016 and NeQuick-2 models, while other optional input parameters are set to default ones. Also, the storm mode of the IRI-2016 model is turned on in this study.

Figure 7 shows the hourly TEC RMSE from the NN, IRI-2016 and NeQuick-2 models with respect to the CODE GIM over the entire Testing set. The first period, i.e., October 19 to December 31, 2014, is shown on the left, and the second period i.e., May 27 to Dec 31, 2016, is shown on the right. The 1st and 2nd hour RMSE errors are shown in the top and bottom rows, respectively, and the hourly Dst index, which represents the condition of geomagnetic activities, is also shown in the top panel. It can be seen that both the 1st /2nd hour TEC RMSE errors from the proposed NN is smaller than those from the IRI-2016 and NeQuick-2 models. The NN prediction errors over the entire Testing set are very low and stable with averaged RMSE within 1.06/1.84 TECU for the 1st /2nd hour, while the RMSE errors from the IRI-2016 and NeQuick-2 models are around 9.21/5.5 TECU, respectively. Therefore, the proposed NN approach is competitive in predicting the global TEC when compared to the traditional IRI-2016 and NeQuick-2 models. One can also see that the time series of the NN RMSE errors remain small during storm periods despite relative increase comparing with quiet times.

4 Conclusion

In this paper, the LSTM NN is successfully applied to forecast the 256 SH coefficients based on historical time series of the SH coefficients data, solar EUV flux and the Dst index, and then the forecasted SH coefficients are used to reproduce the hourly global TEC maps two hours ahead. The above datasets from Jan 1, 2015-May 26, 2016 are used to train the LSTM model, and those from Oct 19-Dec 31, 2014, and May 27-Dec 31, 2016 are selected as the Test set to verify the prediction performance.

Comparing to those models with no external solar or geomagnetic activity drivers concatenated into the LSTM layer, our developed LSTM NN model is able to improve the forecast of the SH coefficients by including the solar and magnetic indices. Therefore, the optimal input data used in the input layer of the LSTM NN architecture are time series of the SH coefficients, solar EUV flux and the Dst index, and time of day. The advantage of this input combination can also be confirmed by the good SH prediction performance in Figure 4, because the Pearson correlation values between the predicted SH coefficients and observations from the Train/Test are generally higher than 0.99, and the predicted errors are also extremely small. We found that the SH predicted error for the 1st hour is slightly smaller than those for the second hour, suggesting the model has better performance in predicting the SH coefficients for the first hour than the second hour.

After using the developed LSTM model to forecast all SH coefficients, the global TEC maps are generated and used for evaluation. The global residuals between the predicted TEC and the CODE GIM TEC for both geomagnetic quiet and storm days are very small, mostly within 1 TECU and rarely reaching 5 TECU. Prominent structures in the ionosphere, such as EIA and SED during storm time, have been reproduced. A comprehensive evaluation of the predicted TEC has been carried out with respect to the independent CODE GIM products, and results show that the 1st/2nd hour TEC residual distributions for both quiet and storm times are concentrated around zero with very small RMSE errors: the 1st/2nd hour TEC RMSE are 0.86/1.51 and 1.27/2.20 TECU during the quiet and storm times, respectively. These indicate that the developed LSTM model performs well during both quiet and storm times. Moreover, the proposed approach also shows competitive TEC forecast performance in terms of RMSE when compared to the persistence model, and the traditional IRI-2016 and NeQuick-2 models. Both the IRI-2016 and NeQuick-2 models are valuable and widely used community models. The conclusion regarding the comparison to both the IRI-2016 and NeQuick-2 models are based on the current version of the model trained and validated using relatively limited datasets. In the future, we will incorporate more data, in particular larger geomagnetic storms, for training and validation and perform a more thorough evaluation.

Currently, the proposed deep learning approach has shown satisfactory short-term TEC prediction performance based on experimental results obtained, and we will extend its prediction horizon to longer term (beyond 2 hours) in further study. At this point, real-time SH coefficients from the CODE center are not available and there is usually a several days' delay for the SH coefficients to become available. In the future, if the SH coefficients become available in real time or near real time, the developed deep learning model can be readily used in forecasting the coefficients and TEC maps.

Acknowledgments

L. Liu acknowledges the support from the Chinese Scholarship Council (No. 201806270175) for visiting Dr. S. Zou at the University of Michigan. Y. Yao is supported by the National Key Research and Development Program of China (No. 2016YFB0501803) and the National Natural Science Foundation innovation research group project (No. 41721003). We also acknowledge the support from the Key Science and Technology Project for the Transport Sector (No. 2019-MS1-013) and the Highway Construction Engineering Technology Project of Zhejiang Provincial Transport Department (No. 2019-GCKY-02). The SH coefficient time series and GIM maps are free available from the CODE analysis center (http://ftp.aiub.unibe.ch/CODE/). The solar EUV data is obtained from the FISM. The geomagnetic data is available from the NASA Goddard Space Flight Center (https://spdf.gsfc.nasa.gov/index .html). The IRI-2016 and NeQuick-2 models are publicly available on http://irimodel.org and https://t-ict4d.ictp.it/nequick2/nequick-2-web-model, respectively.

Reference

[1] Bilitza D, Altadill D, Truhlik V et al (2017). International Reference Ionosphere 2016: from ionospheric climate to real-time weather predictions. Space Weather 15(2):418–429

[2] Boulch, A., Cherrier, N., Castaings, T. (2018). Ionospheric activity prediction using convolutional recurrent neural networks. arXiv preprint arXiv:1810.13273.

[3] Chamberlin, P. C., T. N. Woods, and F. G. Eparvier (2007). Flare Irradiance Spectral Model (FISM): Daily component algorithms and results, Space Weather, 5, S07005, doi:10.1029/2007SW000316.

[4] Chamberlin, P. C., T. N. Woods, and F. G. Eparvier (2008), Flare IrradianceSpectral Model (FISM): Flare component algorithms and results, SpaceWeather, 6, S05001, doi:10.1029/2007SW000372.

[5] Chen, Y., Manchester, W. B., Hero, A. O., Toth, G., DuFumier, B., Zhou, T., ... & Gombosi, T. I. (2019). Identifying solar flare precursors using time series of sdo/hmi images and sharp parameters. Space Weather, 17(10), 1404-1426.

[6] Foster, J. C. (1993). Storm time plasma transport at middle and high latitudes. Journal of Geophysical Research: Space Physics, 98(A2), 1675–1689. https://doi.org/10.1029/92JA02032

[7] Foster, J. C., Coster, A. J., Erickson, P. J., Holt, J. M., Lind, F. D., Rideout, W., et al. (2005). Multiradar observations of the polar tongue of ionization: MULTIRADAR OBSERVATIONS. Journal of Geophysical Research: Space Physics, 110(A9). https://doi.org/10.1029/2004JA010928

[8] García-Rigo, A., Monte, E., Hernández-Pajares, M., Juan, J. M., Sanz, J., Aragón-Angel, A., & Salazar, D. (2011). Global prediction of the vertical total electron content of the ionosphere based on GPS data. Radio science, 46(06), 1-3.

[9] Gers, F. A., Schmidhuber, J., Cummins, F. (1999). Learning to forget: Continual prediction with LSTM. Neural Computation, 12(10), 2451–2471.

[10] Habarulema, J. B., McKinnell, L. A., Opperman, B. D. (2011). Regional GPS TEC modeling; Attempted spatial and temporal extrapolation of TEC using neural networks. Journal of Geophysical Research: Space Physics, 116(A4).

[11] Hernández-Pajares M, Juan JM, Sanz J, Orus R, Garcia-Rigo A, Feltens J, Komjathy A, Schaer SC, Krankowski A (2009). The IGS VTEC maps: a reliable source of ionospheric information since 1998. J Geodesy 83(3):263-275

[12] Hochreiter, S., Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735-1780.

[13] Kleissl, J. (2013). Solar energy forecasting and resource assessment. Academic Press.

[14] Li, Z., Yuan, Y., Wang, N., Hernandez-Pajares, M., Huo, X. (2015). SHPTS: towards a new method for generating precise global ionospheric TEC map based on spherical harmonic and generalized trigonometric series functions. Journal of Geodesy, 89(4), 331-345.

[15] Liu, L., Yao, Y., Zou, S. et al (2019). Ingestion of GIM-derived TEC data for updating IRI-2016 driven by effective IG indices over the European region. J Geod 93, 1911–1930. https://doi.org/10.1007/s00190-019-01291-5

[16] Liu, L., Zou, S., Yao, Y., & Aa, E. (2020). Multi-scale ionosphere responses to the May 2017 magnetic storm over the Asian sector. GPS Solutions, 24(1), 26.

[17] Nava, B., P. Coisson and S.M. Radicella (2008): A new version of the NeQuick ionosphere electron density model, Journal of Atmospheric and Solar-Terrestrial Physics, doi:10.1016/j.jastp.2008.01.015.

[18] Poole, A. W. V., McKinnell, L. A. (2000). On the predictability of foF2 using neural networks. Radio Science, 35(1), 225-234. https://doi.org/10.1029/1999RS900105

[19] Schaer S (1999) Mapping and predicting the earth's ionosphere using the global positioning system. Ph.D. thesis, Ph.D. dissertation. University of Bern, Bern

[20] Srivani, I., Prasad, G. S. V., Ratnam, D. V. (2019). A Deep Learning-Based Approach to Forecast Ionospheric Delays for GPS Signals. IEEE Geoscience and Remote Sensing Letters, 16(8), 1180-1184.

[21] Sun, W., Xu, L., Huang, X., Zhang, W., Yuan, T., Chen, Z., Yan, Y. (2017, July). Forecasting of ionospheric vertical total electron content (TEC) using LSTM networks. In 2017 International Conference on Machine Learning and Cybernetics (ICMLC) (Vol. 2, pp. 340-344). IEEE.

[22] Tulunay, E., Senalp, E. T., Radicella, S. M., Tulunay, Y. (2006). Forecasting total electron content maps by neural network technique. Radio science, 41(4).

[23] Wang, C., Xin, S., Liu, X., Shi, C., Fan, L. (2018). Prediction of global ionospheric VTEC maps using an adaptive autoregressive model. Earth, Planets and Space, 70(1), 18.

[24] Yao Y, Liu L, Kong J, Zhai C (2018) Global ionospheric modeling based on multi-GNSS, satellite altimetry, and Formosat-3/COSMIC data. GPS Solut 22(4):104

[25] Zou, S., Moldwin, M. B., Ridley, A. J., Nicolls, M. J., Coster, A. J., Thomas, E. G., & Ruohoniemi, J. M. (2014). On the generation/decay of the storm-enhanced density plumes: Role of the convection flow and field-aligned ion flow: Generation and Decay of SED Plumes. Journal of Geophysical Research: Space Physics, 119(10), 8543–8559. https://doi.org/10.1002/2014JA020408

[26] Zou, S., Ridley, A. J., Moldwin, M. B., Nicolls, M. J., Coster, A. J., Thomas, E. G., & Ruohoniemi, J. M. (2013). Multi-instrument observations of SED during 24-25 October 2011 storm: Implications for SED formation processes: SED FORMATION PROCESSES. Journal of Geophysical Research: Space Physics, 118(12), 7798–7809. https://doi.org/10.1002/2013JA018860

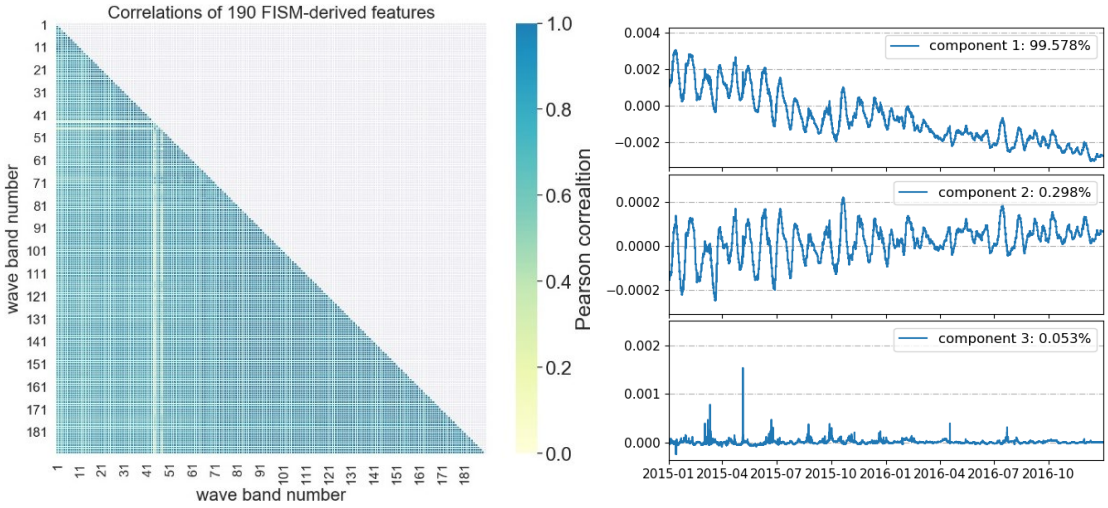


Figure 1 (a) Pearson correlation between the 190 FISM solar EUV fluxes. (b) First three principal components from the 190 FISM wave bands after using the principal component analysis (PCA).

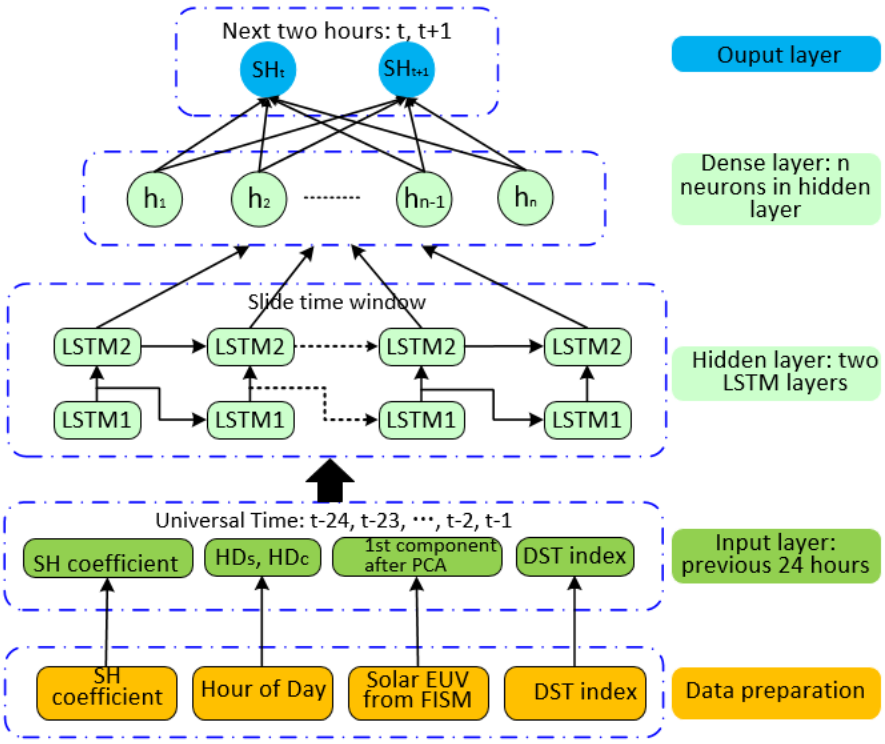


Figure 2 The architecture of the LSTM neural network for the SH coefficient prediction.

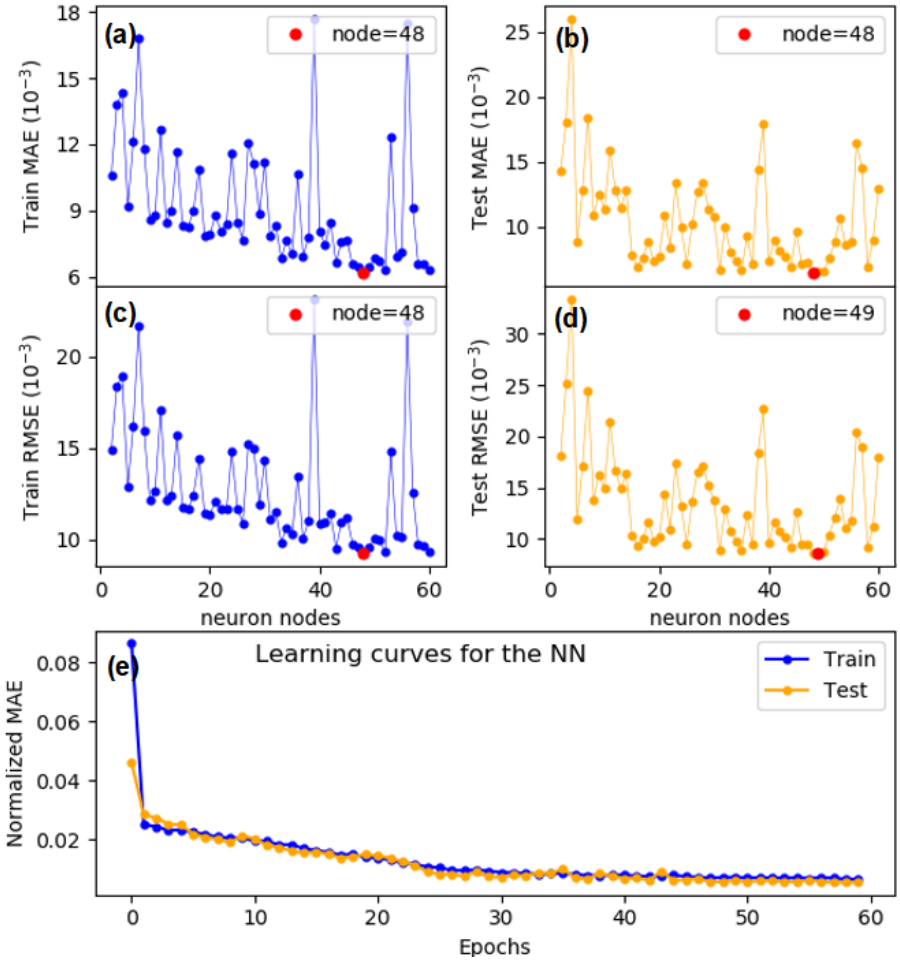


Figure 3 (a-d) MAE and RMSE errors for the 1st SH coefficient prediction results. (e)Normalized MAE from the learning curves. Blue and yellow curves represent corresponding errors from the Train and Test sets, respectively. The input data of the input layer are the 1st SH coefficient, HDc, HDs, the 1st component of the FISM solar irradiance and the Dst index for the previous 24 hours.

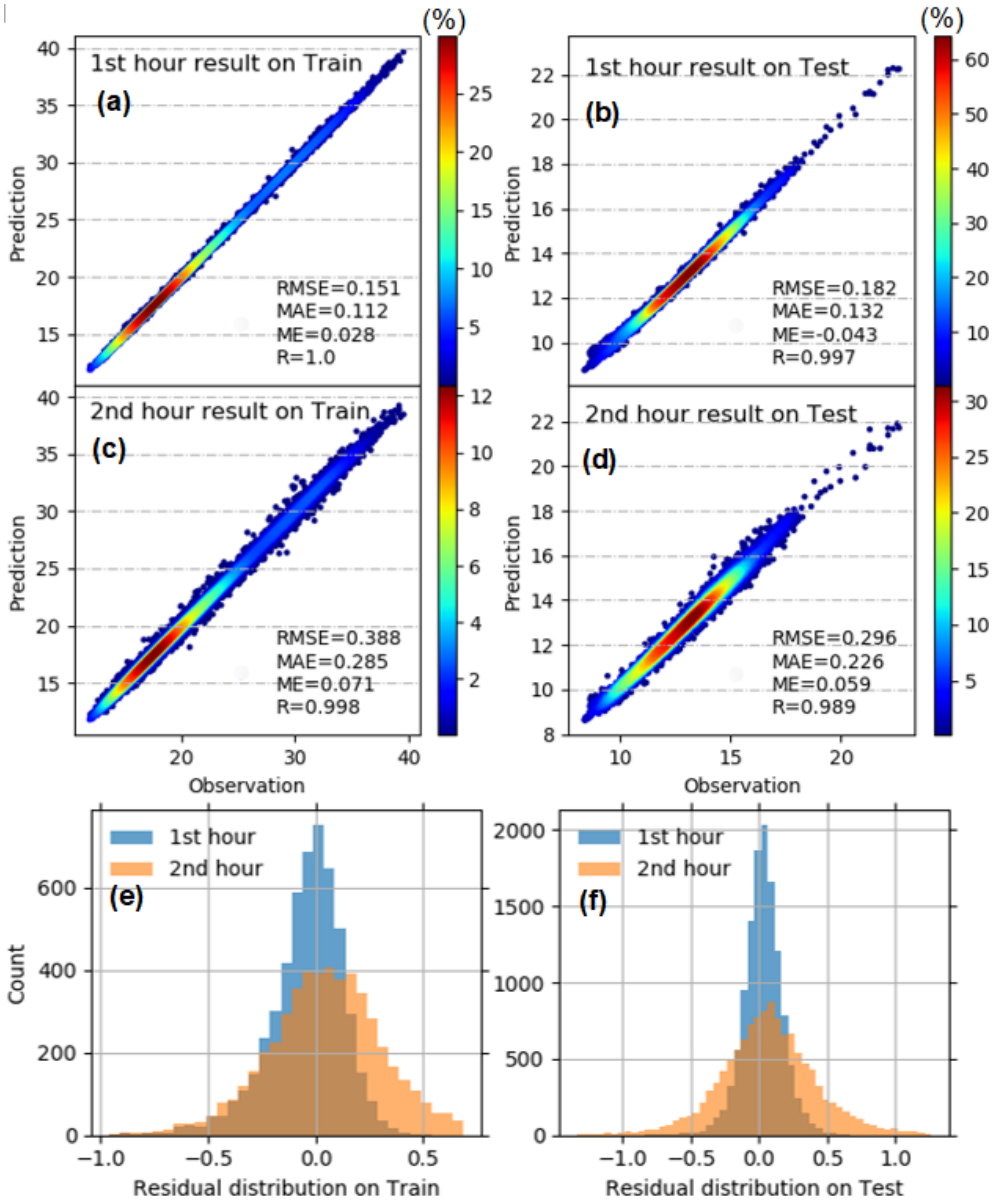
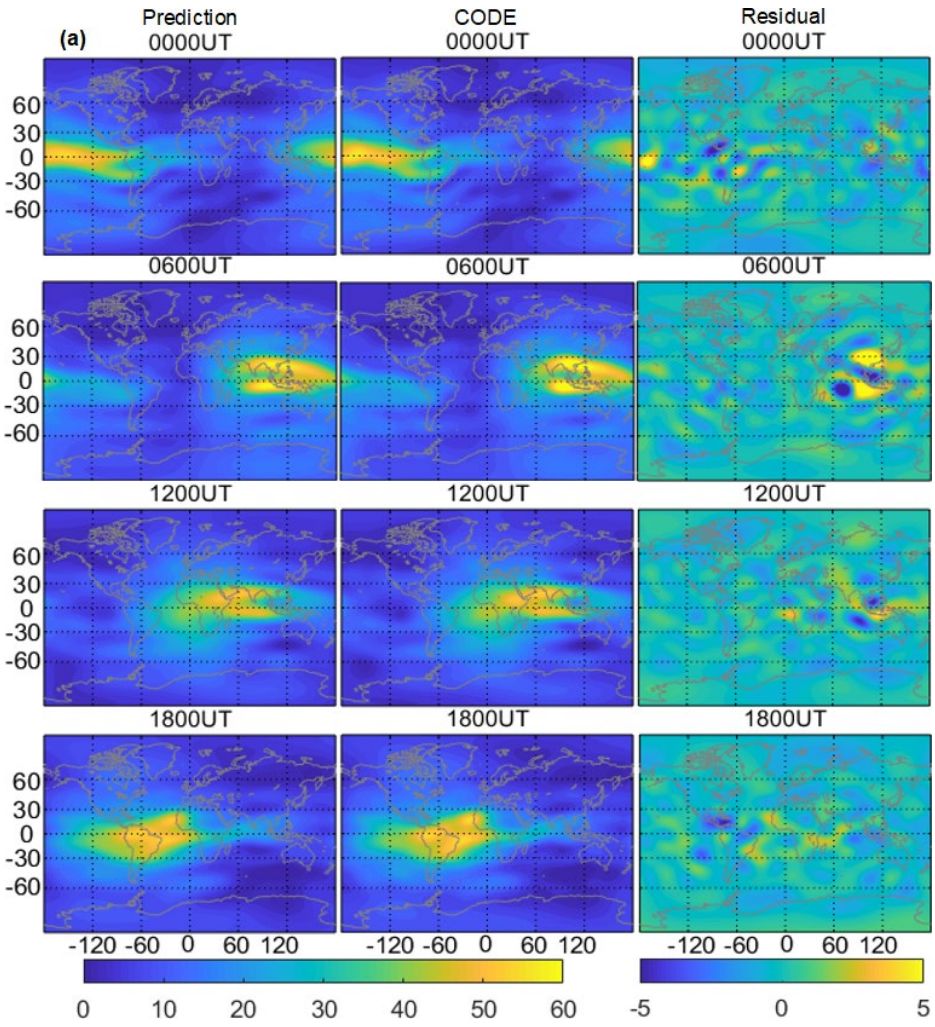


Figure 4 1st SH coefficient prediction results versus Train/Test set for next two hours. 1st hour prediction versus Train (a) and Test set (b). 2nd hour prediction versus Train (c) and Test set (d). The 1st residual distribution for next two hours with respect to Train (e) and Test (f) set. Evaluation metrics, like RMSE, MAE, ME and correlation (R), are also included.



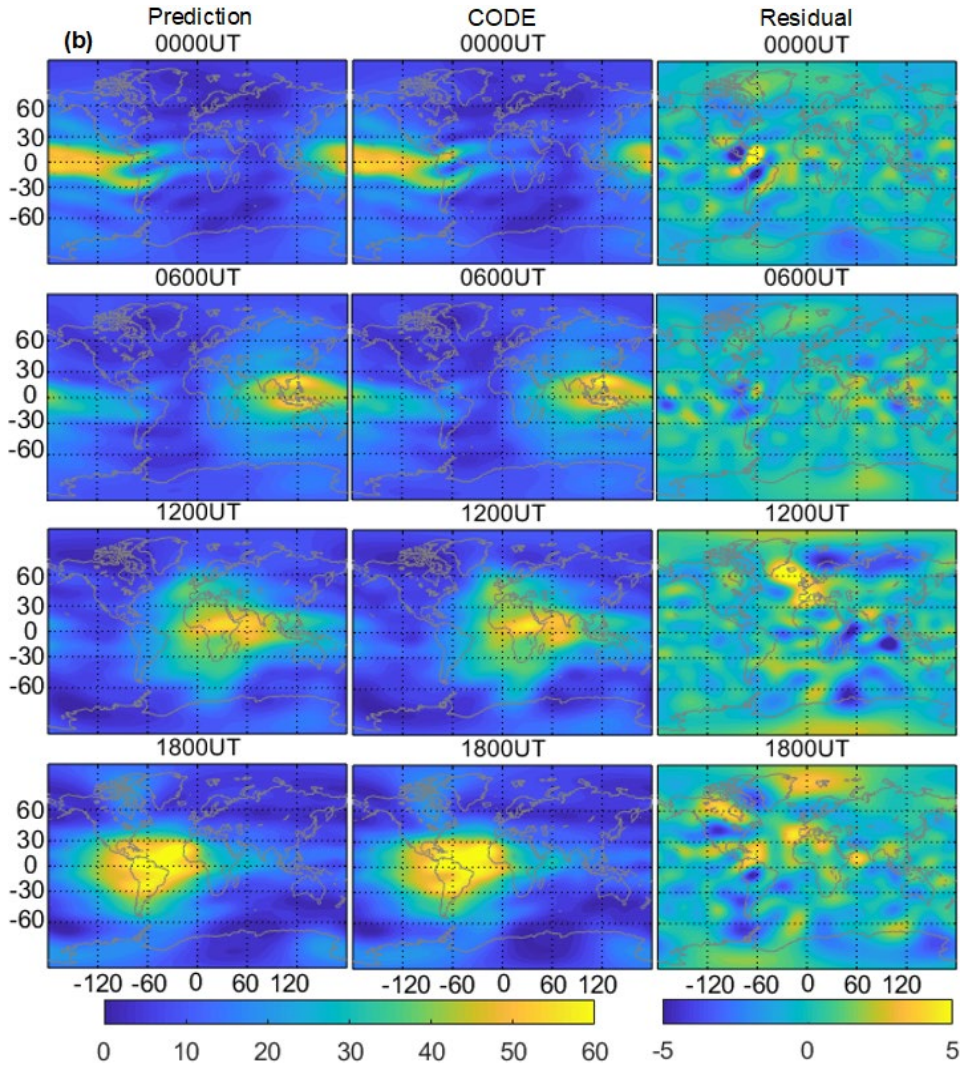


Figure 5 Global TEC maps with six-hour interval under both quiet (October 11, Figure 5a) and storm (October 13, Figure 5b) days in 2016. The predicted TEC maps from our proposed NN and corresponding ones from CODE GIM, along with their differences, are respectively given in the left, middle and right panel of each figure. Full sets of global TEC maps with one-hour interval for both 1st and 2nd hour are available in Movies S1-S12 of the supporting information.

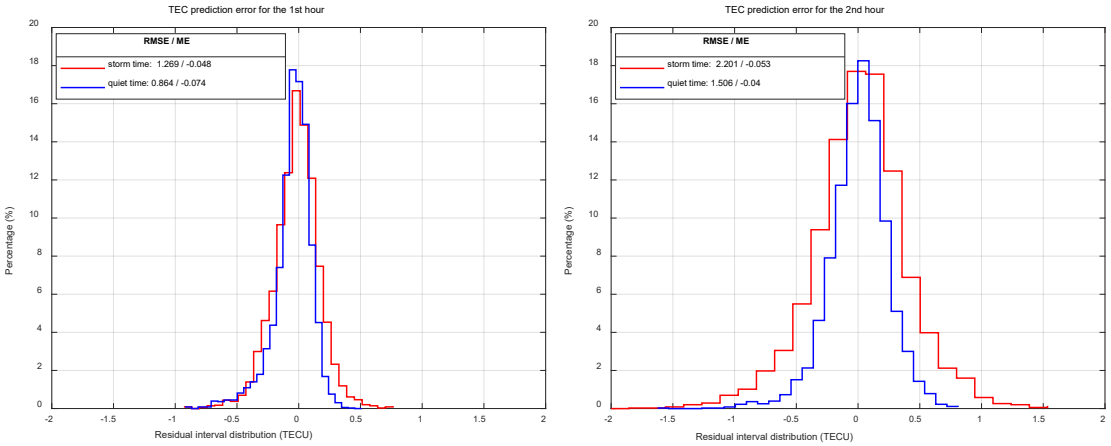


Figure 6 Statistical evaluation of global predicted TEC with respect to the CODE GIM products during both quiet (blue curves) and storm days (red curves) for the Testing set. Left and right panels represent the TEC prediction results for the 1st and 2nd hour, respectively.

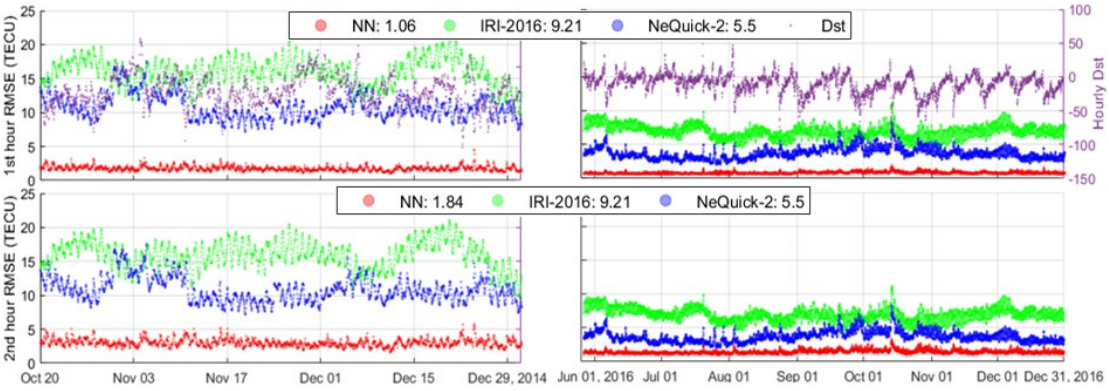


Figure 7 Hourly TEC RMSE errors from the NN (red points), IRI-2016 (green points) and NeQuick-2 (blue points) models with respect to the CODE GIM products over the Testing set (October 19 to December 31, 2014, and May 27 to Dec 31, 2016). The 1st and 2nd hour predicted errors are presented in the top and bottom panel, respectively. Variation of hourly Dst index (purple points) is shown in the right axis of the top panel for reference.

Table 1 Prediction performance in terms of normalized MAE/RMSE for different input parameter combinations

Normalized error	MAE (10-3)		RMSE (10-3)	
Data split	train	validation	train	validation
I	5.869	6.256	8.605	9.031
II	5.919	6.735	9.042	9.853
III	6.103	6.568	9.061	9.486
IV	5.908	6.255	8.974	9.328

Table 2 Global TEC forecasting performance from the proposed NN and persistence model.

Forecast model	Prediction RMSE (TECU)	
	t + 1 (1st hour)	t + 2 (2nd hour)
Persistence	2.69	5.05
NN	1.06 (60.4%)	1.85 (63.3%)