



Article

An Ionospheric TEC Forecasting Model Based on a CNN-LSTM-Attention Mechanism Neural Network

Jun Tang ^{1,2,3} , Yinjian Li ¹ , Mingfei Ding ¹ , Heng Liu ¹, Dengpan Yang ¹ and Xuequn Wu ^{2,*}

¹ School of Transportation Engineering, East China Jiaotong University, Nanchang 330013, China; jtang@ecjtu.edu.cn (J.T.); yinjianli@ecjtu.edu.cn (Y.L.); dingmingfei@ecjtu.edu.cn (M.D.); hengliu@ecjtu.edu.cn (H.L.); yangdengpan@ecjtu.edu.cn (D.Y.)

² Faculty of Land Resources Engineering, Kunming University of Science and Technology, Kunming 650093, China

³ Key Laboratory for Digital Land and Resources of Jiangxi Province, East China University of Technology, Nanchang 330013, China

* Correspondence: xuequnwu@kust.edu.cn

Abstract: Ionospheric forecasts are critical for space-weather anomaly detection. Forecasting ionospheric total electron content (TEC) from the global navigation satellite system (GNSS) is of great significance to near-earth space environment monitoring. In this study, we propose a novel ionospheric TEC forecasting model based on deep learning, which consists of a convolutional neural network (CNN), long-short term memory (LSTM) neural network, and attention mechanism. The attention mechanism is added to the pooling layer and the fully connected layer to assign weights to improve the model. We use observation data from 24 GNSS stations from the Crustal Movement Observation Network of China (CMONOC) to model and forecast ionospheric TEC. We drive the model with six parameters of the TEC time series, Bz, Kp, Dst, and F10.7 indices and hour of day (HD). The new model is compared with the empirical model and the traditional neural network model. Experimental results show the CNN-LSTM-Attention neural network model performs well when compared to NeQuick, LSTM, and CNN-LSTM forecast models with a root mean square error (RMSE) and R^2 of 1.87 TECU and 0.90, respectively. The accuracy and correlation of the prediction results remained stable in different months and under different geomagnetic conditions.

Keywords: ionospheric prediction; total electron content; deep learning; long-short term memory neural network; attention mechanism



Citation: Tang, J.; Li, Y.; Ding, M.; Liu, H.; Yang, D.; Wu, X. An Ionospheric TEC Forecasting Model Based on a CNN-LSTM-Attention Mechanism Neural Network. *Remote Sens.* **2022**, *14*, 2433. <https://doi.org/10.3390/rs14102433>

Academic Editor: Shuanggen Jin

Received: 5 April 2022

Accepted: 17 May 2022

Published: 19 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The ionospheric total electron content (TEC), as a primary ionospheric parameter derived from global navigation satellite system (GNSS) observations, is defined as the total number of electrons within a 1 m^2 column along a path through the ionosphere and is measured in TEC Units (TECU), where $1 \text{ TECU} = 10^{16} \text{ electrons/m}^2$ [1]. The prediction of ionospheric TEC is necessary to indicate adverse space weather conditions for initiating necessary measures in GNSS applications. Significant research studies were carried out to evaluate the utility of using different neural networks to investigate the ionospheric parameters forecasting [2–13]. Francis et al. [3] performed ionospheric parameter prediction and provided a technique to fill in missing data points, while minimizing the impact on data dynamics. Tulunay et al. [4] presented a neural network-based mapping technique for forecasting of ionospheric TEC over Europe. Habarulema et al. [5,6] forecasted ionospheric TEC in southern Africa using feedforward neural networks and investigated both the potential and the limitations of the ionospheric forecast extrapolation capabilities of artificial neural networks (ANNs). Huang et al. [7] implemented TEC forecasting for GPS ground stations using the radial basis function (RBF) neural network technique with the addition of a linear output layer. Ferreira et al. [8] used a neural network model for ionospheric TEC

estimation in the Brazilian region with good results in the presence of low solar activity. Bai et al. [9] took advantage of the extreme learning machine (ELM) with fast learning speed and good generalization and established the ionospheric F2 layer critical frequency (foF2) forecast model. Lee et al. [10] achieved global ionospheric TEC forecasting using a deep learning approach based on conditional generation adversarial networks. Wang et al. [11] proposed a 1-h advance foF2 prediction method based on chaos theory, which has the advantages of simple structure, easy implementation, high prediction accuracy, a small training data set, and no solar and geomagnetic indices. Adolfs et al. [12] developed a fully connected neural network model using global ionosphere maps (GIMs) data from the last two solar cycles and made good predictions for the nighttime winter anomaly (NWA) forecast. Razin et al. [13] proposed a new method for spatial and temporal modeling forecasting of ionospheric TEC during periods of intense solar activity using a support vector machine (SVM) and evaluated the observation data from 37 GPS stations in Iran.

To improve prediction accuracy, some ionospheric researchers performed feature extraction on the input data. Gampala et al. [14] used synchrosqueezing transform (SST) to improve the autoregressive moving average (ARMA) model for ionospheric TEC forecasting. Dabbakuti et al. [15] used singular spectrum analysis to preprocess the input data, which were then fed into an ANN for ionospheric TEC prediction. Meanwhile, they developed a prediction system based on the combination of variational mode decomposition (VMD) and kernel extreme learning machine (KELM) for an ionospheric analysis of the internet of things [16]. Recently, deep learning has inspired a surge of interest in ionosphere research. Ruwali et al. [17] developed a hybrid deep learning forecasted model based on GPS observations from Bengaluru, Guntur and Lucknow stations, which is a combination of long-short term memory (LSTM) and convolutional neural network (CNN). Xiong et al. [18] proposed an extended LSTM neural network model consisting of an encoder–decoder structure to forecast ionospheric TEC, which was verified by using one solar cycle of observation data from 15 GPS stations in China. Kim et al. [19] used an LSTM neural network to design an ionospheric forecasting model suitable for geomagnetic storm periods and forecasted foF2 and fmF2 with good results. Zewdie et al. [20] first used a random forest approach to estimate input parameter importance and then used an LSTM deep recurrent neural network approach to predict ionospheric TEC. Chen et al. [21] used a multi-step auxiliary prediction model of deep learning to predict global ionospheric TEC and found that the method could effectively alleviate increasing errors with prediction time. Although some works have proposed hybrid deep learning methods to capture multiple features including spatial, temporal, and seasonal features for ionospheric parameters prediction, they are usually processed independently. However, the attention mechanism could learn dynamic spatio-temporal relationships in order to capture the relative importance of adjacent monitoring sites and, thus, improve the accuracy of the model [22].

Therefore, we develop a novel CNN-LSTM model with attention mechanism for ionospheric TEC forecasting. The model can optimize the weight distribution of the input information at the full connection layer to predict ionospheric TEC and reflect the spatio-temporal relationships between GNSS stations. In this study, the 9-year (from 2010 to 2018) observation data of 24 GNSS stations from the Crustal Movement Observation Network of China (CMONOC) were selected to predict the value of the TEC 24 h ahead. In addition, the prediction results of the CNN-LSTM-Attention model were compared with those of the NeQuick model, LSTM, and CNN-LSTM neural network model.

2. Data

The experimental data were selected from 2010 to 2018, which are GNSS observations of the CMONOC. The CMONOC is large-scale and high-precision GNSS network which consists of about 260 permanent GNSS continuous reference stations since 2010, and it can produce data with sampling intervals of 30 s. To comprehensively monitor changes of the ionospheric spatial environment in China, we selected 24 GNSS stations covering China for

experimental analysis. The distribution of these stations is shown in Figure 1. Table 1 shows geographic coordinates of 24 GNSS stations from the CMONOC in detail. Table 2 shows the relationship between Universal Time (UT) and Local Time (LT) of 24 GNSS stations. We use the data of 24 GNSS stations in China to calculate the ionospheric TEC by the uncombined precise point positioning (UPPP) method of double-frequency GNSS receivers [23]. The interplanetary magnetic field southward component B_z , the geomagnetic index K_p , the magnetic storm ring current index Dst , and the solar activity index $F10.7$ are also used as auxiliary training data. The TEC time series is derived from GNSS observation data by using the UPPP method and is produced with sampling intervals of one hour. Other auxiliary training data are provided by NASA in 1 h increments.

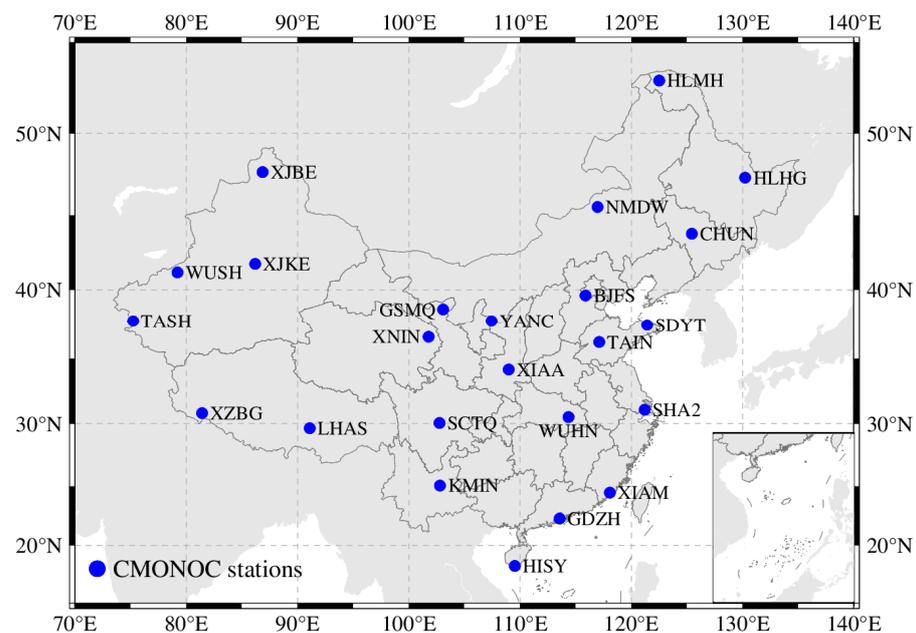


Figure 1. Locations of the 24 GNSS stations from the CMONOC.

Table 1. Geographic coordinates of 24 GNSS stations from the CMONOC.

Stations	Latitude (°)	Longitude (°)	Stations	Latitude (°)	Longitude (°)
BJFS	39.61	115.89	SHA2	31.10	121.20
CHUN	43.79	125.44	TAIN	36.21	117.12
GDZH	22.28	113.57	TASH	37.77	75.23
GSMQ	38.63	103.09	WUHN	30.53	114.36
HISY	18.24	109.53	WUSH	41.20	79.21
HLHG	47.35	130.24	XIAA	34.18	108.99
HLMH	52.98	122.51	XIAM	24.45	118.08
KMIN	25.03	102.80	XJBE	47.69	86.86
LHAS	29.66	91.10	XJKE	41.79	86.19
NMDW	45.51	116.96	XNIN	36.60	101.77
SCTQ	30.07	102.77	XZBG	30.84	81.43
SDYT	37.48	121.44	YANC	37.78	107.44

Table 2. UT to LT conversion for 24 GNSS stations from the CMONOC.

Stations	LT	Stations	LT
BJFS	UT+8	SHA2	UT+8
CHUN	UT+8	TAIN	UT+8
GDZH	UT+7	TASH	UT+5
GSMQ	UT+7	WUHN	UT+8
HISY	UT+7	WUSH	UT+5

Table 2. Cont.

Stations	LT	Stations	LT
HLHG	UT+9	XIAA	UT+7
HLMH	UT+8	XIAM	UT+8
KMIN	UT+7	XJBE	UT+6
LHAS	UT+6	XJKE	UT+6
NMDW	UT+8	XNIN	UT+7
SCTQ	UT+7	XZBG	UT+5
SDYT	UT+8	YANC	UT+7

3. Model and Methodology

3.1. Convolutional Neural Network

The essence of convolution is a weighted sum of data. CNN is a feedforward neural network based on convolutional convolutions, which was initially applied in the field of image recognition and later showed great potential in time series prediction [24]. CNN is composed of convolutional layer, pooling layer, and fully connected layer. The convolutional layer consists of several convolutional units, the parameters of each convolutional unit are optimized by a back-propagation algorithm, and the purpose of the convolutional operation is to extract the different features of the input. The input data are passed through the convolution layer to obtain the feature map. Feature map is the result of the convolution of the input data by the neural network, which characterizes a feature in the neural space. The purpose of the pooling layer is to reduce the feature map, and its main role is to reduce the computational effort by reducing the parameters of the network and to be able to control overfitting to some extent. The fully connected layer combines all local features into global features, which are used to calculate the final score for each category.

In this case, the convolution operation is to multiply the local features by the corresponding weights and then accumulate the sum. The pooling operation is to sample the features extracted from the lower layers to reduce the size of the network and obtain the invariant features of the input data. The calculation formula of feature extraction by CNN is as follows:

$$y = \sigma(W_k \otimes x + b_k) \quad (1)$$

$$W_k \otimes x_{i,j} = \sum_{m=0}^{\alpha-1} \sum_{n=0}^{\beta-1} w_{m,n} \times x_{i+m,j+n} \quad (2)$$

where x is the input data; y is the output data; σ is the activation function; W_k and b_k are the weight coefficient and bias function, respectively; k refers to number of convolution kernels. \otimes denotes the discrete convolution operation; α and β are the size parameters of the convolution kernel, respectively. $w_{m,n}$ represents one of the weight coefficients of convolution kernel which is a weight matrix. m and n represent row and column index of input data, respectively.

3.2. Long-Short Term Memory Neural Network

LSTM neural network is improved on the basis of recurrent neural network (RNN). RNN suffers from gradient disappearance and gradient explosion when processing longer sequences. LSTM neural network provides a new solution to problem of short-term memory and introduces a state value and “gate” control structure based on the recurrent neural network. To selectively filter information, the transmission of data information between different units in the hidden layer is controlled by three thresholds: input gate, forget gate, and output gate [25]. The internal structure of LSTM and the form of data flow are shown in Figure 2. The function of the forget gate is to decide whether to reset the previous information and to multiply it with previous memory information to determine whether the information is retained. The input gate controls the input of the current information. When the information passes through the input unit, it is multiplied by the input gate to

determine whether to write the current information. The output gate controls the output of the current memory information, which is multiplied by current memory information to determine whether to output the information. The LSTM information flow process is calculated as follows:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (3)$$

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (4)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (5)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (6)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (7)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (8)$$

where i_t, f_t, o_t are the input gate, forget gate, and output gate, respectively; σ is the activation functions; W and b are the weight coefficient and bias function, respectively; \tilde{C}_t and C_t are the immediate state and long-term state, respectively; \tanh is the hyperbolic tangent activation function; x and h are the current input and output information, respectively.

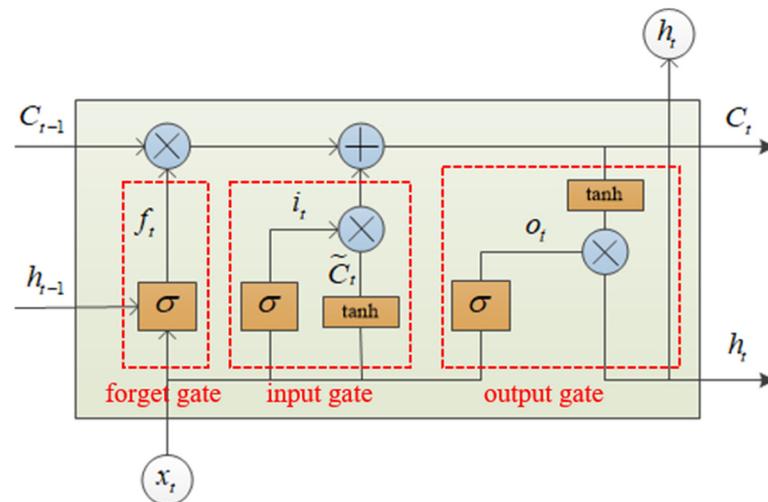


Figure 2. LSTM neural network structure diagram including forget gate, input gate, and output gate.

3.3. Attention Mechanism

When dealing with a large amount of data information, some of the main information will be screened and magnified. Attention mechanism refers to selecting the most appropriate input from numerous alternative information according to observed environmental information. The main purpose of attention mechanism is to process more important information with limited resources and suppress other useless information. The essence is to focus on input weight allocation [26].

In this study, a CNN-LSTM-Attention neural network model is established. The structural framework is shown in Figure 3. The data are first input through the input layer, including the TEC time series and 5 auxiliary training indices of Bz, Kp, Dst, F10.7 and hour of day (HD). The CNN layer is composed of two 1-D convolutions and a maximum pooling layer as encoder based on classic encoder-decoder form. The first convolution reads the input sequence and projects the results onto the feature map. The second convolution performs the same operation on the feature map created at the first layer, trying to magnify its salient features.

The pooling layer is selected for maximum pooling. Maximum pooling means taking the maximum value in the sample as the sample value after sampling. The maximum pooling layer simplifies the feature map, and then it flattens the extracted feature map into a long vector, which is used as the input to the decoding process. Then, the data flow into LSTM model to obtain the TEC forecasted value through three thresholds. Finally, the

final output value is obtained by a weighted sum of attention mechanism, and the TEC forecasted value in the next 24 h is the output.

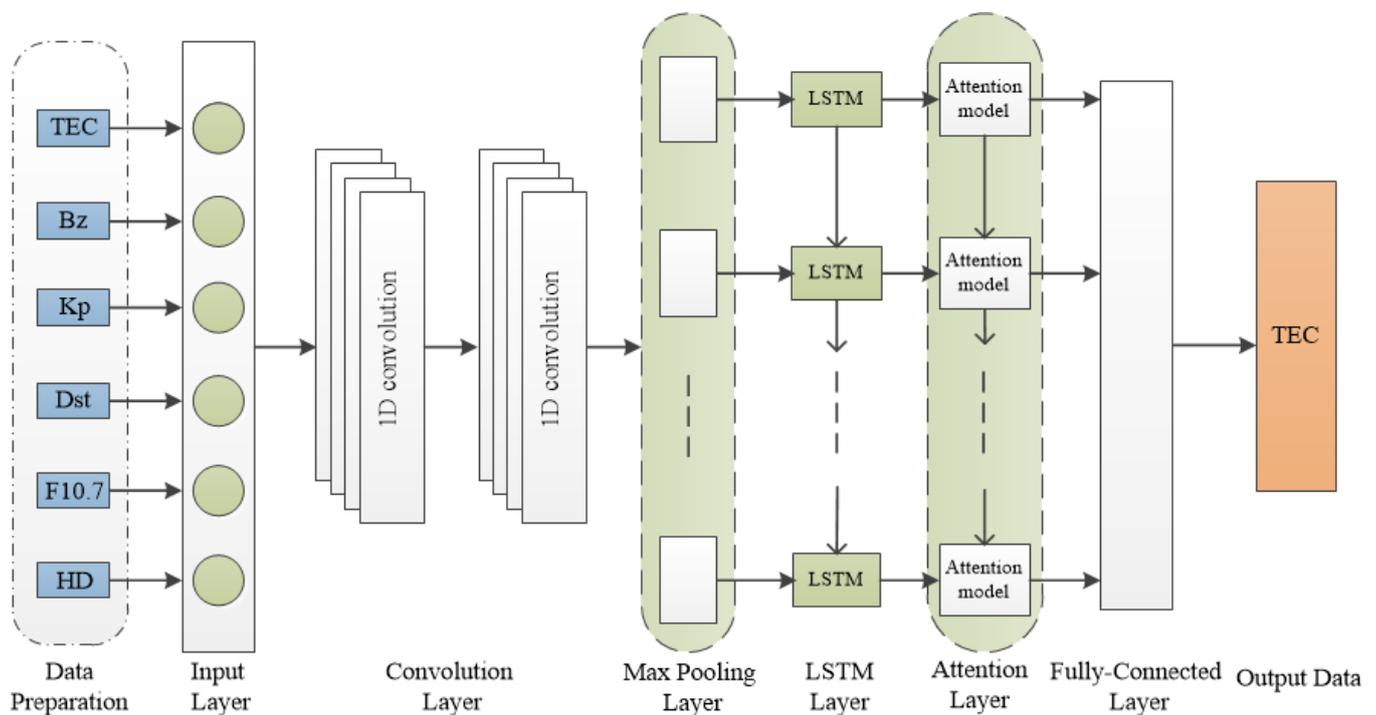


Figure 3. The structural framework of CNN-LSTM-Attention neural network model. The model consists of input layer, convolution layer, pooling layer, LSTM layer, attention layer, and fully-connected layer. Input data, output data, and all layers of data flow in the model are marked.

3.4. Data Organization and Parameter Setting

In this study, experimental data are divided into two parts: the training set and the test set. The data from 2010 to 2017 are used as the training set to train the model, and the data from 2018 are used as the test set. Firstly, the outliers and missing values of the TEC data obtained by the solution are processed. The outliers are eliminated by the triple standard deviation method, and the missing values are filled using the bilinear interpolation method. The variation of TEC is associated with HD and solar and geomagnetic activity. Solar and geomagnetic activity can be characterized by the Bz, Kp, Dst, and F10.7 indices. Then, six features are considered as the training set, which are TEC time series, Bz, Kp, Dst, F10.7 indices, and HD. A total of 192 groups of data from 8 days were used as sample sliding segmentation. Each sample used the data of the first 7 days to forecast the TEC value for the last 1 day, and the training set samples are input into the network training forecast model. Figure 4 shows four input indices Bz, Kp, Dst, and F10.7 and the measured TEC values of four stations at HLMH, BJFS, WUSH, and HISY from 2010 to 2018. The Bz index is uniformly distributed on both sides of 0. Geomagnetic activity shows irregularities, including 2 mega-magnetic storms in 2015. Solar activity peaked in 2014, coinciding with a maximum TEC value.

Hyperparameters are the parameters for which machine learning sets values before starting the learning process. The setting of hyperparameters is the key to the prediction effect of the neural network. We used the training set root mean square error (RMSE) as the evaluation index and chose the grid search method to find optimal hyperparameters. The hyperparameters include batch size, epochs number, and filters number and kernel size. We set the loss function as the mean absolute error (MAE) and used the ‘tanh’ activation function and ‘adam’ optimization algorithm [27] to drive the model.

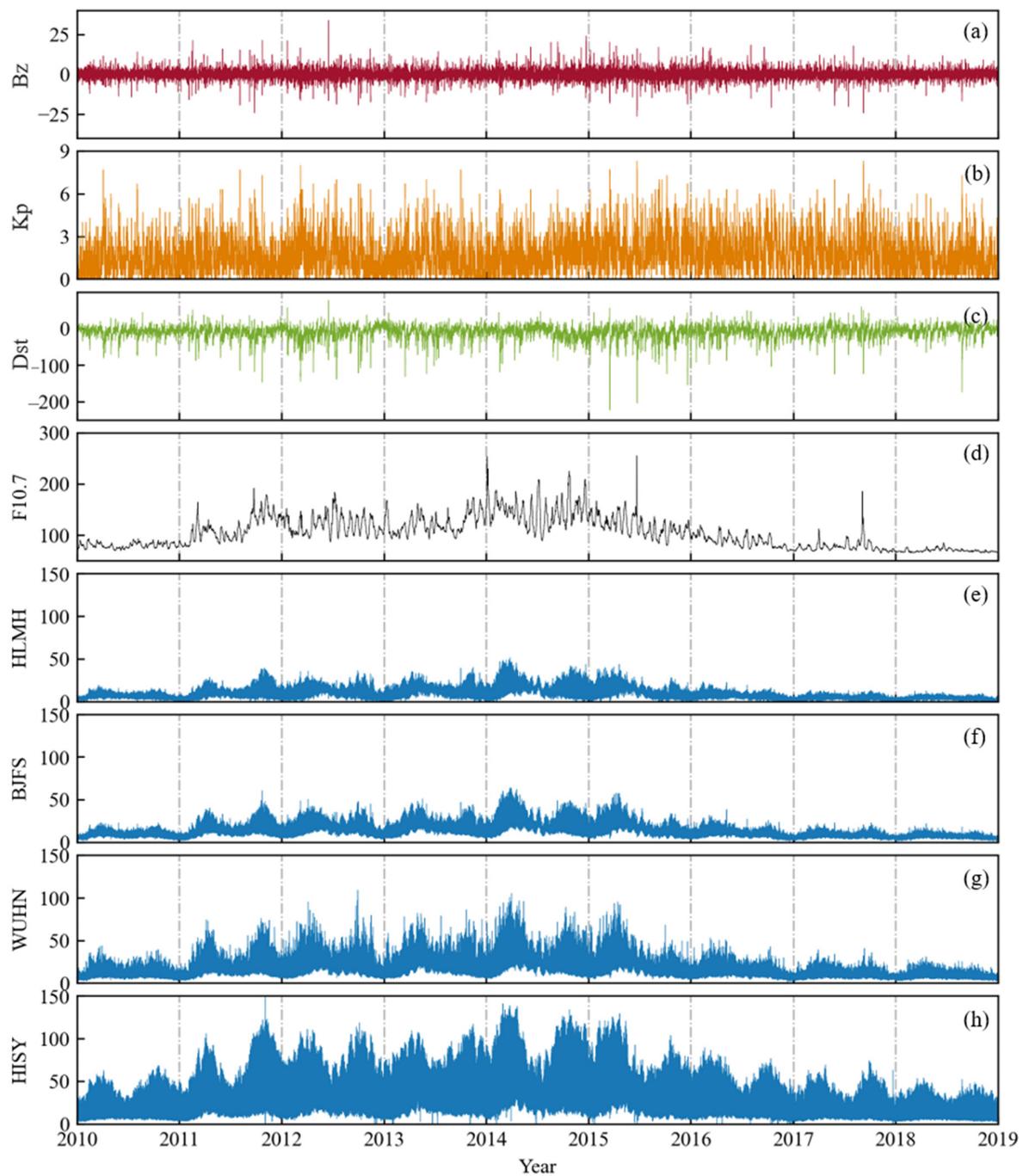


Figure 4. Variations of (a) Bz, (b) Kp, (c) Dst, (d) F10.7 indices and TEC values at (e) HLMH, (f) BJFS, (g) WUSH, and (h) HISY stations from 2010 to 2018.

4. Results and Evaluation

To verify the advantages of the CNN-LSTM-Attention model for ionospheric TEC forecasting, we select the NeQuick model [28], the LSTM neural network [29], and the CNN-LSTM neural network [17] for comparative analysis. The NeQuick model is based on the NeQuick-2 version released by the Abdus Salam International Centre for Theoretical Physics (ICTP), and we used the daily solar radio flux as the input data. The LSTM and CNN-LSTM models have the same input data as the model in this paper. The forecast performance of each model is evaluated according to the station location, time variation, and geomagnetic activity from the test set forecasts. We also used the RMSE, correlation coefficient (R^2), MAE, and mean absolute percentage error (MAPE) to evaluate model

performances. RMSE and MAE indicate the dispersion of errors between forecasted and observed values. R^2 indicates the correlation between forecasted and observed values. MAPE can measure the relative error between the average forecasted value and the observed value. The calculation formulas are as follows:

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

$$R^2 = \left(\frac{\text{cov}(\hat{y}_i, y_i)}{\sigma_{\hat{y}_i} \cdot \sigma_{y_i}} \right)^2 \quad (10)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (11)$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\% \quad (12)$$

where \hat{y}_i is the TEC forecasted value; y_i is the observed value and n is the number of data points.

4.1. Accuracy Assessment of Different Stations

Figure 5 shows the two-dimensional scatterplot of the CNN-LSTM-Attention model predictions and GNSS measurements for HLMH, TASH, SDYT, and HISY stations. The horizontal axis indicates the measured values, and the vertical axis indicates the forecasted values. The color bar indicates the density share. $Y = f(X)$ denotes the equation of the scatter-fitted straight line. Y is the forecasted TEC, and X is the observed TEC. It can be observed that the TEC distributions of different stations show different patterns. The TEC of high latitude station are basically distributed within 10 TECU. The maximum TEC of HISY station located at low latitudes exceeds 50 TECU. The CNN-LSTM-Attention model forecasts and GNSS measurements are highly correlated, with R^2 of 0.87, 0.88, 0.90, and 0.91 for the four stations, respectively.

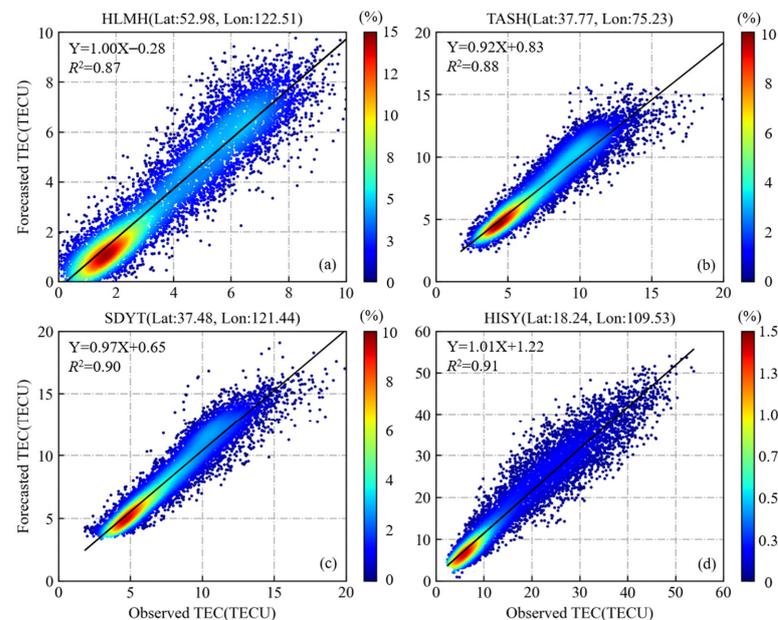


Figure 5. Scatterplot of forecasted TEC as a function of observed TEC for the (a) HLMH, (b) TASH, (c) SDYT, and (d) HISY stations. The horizontal axis is the observed TEC; the vertical axis is the forecasted TEC, and color represents percentages.

Table 3 shows the overall evaluation metrics of the four models. The statistical results are obtained from the data from 24 GNSS stations, 24 h a day and 365 days a year in the

test set. The empirical model NeQuick model has the maximum RMSE and the minimum correlation. The RMSEs of LSTM, CNN-LSTM, and CNN-LSTM-Attention models are 2.25, 2.07, and 1.87 TECU, respectively, and the MAEs are 1.53, 1.36, and 1.17 TECU, respectively. This indicates the neural network model has good forecast performance. The prediction performance of the improved model is the best.

We selected 12 stations covering the range of China in the longitudinal and latitudinal directions for effect evaluation, with 6 stations distributed around 120°E in the longitudinal direction and 6 stations distributed around 30°N in the latitudinal direction. Figure 6 shows the performance evaluation indexes of the annual forecasted results for the four models test set of 12 GNSS stations. The evaluation results of the other 12 stations are shown in Figure A1. Figure 5a,c are the RMSE and R^2 of the observed and forecasted values of the 6 stations distributed in longitude, respectively. Figure 5b,d are the RMSE and R^2 of the observed and forecasted values of the six stations distributed in latitude, respectively. The four colors represent the four models used in the forecast, respectively. It can be observed that the CNN-LSTM-Attention model has better forecasting performance compared to the other three models. RMSE decreases with the increase in latitude near the same longitude, which is about 1 TECU in mid-latitude region and between 1.6 and 3 TECU in six stations around 30°N. The GDZH and HISY stations at low latitudes are affected by ionospheric anomalies near the equator [30]. The main driving forces are due to the interaction of an electric field, which develops close to the magnetic equator, and the horizontal magnetic field causing $E \times B$ drift of the plasma from low altitudes to high altitudes with the plasma then falling under gravity to form the equatorial (or Appleton Anomaly) at approximately plus and minus 20 degrees. The RMSEs of GDZH and HISY stations exceeded 3 TECU. In terms of correlation, the R^2 of all 24 stations of the CNN-LSTM-Attention model exceeded 0.8, and the forecasted values are highly correlated with the measured values.

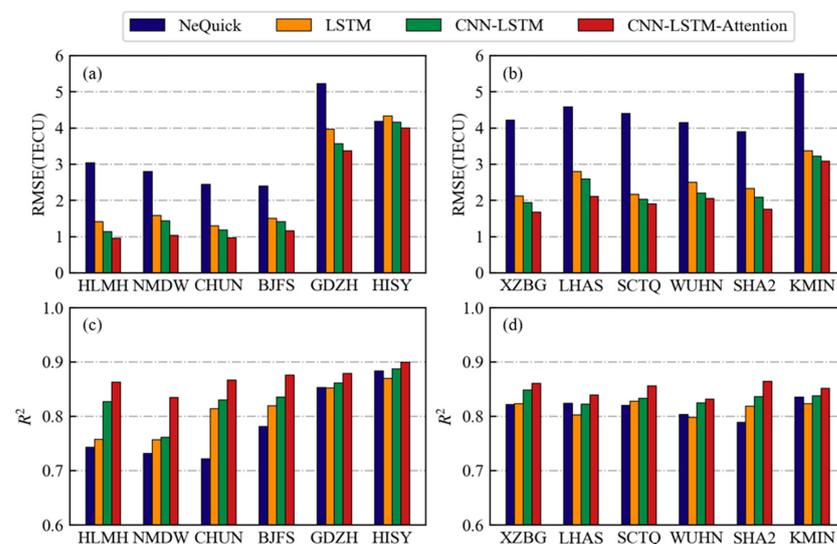


Figure 6. Histogram for RMSE and R^2 of NeQuick (blue), LSTM (yellow), CNN-LSTM (green), and CNN-LSTM-Attention (red) models from 12 stations in the test set. (a,b) are RMSE. (c,d) are R^2 .

Table 3. The overall evaluation indexes of four models in 24 stations in the test set.

Modes	Evaluate Indexes		
	RMSE (TECU)	R^2	MAE (TECU)
NeQuick	3.59	0.81	2.60
LSTM	2.25	0.85	1.53
CNN-LSTM	2.07	0.87	1.36
CNN-LSTM-Attention	1.87	0.90	1.17

Figure 7 shows the single-day RMSE box plots for the 12 GNSS stations. The RMSE box plots for the other 12 stations are shown in Figure A2. Each box represents the 365 RMSEs of a single model throughout the year, in which the red plus sign is an outlier. The top and bottom black bars represent the maximum and minimum values of the 365-day data, respectively. The upper and lower edges of the blue box represent the upper quartile and the lower quartile, respectively. Moreover, the middle red bar represents the median. Comparing the four boxes, it can be observed that the model proposed in this paper performs well in terms of prediction, and the maximum and minimum RMSEs are significantly better than the NeQuick model, which is not significantly different from the performance of LSTM and CNN-LSTM. In terms of median, the performance of the model in this paper is better. GDZH, HISY, and KMIN stations are 2.91 TECU, 3.36 TECU, and 2.66 TECU, respectively. The median of the remaining nine stations is below 2 TECU, which may be due to the active ionosphere at low latitudes.

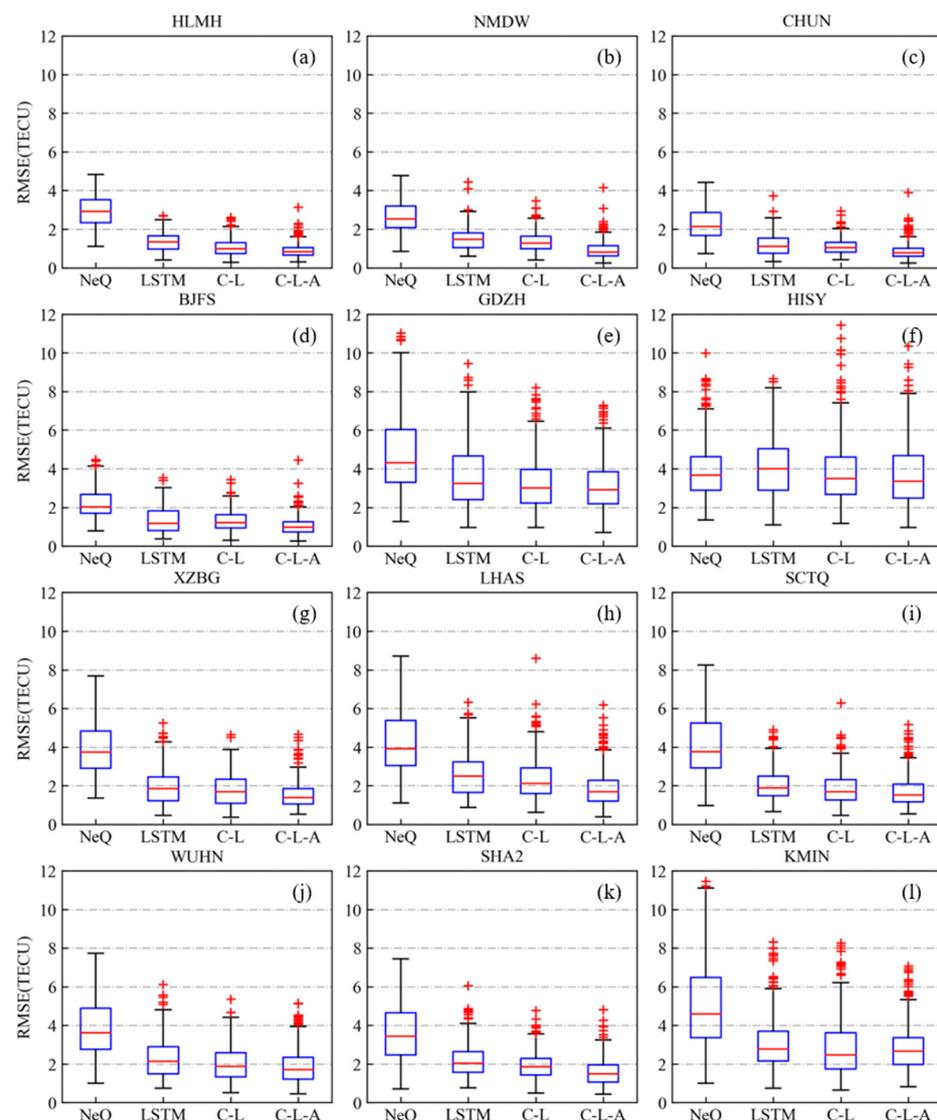


Figure 7. (a–l) The RMSE of NeQuick, LSTM, CNN-LSTM, and CNN-LSTM-Attention models from 12 stations in the test set. The top and bottom black bars represent the maximum and minimum values, respectively. The upper and lower edges of the blue box represent the upper quartile and the lower quartile, respectively. The middle red bar represents the median, and the red plus sign represents the outlier.

4.2. Accuracy Assessment at Different Time Periods

Figure 8 shows the monthly changes of the forecast performance of the four prediction models in different months. The RMSE of the NeQuick model was 2.73, 2.18, 2.85, and 2.94 TECU in June, July, August, and December, respectively, and it was greater than 3 TECU in all other months. The LSTM and CNN-LSTM models perform well, with monthly average RMSE lower than 3 TECU and correlation coefficient higher than 0.8. As observed from the figure, the LSTM neural network model is unstable in some months. The RMSEs for June and July are 2.60 TECU and 2.41 TECU, respectively. The error for July is even higher than the empirical model. The performance of the CNN-LSTM-Attention model is stable, with an RMSE below 2.2 TECU and R^2 above 0.84 in 12 months, which show good stability.

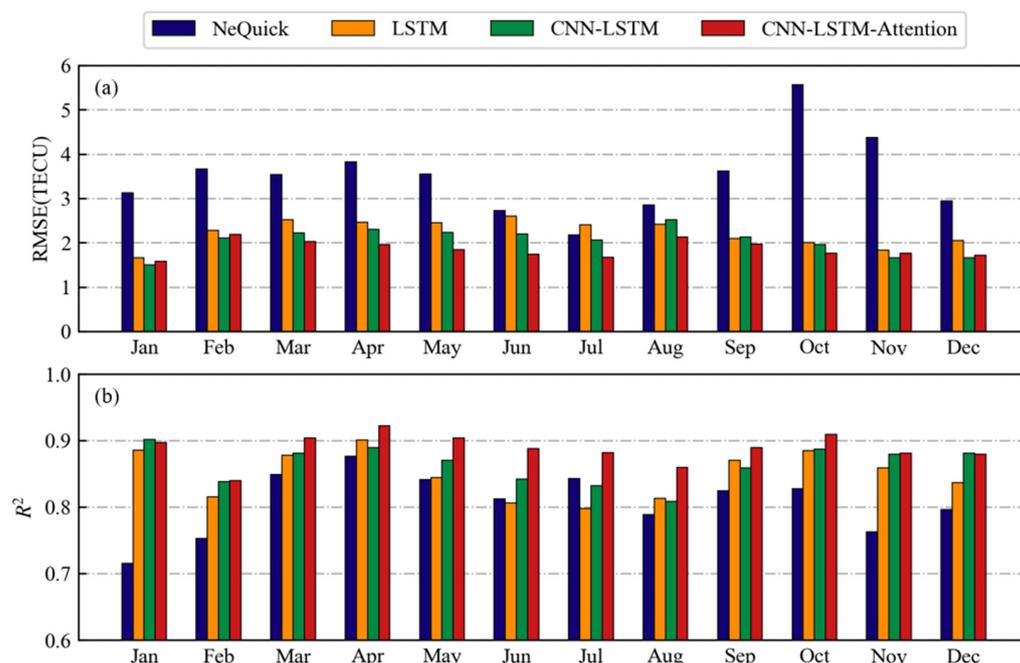


Figure 8. The RMSE and R^2 of NeQuick (blue bars), LSTM (yellow bars), CNN-LSTM (green bars), and CNN-LSTM-Attention (red bars) models from 12 months in the test set. (a,b) are RMSE and R^2 , respectively.

Figure 9 shows the average of the measured values of the 24 stations of the CMONOC and the forecasts of four models for 12 months. The predicted values of the NeQuick model in May, October, and November are quite different from the measured values, and the predicted values of the LSTM and CNN-LSTM models are closer to the measured values. The predicted values of the improved model in this paper are closest to the measured values.

Figure 10 shows the forecast RMSE for each model averaged across all of the stations for the year 2018, where the horizontal axis shows the start time for the forecast in UT. The NeQuick model has the largest RMSE when the forecast is made at 8 UT at 5.97 TECU. The maximum RMSE of the LSTM, CNN-LSTM, and CNN-LSTM-Attention models occurs when the forecast is made at 7 UT with an RMSE of 3.57 TECU, 3.46 TECU, and 3.09 TECU, respectively. The daily ranges of RMSE for these four models are 4.39 TECU, 2.60 TECU, 2.48 TECU, and 2.29 TECU, respectively. The MAPE of the CNN-LSTM-Attention model fluctuates in the range of 12% to 21% in a day. There is no significant fluctuation in MAPE when RMSE reaches its maximum value, which indicates that the large forecasted error at 7 UT and 8 UT is due to the large influence of TEC values. The results indicate that the CNN-LSTM-Attention model exhibits the best performance for all forecast periods, and NeQuick exhibits the worst.

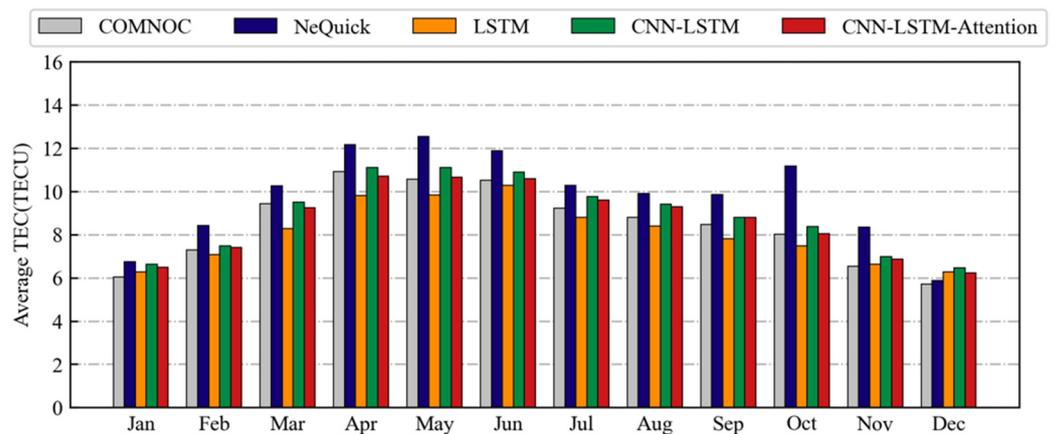


Figure 9. The forecasted averages of GNSS measured values (gray bars), NeQuick (blue bars), LSTM (yellow bars), CNN-LSTM (green bars), and CNN-LSTM-Attention (red bars) models from 12 months in the test set.

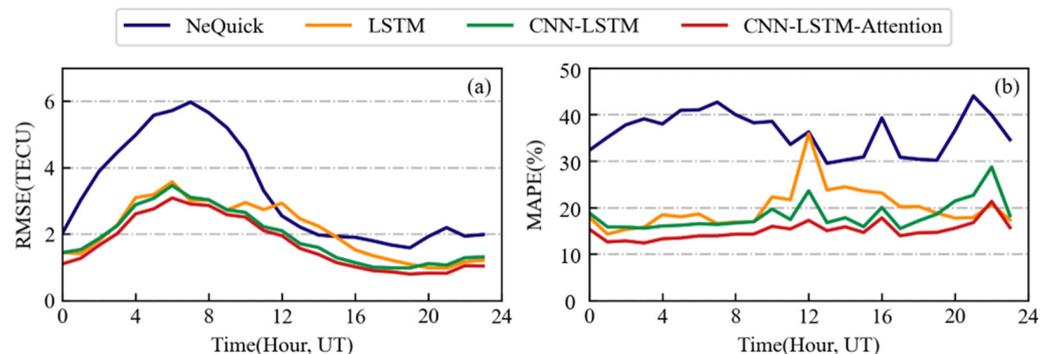


Figure 10. The forecasted (a) RMSE and (b) MAPE for each model, averaged across all of the stations for the year 2018, where the horizontal axis shows the start time for the forecast in UT.

4.3. Accuracy Assessment under Different Geomagnetic Conditions

Geomagnetic activity is one of the main factors affecting the development of ionospheric disturbances [31]. During magnetic storms, the ionospheric TEC at low and middle latitudes shows an enhanced and positive perturbation response. This phenomenon gradually decreases in response intensity with decreasing latitude [32]. To study the prediction ability of the model in this paper under different geomagnetic activity, we divide the data into magnetic quiet days and magnetic storm days for comparative analysis. The magnetic quiet and magnetic storm days are classified according to the daily average of the Kp index. The Kp index is used as an indicator of geomagnetic activity. Kp index presents the index of 3 h ranges in magnetic activity relative to a quiet day. Thus, a higher Kp index means the geomagnetic is more active. Generally, $Kp \geq 3$ can be considered as a magnetic storm day. To further verify the prediction performance of the model under different geomagnetic activities, the GDZH station is selected to analyze the accuracy and the forecasted performance of geomagnetic storms. The GDZH station is located in low-latitude region, and TEC is closely affected by geomagnetic activity, which can well reflect the influence of geomagnetic activity on TEC.

4.3.1. Magnetic Quiet Period

Figure 11 shows the comparison between GNSS measured values and forecasted values of 12 stations on 21 August 2018. The comparison of the other 12 stations is shown in Figure A3. The Kp index on the day is 1.7, and geomagnetic activity is quiet. As observed from the figure, the maximum TEC of the day increases with the decrease in latitude, and there is a great difference between the predicted and measured values of the four stations

north of 40°N using the four models. The predictions of the NeQuick model at some station maxima or minima are inaccurate. The performance of the CNN-LSTM-Attention model is good at 12 stations on the day. The predictions of CNN-LSTM-Attention model for nine stations, which are HLMH, NMDW, CHUN, BJFS, GDZH, LHAS, SCTQ, WUHN, and KMIN, are basically agreement with the observed values. It can accurately present the variation trend of TEC. In general, the forecasted values of the proposed model are consistent with the measured values of GNSS.

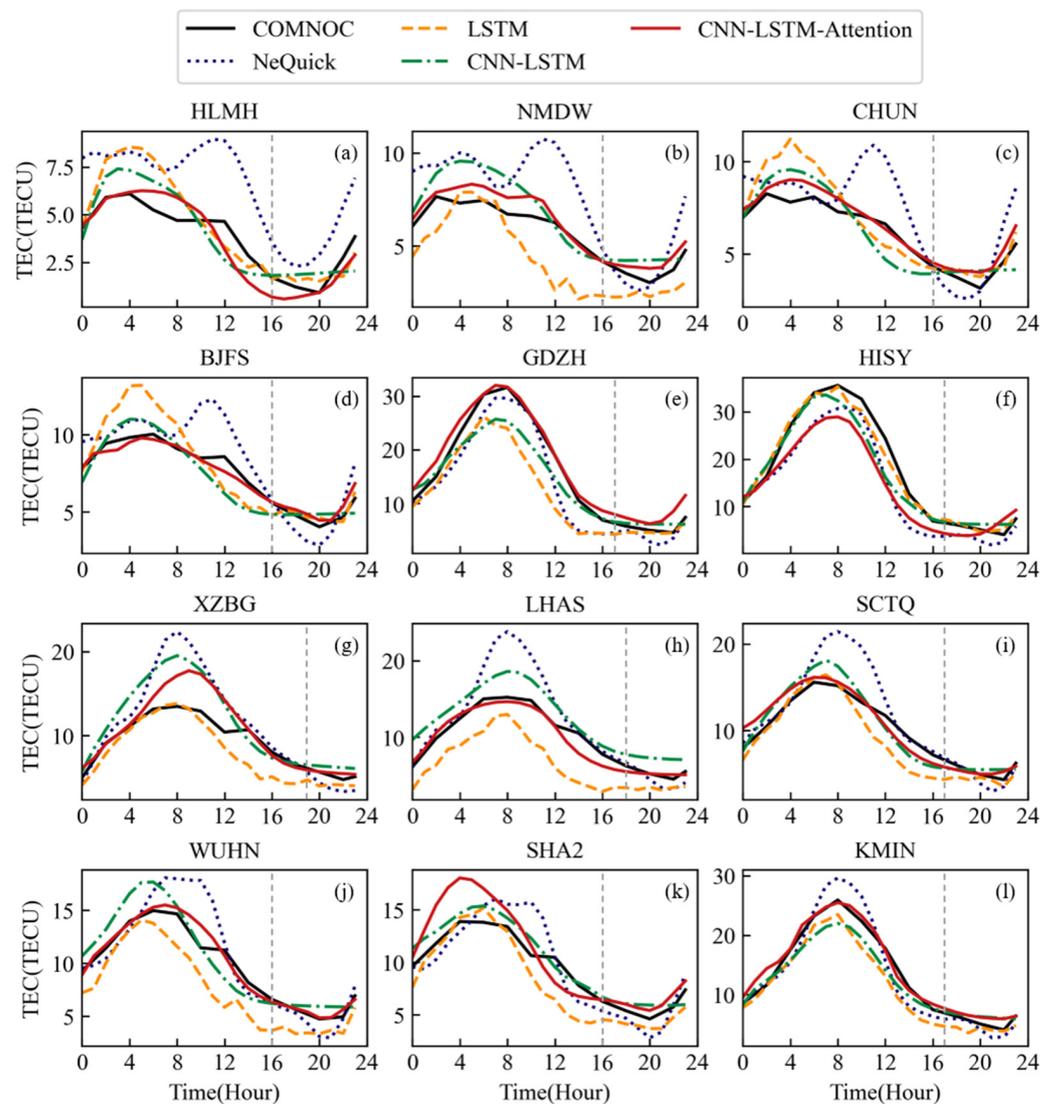


Figure 11. (a–l) Comparison of GNSS measured values (black solid line), NeQuick (blue dotted line), LSTM (yellow dotted line), CNN-LSTM (green dotted line), and CNN-LSTM-Attention (red solid line) models forecasted values for 12 stations on magnetic quiet day. The gray dotted line indicates midnight LT.

Figure 12 shows the distribution of the residuals between the forecasted and measured values of the magnetic quiet period. In general, NeQuick, LSTM, CNN-LSTM, and CNN-LSTM-Attention models forecasted residuals are distributed in the range of -12 to 12 TECU. The prediction residual distribution of CNN-LSTM-Attention model is closer to an unbiased Gaussian distribution. The RMSEs of the four models are 4.14 TECU, 4.29 TECU, 4.14 TECU, and 3.99 TECU, respectively. The MAEs of the four models are 3.16 TECU, 3.00 TECU, 2.98 TECU, and 2.81 TECU, respectively. The NuQuick model has the largest MAE, and the CNN-LSTM-Attention model has the smallest RMSE and MAE.

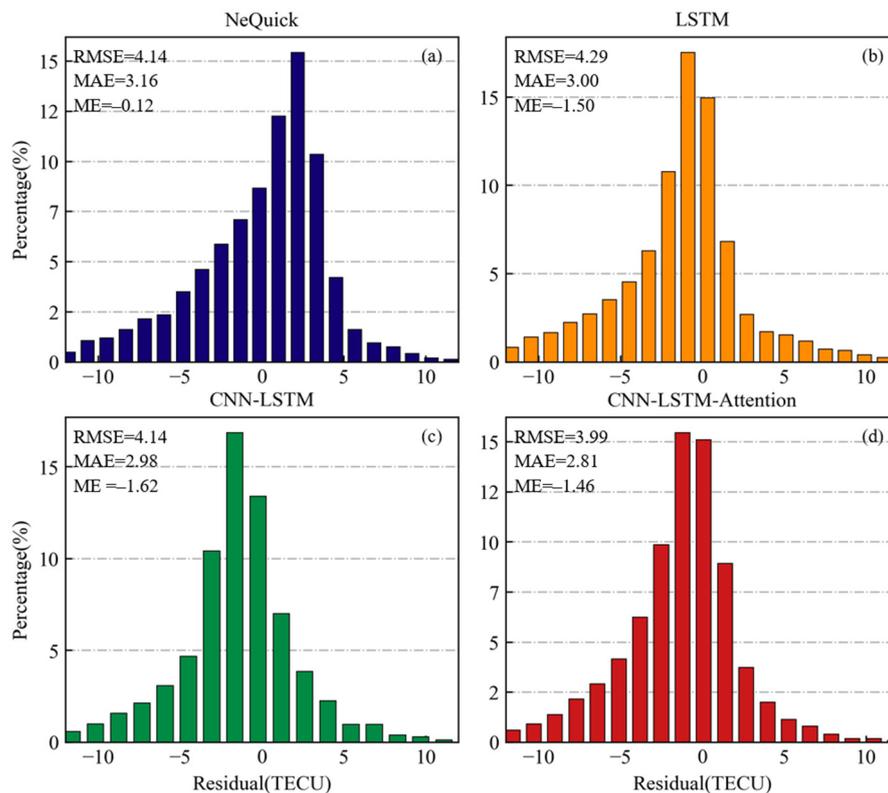


Figure 12. Residual distribution of different models for GDZH station during magnetic quiet period. RMSE, MAE, and ME are also included. (a) NeQuick; (b) LSTM; (c) CNN-LSTM; (d) CNN-LSTM-Attention model.

4.3.2. Magnetic Storm Period

Figure 13 shows the comparison between GNSS measured values and TEC forecasted values of 12 stations on the magnetic storm day of 26 August 2018. The comparison of the other 12 stations is shown in Figure A4. Magnetic storms lead to drastic changes in TEC, and the forecasts of the four forecasting models at the maximum value are all inaccurate. This flaw is particularly evident in the NeQuick model. The LSTM, CNN-LSTM, and CNN-LSTM-Attention models can reflect the impact of magnetic storm on TEC to a certain extent because the geomagnetic indices are added as a training set. The CNN-LSTM-Attention model captures the feature more obviously because of the attention mechanism. The predicted values are the closest to the measured value of GNSS.

Figure 14 shows the distribution of forecasted residuals of the four models during the magnetic storm period. Similarly to the magnetic quiet period, the NeQuick model has the largest residual. The LSTM model forecasted results with RMSE and MAE, and the results are 5.19 TECU and 3.76 TECU, respectively. The forecasted results of the CNN-LSTM model and CNN-LSTM-Attention model include a mean distributed around 0, with RMSEs of 4.43 TECU and 4.11 TECU, respectively, and MAEs of 3.06 TECU and 2.81 TECU, respectively. From the error values, the performance of the model in this paper is better.

Table 4 shows the percentage of residuals between the forecasted and measured values of the four models for the 24 stations in the test set. In the magnetic quiet period and the magnetic storm period, the prediction ratio of residual errors of the improved model in this paper is less than 1 TECU are 62% and 52%, respectively.

Table 4. Residual percentage statistics of 24 stations and 4 models in the test set.

Modes		$\Delta < 1$	$1 \leq \Delta < 2$	$2 \leq \Delta < 3$	$3 \leq \Delta < 4$	$\Delta > 4$
Quiet	NeQuick	27%	24%	19%	11%	19%
	LSTM	48%	28%	13%	5%	6%
	CNN-LSTM	53%	27%	11%	4%	5%
	CNN-LSTM-Attention	62%	24%	7%	3%	4%
Storm	NeQuick	25%	23%	19%	15%	18%
	LSTM	38%	27%	16%	8%	11%
	CNN-LSTM	45%	28%	12%	6%	9%
	CNN-LSTM-Attention	52%	26%	11%	5%	6%

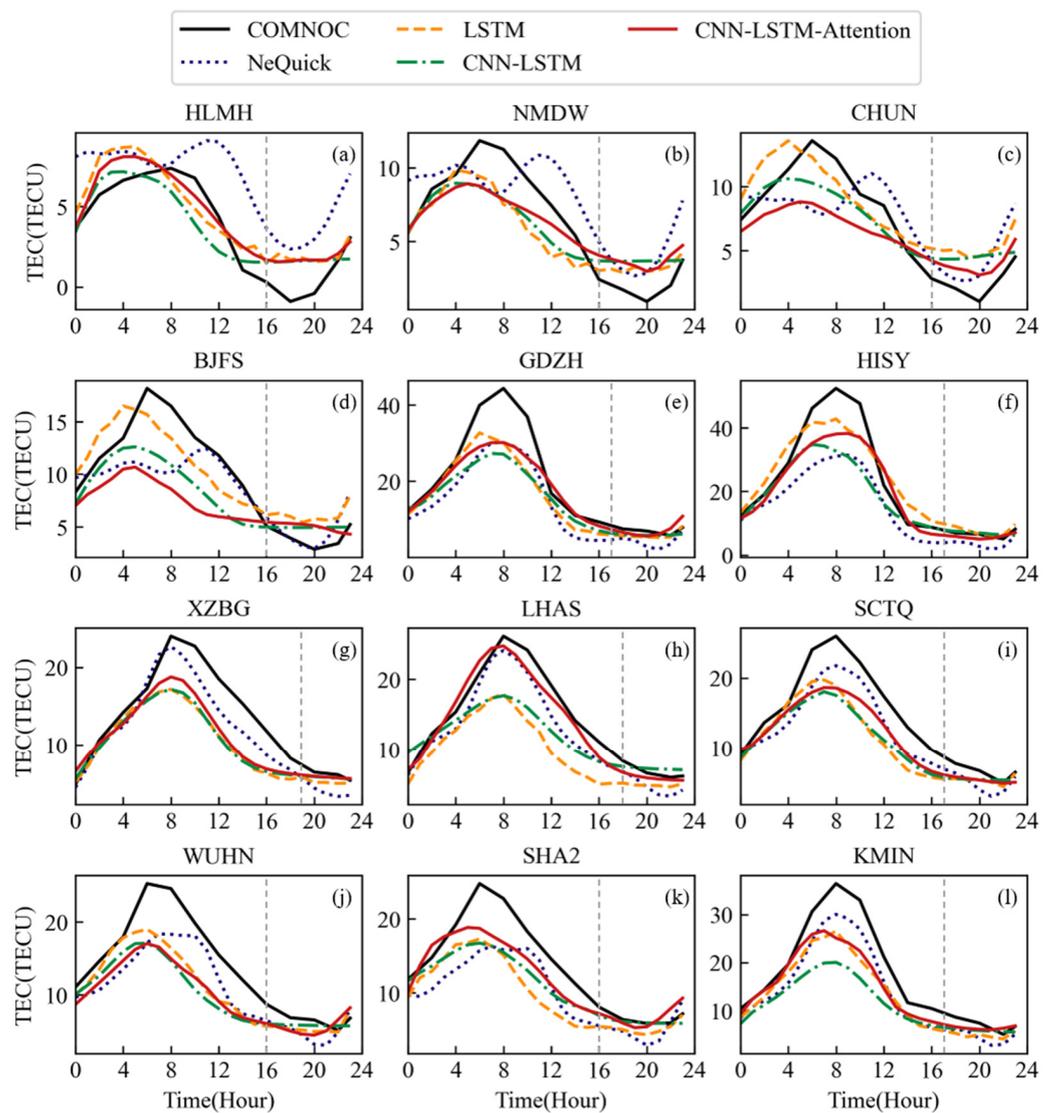


Figure 13. (a–l) Comparison of GNSS measured values (black solid line), NeQuick (blue dotted line), LSTM (yellow dotted line), CNN-LSTM (green dotted line), and CNN-LSTM-Attention (red solid line) models forecasted values for 12 stations on magnetic storm day. The gray dotted line indicates midnight LT.

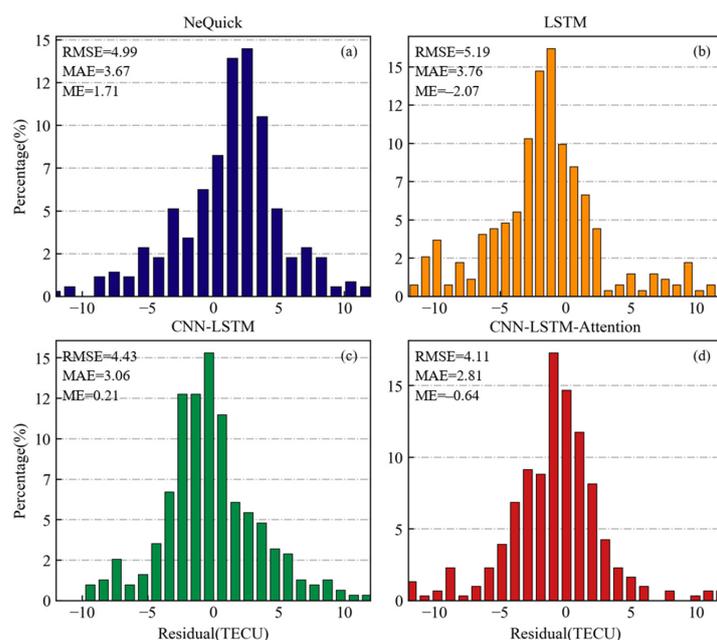


Figure 14. Residual distribution of different models for GDZH station during magnetic storm period. RMSE, MAE, and ME are also included. (a) NeQuick; (b) LSTM; (c) CNN-LSTM; (d) CNN-LSTM-Attention model.

5. Discussion

Deep learning approach has the possibility to find nonlinear functions to forecast their manifestation in the ionosphere. In this paper, a CNN-LSTM neural network model integrating attention mechanism was developed for ionospheric TEC prediction. The evaluation was carried out by using the observation data of 24 GNSS stations from the CMONOC. The observation data from 2010 to 2017 and the five parameters of B_z , K_p , Dst , $F10.7$, and daily hours were used as the training set. The 2018 data were used as the test set to verify the model's performance. We compare the CNN-LSTM-Attention model with NeQuick, LSTM, and CNN-LSTM models to analyze the model's forecasting performance under different conditions. With the decrease in latitude, the forecasted error of the CNN-LSTM-Attention model gradually increases from 1 TECU to 4 TECU. In general, the forecasted errors at the same latitude are the same, which are much lower than that of the empirical model. The monthly average value of TEC increases gradually from January to April, reaches a peak in April, and then gradually decreases. However, the prediction error of the model in this paper does not increase with the increase in monthly average, and it is generally stable at 1.5–2.2 TECU. This shows that the CNN-LSTM-Attention model is stable at different times.

Under different geomagnetic activities, the NeQuick model has RMSEs of 2.88 and 4.62 TECU with correlation coefficients of 0.83 and 0.90 during the magnetic quiet and magnetic storm periods, respectively. The predicted and measured RMSEs for the magnetic quiet and magnetic storm periods are 1.92 and 3.97 TECU, respectively, which are 33.33% and 14.07% lower than those of the empirical model. The correlation coefficients of magnetic quiet and storm periods both are 0.92, which improve 10.84% and 2.22% compared with the empirical model, and the forecasted values are highly correlated with the measured values. The RMSE for the storm period is up to two times higher than for the quiet period. Ensemble learning methods achieve better accuracy than a single learning method, where combining multiple ensembles provides the most optimal results.

6. Conclusions

In this study, an ionospheric forecast model in China is established based on 24 GNSS stations from CMONOC. Compared with empirical models, deep learning model have

higher prediction accuracy, especially during magnetic storms. In this paper, the attention mechanism is added to the LSTM neural network, which inherits the high precision of the deep learning model and effectively improves stability. It shows good performance in different time periods throughout a year in China. The RSME of the CNN-LSTM-Attention model test set for the entire year is 1.87 TECU. As a comparison, the RMSEs of the NeQuick, LSTM, and CNN-LSTM models are 3.59, 2.25, and 2.07 TECU, respectively. The proposed model performs well in both in quiet periods and storm periods of geomagnetic activity. During the quiet period and storm period, the percentages of forecasted residuals greater than 4 TECU are only 4% and 6%, and the percentages of residuals greater than 4 TECU for the NeQuick model are 19% and 18%. However, the time period selected (2010–2018) is limited by the availability of data from the network. Currently, it should be noted that our experiments only discuss a relatively quiet solar period. We acknowledge the increasing solar activity levels and suggest that measurements in the coming years, where activity is expected to be much stronger, will provide a good opportunity for continued model development.

Author Contributions: J.T. and Y.L. designed and performed the experiment; J.T. and Y.L. wrote the paper; M.D. and H.L. analyzed the data; D.Y. drew some of the figures; J.T. and X.W. revised the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the National Key Research & Development Program (No. 2017YFE0131400), the Key Laboratory for Digital Land and Resources of Jiangxi Province, East China University of Technology (No. DLLJ202104), the National Natural Science Foundation of China (No. 42074045), and the National Natural Science Foundation of China (No. 41761089).

Data Availability Statement: OMNI data are obtained from the GSFC/SPDF OMNIWeb interface (<http://omniweb.gsfc.nasa.gov>, accessed on 6 January 2022). NeQuick data are from the Abdus Salam International Centre for Theoretical Physics (ICTP) (<https://t-ict4d.ictp.it/nequick2>, accessed on 6 January 2022).

Acknowledgments: The authors acknowledge the Crustal Movement Observation Network of China (CMONOC) for providing GNSS data, the OMNIWeb for providing OMNI data, and the Abdus Salam International Centre for Theoretical Physics (ICTP) for providing NeQuick data. They would also like to thank the anonymous reviewers for their constructive comments in improving the article.

Conflicts of Interest: The authors declare conflict of interest.

Appendix A

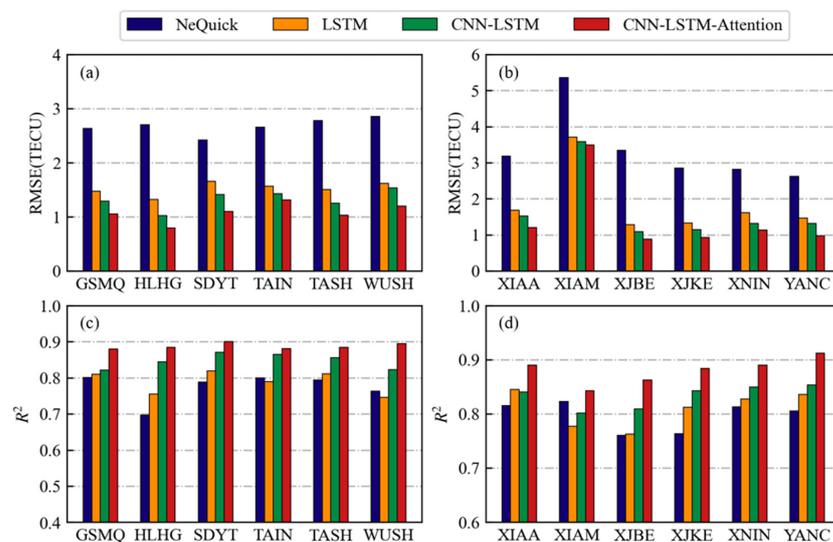


Figure A1. Histogram for RMSE and R^2 of NeQuick (blue), LSTM (yellow), CNN-LSTM (green), and CNN-LSTM-Attention (red) models from other 12 GNSS stations. (a,b) are RMSE. (c,d) are R^2 .

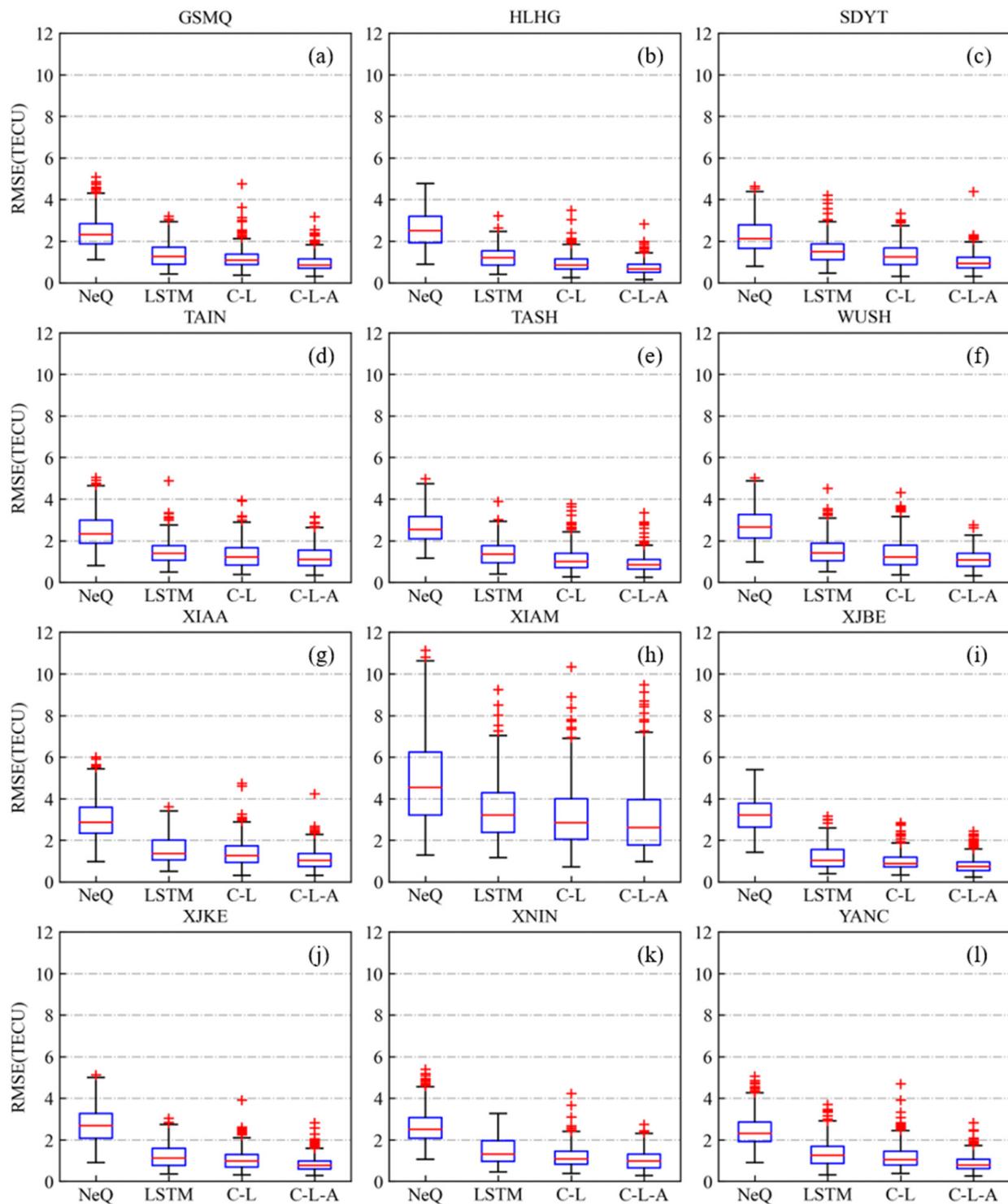


Figure A2. (a–l) The RMSE of NeQuick, LSTM, CNN-LSTM, and CNN-LSTM-Attention models from other 12 GNSS stations. The top and bottom black bars represent the maximum and minimum values, respectively. The upper and lower edges of the blue box represent the upper quartile and the lower quartile, respectively. The middle red bar represents the median, and the red plus sign represents the outlier.

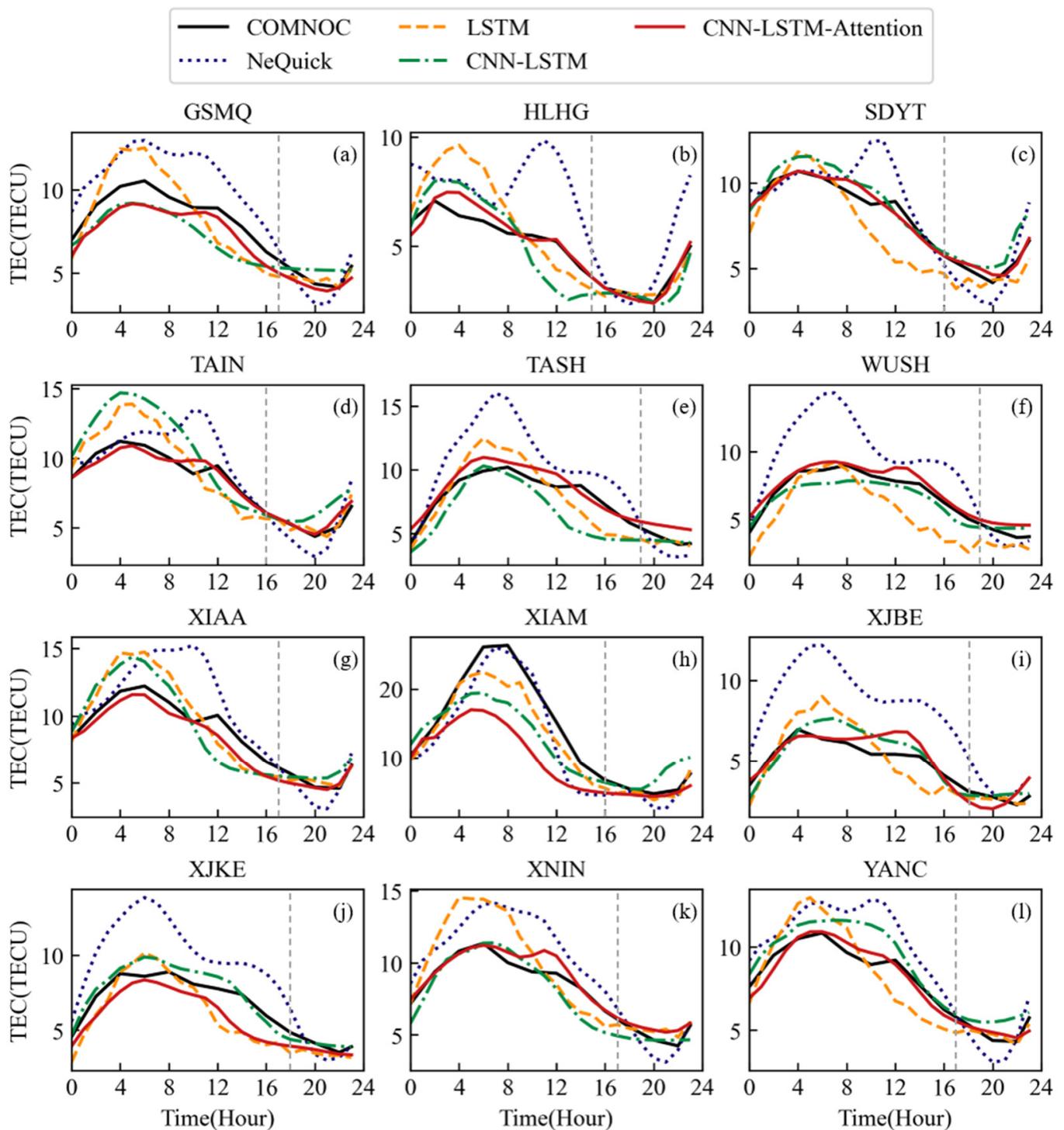


Figure A3. (a–l) Comparison of GNSS measured values (black solid line), NeQuick (blue dotted line), LSTM (yellow dotted line), CNN-LSTM (green dotted line), and CNN-LSTM-Attention (red solid line) models forecasted values for other 12 GNSS stations on magnetic quiet days. The gray dotted line indicates midnight LT.

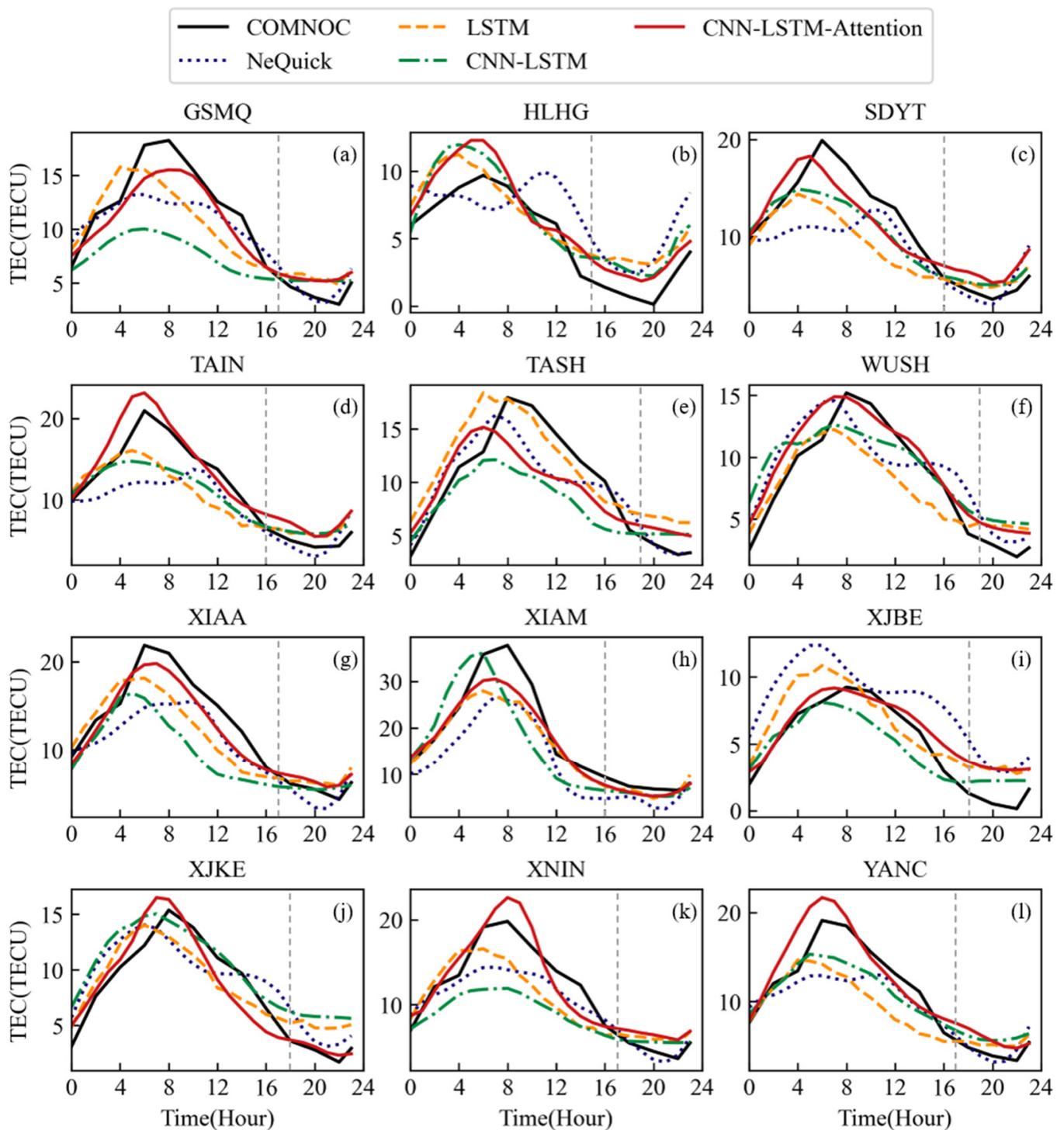


Figure A4. (a–l) Comparison of GNSS measured values (black solid line), NeQuick (blue dotted line), LSTM (yellow dotted line), CNN-LSTM (green dotted line), and CNN-LSTM-Attention (red solid line) models forecasted values for other 12 GNSS stations on magnetic storm day. The gray dotted line indicates midnight LT.

References

- Themens, D.R.; Reid, B.; Jayachandran, P.T.; Larson, B.; Koustov, A.V.; Elvidge, S.; McCaffrey, A.M.; Watson, C. E-CHAIM as a model of total electron content: Performance and diagnostics. *Space Weather* **2021**, *19*, e2021SW002872. [[CrossRef](#)]
- Francis, N.M.; Cannon, P.S.; Brown, A.G.; Broomhead, D.S. Nonlinear prediction of the ionospheric parameter foF2 on hourly, daily, and monthly timescales. *J. Geophys. Res.* **2000**, *105*, 12839–12849. [[CrossRef](#)]

3. Francis, N.M.; Brown, A.G.; Cannon, P.S.; Broomhead, D.S. Prediction of the hourly ionospheric parameter foF2 using a novel nonlinear interpolation technique to cope with missing data points. *J. Geophys. Res.* **2001**, *106*, 30077–30083. [[CrossRef](#)]
4. Tulunay, E.; Senalp, E.T.; Radicella, S.M.; Tulunay, Y. Forecasting total electron content maps by neural network technique. *Radio Sci.* **2006**, *41*, RS4016. [[CrossRef](#)]
5. Habarulema, J.B.; McKinnell, L.A.; Opperman, B.D. Towards a GPS-based TEC prediction model for Southern Africa with feed forward networks. *Adv. Space Res.* **2009**, *44*, 82–92. [[CrossRef](#)]
6. Habarulema, J.B.; McKinnell, L.-A.; Opperman, B.D.L. Regional GPS TEC modeling; Attempted spatial and temporal extrapolation of TEC using neural networks. *J. Geophys. Res.* **2011**, *116*, A04314. [[CrossRef](#)]
7. Huang, L.; Wang, J.; Jiang, Y.; Huang, J.; Chen, Z.; Zhao, K. A preliminary study of the single crest phenomenon in total electron content (TEC) in the equatorial anomaly region around 120°E longitude between 1999 and 2012. *Adv. Space Res.* **2014**, *54*, 2200–2207. [[CrossRef](#)]
8. Ferreira, A.A.; Borges, R.A.; Paparini, C.; Ciralo, L.; Radicella, S.M. Short-term estimation of GNSS TEC using a neural network model in Brazil. *Adv. Space Res.* **2017**, *60*, 1765–1776. [[CrossRef](#)]
9. Bai, H.; Fu, H.; Wang, J.; Ma, K.; Wu, T.; Ma, J. A prediction model of ionospheric foF2 based on extreme learning machine. *Radio Sci.* **2018**, *53*, 1292–1301. [[CrossRef](#)]
10. Lee, S.; Ji, E.; Moon, Y.; Park, E. One-Day Forecasting of global TEC using a novel deep learning model. *Space Weather* **2021**, *19*, e2020SW002600. [[CrossRef](#)]
11. Wang, J.; Feng, F.; Ma, J. An adaptive forecasting method for ionospheric critical frequency of F2 layer. *Radio Sci.* **2020**, *55*, e2019RS007001. [[CrossRef](#)]
12. Adolfs, M.; Hoque, M.M. A neural network-based TEC model capable of reproducing nighttime winter anomaly. *Remote Sens.* **2021**, *13*, 4559. [[CrossRef](#)]
13. Razin, M.R.G.; Moradi, A.R.; Inyurt, S. Spatio-temporal analysis of TEC during solar activity periods using support vector machine. *GPS Solut.* **2021**, *25*, 121. [[CrossRef](#)]
14. Gampala, S.; Devanaboyina, V.R. Application of SST to forecast ionospheric delays using GPS observations. *IET Radar Sonar Navig.* **2017**, *11*, 1070–1080. [[CrossRef](#)]
15. Dabbakuti, J.R.K.K.; Peesapati, R.; Yarrakula, M.; Anumandla, K.K.; Madduri, S.V. Implementation of storm-time ionospheric forecasting algorithm using SSA-ANN model. *IET Radar Sonar Navig.* **2020**, *14*, 1249–1255. [[CrossRef](#)]
16. Dabbakuti, J.R.K.K.; Jacob, A.; Veeravalli, V.R.; Kallakunta, R.K. Implementation of IoT analytics ionospheric forecasting system based on machine learning and ThingSpeak. *IET Radar Sonar Navig.* **2020**, *14*, 341–347. [[CrossRef](#)]
17. Ruwali, A.; Kumar, A.J.S.; Prakash, K.B.; Sivavaraprasad, G.; Ratnam, D.V. Implementation of hybrid deep learning model (LSTM-CNN) for ionospheric TEC forecasting using GPS data. *IEEE Geosci. Remote Sens. Lett.* **2020**, *18*, 1004–1008. [[CrossRef](#)]
18. Xiong, P.; Zhai, D.; Long, C.; Zhou, H.; Zhang, X.; Shen, X. Long short-term memory neural network for ionospheric total electron content forecasting over China. *Space Weather* **2021**, *19*, e2020SW002706. [[CrossRef](#)]
19. Kim, J.; Kwak, Y.; Kim, Y.; Moon, S.; Jeong, S.; Yun, J. Potential of regional ionosphere prediction using a long short-term memory deep-learning algorithm specialized for geomagnetic storm period. *Space Weather* **2021**, *19*, e2021SW002741. [[CrossRef](#)]
20. Zewdie, G.K.; Valladares, C.; Cohen, M.B.; Lary, D.J.; Ramani, D.; Tsidu, G.M. Data-driven forecasting of low-latitude ionospheric total electron content using the random forest and LSTM machine learning methods. *Space Weather* **2021**, *19*, e2020SW002639. [[CrossRef](#)]
21. Chen, Z.; Liao, W.; Li, H.; Wang, J.; Deng, X.; Hong, S. Prediction of global ionospheric TEC based on deep learning. *Space Weather* **2022**, *20*, e2021SW002854. [[CrossRef](#)]
22. Liu, Y.; Zhang, Q.; Song, L.; Chen, Y. Attention-based recurrent neural networks for accurate short-term and long-term dissolved oxygen prediction. *Comput. Electron. Agric.* **2019**, *165*, 104964. [[CrossRef](#)]
23. Zhang, B.; Ou, J.; Yuan, Y.; Li, Z. Extraction of line-of-sight ionospheric observables from GPS data using precise point positioning. *Sci. China Earth Sci.* **2012**, *55*, 1919–1928. [[CrossRef](#)]
24. Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J.; et al. Recent advances in convolutional neural networks. *Pattern Recognit.* **2018**, *77*, 354–377. [[CrossRef](#)]
25. Gers, F.A.; Schmidhuber, J.; Cummins, F. Learning to forget: Continual prediction with LSTM. *Neural Comput.* **2000**, *12*, 2451–2471. [[CrossRef](#)] [[PubMed](#)]
26. Bahdanau, D.; Cho, K.; Bengio, Y. Neural machine translation by jointly learning to align and translate. *Comput. Sci.* **2016**. Available online: <https://arxiv.org/abs/1409.0473> (accessed on 18 January 2022).
27. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *Comput. Sci.* **2014**. Available online: <https://arxiv.org/abs/1412.6980> (accessed on 18 January 2022).
28. Nava, B.; Coisson, P.; Radicella, S. A new version of the NeQuick ionosphere electron density model. *J. Atmos. Sol. Terr. Phys.* **2008**, *70*, 1856–1862. [[CrossRef](#)]
29. Wen, Z.; Li, S.; Li, L.; Wu, B.; Fu, J. Ionospheric TEC prediction using long short-term memory deep learning network. *Astrophys. Space Sci.* **2021**, *366*, 3. [[CrossRef](#)]
30. Song, R.; Zhang, X.; Zhou, C.; Liu, J.; He, J. Predicting TEC in China based on the neural networks optimized by genetic algorithm. *Adv. Space Res.* **2018**, *62*, 745–759. [[CrossRef](#)]

31. Galav, P.; Rao, S.S.; Sharma, S.; Gordiyenko, G.; Pandey, R. Ionospheric response to the geomagnetic storm of 15 May 2005 over midlatitudes in the day and night sectors simultaneously. *J. Geophys. Res. Space Phys.* **2014**, *119*, 5020–5031. [[CrossRef](#)]
32. Tang, J.; Gao, X.; Li, Y.; Zhong, Z. Study of ionospheric responses over China during September 7–8, 2017 using GPS, Beidou (GEO), and Swarm satellite observations. *GPS Solut.* **2022**, *26*, 55. [[CrossRef](#)]