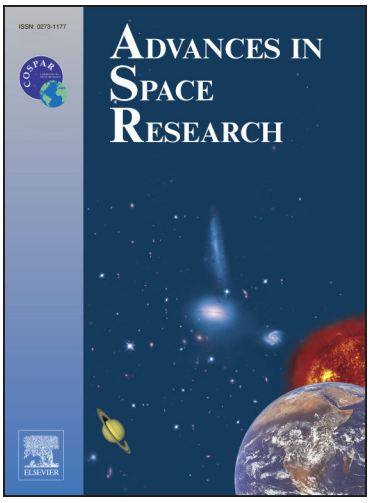
Journal Pre-proofs



A new high-precision short-term ionospheric TEC prediction model based on the DBO-BiLSTM algorithm: A case study of Europe

Qiaoli Kong, Yunqing Huang, Xiaolong Mi, Qi Bai, Jingwei Han, Yanfei Chen, Shi Wang

PII: S0273-1177(25)00230-3
DOI: https://doi.org/10.1016/j.asr.2025.03.012
Reference: JASR 18219

To appear in: *Advances in Space Research*

Received Date: 23 August 2024
Revised Date: 2 March 2025
Accepted Date: 5 March 2025

Please cite this article as: Kong, Q., Huang, Y., Mi, X., Bai, Q., Han, J., Chen, Y., Wang, S., A new high-precision short-term ionospheric TEC prediction model based on the DBO-BiLSTM algorithm: A case study of Europe, *Advances in Space Research* (2025), doi: https://doi.org/10.1016/j.asr.2025.03.012

A new high-precision short-term ionospheric TEC prediction model based on the DBO-BiLSTM algorithm: A case study of Europe

Qiaoli Kong[1*],Yunqing Huang[1],Xiaolong Mi[2],Qi Bai[1],Jingwei Han[1],Yanfei Chen[1],Shi Wang[1]

[1]College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao, China

[2]Department of Land Surveying and Geo-Informatics, Hong Kong Polytechnic University, Hong Kong, China

*Corresponding author: kqlabc3334@163.com

Contributing authors: huangyunqing0901@163.com; xiaolong.mi@polyu.edu.hk; bq974210@163.com; hanjingwei2021@163.com; cyf1217992408@163.com; ws71712@163.com

Abstract: In order to achieve high accuracy of ionospheric total electron content (TEC) short-term prediction for Europe, a hybrid novel deep learning model was established applying the dung beetle optimizer (DBO) algorithm to optimize the bidirectional long short-term memory (BiLSTM) neural network, named DBO-BiLSTM. For evaluating the TEC prediction accuracy of DBO-BiLSTM model, the TEC predicted by this model was compared with TEC computed using GPS observation released by the European Permanent Global Navigation Satellite System network (EPGNSS), and with those predicted by the sparrow search algorithm-based BiLSTM (SSA-BiLSTM), BiLSTM, and long short-term memory (LSTM) neural network models. The test results indicate that the predicted TEC by DBO-BiLSTM has the closest agreement with those solved by GPS data compared with those predicted by the other three models, and the prediction accuracy achieved by DBO-BiLSTM model is the highest with the root mean square error (RMSE) values of 1-h and 2-h predictions reaching 0.57 TECU and 0.92 TECU, respectively. What's more, the optimized hybrid DBO-BiLSTM model can effectively capture the ionospheric characteristics with the spatial-temperal changes, under quiet and moderate disturbed geomagnetic conditions, and during moderate solar activity period. This research provides a valuable hybrid DBO-BiLSTM model for high accuracy short-term prediction of ionospheric TEC for Europe, and gives an important reference for the further comprehensive TEC prediction under more sever disturbed geomagnetic conditions and more violent solar activity periods.

Keywords: Dung beetle optimization algorithm, neural networks, TEC short term prediction, total electron content

1. Introduction

The ionosphere is a part of the Earth's atmosphere that is ionised by solar ultraviolet rays and X-rays approximately within the range of 60–2000 km in latitude (Yuan et al., 2017; Tang et al., 2022). The large number of free electrons generated by ionisation not only can block the direct radiation of high-energy particles in the universe to the Earth, but also can reflect, refract, and absorb electromagnetic wave signals, and thus affect satellite navigation and communication, remote sensing monitoring and other applications (Yao and Gao, 2022; Tang et al., 2024). The degree of impact has direct relation with the total electron content (TEC) in the ionosphere. Therefore, the high-precision prediction of ionospheric TEC is of great significance for the propagation of space electromagnetic wave signals, satellite navigation and communication, and space weather research (Tang et al., 2023a).

The previous prediction models for ionospheric TEC were mainly based on empirical and time series models. Classic empirical models includs the Bent (1972), Klobuchar (1987), and NeQuick models (Giovanni, 1990). In terms of ionospheric prediction, empirical models are characterised by high calculation speeds, but some of them have low prediction accuracy and cannot accurately capture the time-varying characteristics of the ionosphere in local regions (Feng, 2019). Time series models are characterised by the small amounts of input required data and simple structures. Xie et al. (2017) carried out 6-day TEC prediction using Holt-Winters model, and the predicted results were in good agreement with the actual observation. Li et al. (2013) performed the 5-day TEC prediction using the auto-regressive moving average (ARMA) model, and achieved relative accuracies above 90%. Sivavaraprasad and Ratnam (2017) established an ionospheric delay prediction model for India using the ARMA model, and found that this approach could effectively predict the ionospheric delay with a high relative accuracy of 94%. However, with the extension of the forecast time, the prediction accuracy of these time series models decrease seriously, and none of the above models can take the influence of varieties of geomagnetic condition and solar activity into consideration.

In recent years, as the development of deep learning algorithms, neural network models have shown an excellent ability to fit nonlinear data, and increasing numbers of studies applied these models to predict ionospheric TEC. Silva et al. (2023) established a TEC prediction model for Brazil based on deep learning feedforward artificial neural networks (ANN). Li et al. (2021a) developed a three-dimensional ionospheric tomography model based on ANN with multi-source data. The recurrent neural network (RNN) can learn the nonlinear features of the time series data (Lin et al., 2022), therefore, Yuan et al. (2018) built an ionospheric TEC prediction model for single station based on RNN. However, RNN units suffers the gradient dispersion when processing long sequences of data (Hochreiter, 1998). Long short-term memory (LSTM) model can solve the problem of gradient explosion of RNN (Bengio et al., 1994). Xiong et al. (2022) constructed single-station and regional ionospheric TEC prediction models based on the LSTM model, and found that the LSTM model had better performance than other model did. However, LSTM cannot exactly capture the dependencies between the former and later data in the time series. In contrast, BiLSTM can solve the problem by simultaneously inputting sequences in both the forward and backward directions. Shi et al. (2022) achieved short-term prediction of

regional TEC in China using four models, International Reference Ionosphere (IRI-2016), artificial neural network (ANN), LSTM, and BiLSTM, and found that BiLSTM owned the best prediction performance.

When dealing with large amounts of data, neural networks faces the challenges such as low search speeds and easy falling into local optimization. While, by integrating neural network with the optimisation algorithms of strong stability and fast convergence, the weights and thresholds of the neurons can be effectively adjusted to achieve the optimal solution and solve the potential local optimisation problems. Particle swarm optimisation (PSO) and the sparrow search algorithm (SSA) have been used to optimise an Elman neural network for ionospheric TEC prediction (Tang, 2022; Tang, 2024). The genetic algorithm (GA) has been used to optimise a back propagation (BP) algorithm to establish a TEC prediction model (Jiang et al., 2023), and the dingo optimisation algorithm (DOA) has been applied to optimise BP neural network to achieve higher TEC prediction accuracy (Ni et al., 2024). Xue and Shen (2022) proposed a new optimisation algorithm called the dung beetle optimiser (DBO). Duan et al. (2023) applied the DBO algorithm to optimise a combined model involving convolutional neural networks (CNNs) and LSTM for air quality prediction, and found that the prediction results were better than those predicted by the original combined CNN-LSTM model. Li et al. (2024) used the whale optimization algorithm (WOA) to optimise the combined model of convolutional neural networks (CNNs) and LSTM for TEC prediction, and the prediction results owned higher accuracy compared to those predicted by CNN- Gated Recurrent Unit (GRU), BiLSTM, and RNN.

Based on the above analysis, this study propose to construct a novel hybrid DBO-BiLSTM model to perform high accuracy short-term TEC prediction for Europe, applying the great optimizing ability of DBO to break through the limitations of low search speed and easy falling into local optimization of the traditional neural networks. The performance of DBO-BiLSTM was comprehensively investigated using TEC computed by GPS data, predicted results by SSA-BiLSTM, BiLSTM, and LSTM models in different latitude bands, at different time points, under different geomagnetic conditions and in moderate solar activity period. The contents of this study are arranged as follows. Section 2 focuses on the theory and methodology of TEC calculation and construction of prediction model. Section 3 introduces the multi-source data used for TEC prediction. Section 4 analyses and compares the performances of the established TEC prediction models. Section 5 summarizes and analyses the contributions of this research.

2. Theory and Methodology

2.1 Calculation of TEC

The ionospheric TEC is computed using the GPS dual-frequency carrier phase data smoothed by pseudo-range, and can be expressed as (Mannucci et al., 1998):

$$STEC = \frac{f_1^2 f_2^2}{40.3(f_1^2 - f_2^2)} \left\{ \left[(L_1 - L_2) + \frac{1}{N} \sum_{n=1}^{N} (P_2 - P_1) - (L_1 - L_2) \right] + c(d_{DCB}^i + d_{DCB,j}) \right\} \tag{1}$$

where the slant TEC (STEC) is the TEC along the slant path, f_1 and f_2 is the frequencies 1 and 2 of the carrier phase, respectively, N is the number of epochs, P_1 and P_2 are the pseudo-range observations for frequencies 1 and 2, respectively, L_1 and L_2 are the carrier phase observations of frequencies 1 and 2, respectively, d_{DCB}^i and $d_{\mathrm{DCB},j}$ are the differential code bias of the satellite and receiver, respectively, and c is the speed of light.

The vertical TEC (VTEC) is applied to avoid the impact of the susceptibility of STEC to the signal propagation path and satellite orbit inclination. VTEC can be obtained by STEC (Schaer, 1999) as shown in Equation (2):

$$VTEC = \frac{STEC}{MF(z)} \tag{2}$$

$$MF(z) = \frac{1}{\cos z'} = \frac{1}{\sqrt{1 - \left(\frac{R}{R + H_{opt}} \sin(\alpha z) \right)^2}} \tag{3}$$

where z is the zenith distance of the receiver, z' is the zenith distance of the ionospheric pierce point (IPP), R is the radius of the Earth (6371 km), H_{opt} is the height of the ionosphere thin layer, and α is the zenith distance coefficient. In this study, we set H_{opt} =450 km and α = 0.9782.

Based on formula (2), the VTEC of all grid nodes can be obtained using the spherical harmonic function with 6 orders and 4 degrees.

2.2 Construction of the TEC Prediction Model

Neural network algorithms are widely applied in various fields, such as classification, image recognition, speech recognition, linear/nonlinear data fitting, and time series data prediction (Li et al., 2021b). In this research, a multiple regression model is established using a single-layer neural network to predict ionospheric TEC. LSTM is a widely used

gated RNN. It can transfer relevant information to long sequences for prediction, and can solve the problems of short-term memory and gradient disappearance of the traditional RNN (Bengio et al., 1994). BiLSTM is an extension model of LSTM, in which an additional reverse loop layer is introduced to enable the LSTM to effectively abstract and capture the forward and backward association characteristics of the TEC time series, thus making the model more robust (Tang et al., 2023). This study sets the BiLSTM neurons to 200, the maximum training period to 50, the initial learning rate to 0.01, and the number of regularization to 0.001. The structure of the BiLSTM model constructed in this study is shown in Figure 1.

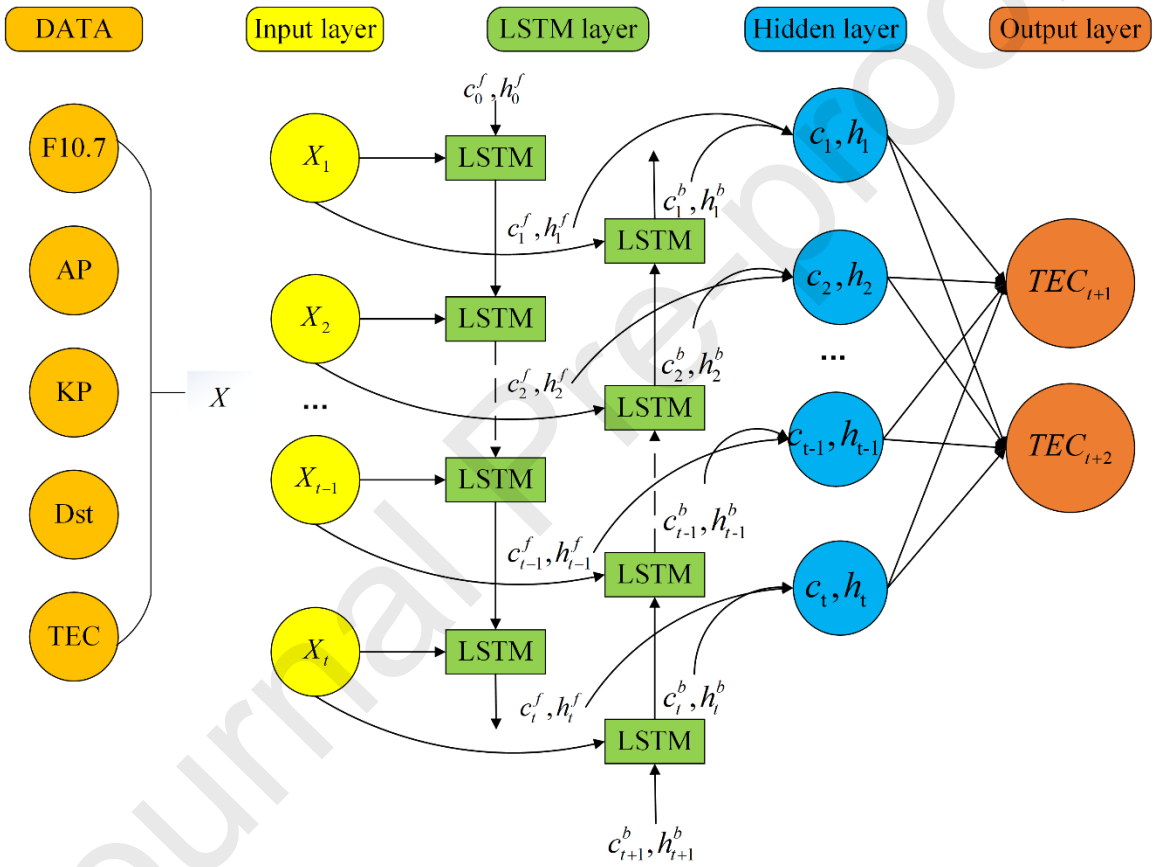


Figure 1. The BiLSTM model structure

In Figure 1, X_t is the independent variable in the *t*-th epoch, and *h* and *c* represent the temporary results and storage units, respectively. In this study, the root mean square error (RMSE) is used as a loss function for the LSTM model and the Adam optimiser is applied to the BiLSTM model.

DBO is a new population intelligence algorithm that is characterised by its great optimisation ability and high convergence speed (Xue and Shen, 2022). The DBO algorithm used in this study is based on four main processes: rolling, breeding, foraging, and stealing, as described below.

(1) Rolling: We assume that the light intensity affects the position of the dung beetle and there are no obstacles. The formula used to update the positions of the dung beetles is as follows:

$$\begin{aligned} x_i(\mathrm{t}+1) &= x_i(\mathrm{t}) + \alpha \times k \times x_i(\mathrm{t}-1) + b \times \Delta x \\ \Delta x &= |\, x_i(\mathrm{t}) - X^W \,| \end{aligned} \tag{4}$$

where t is the current iteration number, $x_i(\mathrm{t})$ is the position information of the *i*-th dung beetle at the *t*-th iteration, α is a natural coefficient with a value of −1 or 1, $k \in (0, 0.2]$ is the constant of the deflection coefficient, b is a constant with a value of (0,1), X^W is the global worst position, and Δx is used to simulate changes of light intensity.

(2) Breeding: the DBO algorithm adopts an edge selection strategy to simulate the spawning area of dung beetles, which is defined as:

$$\begin{aligned} Lb^* &= \max(\mathrm{X}^* \times (1-R), Lb) \\ Ub^* &= \min(\mathrm{X}^* \times (1+R), Ub) \end{aligned} \tag{5}$$

where X^* is the current local optimal position, Lb^* and Ub^* are the lower and upper limits of the spawning area, respectively, $R = 1 - \frac{t}{T_{\max}}$, $T_{\max}$ is the maximum number of iterations, and Lb and Ub are the lower and upper limits of the optimisation problem, respectively.

(3) Foraging: When dung beetles forage, the boundary of the optimal foraging area is determined based on the changes of their positions during the foraging process, and is defined as:

$$\begin{aligned} Lb^b &= \max(X^b \times (1-R), Lb) \\ Ub^b &= \min(X^b \times (1+R), Ub) \end{aligned} \tag{6}$$

where X^b is the global optimal position, and Lb^b and Ub^b are the upper and lower bounds in the optimal foraging area, respectively. The updated position of the dung beetle

is:

$$x_i(\mathrm{t}+1) = x_i(\mathrm{t}) + C_1 \times (x_i(\mathrm{t})\text{-}Lb^b) + C_2 \times (x_i(\mathrm{t})\text{-}Ub^b) \quad (7)$$

where $x_i(\mathrm{t})$ is the position of the *i*-th dung beetle in the *t*-th iteration, C_1 is a random number following the normal distribution, and C_2 is a random vector in the range (0,1).

(4) Stealing: When a dung beetle steals, the location of the thief beetle is updated as follows:

$$x_i(\mathrm{t}+1) = X^b + S \times g \times (|x_i(\mathrm{t}) - X^*| + |x_i(\mathrm{t}) - X^b|) \quad (8)$$

where $x_i(\mathrm{t})$ is the position of the *i*-th thief beetle in the *t*-th iteration, X^b is the global optimal position, *g* is a random vector of size $1 \times D$ following the normal distribution, and S is a constant.

In order to fully take the advantage of fast convergence of the DBO algorithm, the DBO-BiLSTM neural network model is established by combining BiLSTM model. The maximum number of iterations of DBO algorithm is set to 20 and the number of dung beetle population is set to 30. The optimal parameters can be obtained in a short time by this parameter setting. The range of BiLSTM neuron number is from 50 to 300, that of the initial learning rate is from 0.001 to 0.01, and that of the maximum training period is from 50 to 80.

The detailed steps of TEC prediction using the DBO-BILSTM model are as follows:

Step 1: TEC data is combined with KP, AP, Dst, F10.7 index of one-hour interval resolution to form the data set, and 80% of the data set is used for model training and 20% for testing.

Step 2: The initial hyperparameters including learning rate, number of neurons and training period are set for BiLSTM model.

Step 3: The hyperparameter matrix is composed and input into the DBO algorithm, and then the position and fitness of dung beetles is initialized and calculated.

Step 4: According to the total population number of dung beetles, the population number is allocated, and according to different population location update strategies, the location is updated.

Step 5: After the location update of all dung beetle populations, the optimal hyperparameter combination is calculated and sorted according to the optimal degree from top to bottom, and the hyperparameters exceeding the set range will be deleted.

Step 6: Repeat step 5 until the maximum number of iterations is reached, the optimal combination of hyperparameters is obtained, and BiLSTM training is finished.

Step 7: BiLSTM model is established with the optimal hyperparameters from step 6 and TEC prediction results achieved.

The TEC prediction process of the DBO-BiLSTM model is shown in Figure 2.

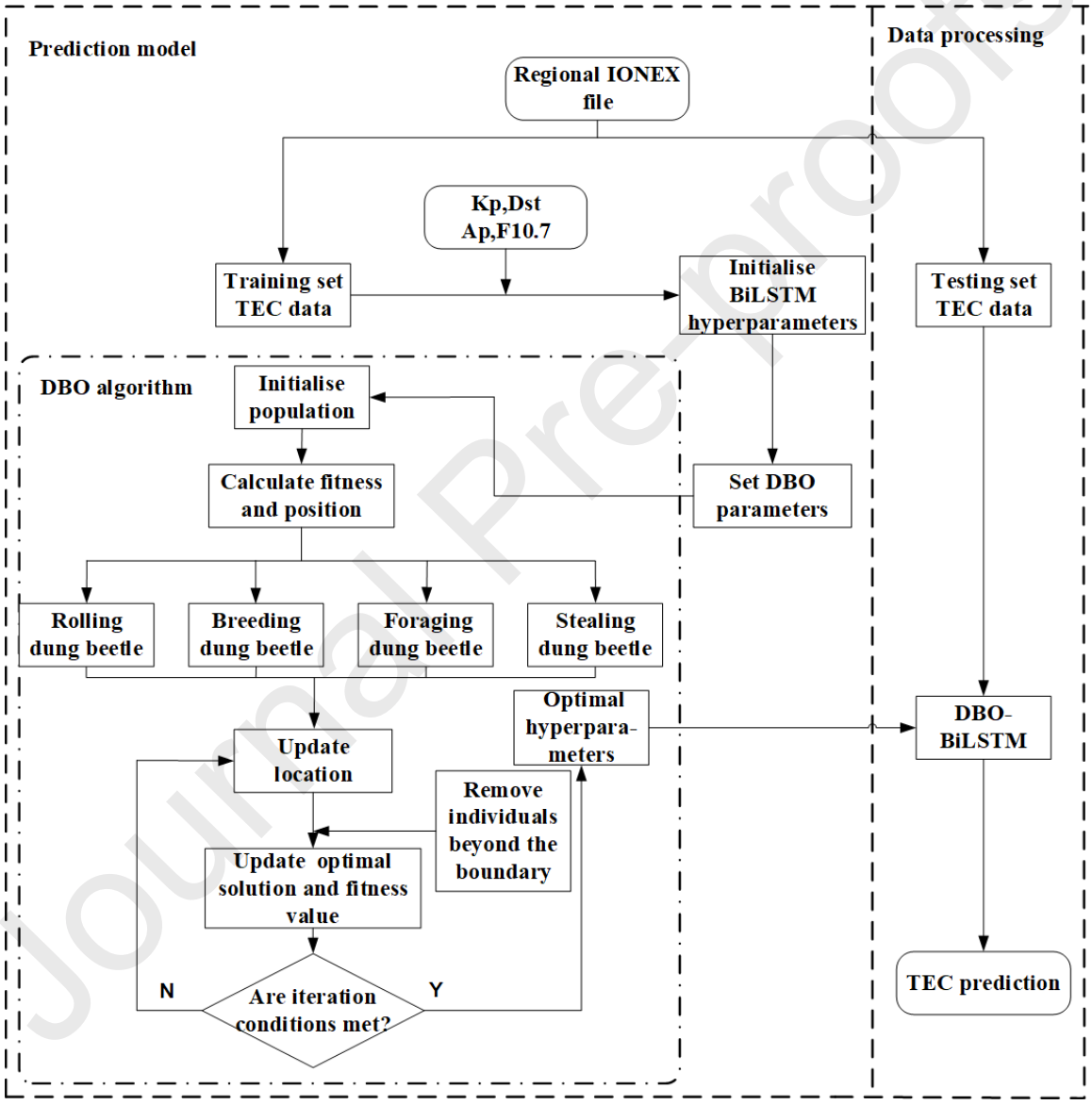


Figure 2. Ionospheric TEC prediction process of the DBO-BiLSTM model

3. Data Source

The European Permanent Global Navigation Satellite System Network is a voluntary consortium of over 100 self-funded institutions, universities, and research institutions from over 30 European countries. In 2022, this network consisted of 360 continuous tracking GNSS stations, as shown in Figure 3, and provided large amounts of high-precision, high-stability, and real-time GNSS observation data. These observations are widely used for monitoring surface deformation, sea level changes and space weather research. In this study, GPS data of 360 GNSS stations from 1 February 2022 to 31 August 2022 were used for the TEC calculations.

The Ap, Kp, and Dst indices are important indicators for measuring the degree of geomagnetic activity, and the solar radio flux index F10.7 is an important indicator for characterising the degree of solar activity. These four indices together reflect the extent of the space weather disturbances which can induces TEC changes effectively. Therefore, in this study, these four indices are used as variables in deep learning models to establish the relationship between ionospheric TEC and space weather environment changes.

German Researth Center For Geosciences (GFZ) provides Kp and Ap with time resolution of 3 h, F10.7 and Dst with time resolution of 1 d and 1 h, respectively. The time resolution of TEC computed by GPS data used in this study is 1h. In order to unite the time resolution to 1 h, the values of Ap for each hour during 3 hours are set to the same value with that provided by GFZ, and the value of F10.7 of each hour is set to the same value of that day. Similarly, the values of Kp for each hour are achieved by the same way of Ap. In the DBO-BiLSTM model, 24-h values of the Ap, Kp, Dst, F10.7 and TEC at the same point are used as independent variables for the 1-h or 2-h TEC prediction.

In this research, the TEC dataset of 187 ionospheric grid points in Europe is calculated from February 1 to August 31, 2022. The training dataset covers the period from February 1 to July 20, 2022, and the test dataset from July 21 to August 31, 2022. Setting Kp>3 as the geomagnetic disturbance, the data under the quiet geomagnetic condition accounts for 93% and the one under the disturbed condition accounts for 7%.

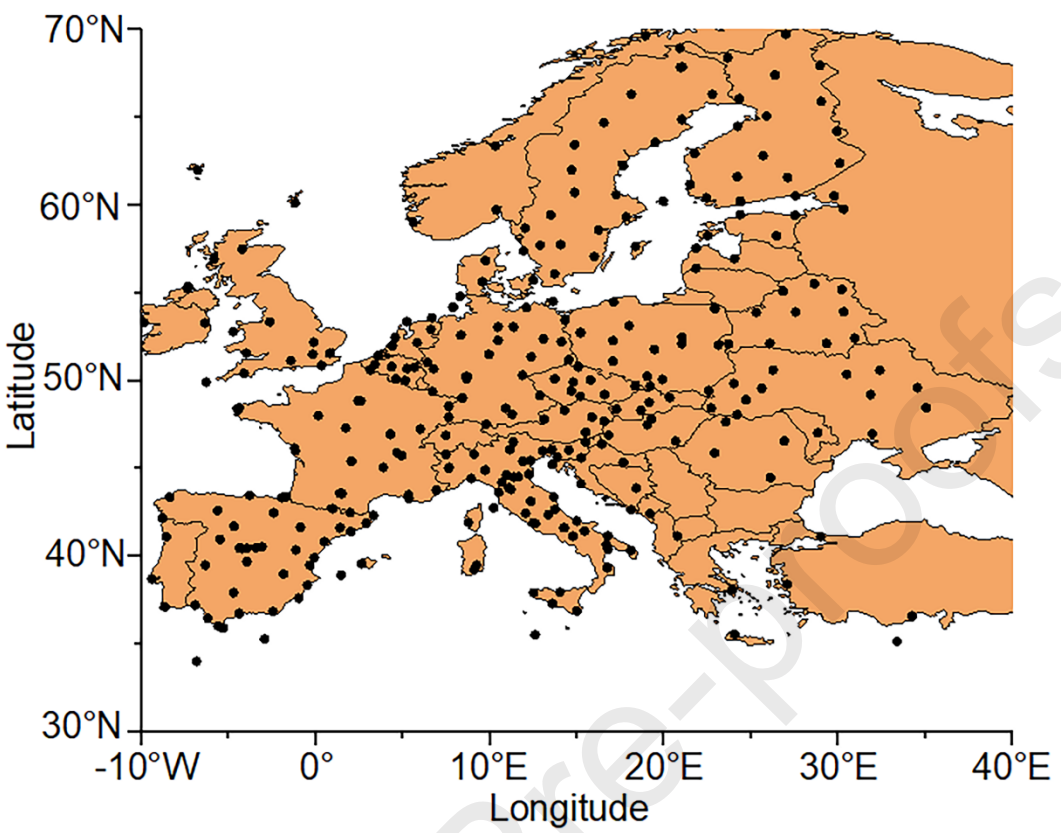


Figure 3. Distribution of GNSS stations of the European Permanent Global Navigation Satellite System Network

4. Analysis of Experimental Results

The TEC prediction accuracy of the DBO-BiLSTM model was compared with those predicted by the SSA-BiLSTM, BiLSTM, and LSTM models. The root mean square error (RMSE), fitting coefficient (R^2), Pearson correlation coefficient (Corr) and mean absolute error (MAE) defined as follows were used to evaluate the accuracy of the TEC prediction results.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(\mathrm{X}_{s_i} - \mathrm{X}_{p_i})^2}{N}} \tag{9}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(\mathrm{X}_{p_i} - \mathrm{X}_{s_i})^2}{\sum_{i=1}^{N}(\mathrm{X}_{s_i} - \bar{\mathrm{X}}_{s_i})^2} \tag{10}$$

$$Corr = \frac{\sum_{i=1}^{N}(\mathrm{X}_{s_i} - \bar{\mathrm{X}}_{s_i})(\mathrm{X}_{p_i} - \bar{\mathrm{X}}_{p_i})}{\sqrt{\sum_{i=1}^{N}(\mathrm{X}_{s_i} - \bar{\mathrm{X}}_{s_i})^2 \sum_{i=1}^{N}(\mathrm{X}_{p_i} - \bar{\mathrm{X}}_{p_i})^2}} \tag{11}$$

$$MAE = \frac{\sum_{i=1}^{N}\left|X_{s_i} - X_{p_i}\right|}{N} \tag{12}$$

where X_{s_i} is the TEC value calculated by GPS data, X_{p_i} is the predicted ionospheric TEC, and $\bar{\mathrm{X}}_{s_i}$ and $\bar{\mathrm{X}}_{p_i}$ are the mean values of TEC computed by GPS data and the mean predicted values, respectively.

4.1 Analysis of the Overall Prediction Accuracy of the Model

(1) Correlation Analysis of Prediction Results

The correlation between the calculated TEC test dataset by GPS data and those predicted for 1-h and 2-h by the DBO-BiLSTM, SSA BiLSTM, BiLSTM, and LSTM models are shown in Figure 4. In each subgraph, the black line is a linear function obtained by fitting the TEC, and three statistical indicators are RMSE, R^2, and Corr.

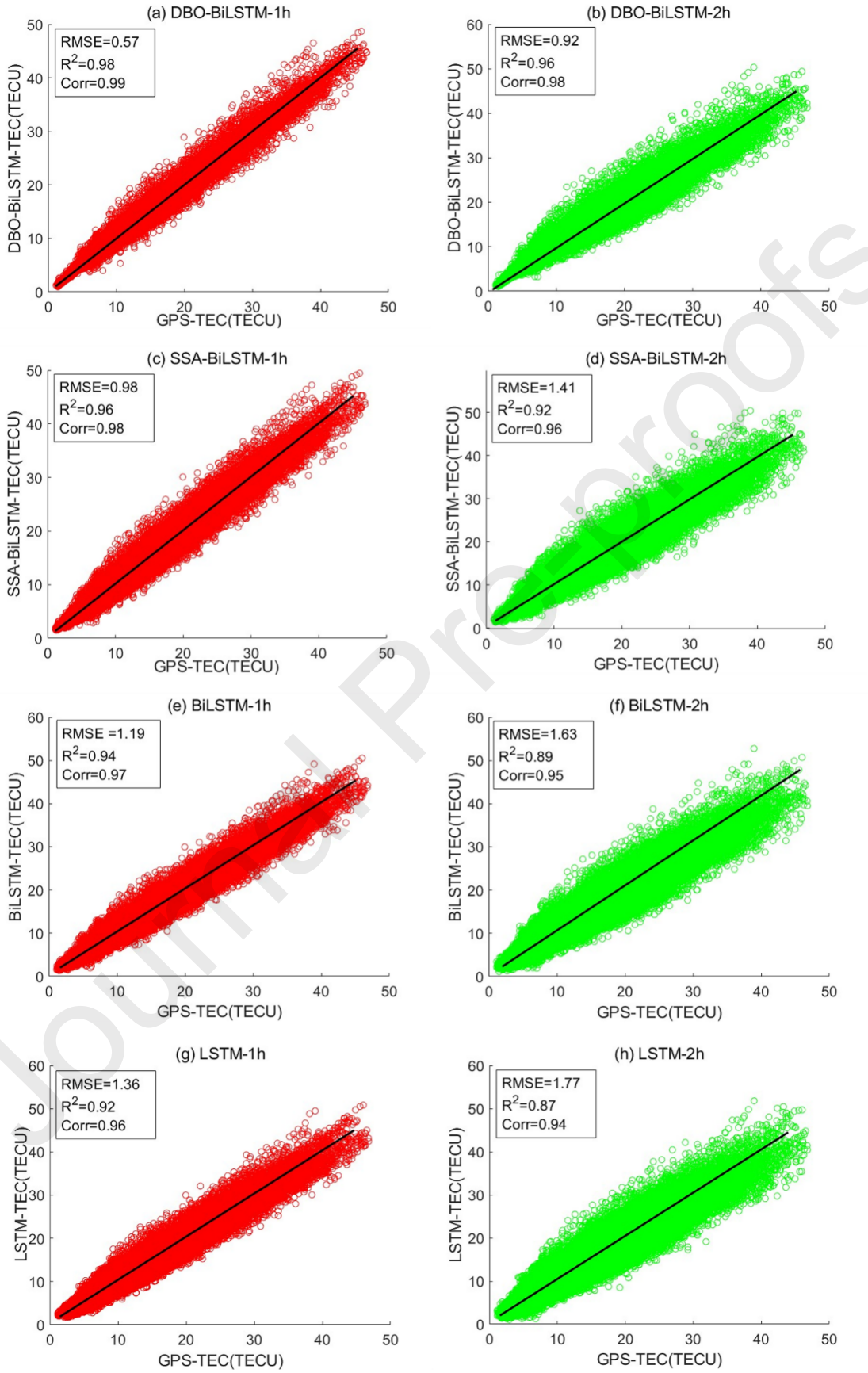


Figure 4. The correlations between the TEC estimated using GPS data and the predicted ones for 1-h and 2-h by DBO-BiLSTM, SSA BiLSTM, BiLSTM, and LSTM models

From the values of RMSE, R^2 and Corr in Figure 4, it can be noticed that the predicted TEC in the 1-h and 2-h by the four models achieves close correlation with the estimated one using GPS data. The DBO-BiLSTM model owns the highest consistency with the TEC estimated using GPS data, no matter for 1-h or 2-h, followed by SSA-BiLSTM, BiLSTM, and LSTM. From the scatter distributions and the statistical correlation indicators in Figure 4, it can be seen that the accuracy of the predicted results for 1 h are better than those for 2 h. Taking RMSE as an example, the RMSE values for the 1-h predicted results of the DBO-BiLSTM, SSA-BiLSTM, BiLSTM and LSTM models are reduced by 38%, 30%, 26% and 23% compared to those of the 2-h predicted results, respectively.

(2) Analysis of TEC Prediction Accuracy in Different Latitude Bands

TEC has a close relationship with latitude, therefore, in this study, the European is divided into four latitude bands of 30°N–40°N, 40°N–50°N, 50°N–60°N and 60°N–70°N. Tables 1 listed the statistics of the accuracy indices between the TEC of the test datasets estimated from GPS data and the ones predicted by the four models (DBO-BiLSTM, SSA-BiLSTM, BiLSTM and LSTM) in these four latitude bands for 1-h.

Table 1. Accuracy indices statistics between the TEC values estimated using GPS data and the ones predicted by four models in different latitude bands for a 1-h period

Latitude range	Evaluation index	DBO-BiLSTM	SSA- BiLSTM	BiLSTM	LSTM
30°N–40°N	RMSE/MAE(TECU)	0.93/0.59	1.47/1.09	1.70/1.35	1.86/1.55
	R^2/Corr	0.98/0.99	0.96/0.98	0.94/0.97	0.93/0.97
40°N–50°N	RMSE/MAE(TECU)	0.54/0.36	0.99/0.73	1.21/0.97	1.38/1.18
	R^2/Corr	0.98/0.99	0.96/0.98	0.94/0.97	0.92/0.96

50°N–60°N	RMSE/MAE(TECU)	0.41/0.29	0.77/0.55	0.96/0.78	1.13/0.98
	R^2/Corr	0.99/0.99	0.97/0.98	0.95/0.98	0.93/0.97
60°N–70°N	RMSE/MAE(TECU)	0.36/0.27	0.66/0.47	0.85/0.68	1.01/0.88
	R^2/Corr	0.99/0.99	0.96/0.98	0.94/0.97	0.91/0.96

Tables 1 show that the overall prediction accuracy of the DBO-BILSTM model is higher than those of the other three models. For the 1-h prediction results of the 30°N–40°N band, the RMSE values for the DBO-BiLSTM, SSA-BiLSTM, BiLSTM, and LSTM models are 0.93, 1.47, 1.70, and 1.86 TECU, respectively. The corresponding correlation coefficients were 0.99, 0.98, 0.97, and 0.97, respectively. The RMSE values for the 1-h predictions of the DBO-BiLSTM model were reduced by 37%, 45% and 50%, respectively, compared with SSA-BiLSTM, BiLSTM, and LSTM models, respectively. Similar performance can also be seen in the other latitude bands and in the 2-h predicted results of the four models (listed in the attachment).

4.2 Analysis of TEC Prediction Accuracy with Local Time variation

(1) Analysis of Average Daily Change in TEC with Local Time variation

In order to comprehensively investigate the characteristics of ionospheric TEC changes in the whole European region, nine uniformly distributed points were studied for the more detailed analysis, namely P_1 (30°N, 0°), P_2 (30°N, 15°E), P_3 (30°N, 30°E), P_4 (50°N, 0°), P_5 (50°N, 15°E), P_6 (50°N, 30°E), P_7 (70°N, 0°), P_8 (70°N, 15°E), and P_9 (70°N, 30°E). Figure 5 shows the average daily changes of TEC and those of the 1-h predicted TEC by the DBO-BiLSTM, SSA-BiLSTM, BiLSTM, and LSTM models (circle lines), and the RMSE values (columns) for each hour of the day.

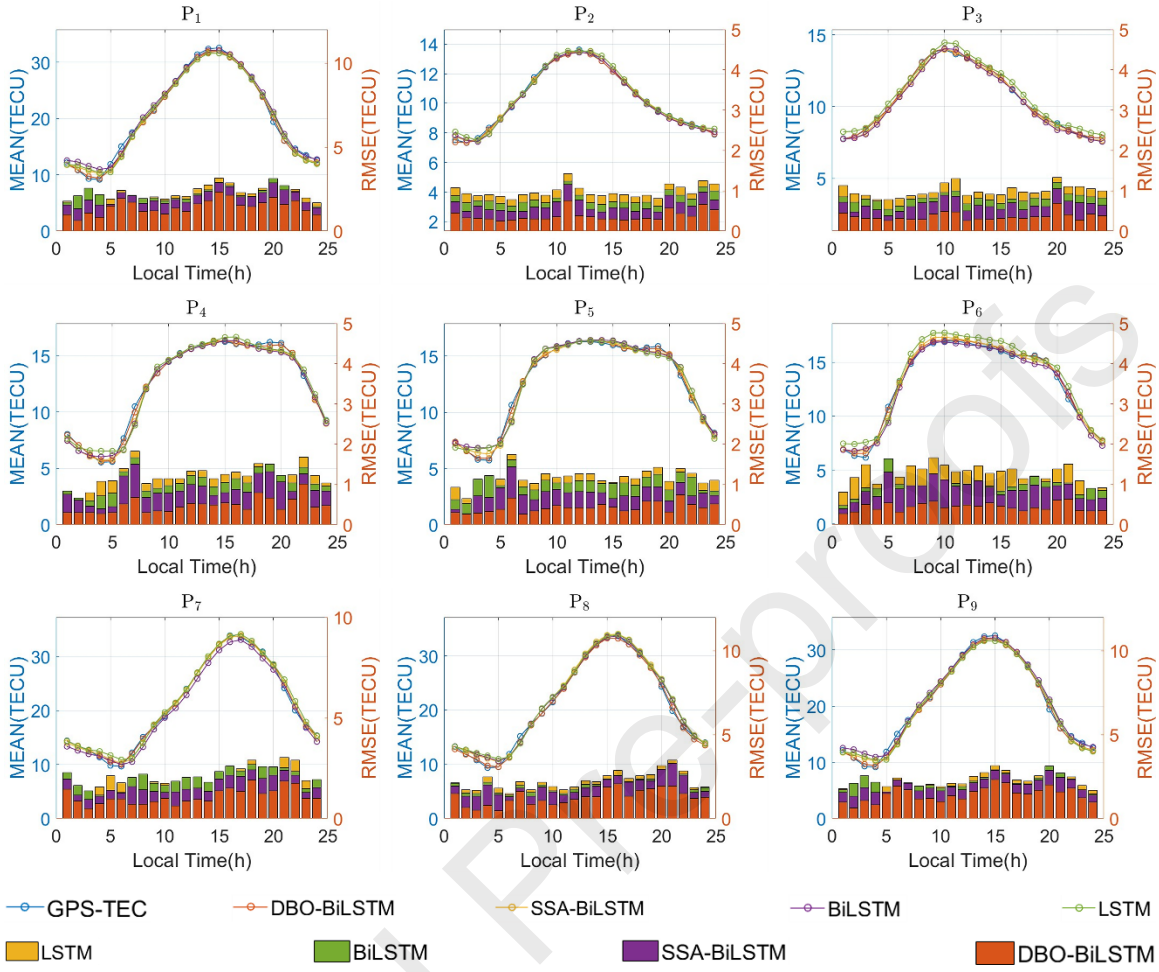


Figure 5. Average daily changes of GPS-TEC and 1-h predicted TEC with the DBO-BiLSTM, SSA-BiLSTM, BiLSTM and LSTM models, and the RMSE values of the predicted results by these four models at nine grid points

From Figure 5, it can be seen that the average daily TEC variation predicted by the four models has a close agreement with those computed using GPS data at different grid points. It is should be noticed that the most RMSE values of the DBO-BiLSTM model for 1-h predictions remains below 2 TECU at the 9 grid points, and the values are the smallest compared with those of other three models, and the followed ones are SSA-BiLSTM, BiLSTM and LSTM. It also can be observed that, no matter for RMSE or MAE, the values of them in lower latitudes are smaller than those in the higher latitudes. The main reason is that the ionosphere suffers more serious fluctuation as the rise of latitude, correspondingly, TEC are more difficult to be modelled and predicted. Therefore, the high accuracy TEC prediction is worth deep and thorough research in the future.

The optimal hyperparameters achieved by DBO optimization of nine grid points are

listed in Table 2.

Table 2. The optimal hyperparameters of the DBO-BiLSTM model with nine grid points

Hyperparameters \ Grid point	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9
Number of Neurons	127	101	169	208	112	12	124	208	136
Maximum training period	71	60	61	79	53	47	51	39	45
Initial learning rate	0.003564	0.004523	0.004152	0.004307	0.005232	0.005126	0.009272	0.007426	0.003654

(2) Analysis of TEC Prediction Accuracy for Daytime and Night Time of Local Time

In order to evaluate the TEC prediction ability of the four models during the day time and night time, local time periods of test dataset were set as 10:00–18:00 and 19:00–9:00 to represent day time and night time, respectively. The average RMSE values for the 1-h prediction results of DBO-BiLSTM, SSA-BiLSTM, BiLSTM, and LSTM models during the day time and night are listed in Table 3.

Table 3. Average RMSE values for 1-h prediction results of the DBO-BiLSTM, SSABiLSTM, BiLSTM, and LSTM models during the day time and night time (TECU)

Grid point	DBO-BiLSTM		SSA-BiLSTM		BiLSTM		LSTM	
	Day	Night	Day	Night	Day	Night	Day	Night
P_1	1.31	1.08	1.93	1.53	2.36	1.84	2.31	1.95

P_2	1.63	1.16	2.25	1.84	2.31	1.95	2.52	2.14
P_3	1.78	1.31	2.36	1.84	2.47	2.01	2.64	2.06
P_4	0.55	0.48	1.00	0.87	1.21	1.02	1.31	1.19
P_5	0.45	0.44	0.78	0.81	1.07	1.04	1.20	1.18
P_6	0.46	0.44	0.93	0.89	1.02	0.97	1.30	1.28
P_7	0.42	0.36	0.72	0.69	0.86	0.85	1.03	1.01
P_8	0.42	0.35	0.69	0.61	0.84	0.76	1.02	0.94
P_9	0.38	0.37	0.67	0.66	0.83	0.81	1.00	0.99

It can be observed from Table 3 that the average RMSE values of the predicted results by the four models in the daytime are all larger than those in the night time, and the RMSE values at the three points P_1, P_2 and P_3 in low latitudes are larger than other points with same longitudes in high latitudes no matter in day time or in night time. Table 3 also indicated that, compared with the other three models, DBO-BiLSTM achieved the highest prediction accuracy no matter in the day time or night time.

(3) Comparison and Analysis of TEC Map computed using GPS data and Predicted Ones

In order to assess the predictive ability of the proposed TEC prediction model in the European region, TEC map difference between the map computed using GPS data and those predicted by the four prediction models were computed at four time points on 24th

August 2022 (the quiet geomagnetic period: Kp<1) at 0:00, 6:00, 12:00, and 18:00, as shown in Figure 6.

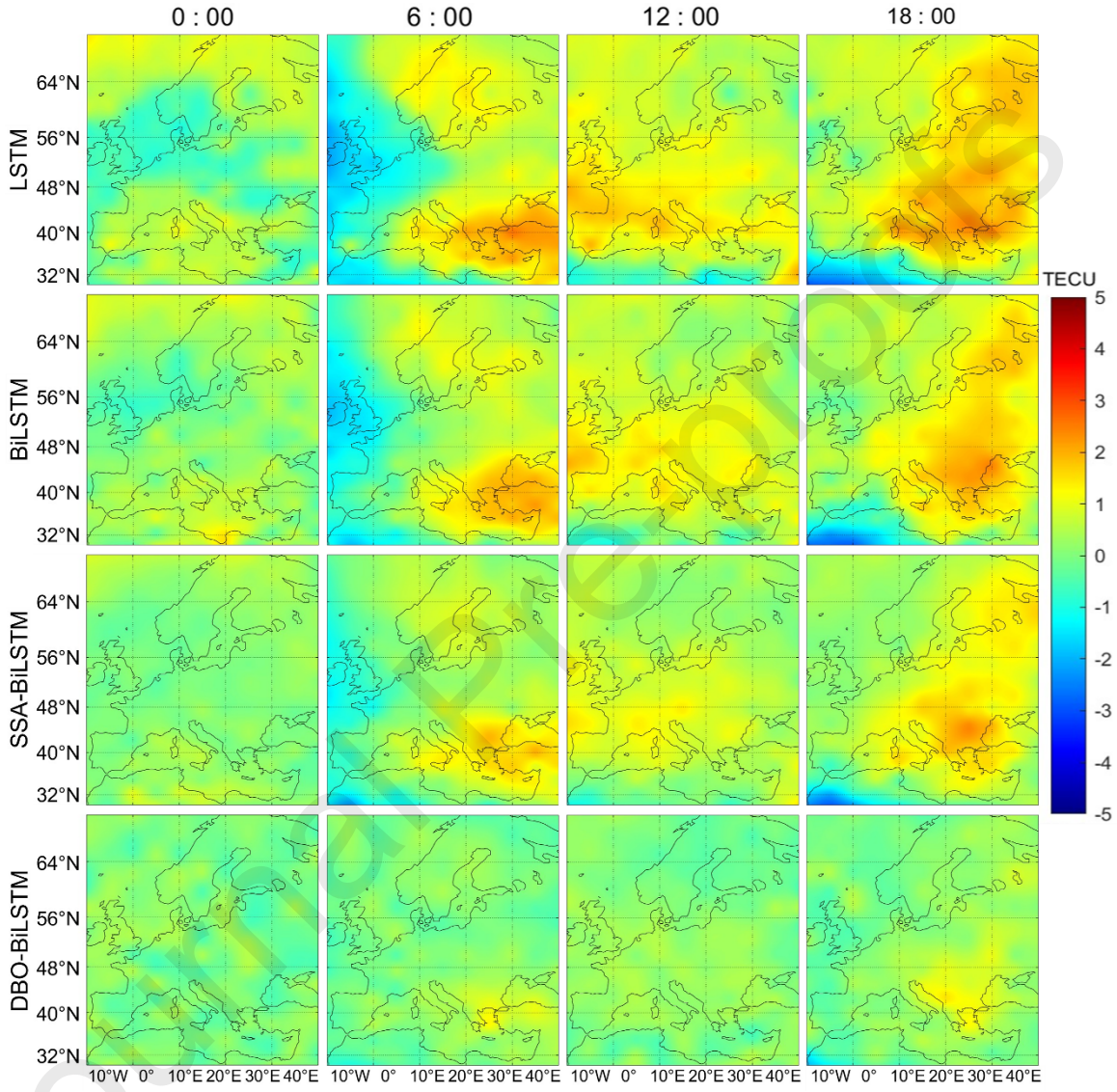


Figure 6. Distribution maps of the differences between TEC computed by GPS data and those predicted by four models for the European region at different time points on 24th August, 2022

As can be seen from Figure 6, DBO-BiLSTM model achieved the smallest difference between the TEC solved by GPS data and the TEC predicted by this model, and SSA-BiLSTM, BiLSTM and LSTM followed in order. On the whole, the region with large TEC differences are mostly concentrated in marine regions or the regions with fewer GNSS receivers, such as the southeast and the western marine region at 6:00 local time, and the

southwest and northeast at 18:00 local time in European region. On August 24, 2022, both geomagnetic and ionospheric activities are quite, so even for LSTM forecast model, performing worst compared with the other models, the largest difference between the TEC forecasted and that gained by GPS is no more than 5TECU.

4.3 Analysis of the TEC Prediction Accuracy of the Built Models under moderate disturbed Geomagnetic condition

During the time period of the GPS observation data used in the study, moderate magnetic storm events occurred. Both moderate geomagnetic storm periods, from 21:00 on 7th August to 8:00 on 8th August 2022, and from 15:00 on 17th August to 22:00 on 18th August 2022, were selected to carry out the evaluation of the prediction ability for the four models. In both periods, Dst range was from -30nT to -60nT and Kp range was from 4 to 7. According to the values of the geomagnetic indices, the rest test dataset were all under the quiet geomagnetic condition. Table 4 lists the RMSE values of the 1-h prediction results at nine grid points under moderate disturbed and quiet geomagnetic conditions.

Table 4. RMSE values of 1-h TEC predictions of DBO-BiLSTM, SSA-BiLSTM, BiLSTM, and LSTM models under moderate disturbed and quiet geomagnetic conditions (TECU)

Grid point	DBO-BiLSTM		SSA-BiLSTM		BiLSTM		LSTM	
	Disturbed	Quiet	Disturbed	Quiet	Disturbed	Quiet	Disturbed	Quiet
P_1	1.37	1.14	1.96	1.66	2.50	2.00	2.28	2.08
P_2	1.65	1.32	2.36	1.96	2.47	2.05	2.60	2.25
P_3	1.82	1.46	2.45	2.00	2.48	2.16	2.71	2.24
P_4	0.85	0.46	1.34	0.86	1.44	1.05	1.57	1.19
P_5	0.73	0.40	1.11	0.76	1.32	1.02	1.53	1.15

P_6	0.59	0.43	1.18	0.87	1.27	0.96	1.45	1.27
P_7	0.77	0.33	1.63	0.65	1.32	0.80	1.45	0.96
P_8	0.64	0.36	1.02	0.61	1.16	0.77	1.37	0.94
P_9	0.54	0.35	0.94	0.63	1.13	0.78	1.24	0.97

Table 4 indicates that, overall, the RMSE values under the disturbed geomagnetic conditions are larger than those under the quiet conditions. This is mainly because, the ionosphere is more active and the TEC change is more drastic under the disturbed geomagnetic condition than under quiet condition. It can also be noticed that the smallest TEC RMSE are achieved by the DBO-BiLSTM model, no matter under disturbed or quiet geomagnetic conditions, and SSA-BiLSTM, BiLSTM and LSTM followed. Therefore, under the same geomagnetic environment, the DBO-BiLSTM model can gain the highest TEC prediction accuracy compared to the other three models.

5. Conclusion and Discussion

This study aims to establish a hybrid novel deep learning model to achieve high accuracy for short-term TEC prediction in European region. To achieve this goal, a practical and optimized DBO-BiLSTM model for high accuracy short-term TEC prediction was developed for the study region. This hybrid model applies DBO algorithm to optimize BiLSTM neural network and taking the geomagnetic and solar activity indices into consideration, thereby to break through the limitations of traditional deep learning algorithms and the traditional TEC prediction models. The predicted results of DBO-BiLSTM are assessed by the TEC computed with GPS observation, and the performance of the novel optimized model is compared with those of SSA-BiLSTM, BiLSTM and LSTM models. The test results demonstrated that DBO-BiLSTM model outperformed the other three models in prediction accuracy for both 1-h and 2-h predictions. The main findings and contributions are as follows:

DBO-BiLSTM model can best capture the characteristic of ionospheric in the study region, and gains the closest relationship with TEC computed by GPS data compared with SSA-BiLSTM, BiLSTM and LSTM models evaluated by RMSE, R^2, Corr and MAE indices.

The prediction performance of DBO-BILSTM was evaluated in different latitude bands, at different local time points and under different geomagnetic conditions at 9 grid sites. The test results show that the DBO-BILSTM model can effectively capture the TEC characteristic with diurnal variation, latitude variation and geomagnetic condition variation.

In this research, the DBO-BiLSTM model was established under the moderate solar activity condition with the average F10.7 index value around 120, and the index had an obvious rising tendency. While, the TEC prediction accuracy has positive relationship with time series length of the training dataset (Shi et al. 2022), and has close relationship with the tense of solar activity (Nath et al. 2023). In this study, DBO-BiLSTM model showed good performance in short-term TEC prediction, and achieved higher accuracy than those predicted by Shi et al. (2022) with BiLSTM model, but the prediction accuracy gradually decreased with the extension of the prediction time length as proved by Xiong et al. (2021). Therefore, in order to improve the accuracy of the DBO-BiLSTM model, it is necessary to evaluate the performance of the model using longer time series of training datasets and under different intense of solar activity for longer prediction time period.

The above analysis and discussion verified that DBO-BiLSTM model owns superiority in short-term ionospheric TEC prediction, particularly in capturing spatiotemporal variation characteristics of TEC. The optimized novel hybrid model owns a promising future in TEC prediction and can provide important reference to the time series prediction of other atmospheric phenomena in the future.

Acknowledgments: This research is funded by the National Natural Science Foundation of China (Grant NO. 42374038 and 42271436), the Shandong Natural Science Foundation of China (Grant NO. ZR2023MD072 and ZR2021MD030), a Project of Shandong Province Higher Education Science and Technology Program (Grant NO. J17KA077) and Talent introduction plan for Youth Innovation Team in universities of Shandong Province (innovation team of satellite positioning and navigation)

Data Availability Statement

We are grateful to the providers of all the data used in this work for making their data available. The ground-based GPS data are available from the European Permanent Global Navigation Satellite System network (EPGNSS) (ftp://epncb.oma.be/). The geomagnetic activity index Kp, AP and solar activity index F10.7 are available from the GFZ German Research Center For Geosciences (ftp://ftp.gfz-potsdam.de/). The geomagnetic activity index Dst data is available from the National Centers For Environmental Information (ftp://ftp.ngdc.noaa.gov).

References

Bengio,Y., Simard,P.Y.,& Frasconi,P. (1994). Learning long-term dependencies with gradient descen t is difficult. *IEEE transactions on neural networks*,5(2),157-66.https://doi.org/10.1109/72.279181

Bent,R.B.,Llewellyn,S.K.,& Walloch,M.K. (1972). Description and evaluation of the Bent ionospheri c model. DBA Systems.

Duan,J.,Gong,Y.,Luo,J.,& Zhao,Z. (2023). Air-quality prediction based on the ARIMA-CNN-LSTM combination model optimized by dung beetle optimizer. *Scientific Reports*,13(1),1-16. https://doi.org/ 10.1038/s41598-023-36620-4

Feng,J. (2019). Research and implementation of empirical TEC models. *Acta Geodaetica et Cartog raphica Sinica*,48(10),1339. https://doi.org/10.11947/j.AGCS.2019.20180599

Giovanni,G.D.,& Radicella,S.M. (1990). An analytical model of the electron density profile in the i onosphere. *Advances in Space Research*,10(11),27-30. https://doi.org/10.1016/0273-1177(90)90301-F

Hochreiter, S.(1998). The vanishing gradient problem during learning recurrent neural nets and prob lem solutions. I*nternational Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*,6(02),1 07-116. https://doi.org/10.1142/S0218488598000094

Jang,L.,Sun,R.,Lu,Z.,Xu,C.,Liang,D.,Hu,D. (2023). Modeling and accuracy analysis of GNSS ionosph eric error inEU-China based on GA-BP. *Journal of Beijing University of Aeronautics and Astronau tics*,49(06),1533-1542. https://doi.org/10.13700/j.bh.1001-5965.2021.0476

Klobuchar,J.A. (1987). Ionospheric Time-Delay Algorithm for Single-Frequency GPS Users. *IEEE T ransactions on Aerospace and Electronic Systems*, AES-23(3),325-331. https://doi.org/10.1109/TAES. 1987.310829

Li,L.,Zhang,S.,Wang,Y.,Hu,Q.,Ying,S.(2013)Ionospheric Total Electron Content Prediction Based on A RMA Model. *Journal of Basic Science and Engineering*,21(05),814-822. https://doi.org/10.3969 /j.iss n.1005-0930.2013.05.002

Li,W.,Zhao,D., He,C., Shen,Y., Hu,A., & Zhang,K.(2021a). Application of a Multi-Layer Artificial Neural Network in a 3-D Global Electron Density Model Using the Long-Term Observations of C OSMIC, Fengyun-3C, and Digisonde. *Space Weather*,19(3), 1-19.https://doi.org/10.1029/2020SW0026 05

Li,W.,He,C.,Hu,A.,Zhao,D.,Shen,Y., & Zhang,K.(2021b). A new method for improving the performa nce of an ionospheric model developed by multi-instrument measurements based on artificial neural network. *Advances in Space Research*,67(1),20-34. https://doi.org/10.1016/j.asr.2020.07.032

Li.W.,Zhu,H.,Shi,S.,Zhao,D.,Shen,Y., & He,C.(2024). Modeling China's Sichuan-Yunnan's ionosphere based on multi-channel WOA-CNN-LSTM algorithm. *IEEE Transactions on Geoscience and Remot e Sensing*. https://doi.org/10.1109/TGRS.2024.3403684

Lin,X.,Wang,H.,Zhang,Q.,Yao,C.,Chen,C.,Cheng,L.,& Li,Z.(2022).A spatiotemporal network model for global ionospheric TEC fore-casting. *Remote Sensing*,14(7),1717. https://doi.org/10.3390/rs14071717

Mannucci,A.J.,Wilson,B.D.,Yuan,D.N.,Ho,C.H.,Lindqwister,U.J.,&Runge,T.F.(1998). A global mapping technique for GPS-derived ionospheric total electron content measurements. *Radio Science*, 33(3), 5 65–582. https://doi.org/10.1029/97rs02707

Ni,Y.,Yan,M.,Lu,R.(2023).Short-term prediction of ionospheric TEC based on DOA-BP neural netwo rk. *Acta Aeronautica et Astronautica Sinica*,45(3),1-15. https://doi.org/10.7527/S1000-6893.2023.28707

Nath,S.,Chetia,B.,Kalita,S.(2023). Ionospheric TEC prediction using hybrid method based on ensemb

le empirical mode decomposition (EEMD) and long short-term memory (LSTM) deep learning mod el over India. *Advances in Space Research,*71(5),2307-2317. https://doi.org/10.1016/j.asr.2022.10.067

Schaer, S.(1999). Mapping and predicting the earth's ionosphere using the Global Positioning Syste m, (Doctoral dissertation). Bern University.

Shi,S.,Zhang,K.,Wu,S.,Shi,J.,Hu,A.,Wu,H.,& Li,Y.(2022). An investigation of ionospheric TEC predic tion maps over China using bidirectional long short-term memory method. *Space Weather*，20(6)， 1-20. https://doi.org/10.1029/2022SW003103

Silva,A.,Moraes,A.,Sousasantos,J., Maximo,M.,Vani,B.,& Faria,C.J.(2023).Using Deep Learning toMap Ionospheric Total Electron Content over Brazil. *Remote Sensing*,15(2),412. https://doi.org/10.3390/rs1 5020412

Sivavaraprasad,G.,& Ratnam,D.V. (2017). Performance evaluation of ionospheric time delay forecast ing models using GPS observations at a low-latitude station. *Advances in Space Research*,60(2),475- 490. https://doi.org/10.1016/j.asr.2017.01.031

Tang,J.,Yang,D.,&Ding,M.(2023). Forecasting ionospheric foF2 using bidirectional LSTM and attenti on mechanism. *Space Weather*, 21(11).https://doi.org/10.1029/2023SW003508

Tang,L., &Chen,G.(2022). Equatorial Plasma Bubble Detection Using Vertical TEC From Altimetry Satellite.Space Weather,20(8), 1-5. https://doi.org/10.1029/2022SW003142

Tang, L. (2023a). Ionospheric disturbances of the January 15, 2022, Tonga volcanic eruption obser ved using the GNSS network in New Zealand. *GPS Solution*,27(1), 53. https://doi.org/10.1007/s102 91-023-01395-8

Tang,L., Li,J., Zhang,F.,& Zou,F. (2023b). Comparative study of ionospheric sporadic E occurrence rates using SNRnstd and S4 derived from GPS radio occultation. *GPS Solution*,27(3), 120.https://d oi.org/10.1007/s10291-023-01464-y

Tang,L.,Zhang,F., Li,P.,Deng,Y.,& Chen,W.(2024).Effects of equatorial plasma bubble-induced ionos pheric gradients on GNSS PPP-RTK. *GPS Solution*,28(3), 124. https://doi.org/10.1007/s10291-024-01 664-0

Tang,J.,Zhong,Z.,Li,Y.,Gao,X.(2022).Short-Term Prediction Model of lonospheric TEC Based on SS A-Elman Neural Network. *Journal of Geodesy and Geodynamics*,42(04),378-383. https://doi.org/10.1 4075/j.jgg.2022.04.009

Tang,J.,Zhong,Z.,Ding,M.,Wu,X.(2023).A Prediction Model of lonospheric TEC in China Based on Elman Neural Network lmproved by Particle Swarm Optimization Algorithm. *Geomatics and Infor mation Science of Wuhan University*. https://doi.org/10.13203/j.whugis20220254

Wang,Z.,Huang,G.,Du,S.,Xie,W.(2022).Modeling and accuracy assessment of regional ionospheric TE C over Shaanxi province based onBDS-3. *Journal of Nanjing University of Information Science & Technology (Natural Science Edition)*,14(06),737-743. https://doi.org/10.13878/j.cnki.jnuist.2022.06.014

Xie,S.,Chen,J.,Hunag,L.,Wu,P.,Qin,X.,& Liu,L.(2017).Ionospheric TEC Prediction Based on Holt-Win ters Models.*Journal of Geodesy and Geodynamic*,37(01),72-76. https://doi.org/10.14075/j.jgg.2017.01. 016

Xiong,B.,Li,X.,Wang,Y.,Zhang,H.,Liu,Z.,Ding,F.,& Zhao,B.(2022).Prediction of ionospheric TEC over

China based on long and short-term memory neural network. *Chinese Journal of Geophysics*,65(07), 2365-2377. https://doi.org/10.6038/cjg2022P0557

Xiong,P.,Zhai,D.,Long,C.,Zhou,H.,Zhang,X.,&Shen,X.(2021). Long short-term memory neural network for ionospheric total electron content forecasting over China. *Space Weather*,19(4). https://doi.org/10.1029/2020SW002706

Xue,J.,& Shen,B. (2022). Dung beetle optimizer: a new meta-heuristic algorithm for global optimization. *The Journal of Supercomputing*,79(7),7305-7336. https://doi.org/10.1007/s11227-022-04959-6

Yao,Y.,& Gao,X.(2022).Geomatics and Information Science of Wuhan University.*Geomatics and Information Science of Wuhan University*, 47(10),1728-1739. https://doi.org/10.13203/j.whugis20220364

Yuan,Y.,Huo,X.,& Zhang,B.(2017).Research Progress of Precise Models and Correction for GNSS Ionospheric Delay in China over Recent Years.*Acta Geodaetica et Cartographica Sinica*,46(10),1364-1378.https://doi.org/10.11947/j.AGCS.2017.20170349

Yuan,T.,Chne,Y., Liu,S.,Gong,J.(2018).Prediction Model for lonosphericTotal Electron Content Based on Deep Learning Recurrent Neural Network. *Chin. J. Space Sci*,38(01),48-57.https://doi.org/10.11728/cjss2018.01.048

Declaration of Interest Statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.