



OPEN ACCESS

EDITED BY

Daniel Okoh,
The National Space Research and
Development Agency (NASRDA), Nigeria

REVIEWED BY

Sampad Kumar Panda,
KL University, India
Yun Gong,
Wuhan University, China

*CORRESPONDENCE

Zan-Yang Xing,
✉ xingzanyang@sdu.edu.cn
Qing-He Zhang,
✉ zhangqinghe@sdu.edu.cn

RECEIVED 17 April 2024

ACCEPTED 09 August 2024

PUBLISHED 04 September 2024

CITATION

Qiao F, Xing Z-Y, Zhang Q-H, Zhang H-B,
Zhang S-R, Wang Y, Ma Y-Z, Zhang D, Lu S
and Varghese M (2024) A long time-series
forecasting informer architecture-based
ionospheric foF2 model in the low-latitude
region.
Front. Astron. Space Sci. 11:1418918.
doi: 10.3389/fspas.2024.1418918

COPYRIGHT

© 2024 Qiao, Xing, Zhang, Zhang, Zhang,
Wang, Ma, Zhang, Lu and Varghese. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with
these terms.

A long time-series forecasting informer architecture-based ionospheric foF2 model in the low-latitude region

Feng Qiao^{1,2}, Zan-Yang Xing^{2*}, Qing-He Zhang^{2,3*},
Hong-Bo Zhang⁴, Shun-Rong Zhang⁵, Yong Wang²,
Yu-Zhang Ma², Duan Zhang^{2,3}, Sheng Lu² and Manu Varghese²

¹School of Information Science and Engineering, Shandong University, Qingdao, China, ²Shandong Provincial Key Laboratory of Optical Astronomy and Solar-Terrestrial Environment, Institute of Space Sciences, Shandong University, Weihai, China, ³State Key Laboratory of Space Weather, Center for Space Science and Applied Research, Chinese Academy of Sciences, Beijing, China, ⁴China Research Institute of Radiowave Propagation, Qingdao, China, ⁵MIT Haystack Observatory, Westford, MA, United States

Deep learning models have made great accomplishments in space weather forecasting. The critical frequency of the ionospheric F2 layer (foF2) is a key ionospheric parameter, which can be understood and predicted by some advanced new deep learning technologies. In this paper, we utilized an Informer architecture model to predict foF2 for several hours up to 48 h and analyzed its variations during periods of quiet, moderate, and intense geomagnetic conditions. The Informer method forecasts the temporal variations of foF2 by processing and training the past and present foF2 data from the Haikou station, China, during 2006–2014. It is evident that the Informer–foF2 model achieves better prediction performance than the widely used long short-term memory model. The Informer–foF2 model captures the correlation features within the foF2 time series and better predicts the variations ranging for hours up to days during different geomagnetic activities.

KEYWORDS

Informer, foF2, ionosphere, long short-term memory, long sequence time-series forecasting

1 Key Points

- An Informer architecture-based model is used to forecast ionospheric foF2 at low latitudes.
- The Informer–foF2 model has advantages in predicting variations from several hours up to 48 h.
- The Informer–foF2 model forecast for geomagnetic storms is in good agreement with the observations.

2 Introduction

The F2 layer of the ionosphere has the highest degree of ionization, which is closely linked to the Global Positioning System (GPS) and other navigation systems as well as

long-range, high-frequency (HF) communications. The critical frequency of F2 layer (foF2), as one of the most important parameters of radio science, controls the electromagnetic wave propagation through the ionosphere. Therefore, accurate prediction of foF2 is important and a difficult problem in space weather forecasting, especially for long time series at low latitudes (Cander et al., 1998; Rao et al., 2018; Zhang et al., 2022; Fan et al., 2019). There are many precedents for researchers to combine ionospheric phenomena with deep learning (Chen et al., 2019; Zhou et al., 2021; Wang et al., 2022; Yang et al., 2024; Hu and Zhang, 2018). Although many models could simulate and predict the ionospheric F2 layer variations, the complex changes in solar activities make the spatial and temporal variability of foF2 difficult to be predicted and even more difficult to achieve the expected performance at low latitudes. Other ionosphere models such as the International Reference Ionosphere (IRI) (empirical model) (Wichaiapanich et al., 2017), random forest (machine learning model), and autoregressive integrated moving average model (ARIMA, time-series model) cannot predict the sudden changes in foF2 caused by a geomagnetic storm. These models do not perform well in long-term forecasting since they are restricted by long-range dependencies. In order to achieve improvement in the forecasting accuracy over a longer time span, deep learning techniques are implemented to forecast foF2 for a sufficiently long duration. Long sequence time-series forecasting means the forecast of high-resolution ionospheric parameters continuously ahead of a few days.

There are several studies on ionospheric modeling using the basic deep learning approach of artificial neural networks (ANNs), which began in the mid-1990s (Williscroft and Poole, 1996; Altinay et al., 1997; Sai Gowtam and Tulasi Ram, 2017; Wang et al., 2013; Kim et al., 2020; Moon et al., 2020; Li et al., 2021; Zhou et al., 2021). Williscroft and Poole (1996) used 10 years of foF2 data observed from the Grahamstown ionosonde and the sunspot number and Ap index to train a simple ANN. Altinay et al. (1997) developed a 1-h prediction model for foF2 using 10 years of foF2 data observed from the Poitiers ionosonde in Central Europe and the Kp index to train a multilayer perceptron model. Since these ANN-based models do not consider past data for the current specific period, the predictive performance cannot be applicable to the phenomena affected by the prior states that are older than the specified period and are naturally not competent for long time-series prediction tasks.

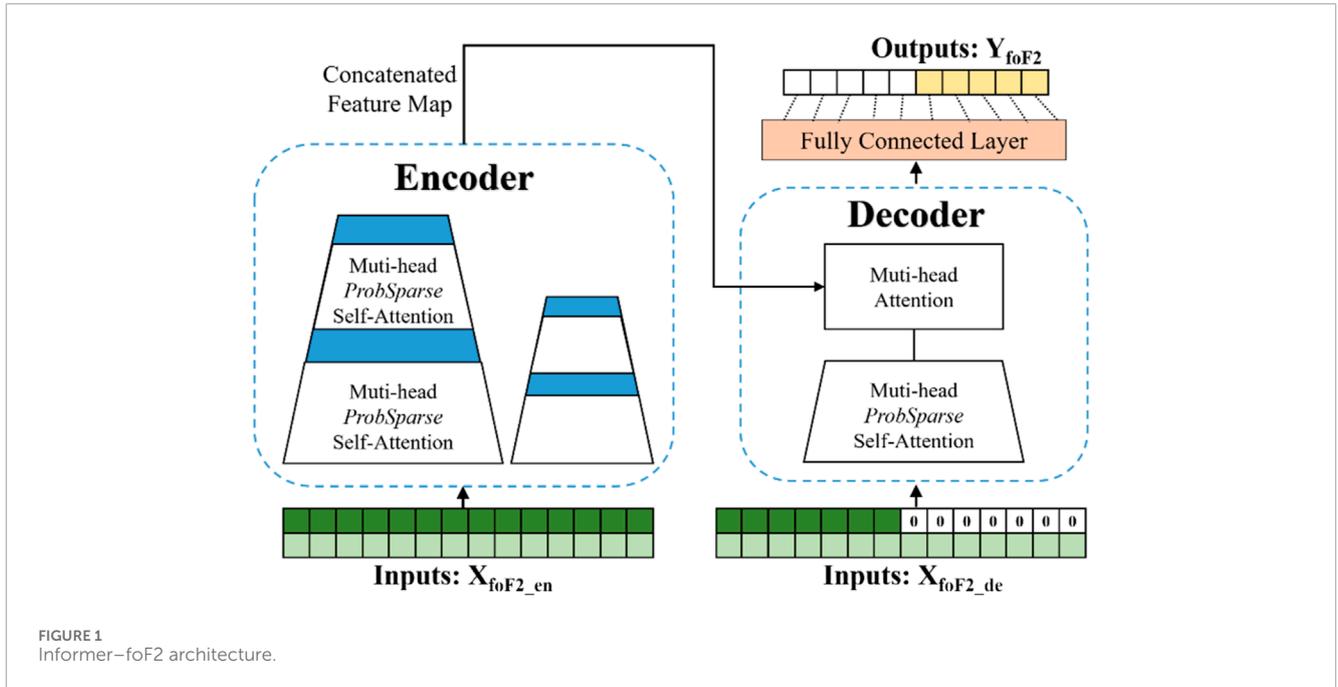
To overcome the shortcomings of the above algorithms, the long short-term memory (LSTM) algorithm, a technique that remembers past data and reflects it in predictions, was developed (Ergen and Kozat, 2018). Earlier, foF2 and hmF2 parameters were predicted by Moon et al. (2020) using LSTM models. Kim et al. (2020) used a physics-based model, which could predict up to 24 h of mid-latitude ionospheric data by inputting the parameters predicted by Moon et al. (2020). Although the LSTM-based model showed reasonably good prediction performance on geomagnetically quiet days, the model predictions are not correlated well during geomagnetically disturbed periods (Kim et al., 2021). The current LSTM model significantly improves the prediction accuracy and time span of the long time series, but its response to abrupt changes was not sufficiently rapid due to the retention of memory states (Lissa et al., 2020).

The LSTM models incorporating the attention module have been applied in long-term time-series prediction studies. The attention mechanism offers the possibility to focus on the response triggered by a certain factor Xia et al., 2022, such as the foF2 anomalies caused by geomagnetic perturbations (Liu and Guo, 2019; Rao et al., 2021). Tang et al. (2023) combined BiLSTM with Attention to predict foF2. With the widespread use of the transformer model (Vasmani et al., 2017), the fully attention-based model is applied to more and more fields, and its excellent performance has led to its use for solving long sequence time-series forecasting (LSTF) problems, but the computational complexity limitation of the transformer itself has led to unsatisfactory results. The transformer-like model Informer proposed by Zhou et al. (2021), which uses the *ProbSparse* self-attention mechanism, has shown excellent capability in the long-term time-series prediction problem, and at the same time, the high computational speed, low complexity, and high accuracy demonstrated by the Informer are very attractive. The long-term time-series prediction of the ionospheric foF2 has been implemented in this study based on the foF2 data from the Haikou station at a low-latitude region. We achieved better long-term forecasting performance and avoided self-correction due to geomagnetic activities (Gao et al., 2020). We present the prediction results of foF2 at low latitudes using the Informer model and discuss its advantages and limitations. The Informer-based model shows great potential in solving LSTF problems and saving computing resources. From an experimental standpoint, the Informer-based model has excellent performance in time-series prediction and has great value for further research and application. In this study, we refer to the 48-h prediction as a long-term time-series prediction and attempt to find the most suitable model for long-term time-series prediction on foF2. The models constructed in this paper are applicable to the quiet, moderate, and intense geomagnetic periods, and the comparisons of these models in different periods are discussed.

3 Informer model architecture

Due to the significant correlation with the foF2 time series, recurrent neural networks (RNNs) have been extensively used in foF2 prediction. However, typical RNNs are limited in their long-term dependence due to gradient issues, which do not represent abrupt events. The transformer module by using self-attention can solve those problems to some extent, but the module is limited by computation complexity and error accumulation in the decoding process directly in long sequence time-series forecasting. So the Informer-foF2 model is developed to perform foF2 forecasting using *ProbSparse* self-attention to simplify the calculation and the generative decoder to output the forecast results directly (Bi et al., 2022).

The Informer model was proposed by Zhou et al., in 2021, and we fine-tuned the model to make it better at predictive tasks of foF2. Figure 1 shows the Informer-foF2 architecture. It has two critical components denoted as encoder and decoder modules. The encoder converts input information into a dense vector of fixed dimensionality and extracts features from elements to generate feature mapping. Inversely, the



decoder combines the information and feature mapping to the predicted outputs.

ProbSparse self-attention differs in finding a more active query to simplify corresponding attention calculation from traditional multi-head self-attention. The sparsity measurement of *ProbSparse* self-attention is shown below Equation 1.

$$\overline{M}(q_i, K) = \max_j \left\{ \frac{q_i k_j^T}{\sqrt{d}} \right\} - \frac{1}{L_K} \sum_{j=1}^{L_K} \frac{q_i k_j^T}{\sqrt{d}}, \quad (1)$$

where L_k is $1/q(k_j | q_i)$, the i times query of sparsity measurement is defined as q_i ; and q_i and k_j from Q and K , generated in the encoder training process, represent query and key vectors, respectively. *ProbSparse* self-attention calculates the corresponding attention between the active queries and keys, replacing the other by a uniform distribution. Attention calculation is defined as follows Equation 2, where \overline{Q} is a sparse matrix and it contains the sparsity measurement:

$$\overline{A}(Q, K, V) = \text{Softmax} \left(\frac{\overline{Q}K^T}{\sqrt{d}} \right) V. \quad (2)$$

The distilling block is using 1-d max pooling operation to accomplish feature downsampling in order to make encoder feature extraction quick and simple. The *ProbSparse* self-attention and distilling operations alternately stack. The encoder output feature mapping is acquired by a two-channel stack.

For the decoder architecture, the generative decoder generates all predicted outputs at once to replace the transformer's decoder, in which it avoids the time-consuming dynamic decoding process in the encoder-decoder architecture. A decoder is composed of two decoder layers, each with a self-attention, a cross-attention, a three-layer norm, and a dropout. The decoder input is a truncation of the later part of the encoder input and a matrix with the

same shape as the predicted target. After passing through the decoder, each placeholder (position to be predicted) has a vector, which is then input into a fully connected layer to obtain the predicted results.

A detailed explanation of the Informer architecture is given by Zhou et al. (2021). The specific model composition of the Informer for the foF2 prediction is shown in Table 1.

The LSTM-based model and IRI will appear as contrast models to demonstrate the advantages of the Informer-based model. The LSTM model is a widely used time-series forecasting model in various fields. In previous studies, LSTM has been shown to significantly improve model prediction RMSE (root mean squared error) over empirical models such as IRI and pure time-series models such as ARIMA. LSTM is an excellent model for time-series problems with multivariate effects, especially for time-series data prediction tasks such as foF2, where the fluctuations are severe and perturbations are diverse. The principle of the BiLSTM model is similar to that of the LSTM model, and the improvement is less significant than that of the Informer. This work mainly discusses the effect of the self-attention model on the long-term time-series prediction of foF2, so LSTM is chosen as the representative model to verify the effect of the Informer. The LSTM used in this comparison has been tuned for several experiments and can basically reach the average level of the LSTM prediction models for low-latitude regions, as found in previously published results by several researchers. In order to better demonstrate the model enhancement, we also compare the model outputs with the commonly used empirical model, IRI-2016. We trained 1-h, 5-h, 12-h, 24-h, and 48-h models based on LSTM and Informer. In addition, we perform the same procedure for the corresponding IRI output for foF2 values at the Haikou station (20.0° N and 110.1° E), China, and use these outputs for comparisons with the actual foF2 measurements from Haikou.

TABLE 1 Specific model composition of the Informer.

Encoder			N
Inputs	1×3Conv1d	Embedding ($d = 512$)	
ProbSparse Self-attention Block	Multi-head ProbSparse attention ($h = 8$)		4
	Add, LayerNorm, dropout ($p = 0.05$)		
	Pos-wise FNN ($d_{inner} = 2048$), GELU		
	Add, LayerNorm, Dropout ($p = 0.05$)		
Distilling	1×3Conv1d, ELU		
	Max pooling (stride = 2)		
Decoder			N
Inputs	1×3Conv1d	Embedding ($d = 512$)	
Masked PSB	Add mask on attention block		
ProbSparse Self-attention Block	Multi-head ProbSparse attention ($h = 8$)		2
	Add, LayerNorm, Dropout ($p = 0.05$)		
	Pos-wise FNN ($d_{inner} = 2048$), GELU		
	Add, LayerNorm, Dropout ($p = 0.05$)		
Batch-Size	32		
Epochs	50		

4 Data and model inputs

The ionospheric foF2 measurements are manually scaled from the ionograms of the Haikou station (20.0° N and 110.1° E) in 1-h resolution. The solar radio flux F10.7 is available from the Space Weather website <https://www.spaceweather.gc.ca/index-en.php> in 1-day resolution. The geomagnetic indices such as Dst and Kp are available from the Kyoto World Data Center, <https://wdc.kugi.kyoto-u.ac.jp/dstae/index.html>, having resolutions of 1 h and 3 h, respectively.

The foF2 data from 1 January 2006 to 31 May 2014 are the input and target of the Informer–foF2 model. In order to ensure the input data continuity of the time-series model, the missing data of foF2 have been interpolated by the linear interpolation method. The model output at the time of the geomagnetic disturbances is generated considering F10.7, Dst, and Kp indices. The data have been processed at 1-h intervals and are shown in Figure 2. Due to foF2 exhibiting outstanding time-auto correlation, the 24-h prediction window was chosen for the study. We use the foF2 of the past 96 h to predict the foF2 of the future 48 h, which are all multiples of 24.

In order to make the Informer–foF2 model fit better with the prepared training datasets, we changed the numbers of the encoder and decoder layers, multi-head attention, dropout,

minimum batch size, and embedding to optimize the configuration of the proposed model. The total number of foF2 samples is split as 70% for training, 10% for validation, and 20% for testing and prediction.

To quantitatively evaluate the performance of our model, we calculated the RMSE (root mean square error), MAE (mean absolute error), MSE (mean square error), MAPE (mean absolute percentage error), and MSPE (mean squared percentage error) as evaluating indicators. The calculation methods are as follows:

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

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right|, \quad \text{and} \quad MSPE = \frac{100\%}{n} \sum_{i=1}^n \left(\frac{\hat{y}_i - y_i}{y_i} \right)^2.$$

Here, y_i is the ground truth and \hat{y}_i is the model prediction. The primary reference metric is RMSE.

5 Results

Table 2 summarizes the performance of the models used in the study. The Informer–foF2 model achieves the best performance among different models. The results have shown that the Informer–foF2 model shows better predictive performance for lower RMSE by approximately 17%, 47%, 29%, 35%, and 27% compared to the LSTM–foF2 model. Moreover, the Informer model significantly outperforms the best level of LSTM models, as shown in Table 2. Figure 3 shows the RMSE of LSTM–foF2 and Informer–foF2 models. Informer has a significant improvement in prediction accuracy and is more accurate for long-term (5–48 h) forecasting. It is more suitable for long-term forecasting than LSTM and is far superior to the IRI-2016 model. Therefore the IRI-2016 model will not be referenced for comparison in the latter picture.

Figure 4 shows the RMSE of Informer–foF2 for each hour during the entire 48-h prediction. The forecast errors corresponding to the 24th hour and 48th hour are marked in red, with a significant decrease compared to other points in the forecast. It is obvious that as time goes on, the errors increase gradually but not linearly; after approximately 15 h, the increase in RMSE becomes gentle. It shows that the Informer–foF2 model has enormous advantages in long-term time-series prediction, and we believe that the correlation among foF2 data also plays an important role in this work. The results have also shown clear 24-h diurnal periodicities at the Haikou station.

Figure 5 shows that the diurnal variations of foF2 are contained within 2-day forecasts for the low-latitude region, where the predictive output of the Informer and LSTM models is benchmarked against actual measurement values. The blue solid line denotes the forecasts by the Informer model, the black dashed line indicates LSTM outputs, and the red dotted line is the actual foF2 measurements. During the spring equinox, the prediction curve of the Informer-48 model aligns more closely with the actual values. At the summer solstice, the LSTM-48 model prediction deviates significantly at approximately 10 h, and in the latter half of the prediction window, both model prediction curves diverge from the actual values to some extent; however, the difference curve shows that the Informer-48 model performs relatively better. Around the autumn equinox, both models perform

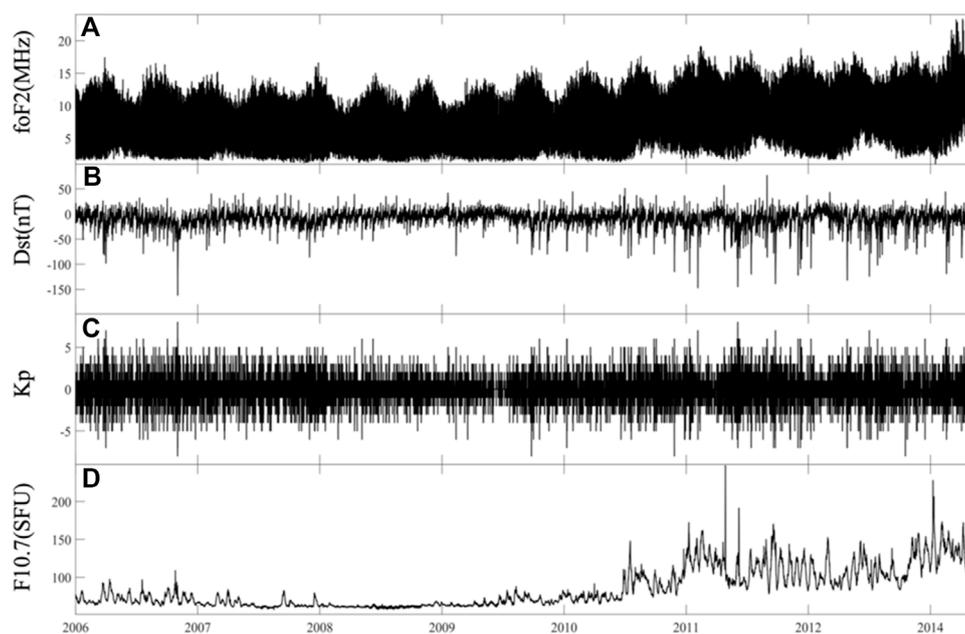


FIGURE 2 Actual values of (A) foF2, (B) Dst, (C) Kp, and (D) F10.7 to the Informer–foF2 model input during January 2006–May 2014.

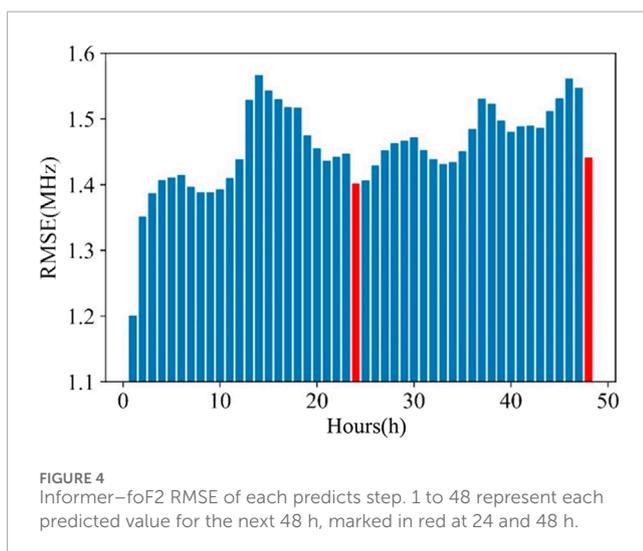
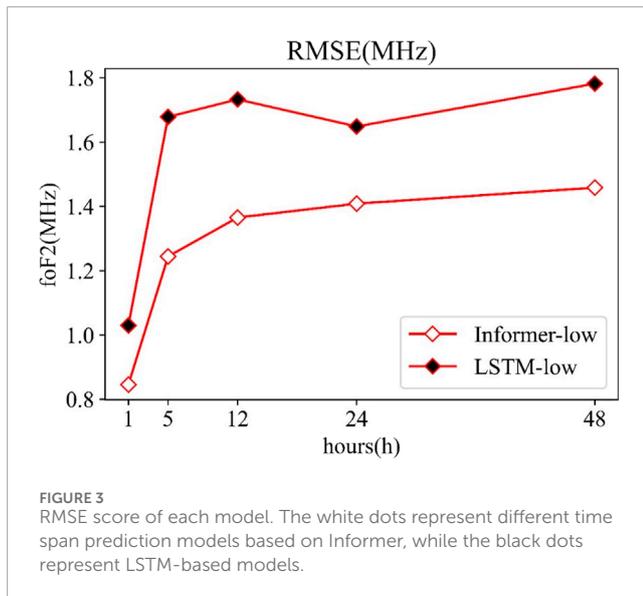
TABLE 2 Accuracy of foF2 forecasting compared to the IRI-2016 model and other LSTM observations.

Model	Predict time	MAE	MSE	RMSE	MAPE	MSPE
IRI-2016	1 h	1.21	2.62	1.62	0.16	0.045
Informer–foF2	1 h	0.61	0.71	0.84	0.07	0.011
	5 h	0.91	1.54	1.24	0.10	0.023
	12 h	1.01	1.86	1.36	0.11	0.032
	24 h	1.04	1.98	1.41	0.12	0.035
	48 h	1.09	2.12	1.45	0.13	0.038
LSTM–foF2	1 h	0.75	1.06	1.03	0.09	0.022
	5 h	1.23	2.81	1.67	0.14	0.047
	12 h	1.29	3.01	1.73	0.15	0.052
	24 h	1.24	2.79	1.67	0.14	0.053
	48 h	1.32	3.17	1.78	0.16	0.064

similarly for the initial 36 h, but beyond this point, the LSTM-48 model prediction curve noticeably deviates from the actual data, demonstrating the Informer-48 model's advantage in longer-term predictions. During the winter solstice, it is observed that the model performance deteriorates, starting approximately 36 h toward the end of the prediction window, with the LSTM-48 model exhibiting significantly more fluctuations than the Informer-48 model. Thus, the Informer-based ionospheric F2 layer critical

frequency prediction model exhibits superior performance in low-latitude ionospheric foF2 forecasting, showing substantial advantages over the LSTM model.

In order to show the performance of the Informer–foF2 model in the storm period, two storm events have been selected, which occurred on 17 March 2013 and 28 March 2013 (Figures 6, 7). The storm on 17 March had a Dst minimum of -131 nT (03 LT, 19 March), while the storm on 28 March had -59 nT (20 LT, 28



March) (Shim et al., 2018). The blue line represents the results of the Informer 48-h prediction model of the event (the labels in the figures is Informer-48); the (Rao et al., 2021) black line represents the results of the LSTM 48-h prediction model (the labels in the figures is LSTM-48); the pink and green lines represent the results of the Informer and LSTM 1-h prediction models (the labels in the figures are Informer-1 and LSTM-1), respectively; and the red line is the real foF2 value (the labels in the figures is REAL). The gray region needs special attention as it can more clearly distinguish the performance differences of the models.

As shown in Figure 6, the measured foF2 values at the Haikou station are presented alongside the predicted foF2 values by both the Informer-foF2 and LSTM-foF2 models. During the intense geomagnetic storm, the Dst index plummeted to approximately -120 nT and persisted for an extended period, inducing fluctuations in foF2, which in turn affect the performance of predictive models, particularly those spanning longer durations. It is evident that

prior to the storm, the actual measurements closely align with the predictions from the Informer-foF2 model. As the storm commences, both models continue to approximate the actual foF2 values well, and upon examining the differences between the predicted and measured values, the results from the Informer-48 and LSTM-48 models are remarkably similar, suggesting that both the LSTM and Informer models provide a good fit for the data to be forecasted in the initial phase of long sequence time-series predictions.

Notably, as the Dst index hovers at approximately -100 nT for approximately 10 h before continuing its descent, a significant deviation is observed in the LSTM 48-h prediction model from the actual values during this period, as shown in the gray-shaded area 1, whereas the Informer-48 model maintains a good performance. A shortcoming of the 48-h models is their inability to accurately grasp fluctuations that oscillate within smaller time intervals; the attention-based Informer model exhibits a pronounced advantage in unpredictable fluctuation events such as geomagnetic storms. As the storm gradually abates, a brief resurgence in the predictive accuracy of both models is observed. However, it is clear that the predictions from the Informer-48 and LSTM-48 models begin to diverge from the actual ionospheric foF2 values, particularly toward the end of the forecasting window, as depicted in gray-shaded area 2, where both models' predictions substantially veer off the true foF2 curve, with the discrepancies significantly increasing. In addition to the objective influence of time and the cumulative effect of predictive errors, the persistent impact of the storm's incomplete recovery should be considered, with all three factors contributing to the substantial divergence in the predictive models. The aforementioned temporal objective influence refers to the noticeable decline in model performance during nighttime (Feng et al., 2021). Both gray areas 1 and 2 occur during the local time interval of approximately 22:00 to 4:00, where a considerable drop in model performance is evident. Overall, the Informer-48 ionospheric foF2 prediction model exhibits a certain superiority over the LSTM-48 model, clearly closer to predicting the true values.

Figure 7 shows that, compared with the characterization of the strong magnetic storm described in Figure 6, the prediction accuracy of the model for the moderate magnetic storm is significantly closer to the measured value. In the 10 h preceding the forecast, both the Informer-48 and LSTM-48 models exhibit deviations from the actual foF2 measurements, with the Informer-48 predictions more closely mirroring the actual data. At the minimum of the Dst index within gray-shaded area 1, the Informer-48's forecasted values are seen to align almost congruently with the actual foF2 measurements, whereas the LSTM-48's forecasts display a discernible discrepancy, highlighting the LSTM model's difficulties in capturing the abrupt shifts associated with geomagnetic activity. In the waning hours of this event, as denoted by gray-shaded area 2, the Informer-48 model's performance remains superior to that of the LSTM-48 model. This advantage in the predictive capability at the end of the forecast window further underscores the Informer-48 model's proficiency in accurately undertaking long-duration foF2 forecasting tasks.

Figure 8 describes the performance of the low-latitude forecasting model during a geomagnetic quiet period, providing a lucid illustration of the model's capabilities when the perturbations of magnetic storms are absent. At the inception of the forecast, as

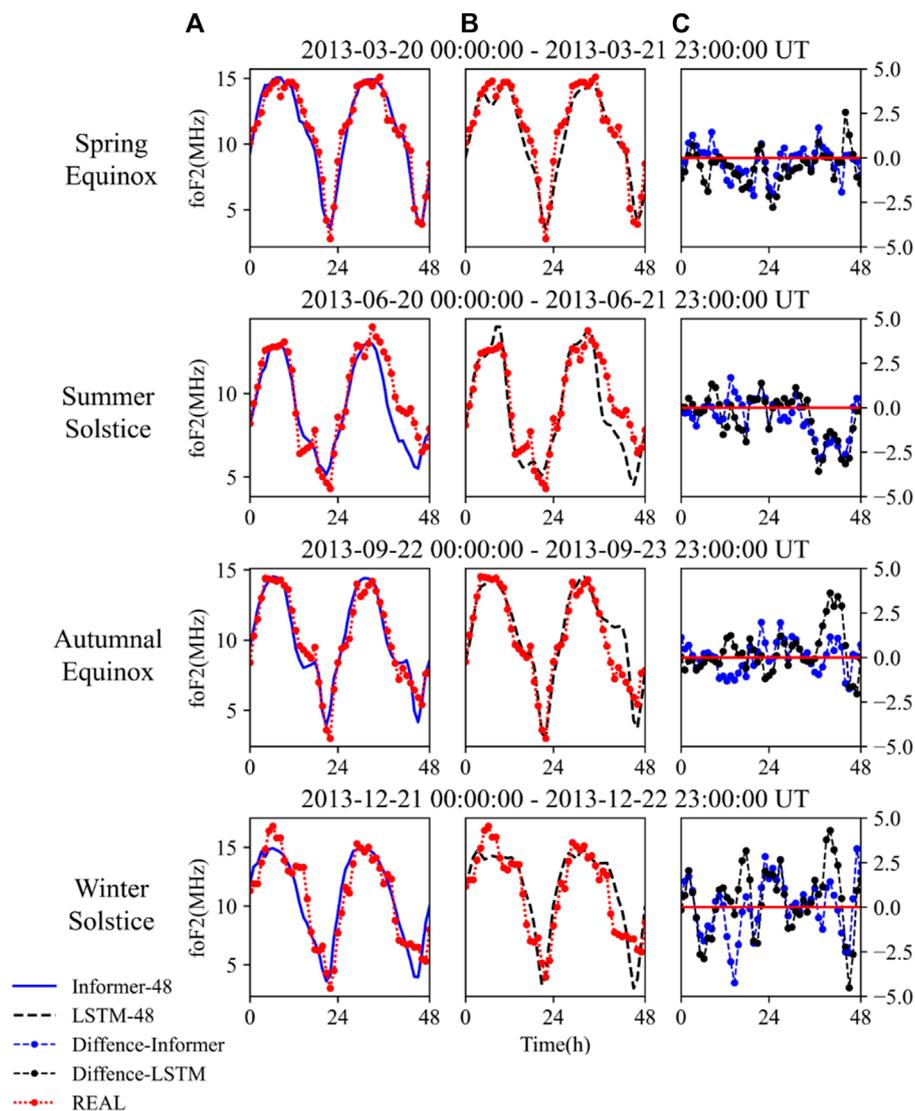


FIGURE 5

Figure shows the comparison between the Informer-48 and LSTM-48 model predictions against the actual values around the spring equinox, summer solstice, autumnal equinox, and winter solstice. (A) shows the comparison of the Informer model with the actual observations of foF2. (B) shows the comparison between the predicted values of the LSTM model and the measured values of foF2. (C) shows the difference between the predicted and measured values of the two models.

highlighted by the gray-shaded area 1 in the diagram, the LSTM-48 model's predictions manifest a notable deviation from the actual measurements. Similarly, within gray-shaded area 2, the LSTM-48 model continues to demonstrate suboptimal performance, starkly contrasted by the Informer-48 model's precise mirroring of the actual foF2 trajectory. These specific intervals correspond to the periods of gradual decline from the day's peak foF2 values, where the LSTM's retention of historical temporal data may be the culprit of the observed predictive inaccuracies. This underscores the advantage of utilizing the one-step generative decoder inherent to the Informer-48 model.

We calculate the RMSE of models, as shown in Figures 6–8, to more visually display the performance of the mode during geomagnetic storm events and quiet period. The results are shown in Table 3. The calmer the geomagnetic activity, the

smaller the model prediction error will be. This is because the occurrence of geomagnetic storms is very sudden, and their effects are difficult to predict, and even introducing many parameters to assist in prediction will inevitably have a certain lag, affecting the model prediction results.

Combining various factors, the Informer model demonstrates great performance in the foF2 prediction task, while showing good stability and prediction accuracy in the long duration time series, and it captures general trends of variations during the storm.

6 Conclusion

In this paper, the Informer long-term time-series architecture is implemented to forecast the ionospheric foF2 variation in the

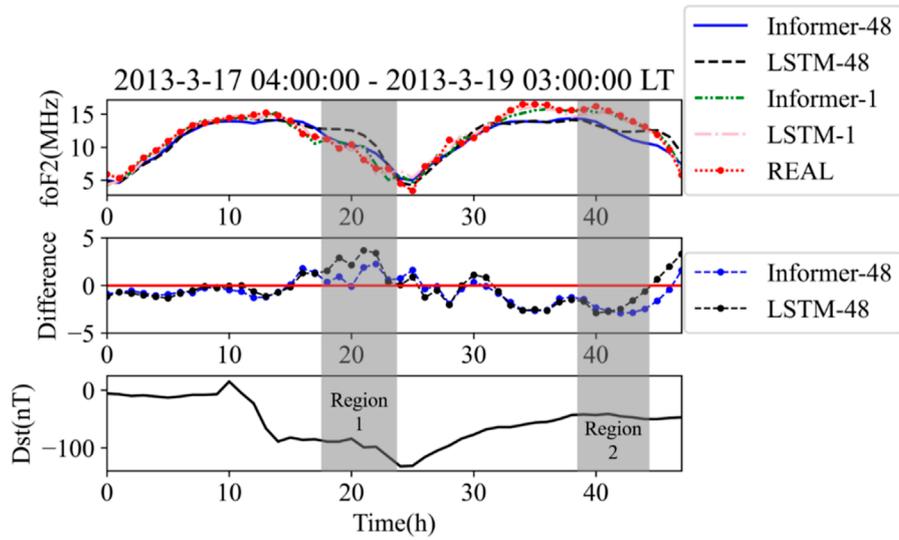


FIGURE 6 Comparison between the observed values and predicted foF2 values during the intense geomagnetic storm on 17 March 2013.

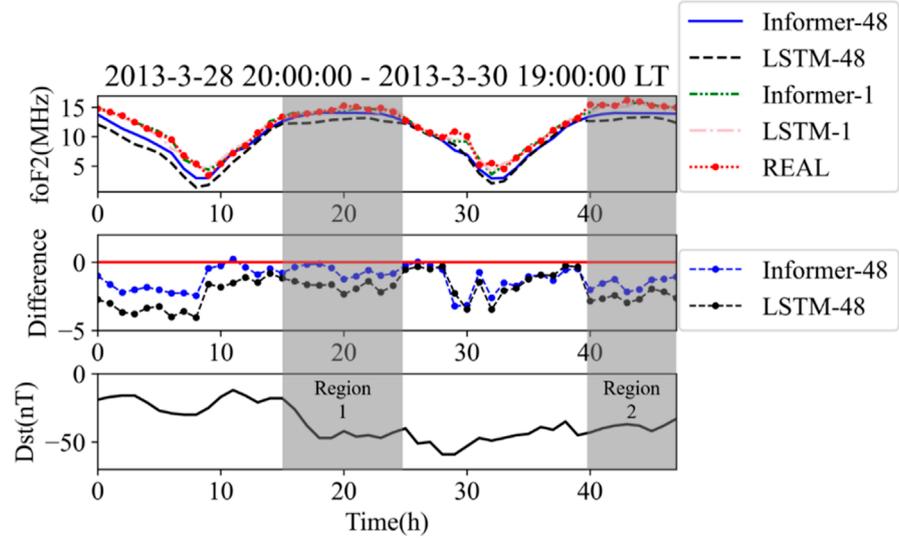


FIGURE 7 Comparison between the observed values and predicted foF2 values during the moderate geomagnetic storm on 28 March 2013.

low latitude. The model input is the ionospheric foF2 data at the Haikou station, China, along with space weather parameters Kp, F10.7, and Dst from 2006 to 2014. Prediction results are compared with the output values of traditional LSTM models and are analyzed in detail during moderate and intense geomagnetic storm events, separately. The overall performance of the model for all the events was discussed, and the conclusion was drawn that the prediction performance of the model is degraded at night. The analysis of events also demonstrated the superior advantage of the informer model in long-term (from hours to within 2 days) sequence prediction tasks. In the measurement examples used for verification, the Informer-foF2-48 h model

predicts foF2 parameters with the RMSE of 1.245 MHz and shows that the prediction accuracy maintains stability as the prediction window expands. The RMSE demonstrates that the proposed method performed well in long-term time-series prediction compared with other models and captured some variations in the ionospheric foF2. The prediction results of the proposed Informer-foF2 model provide insights into ionospheric foF2 prediction at low latitudes and long-term time-series prediction from hours up to 2 days. The Informer-based ionospheric foF2 prediction model proved that it can forecast the continuous change in ionospheric foF2 more accurately and more reliably than existing methods.

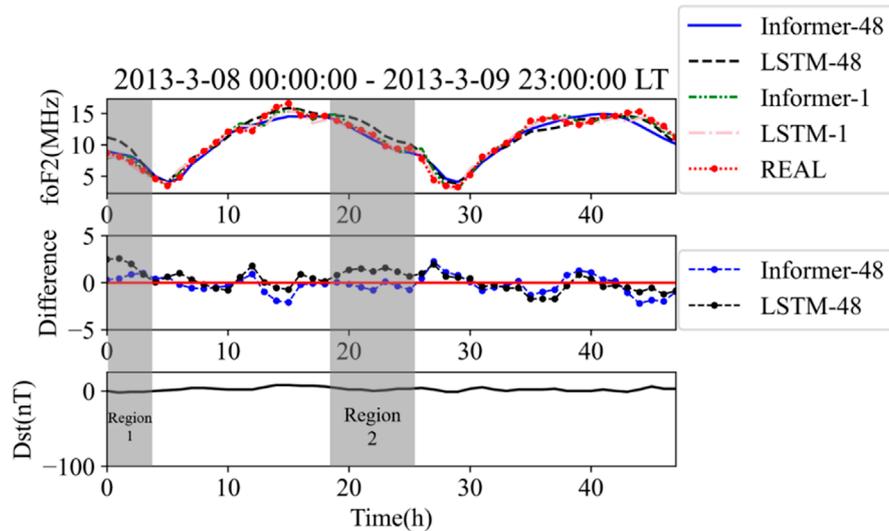


FIGURE 8 Comparison between the observed values and predicted foF2 values during the geomagnetic quiet period.

TABLE 3 RMSE score of 48-h models in events.

Model	RMSE—strong storm	RMSE—medium storm	RMSE—quiet period
LSTM-48	1.68	2.28	1.08
Informer-48	1.46	1.45	0.96
LSTM-1	0.81	0.63	0.71
Informer-1	0.80	0.58	0.69

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

FQ: writing—original draft and writing—review and editing. Z-YX: writing—review and editing. Q-HZ: writing—review and editing. H-BZ: writing—review and editing. S-RZ: writing—review and editing. YW: writing—review and editing. Y-ZM: writing—review and editing. DZ: writing—review and editing. SL: writing—review and editing. VM: writing—review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The work was supported by the National Natural Science Foundation of China (Grants 42325404, 42120104003, and 42204164), the Stable-

Support Scientific Project of China Research Institute of Radiowave Propagation (Grant A132312191), the Chinese Meridian Project, the China Postdoctoral Science Foundation (Grant 2021M701974), and the foundation of National Key Laboratory of Electromagnetic Environment (Grant No. 6142403230102).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Altinay, O., Tulunay, E., and Tulunay, Y. (1997). Forecasting of ionospheric critical frequency using neural networks. *Geophys. Res. Lett.* 24 (12), 1467–1470. doi:10.1029/97gl01381
- Bi, C., Ren, P., Yin, T., Zhang, Y., Li, B., and Xiang, Z. (2022). An informer architecture-based ionospheric foF2 model in the middle latitude region. *IEEE Geoscience Remote Sens. Lett.* 19, 1–5. doi:10.1109/lgrs.2022.3160422
- Cander, L. R., Milosavljevic, M. M., Stankovic, S. S., and Tomasevic, S. (1998). Ionospheric forecasting technique by artificial neural network. *Electron. Lett.* 34, 1573–1574. doi:10.1049/el:19981113
- Chen, Z., Jin, M., Deng, Y., Wang, J., Huang, H., Deng, X., et al. (2019). Improvement of a deep learning algorithm for total electron content maps: image completion. *J. Geophys. Res.* 124 (1), 790–800. doi:10.1029/2018ja026167
- Ergen, T., and Kozat, S. S. (2018). Online training of LSTM networks in distributed systems for variable length data sequences. *IEEE Trans. Neural Netw. Learn. Syst.* 29 (10), 5159–5165. doi:10.1109/tnnls.2017.2770179
- Fan, J., Liu, C., Lv, Y., Han, J., and Wang, J. (2019). A short-term forecast model of foF2 based on Elman neural network. *Appl. Sci.* 9 (14), 2782. doi:10.3390/app9142782
- Feng, J., Zhou, Y., Zhou, Y., Gao, S., Zhou, C., Tang, Q., et al. (2021). Ionospheric response to the 17 March and 22 June 2015 geomagnetic storms over Wuhan region using GNSS-based tomographic technique. *Adv. Space Res.* 67 (1), 111–121. doi:10.1016/j.asr.2020.10.008
- Gao, C., Jin, S., and Yuan, L. (2020). Ionospheric responses to the June 2015 geomagnetic storm from ground and LEO GNSS observations. *Remote Sens.* 12 (14), 2200. doi:10.3390/rs12142200
- Hu, A., and Zhang, K. (2018). Using bidirectional long short-term memory method for the height of F2 peak forecasting from ionosonde measurements in the Australian Region. *Remote Sens.* 10 (10), 1658. doi:10.3390/rs10101658
- Kim, J.-H., Kwak, Y.-S., Kim, Y., Moon, S.-I., Jeong, S.-H., and Yun, J. (2021). Potential of regional ionosphere prediction using a long short-term memory deep-learning algorithm specialized for geomagnetic storm period. *Space weather.* 19, e2021SW002741. doi:10.1029/2021sw002741
- Kim, J. H., Kwak, Y. S., Kim, Y. H., Moon, S. I., Jeong, S. H., and Yun, J. Y. (2020). Regional ionospheric parameter estimation by assimilating the LSTM trained results into the Sami2 model. *Space weather.* 18 (10). doi:10.1029/2020sw002590
- Li, X., Zhou, C., Tang, Q., Zhao, J., Zhang, F., Xia, G., et al. (2021). Forecasting ionospheric foF2 based on deep learning method. *Remote Sens.* 13 (19), 3849. doi:10.3390/rs13193849
- Lissa, D., Srinivasu, V. K. D., Prasad, D. S. V. V. D., and Niranjana, K. (2020). Ionospheric response to the 26 August 2018 geomagnetic storm using GPS-TEC observations along 80° E and 120° E longitudes in the Asian sector. *Adv. Space Res.* 66 (6), 1427–1440. doi:10.1016/j.asr.2020.05.025
- Liu, G., and Guo, J. (2019). Bidirectional LSTM with attention mechanism and convolutional layer for text classification. *Neurocomputing* 337, 325–338. doi:10.1016/j.neucom.2019.01.078
- Moon, S., Kim, Y. H., Kim, J. H., Kwak, Y. S., and Yoon, J. Y. (2020). Forecasting the ionospheric F2 parameters over jeju station (33.43°N, 126.30°E) by using long short-term memory. *J. Korean Phys. Soc.* 77 (12), 1265–1273. doi:10.3938/jkps.77.1265
- Rao, S. S., Chakraborty, M., and Pandey, R. (2018). Ionospheric variations over Chinese EIA region using foF2 and comparison with IRI-2016 model. *Adv. Space Res.* 62 (1), 84–93. doi:10.1016/j.asr.2018.04.009
- Rao, T. V., Sridhar, M., Ratnam, D. V., Harsha, P. B. S., and Srivani, I. (2021). A bidirectional long short-term memory-based ionospheric foF2 and hmF2 models for a single station in the low latitude region. *IEEE Geoscience Remote Sens. Lett.* 19 (99), 1–5. doi:10.1109/LGRS.2020.3045702
- Sai Gowtam, V., and Tulasi Ram, S. (2017). An artificial neural network-based ionospheric model to predict N_mF_2 and h_mF_2 using long-term data set of FORMOSAT-3/COSMIC radio occultation observations: preliminary results. *J. Geophys. Res. Space Phys.* 122 (11), 11743–11755. doi:10.1002/2017ja024795
- Shim, J. S., Tsagouri, I., Goncharenko, L., Rastaetter, L., Kuznetsova, M., Bilitza, D., et al. (2018). Validation of ionospheric specifications during geomagnetic storms: TEC and foF2 during the 2013 March storm event. *Space weather.* 16 (11), 1686–1701. doi:10.1029/2018sw002034
- Tang, J., Yang, D., and Ding, M. (2023). Forecasting ionospheric foF2 using bidirectional LSTM and attention mechanism. *Space weather.* 21, e2023SW003508. doi:10.1029/2023SW003508
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention is all you need. *Adv. neural Inf. Process. Syst.* doi:10.48550/arXiv.1706.03762
- Wang, R., Zhou, C., Deng, Z., Ni, B., and Zhao, Z. (2013). Predicting foF2 in the China region using the neural networks improved by the genetic algorithm. *J. Atmos. Solar-Terr. Phys.* 92, 7–17. doi:10.1016/j.jastp.2012.09.010
- Wang, P., Li, J., Han, B., ZeJun, H. U., Gao, X.Bo, Liu, J. J., et al. (2022). Modeling the polar ionospheric convection velocity vectors using shadow neural networks. *Chin. J. Geophys. (in Chinese)* 65 (4), 1197–1213. doi:10.6038/cjg2022P0255
- Wichaipanich, N., Hozumi, K., Supnithi, P., and Tsugawa, T. (2017). A comparison of neural network-based predictions of foF2 with the IRI-2012 model at conjugate points in Southeast Asia. *Advances in Space Research* 59 (12), 2934–2950. doi:10.1016/j.asr.2017.03.023
- Williscroft, L. A., and Poole, A. W. V. (1996). Neural networks, foF2, sunspot number and magnetic activity. *Geophysical Research Letters* 23 (24), 3659–3662. doi:10.1029/96gl03472
- Xia, G., Liu, M., Zhang, F., and Zhou, C. (2022). CAiTST: conv-attentional image time sequence transformer for ionospheric TEC maps forecast. *Remote Sensing* 14 (17), 4223. doi:10.3390/rs14174223
- Yang, Q., Xie, M., Su, H., Han, D., and He, Q. (2024). Statistical characteristics of multi-scale auroral arc width based on machine learning. *Journal of Geophysical Research Space Physics* 129, e2023JA031954. doi:10.1029/2023ja031954
- Zhang, B., Wang, Z., Shen, Y., Li, W., Xu, F., and Li, X. (2022). Evaluation of foF2 and hmF2 parameters of IRI-2016 model in different latitudes over China under high and low solar activity years. *Remote Sensing* 14 (4), 860. doi:10.3390/rs14040860
- Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., et al. (2021). Informer: beyond efficient transformer for long sequence time-series forecasting. *Proc. 35th AAAI conf. artif. intell.* 35, 11106–11115. doi:10.1609/aaai.v35i12.17325