Proceedings of the 2017 International Conference on Machine Learning and Cybernetics, Ningbo, China, 9-12 July, 2017

FORECASTING OF IONOSPHERIC VERTICAL TOTAL ELECTRON CONTENT (TEC) USING LSTM NETWORKS

WENQING SUN[1, 2], LONG XU[1], XIN HUANG[1], WEIQIANG ZHANG[2], TIANJIAO YUAN[3], ZHUO CHEN[4], YIHUA YAN[1]

[1]Key Laboratory of Solar Activity, National Astronomical Observatories, Chinese Academy of Sciences, Beijing, China
[2]College of Mathematics and Statistics, Shenzhen University, Shenzhen, China
[3]National Space Science Center, Chinese Academy of Sciences, Beijing, China
[4]School of Computer Science and Engineering, Nanyang Technological University, Singapore
EMAIL: 442305153@qq.com, lxu@nao.cas.cn, xhuang@bao.ac.cn, wqzhang@szu.edu.cn, yuantj19900407@163.com, chenzhuo.zoom@gmail.com, yyh@nao.cas.cn

Abstract:

Ionosphere is an important space environment near the earth. Its disturbance would result in severe propagation effects to radio information system, thus causing bad influences on communication, navigation, radar and so on. The total electron content (TEC) is an important parameter to present the disturbance of ionosphere, so TEC forecast is very meaningful in scientific research field. In this paper, we propose a long short-term memory (LSTM) based model to predict ionospheric vertical TEC of Beijing. The input of our model is a time sequence consisting of the vector of daily TECs and other closely related parameters. The output is TECs of future 24 hours. The result shows the root of mean square (RMS) error of test data can reach 3.50 and RMS error is less than this number during the period of low solar activity. Compared to multilayer perceptron network, LSTM is more promising and reliable to forecast ionospheric TEC.

Keywords:

Ionospheric; TEC; LSTM; Forecast

1. Introduction

The ionosphere is an important space environment near the earth. Its disturbance can cause severe propagation effects to radio information system, especially during ionospheric storm, thus influencing radio information system negatively, such as communication, navigation, radar and so on. [1, 2, 3] Total electron content (TEC) is an important parameter to describe the properties of ionosphere, and it is directly related to propagation effects, so TEC forecast is important in scientific research. TEC is defined by the integral of electron density in a column of $1\,m^2$ cross section along the single transmission path. The researchers have developed different models, such as the nonlinear radial basis function (RBF) neural network [4, 5], artificial neural network with a genetic algorithm [6,7] to simulate the ionospheric morphology. However, these models do not well consider the time sequential feature of TEC. Fortunately, recurrent neural networks (RNNs) are designed to process sequential data in deep learning.

Long Short-Term Memory (LSTM) is a specific recurrent neural network (RNN) architecture that was designed to model temporal sequences and their long-range dependencies [8]. The LSTM has been found extremely successfully in various sequence prediction and sequence labeling tasks, such as speech recognition [9, 10], handwritten generation [11, 12], machine translation [13]. In this paper, our objective is to forecast the TECs of future 24 hours from historical sequence data consisting of TEC and related parameters such as solar radio flux at 10.7 cm (F107) and magnetic activity index (ap). In another words, we want to figure out the correlation between sequences and sequences. Motivated by the extraordinary ability of LSTM to learn long dependencies of sequences, we try to apply LSTM to forecast ionospheric TEC.

The rest of this paper is organized as follows. In Section II, we provide a brief review of the LSTM and introduce the proposed network architecture. In Section III, we describe the experimental design, including configurations and results. Final section concludes this paper.

2. Proposed network

The proposed network architecture for predicting TEC consists of input layer, one LSTM layer, one hidden layer

978-1-5386-0408-3/17/$31.00 ©2017 IEEE

and output layer as shown in Fig. 1. In Fig. 1, every element of the input layer is a time sequence consisting of the vectors of TEC, ap and F107.

inner structure is shown in Fig. 2.

A LSTM layer map the input sequence $x=\left(x_1,x_2,\cdots,x_T\right)$ to an hidden output sequence $h=\left(h_1,h_2,\cdots,h_T\right)$. Its mapping process can be presented by following equations:

$$i_t=\sigma\left(W_{ix}\,x_t+W_{ih}h_{t-1}+b_i\right), \tag{1}$$

$$f_t=\sigma\left(W_{fx}\,x_t+W_{fh}h_{t-1}+b_f\right), \tag{2}$$

$$o_t=\sigma\left(W_{ox}\,x_t+W_{oh}h_{t-1}+b_o\right), \tag{3}$$

$$c_t=f_t*c_{t-1}+i_t*tanh\left(W_{ox}x_t+W_{ch}h_{t-1}+b_c\right), \tag{4}$$

$$h_t=o_t*tanh\left(c_t\right), \tag{5}$$

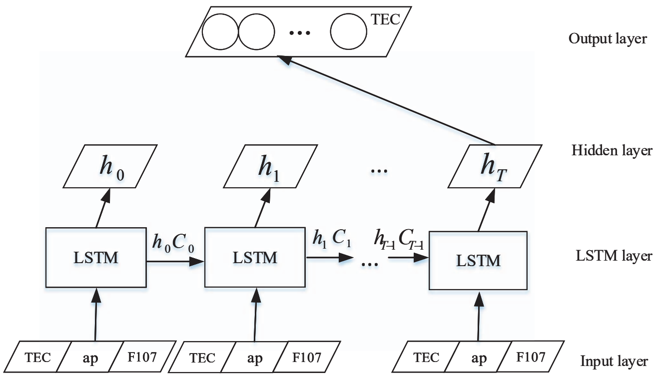


FIGURE 1. Network Architecture

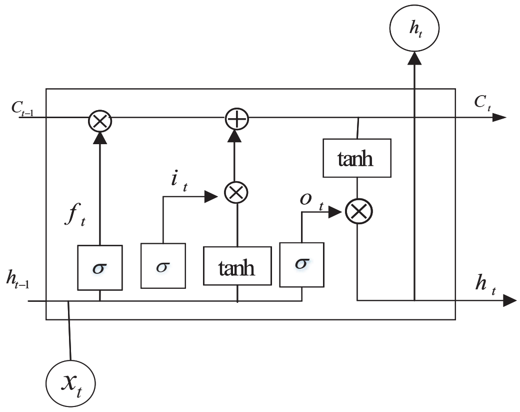


FIGURE 2. LSTM block

where i is the input gate which controls how much of the input to each hidden unit is written to the internal state vector C_t. f is the forget gate which determines how much of the previous internal state C_{t-1} is preserved. This combination of write and forget gates allows the network to control what information should be stored and overwritten across each time-step. o is the output gate controls how much of each unit's activation is preserved, it allows the LSTM cell to keep information that is not relevant later. C_t is the cell state which learn the long dependencies of historical sequential data. Final hidden state h can be the representation of sequence. After the processing of LSTM, we give the final outputs which are sequential vectors of TEC in future 24 hours. Time step in this model is five. In another words, we explore the relationship between historical five-day's sequence and future sequence. LSTM can is design to find out the long dependencies of sequence, that makes our model reasonable to our task.

3. Experiments

A. Configurations

The performance of network model depends on a good network structure setting. Thus, the primary task is to identify suitable input parameters and training process parameters for the TEC forecast model. Considering the influences of solar and geomagnetic activities to TEC, solar radio flux at 10.7 cm (F107), magnetic activity index (ap) and TEC index compose the input vector with dimension 33. TEC is collected by the Center for Orbit Determination in Europe (CODE) Grid Ionospheric Model (GIM) with the time resolution of an hour. F107 is provided by National Geophysical Data Center of National Oceanic and atmospheric administration (NOAA) with the time resolution of one day. (ftp://ftp.ngdc.noaa.gov/) ap index is provided by World Data Center for Geomagnetism, Kyoto with time resolution of 3 hours. (http://wdc.kugi.kyoto-u.ac.jp/index.html) The dataset is collected from Jan. 1, 1999 to Sept. 1, 2016. Of all data, the data of the years of 2001 and 2015 is used for testing, and the rest is used for training. A total of 6300 samples, 5800 of which are training data, and 500 of which are testing data.

As description above, model inputs are sequence vectors consisting of TEC, ap and F107, model outputs are sequence vectors of TEC whose dimension is 24. In order to evaluate the model performance, we compare LSTM and multilayer perceptron in three aspects. Firstly, we compare their training process, then we use root of mean square （RMS） error to evaluate the prediction accuracy, at last, we give the curves of predictions and observations to analyze their learning ability in different conditions.

B. *Performance Comparison*

We compare the performance of LSTM and multilayer perceptron (MLP) with one hidden layer in terms of training process, prediction accuracy and learning ability.

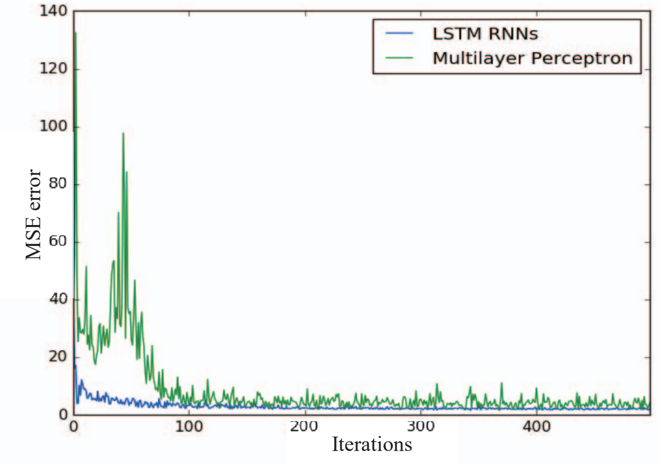


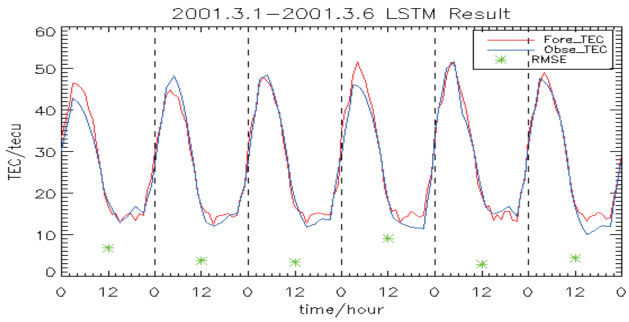
FIGURE 3. Learning curve

In Fig.3, horizontal and vertical coordinates are respectively the iterations and MSE error. And, the learning curves of them are illustrated in Fig. 3. Observing Fig. 3, it is obvious that the learning curve of LSTM decays faster than that of MLP. In addition, the convergence profile of LSTM is more stable than the other. Furthermore, the loss (RME error) of LSTM is always less than that of MLP.

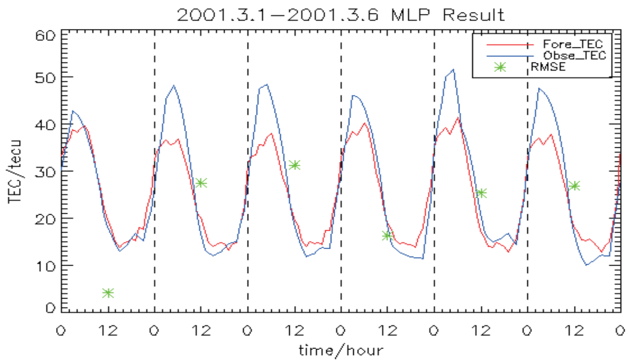
TABLE 1. RMS eroor

model	LSTM	MLP
RMS error	3.47998	5.02391

From Table I, we can observe that RMS error of LSTM is 3.47998, which is less than the MLP of 5.02391. It means LSTM is better than MLP in prediction. Learning ability analysis can be demonstrated by the curve comparison of forecast and observation TEC. We presented the prediction of six days from 2001/3/1 to 2001/3/6 in Fig. 4. It can be observed that the TEC curve forecasted by LSTM is closely fit for the observed TEC curve which is the ground truth, while there is considerable gap between the curve forecasted by MLP and the ground truth. It indicates MLP is inferior to LSTM in TEC forecast. Nevertheless, the tendency of forecast is right.



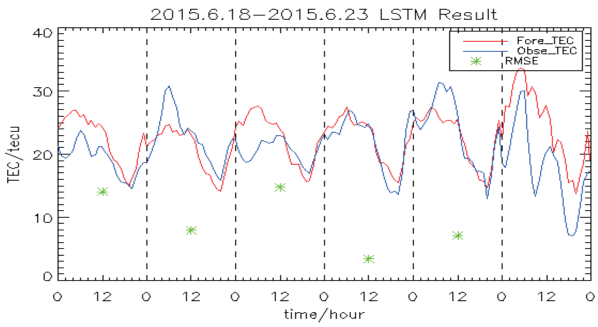
(a) Forecast of LSTM



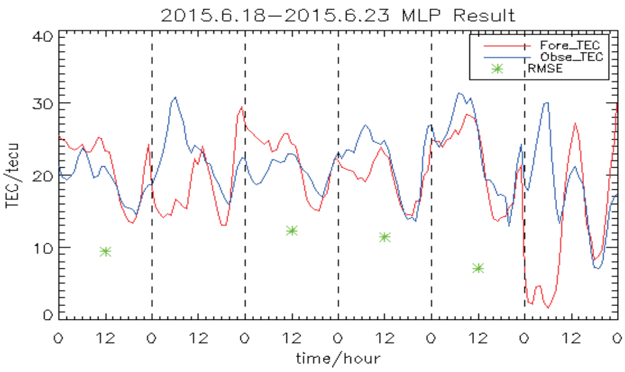
(b) Forecast of MLP

FIGURE 4. TEC prediction for 2001/3/1-2001/3/6

From 2015/6/18-2015/6/23, variation of TEC is very complicated. Beside the periodic variation, there are strong disturbance during one period. On this occasion, LSTM can capture the cyclic changes, and therefore give a better forecast, however, MLP is totally wrong with an inverse direction. Comparing to MLP, LSTM have the advantage of predicting long sequence data due to the memory cell. LSTM can learn the long dependencies of sequential data, not only the very past moment, but also a segment of history is taken into account. While MLP does not utilize history information, so it may fail in case of turbulence of TEC.



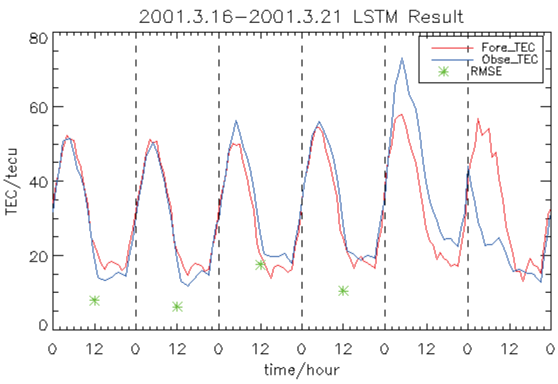
(a) Forecast of LSTM



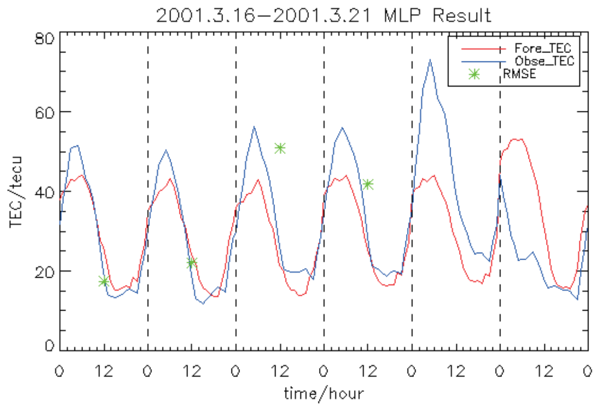
(b) Forecast of MLP

FIGURE 5. TEC prediction for 2015/6/18-2015/6/23

In case TEC varies suddenly, such as the case illustrated in Fig. 6, where the peaks of observed TEC in the five past days are large, while it dramatically becomes low in the next day. In this situation, RMS error becomes significant.



(a) Forecast of LSTM



(b) Forecast of MLP

FIGURE 6. TEC prediction as dramtic variation of peak

Above, we discussed three cases of TEC changes, where LSTM all behave better than MLP. The achievement owes to the special design of LSTM, so that it can learn long dependences of sequential data. Thus, the interaction and relationship of the elements of sequential data can be learnt, and the better representation of input data can be obtained.

4. Conclusions

This paper presented a model of LSTM RNN based for TEC forecast of the ionosphere. The experimental results showed that the proposed model outperforms existing methods with more stable convergence tendency, less RMS error and can give good prediction in different condition. This outperformance of the proposed model on TEC forecast once again verified the superiority of LSTM on processing time sequence. In the near future, we will go further for improving the performance on TEC forecast during ionospheric storm.

Acknowledgements

This work was partially supported by a grant from the National Natural Science Foundation of China under Grant 61572461, 61472257, 11433006, 11303051 and CAS 100-Talents (Dr. Xu Long)

References

[1] Schaer S. Mapping and predicting the Earth's ionosphere using the Global Positioning System.

Dissertation of Doctoral Degree. Bern: Astronom Inst, Univ Bern, 1999

[2] Tan S S, Zhuo B, Guo S T, et al. Research on COMPASS navigation signals of China. China Space Sci Tech, 2011, 31: 9-14

[3] Tan S S, Development and thought of compass navigation satellite system J Astronautics, 2008, 29: 391-396

[4] Yilmaz A Akdogan K E, Gurun M. Regional TEC mapping using neural networks. Radio Sci, 2009, 44: RS3007

[5] Habarulema J B, McKinnell L A, Cilliers P J, et al. Application of neural networks to South African GPS TEC modeling. Adv Space RES, 2009, 43: 1711-1720

[6] Liu L,WantedW,Chen Y, et al. Solar activity effects of the ionosphere: A brief review. Chin Sci Bull, 2011, 56: 1202-1211

[7] Huang, Z., Yuan, H., 2014. Ionospheric signal-station TEC short-term forecast using RBF neural network. Radio Sci. 49, 283-292. http://dx.doi.org/10.1002/2013RS005247

[8] S. Gochreiter and J. Schmidbuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735-1780, Nov. 1997

[9] Alex Graves and Juergen Schmidhuber, "Offline Handwriting Recognition with Multidimensional Recurrent Neural Networks," in NIPS. 2009.

[10] A. Graves and N. Jaitly. Towards end-to-end speech recognition with recurrent neural networks. In Proceedings of the 31st International Conference on Machine Learning (ICML-14), pages 1764–1772, 2014. 2, 3

[11] Mohamed, Abdel-rahman, George E. Dahl, and Geoffrey Hinton. "Acoustic modeling using deep belief networks." Audio, Speech, and Language Processing, IEEE Transactions on 20.1 (2012): 14-22.

[12] Collobert, Ronan, et al. "Natural language processing (almost) from scratch." The Journal of Machine Learning Research 12 (2011): 2493-2537.

[13] Cho, Kyunghyun, van Merrienboer, Bart, Gulcehre, Caglar, Bougares, Fethi, Schwenk, Holger, and Bengio, Yoshua. Learning phrase representations using rnn encoder-decoder for statistical machine translation. ArXiv preprint arXiv:1406.1078, 2014.

[14] Y bengio, P.Simard, and P. Frasconi, "Learning long-term dependencies with gridient descent is difficult," IEEE Trans. Neural Networks, vol. 5,pp. 157-166, 1994