Prediction high frequency parameters based on neural network

Wenhe Zhang¹, Guoce Huang¹, Guisheng Wang¹ and Ye Wang¹*

¹Graduate College, Air Force Engineering University, Xi’an 710058, China
²Information and Navigation College, Air Force Engineering University, Xi’an 710058, China
³Corresponding author’s e-mail: 1065262386@qq.com

Abstract. Aiming at the shortcomings of the current high frequency communication frequency parameter prediction method, the frequency parameter prediction method based on Gated Recurrent Unit Recurrent Neural Networks (GRU RNN) is proposed. Through the analysis of the ionospheric parameter f0F2 data, the GRU can predict the f0F2 value by long-term memory of the historical data when processing the time series related data. Compared with other prediction methods, the error between the predicted value and the true value is only 2%. The research results show that the model to predict the f0F2 value in advance is feasible.

1. Introduction

High frequency communication has the advantages of over-the-horizon, low-cost, strong invulnerability. Due to the time-varying characteristics of shortwave channel, its availability and reliability are greatly affected by frequency variation, so the selection of frequency is crucial. By predicting the ionospheric parameters, it can provide guidance for selecting the appropriate shortwave communication frequency, thus improving the short-wave communication quality.

Currently frequency prediction methods are divided into long-term prediction and short-term prediction. The long-term prediction models is established based on early historical data, they use mathematical formulas to calculate ionospheric parameters. Representative methods include the ITU CCIR Early Recommendations (340 Report) and ITU-RP.533 methods, as well as the AGA prediction method [1]. Based on these methods, some institutions developed software to predict ionospheric parameters such as REC533 and VOACAP [2]. The short-term forecast is based on recent historical data, and the predicted frequency value is obtained by mathematical methods such as averaging. With the development of artificial neural networks [3, 4, 5], artificial neural networks show good characteristics in predicting nonlinear data [6]. Therefore, researchers use artificial neural networks to predict nonlinear short-wave ionospheric parameters. The earliest method is to use the back propagation neural network [7,8] to predict the mid-month value of f0F2, the minimum error is about 0.3MHz; A method combining chaos and neural network is proposed to predict the frequency parameters of short-wave communication. The chaotic method is used to reconstruct the attractor of phase space system [9]; The literature [10] proposed a prediction method based on the evolution of wavelet neural network and chaos theory, using the smoothing effect of wavelet algorithm to reduce the error. Literature [11] proposed a technique combining phase space reconstruction and fuzzy wavelet neural network [12-15], and using singular value decomposition to denoise historical data in the data preprocessing stage, the prediction accuracy is greatly improved. However, these methods...
only consider the monthly law value change of $f_0F_2$. The law of $f_0F_2$ mid-month changes greatly at dawn and dusk, and the change of the adjacent days are not involved; at the same time, back-propagation neural networks have local minimization, slow convergence, gradient disappearance or explosion, which all cause interference to the final prediction effect and accuracy.

For the above problems, considering the relation between sequential data and the disappearance of gradient in back propagation, this paper proposes an improved approach of Gated Recurrent Unit (GRU) neural network to predict ionospheric parameters.

2. GRU model construction

Recurrent Neural Networks (RNN) is a widely used artificial neural network model, which is good at processing sequential data. The basic structure is composed of three parts: input layer, hidden layer and output layer. Its structure is shown in figure 1. Unlike traditional neural networks, RNN has "memorability", in which the loop in the hidden layer can transmit historical information backward, making the current moment closely related to historical information. Using this feature, RNN can process sequences composed of any moment. Neurons in RNN are connected by weight $U$ between neurons, they form a loop to train sequence data and predict output. Different from traditional neural networks, each layer in RNN shares parameters $W$, $U$ and $V$, which reduces the number of parameters to be learned in the network, but the input of each step is different. When new data is input, the information from time $t-1$ is calculated to become the output at the current moment, and each data sample is processed according to this process.

![Figure 1. The structural model of RNN](image1)

The mathematical description of RNN is shown in equations (1) and (2). In the formula, $W$, $U$, $V$ are the weight parameters, $s_t$ is the $t$ step state of the hidden layer, it is the memory unit of the network, $y_t$ is the output of the $t$ step, $\phi$ is the nonlinear activation function, and generally it is the tanh function.

$$s_t = \phi(Wx_t + Us_{t-1}) \quad (1)$$

$$y_t = \text{soft max}(Vs_t) \quad (2)$$

![Figure 2. The hidden layer structure of GRU](image2)

The external structure of GRU is the same as the traditional RNN structure, with only the improvement of its hidden layer, which has the function of long-term memory of important historical information. In this way, when back propagating, the gradient will not disappear. Its hidden layer structure is shown in figure 2. The GRU network has only two gate structures: the update gate and the reset gate. Compared with other improved circular neural networks, GRU structure is simpler and faster.
At time t, the input data and the hidden state transmitted from the previous moment are updated through the gate mechanism to obtain the hidden state and memory contents of the current moment. In this way, according to the weight parameters at each moment, how much information at the previous moment needs to be retained and transmitted can be determined. The mathematical description is as follows:

We input data $x_t$ and hidden state $h_{t-1}$ to get state $r_t$ by the reset gate, as shown in (3).

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t])$$  \hspace{1cm} (3)$$

We input data $x_t$ and hidden state $h_{t-1}$ to get state $z_t$ by the update gate, indicating how much of the previous hidden state information needs to be transferred to the current hidden state $h_t$, as shown in equation (4).

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t])$$  \hspace{1cm} (4)$$

After that, the network determines the memory content of the current moment, that is, the important information of the previous moment is recorded, as shown in equation (5).

$$\tilde{h}_t = \tanh(W_r \cdot [r_t \cdot h_{t-1}, x_t])$$  \hspace{1cm} (5)$$

Finally, the network computes the information vector $h_t$ to be passed to the next unit, as shown in equation (6).

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t$$  \hspace{1cm} (6)$$

After iterative calculation, the final output of the network is $y_t$, as shown in equation (7).

$$y_t = \sigma(W_o \cdot h_t)$$  \hspace{1cm} (7)$$

Through this algorithm, the problem of gradient disappearance or explosion in traditional neural network training is solved, and long-term learning of sequence data is realized. In this paper, the two-layer GRU network is used to predict the ionospheric parameters, with 64 neurons in each layer, which greatly reduces the training time. The input dimension of the first layer network is set as 1, the output dimension is 50, the input dimension of the second layer network is 50, and the activation functions of both layers are tanh functions. Finally, the network outputs a one-dimensional solution. MSE is used as loss function and RMSprop as optimization function to make the model converge as soon as possible. GRU network model is divided into training process and prediction process. Firstly, the program reads data and preprocesses the data, which is divided into test set and training set, and the training set is used to train the model. After training the model, start the prediction, and the results are compared with the real values in the test set to analyze the prediction effect.

3. Prediction of Ionospheric Parameters

The main factors which affect the selection of shortwave communication frequency are $f_0F2$ and $M(3000)F2$ of the ionospheric parameters, and the variation law has nonlinear characteristics. In this regard, we use the actual observed data to train the GRU network to learn the optimal network parameters, and determine the prediction model. In order to verify the usability and prediction accuracy of the model, the simulation experiment uses the monthly median $f_0F2$ data which measured in Beijing in January 2011. The data set has a total of 775 hours as shown in Figure 3. When the data is pre-processed, the missing data points are filled with the corresponding monthly average value. The first 80% of the total data is 620 hours as the training set, and the last 20% is 155 data as the test set, as shown in Table 1. First we normalize the data to enhance the stability of the data set and convert the data into a numpy array for use by the model. Secondly, we started to train the model, inputting 128 data each time, every 24 data as a sequence, and the number of iterations is 400 generations. Finally, after the training, the trained model is used to predict the results, and the predicted values and the true values in the test set are compared to test the accuracy of the model.

Table 1 Training and test data distribution
Table 1. Assignment of the training and testing data

| Data number | Training number | Testing number |
|-------------|-----------------|----------------|
| 778         | 620             | 155            |

Figure 3. Observation of f0F2 in Beijing

4. Simulation analysis

Footnotes should be avoided whenever possible. If required they should be used only for brief notes that do not fit conveniently into the text.

The GRU neural network prediction results and real measured values are shown in Fig. 4. The vertical axis is the frequency value of the predicted sample in MHz, and the horizontal axis is the predicted time point in hours/h. It can be seen that the model has a good prediction effect, especially in the previous predictions and the change trend is well predicted at the position where the change is faster. It can be clearly seen that the shorter the prediction time is, the smaller the error is. On the contrary, the longer the prediction time is, the larger the relative error is, and the prediction value is less reliable.

The mean square error (MSE) between the predicted and actual values are used to describe the accuracy of the prediction,

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (observed_i - predicted_i)^2$$  \hspace{1cm} (8)

Figure 4 GRU Neural network model prediction results

In order to reflect the advantages of the model, the f0F2 value is predicted by using GRU network, BP neural network and Fuzzy Wavelet Neural Network (FWN). By comparing the model complexity and precision, it is found that the two-layer GRU network model has higher prediction accuracy, faster convergence rate and better global convergence, which is superior to the other two models in all aspects. The parameters of models are shown in Table 2.

Table 2. Comparison of various algorithms

| Algorithm | MSE  | convergence            | convergence rate |
|-----------|------|------------------------|------------------|
| BP        | 0.096| Bad global convergence | Training 2000 times |
5. Conclusion
The deep learning algorithm is a breakthrough in the field of artificial neural networks. By combining low-level features to form more abstract high-level features, the data distribution characteristics are found, which show good prediction results in various data predictions. In order to predict the shortwave communication frequency, the researchers have proposed many prediction methods, this paper proposes to use the GRU network to construct the model to predict the ionospheric parameters, and use the f0F2 data to train and to test, and the experimental results verify the feasibility of the proposed method. The proposed method has the advantages of fast convergence and high prediction accuracy, and provides a new method for predicting short-wave frequencies.

Because many factors can change the ionospheric parameters, this paper does not summarize other factors. In this paper, the GRU model only predicts and verifies the change of median value f0F2, and the subsequent work needs to add more factors affecting the short-wave frequency and further improve the prediction accuracy of the model.

Acknowledgments
I appreciate my college which gives me a comfortable learning atmosphere. And I am also greatly indebted to all my co-authors who have helped me to develop the fundamental and essential academic competence.

References
[1] LI Z Q, SUN X R. (1992) YaDa prediction compared with CCIR recommendation method (340 report) [J]. Chinese Journal of Radio Science, 7(2):17-29.
[2] BARCLAY L, BEHM C, CARROLL S. (2019) Digitally modulated HF communications reliability: modifications to ITU-R Rec.P.533 propagation model and the associated computer program REC533. The Institution of Engineering and Technology, International Conference on Ionospheric Radio Systems and Techniques.
[3] Henry Leung, Wang L. (2001) Prediction of noisy chaotic time series using an optimal radial basis function neural network. IEEE Transactions on Neural Networks.
[4] NATH S, KOTAL S D,KUNDU P K. (2015) Seasonal prediction of tropical cyclone activity over the North Indian Ocean using the neural network model. Atmosfera.
[5] F. Azmat, Y. Chen, and N. Stocks, (2016) Analysis of spectrum occupancy using machine learning algorithms, IEEE Trans. Veh. Technol., vol. 65, no. 9, pp. 6853-6860, Sep.
[6] HAJIHASSANI M,ARMAGHANI D J,MARTO A, et al. (2015) Ground vibration prediction in quarry blasting through an artificial neural network optimized by imperialist competitive algorithm. Bulletin of Engineering Geology and the Environment.
[7] XUE J D, JI X Z, TAO K, et al. (2014) Prediction of cement filling materials performance using improved BP neural network [J]. Journal of Chemical & Pharmaceutical Research, 30: 207 -214.
[8] ZHANG X, LI M Q, WANG C. (2017) A prediction method for HF communication channel based on BP neural network [J]. Industrial Control Computer, 30(10):55-56+59.
[9] JIAN X C, ZHANG J L. (2001) Prediction of frequency parameters in short wave radio communication based on chaos and neural networks [J]. Journal of Tsinghua University (Science and Technology), 41 (1):16-19.
[10] ZHAO D Q, CHEN P G, SUN G M, DUAN J Y. (2018) Prediction of best Frequency parameters of HF communication based on MEA-WNN [J]. Journal of Beijing University of Technology, 44(02): 215-219.
[11] REN S T, GUO L. (2011) A prediction method for HF radio communications frequency based on FWNN [J]. Communication Technology, 44(4): 37-39.

[12] LAI J, WANG H, LIU X, et al. (2007) A quantitative prediction method of network security situation based on wavelet neural network[C] The First International Symposium on Data, Privacy, and E-Commerce, New York: IEEE, 2007: 197 -202.

[13] R. VENKATA RAMANA, B. KRISHNA, S. R. KUMAR, N. G. PANDEY. (2013) Monthly Rainfall Prediction Using Wavelet Neural Network Analysis. Water Resources Management.

[14] LAI J B, WANG H Q, LIU X W, LIANG Y. (2007) A Quantitative Prediction Method of Network Security Situation Based on Wavelet Neural Network. First International Symposium on Data, Privacy and E-Commerce.

[15] Che G, Luh, P. B., Michel, L.D, Wang Y Q, Friedland, P.B. (2013) Very Short-Term Load Forecasting: Wavelet Neural Networks With Data Pre-Filtering. Power Systems, IEEE Transactions on.