International Journal of Scientific Engineering and Research (IJSER)
ISSN (Online): 2347-3878
Impact Factor (2022): 7.741

Short-Term Wind Speed Prediction Based on MI-BI-LSTM

Baolin Ma[1]

[1]North China Electric Power University, School of Control and Computer Engineering
No. 2 Beinong Road, Changping District, Beijing, China
1502945751 [at]qq.com

Abstract: ***Wind speed has randomness, volatility and uncontrollable factors, and with accurate wind speed prediction, wind turbines can be more efficiently controlled, and the grid-connection effects of wind generators can be reduced. The bidirectional long short-term memory neural network can provide the output layer with complete contextual information of the past and future of each point in the input sequence, and the wind speed is a time series signal coupled with some simple signals that change with time. Based on the above information, this paper proposes a short-term wind speed prediction method based on Multiple Input Bidirectional Long Short Time Memory (MI-BI-LSTM). Pearson correlation coefficient was used to verify the correlation of different data features, and the multi-dimensional data features with strong correlation were used as the input of the model to establish the wind speed prediction model. The simulation results show that this network can better fit the variation trend of actual wind speed series and has better prediction performance.***

Keywords: wind speed prediction, Long and short term memory neural network, Multidimensional features

1. Introduction

China has pledged to achieve the long-term goal of "carbon peak 2030 and carbon neutral 2060", focusing on the development of clean energy industry and speeding up the green economic and social transformation, and building a new power system with new energy as the main body. As a kind of clean energy, wind power has far-reaching significance for promoting energy transformation [1]- [3]. Wind speed is affected by many external factors, such as ambient temperature, atmospheric pressure, ground roughness, etc. Due to the random fluctuation of wind speed, large-scale wind turbine grid integration will have a great impact on the power grid, and even affect the stable operation of the power grid in serious cases. Accurate wind speed prediction can timely regulate the yaw and paddling of wind turbines, improve power quality of power grid, and effectively ensure large-scale grid connection of renewable energy power generation [4]- [8].

The main wind speed prediction models are physical model and statistical model. There are two types of wind speed predictions: long-term and short-term. Long-term prediction is usually used for site selection, wind resource measurement and evaluation, and physical model method is suitable for long-term prediction. Short-term wind speed prediction is for wind turbine control and power system scheduling, and the prediction duration is generally a few minutes in the future. The short-term wind speed prediction can be achieved by establishing statistical models based on historical wind speed and direction data of wind farms and data of nearby weather stations [9-10].

Currently, the commonly used short-term wind speed prediction methods include Kalman filter [11], time series method [12-13], spatial correlation method [14]- [16], chaos theory [17] and artificial neural network [18]- [21]. Due to the continuous development of artificial intelligence, the complex nonlinear relationship between meteorological factors in traditional physical models has been improved by building machine learning and deep learning models. For example, support vector machine [22], multilayer perceptron and extreme learning machine can extract complex uncertain features in wind speed series. Deep learning model automatically obtains abstract features and hidden structures inherent in data analysis from the bottom to the top through distributed hierarchical feature extraction. Cyclic neural network is a kind of neural network specially used to solve the problem of time series data analysis. It has higher nonlinear mapping ability and can extend the model to various forms and have certain generalization ability by using parameter sharing. LSTM network is an improved network of RNN. It replaces neurons in the hidden layer of RNN with memory units to control the selective passing of information and realize the memory of past information.

For single input model can extract information from a one-dimensional wind speed time series, and unidirectional short-term memory neural network can only extract before to the spread of the time series data, this paper proposes a multiple input and two-way short-term memory neural network based model to realize the short-term forecast wind speed, the change laws of the current wind speed before and after the point in time. By comparing with different neural network models, experimental results verify the accuracy of wind speed prediction and improve the ability of short-term wind speed prediction.

2. Basic Model

2.1 Long and short term memory model (LSTM)

Time series is a random variable, which is a complex signal coupled with simple signals that change over time. In essence, it reflects the trend of variables changing over time, and each data node has sequential correlation in time. In traditional forward neural network, the network structure itself can't reflect the data node with the correlation of time sequence, and loops neural network model [22] through the connection between the hidden layer nodes, can good description data in time series, the order of the connection between the structurally more conform to the requirements of the time series prediction. However, RNN lacks selectivity in the use of sequence information, and too long sequence length of input features is prone to the problem of gradient dispersion.

LSTM network model [23-24] is an improved RNN model, which can automatically save and delete part of time states and extract more complex feature relations in time series. It is suitable for solving complex time series prediction problems. The basic elements of the LSTM network model are shown in figure 1.

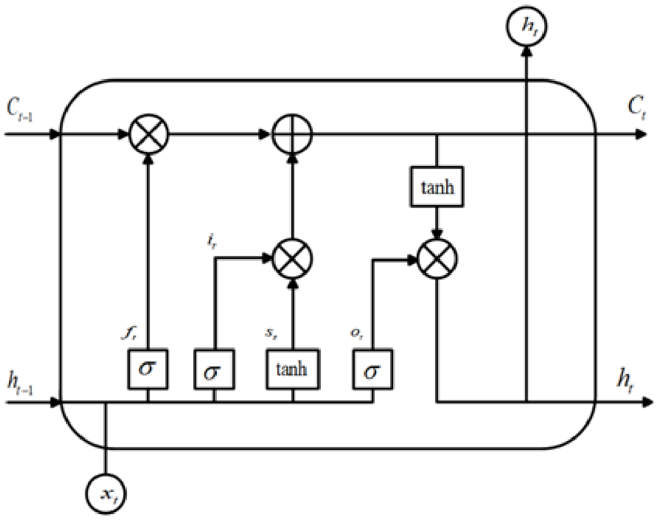

Figure 1: LSTM network structure

In figure 1, f_t, i_t, o_t are the states of forgetting gate, input gate and output gate respectively; s_t is the status of the input node; h_t, h_{t-1} is the output of the current time and the previous time respectively; x_t is the input of the current sequence; C_t, C_{t-1} is the state information of the memory unit at the current time and the previous time respectively. The basic unit of LSTM network includes gate structure and memory unit. The information of memory unit can be added or deleted by using forgetting gate, input gate and output gate. The inputs of the basic unit of the LSTM network include x_t, C_{t-1} and h_{t-1}. In the forgetting gate, x_t and h_{t-1} jointly determine the forgetting part of C_{t-1}. In the input gate, x_t, h_{t-1} determines the update part of the current memory unit state information after Sigmod and tanh function transformation. Update C_{t-1} to C_t by forgetting gate and input gate. In the output gate, x_t and h_{t-1} are transformed into o_t by Sigmod function, and C_{t-1} is transformed into o_t by tanh function. The specific calculation process of this unit is as follows:

$$f_t = \sigma(W_f \bullet [h_{t-1}, x_t] + b_f) \quad (1)$$
$$i_t = \sigma(W_i \bullet [h_{t-1}, x_t] + b_i) \quad (2)$$
$$s_t = \tanh(W_s \bullet [h_{t-1}, x_t] + b_s) \quad (3)$$
$$C_t = f_t \circ C_{t-1} + i_t \circ s_t \quad (4)$$
$$o_t = \sigma(W_o \bullet [h_{t-1}, x_t] + b_o) \quad (5)$$
$$h_t = o_t \bullet \tanh(C_t) \quad (6)$$

Input gate, input node, forgetting gate and output gate, respectively, are represented by matrices W_f, b_f, W_i, b_i, W_s, b_s and W_o, b_o. $\sigma(\cdot)$ and $\tanh(\cdot)$ are Sigmod functions and hyperbolic tangent functions respectively.

2.2 Bidirectional LSTM network model

LSTM and RNN network models only transmit information from the front to the back, only realize model prediction through one-way data law, can not fully extract the time characteristic information in the time series, two-way LSTM model uses forward and backward data to predict the current data, strengthen the connection between characteristic information and predicted results. It can improve the prediction accuracy of the model [25]. The hidden layer of BI-LSTM is composed of two unrelated LSTM, namely, the forward hidden vector $\overrightarrow{h}_t$ and the backward hidden vector $\overleftarrow{h}_t$. The forward output sequence and reverse output sequence of LSTM model are taken as the input of BI-LSTM output layer and transferred to the node of the next layer of the network after the splicing of the output layer. y_t is the final output value. The model structure is shown in figure 2.

$$\overrightarrow{h}_t = H(W_{x\overrightarrow{h}} x_t + W_{\overrightarrow{h}\overrightarrow{h}} \overrightarrow{h}_{t-1} + b_{\overrightarrow{h}}) \quad (7)$$
$$\overleftarrow{h}_t = H(W_{x\overleftarrow{h}} x_t + W_{\overleftarrow{h}\overleftarrow{h}} \overleftarrow{h}_{t-1} + b_{\overleftarrow{h}}) \quad (8)$$
$$y_t = \Phi(W_{\overrightarrow{h}y} \overrightarrow{h}_t + W_{\overleftarrow{h}y} \overleftarrow{h}_t + b_y) \quad (9)$$

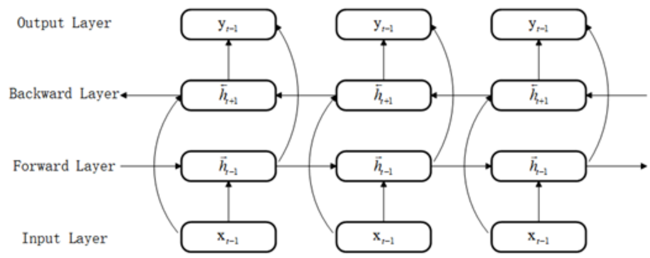

Figure 2: Bi-LSTM model structure

3. Short-term wind speed prediction model based on MI-BL-LSTM

When evaluating wind energy resources, air density, average wind speed, turbulence intensity, wind shear index and other indicators need to be considered. Wind shear index refers to the change trend of wind speed increasing with the increase of height, which is caused by the effect of ground roughness on wind speed.

Wind measuring device in wind farm can collect environmental temperature, pressure, relative humidity and wind speed size and direction at different heights at the same time, and select some features with strong correlation as the input data of the model by using Pearson correlation principle. Assume that a set of wind speed data measured by the wind measuring device within a period of time is $X_N = [x_1, x_2, \Lambda\, x_m, \Lambda, x_N]$ (N is the number of sampling points), where $x_m = [x_{m1}, x_{m2}, x_{m3}, \cdots, x_{mM}]$ (M is the number of features of sampling data), the first K data are determined as the training set of the network model, and the last N-K data are used to test the training results of the network model.

$$X_{tr} = \begin{bmatrix} x_1 & x_2 & \cdots & x_L \\ x_2 & x_3 & \cdots & x_{L+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{K-L} & x_{(K-L)\times L} & \cdots & x_{K-1} \end{bmatrix}_{(K-L)\times L} \tag{10}$$

$$Y_{tr} = \begin{bmatrix} x_{L+1} \\ x_{L+2} \\ \vdots \\ x_K \end{bmatrix}_{(K-L)\times 1} \tag{11}$$

$$X_T = \begin{bmatrix} x_{K-L+1} & x_{K-L+2} & \cdots & x_K \\ x_{K-L+2} & x_{K-L+3} & \cdots & x_{K-L+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N-L} & x_{N-L+1} & \cdots & x_K \end{bmatrix}_{(N-K)\times L} \tag{12}$$

Where L is the time step required by the network model, X_{tr} is the input data of the model, Y_{tr} is the target output of the model, and X_T is the model test set.

For forward training of BI-LSTM model, input $X_{tr}(i)$ is substituted into Equations (1)- (6) to obtain the first group of state output $\{h_i^f, h_{i+1}^f, \cdots, h_{i+L}^f\}$; Similarly, for reverse training, input $X_{tr}(i)$ is substituted into Equations (1)- (6) in reverse order, and a set of state output is obtained as $\{h_{i+L}^b, h_{i+L-1}^b, \cdots, h_i^b\}$. The forward and reverse output results are spliced to obtain $H_i \in R^{2\times L}$.

$$Hi = \left\{\{h_i^f, h_{i+L}^b\}, \{h_{i+1}^f, h_{i+L-1}^b\}, \cdots, \{h_{i+L}^f, h_i^b\}\right\} \tag{13}$$

The training and testing process of MI-BI-LSTM model is shown in Figure 3. Firstly, the data features with strong correlation with wind speed were selected, the data were regularized, and the network weight and bias vector were updated by gradient descent algorithm. The forward and backward operations of the network are carried out, and the output of weight training is obtained by splicing and activation functions. The model parameters are updated by back propagation algorithm, and the optimal results are obtained. Finally, the network model is tested with the test set. The wind speed predicted by the network model is compared with the actual wind speed, and the network model is evaluated with the error evaluation parameters.

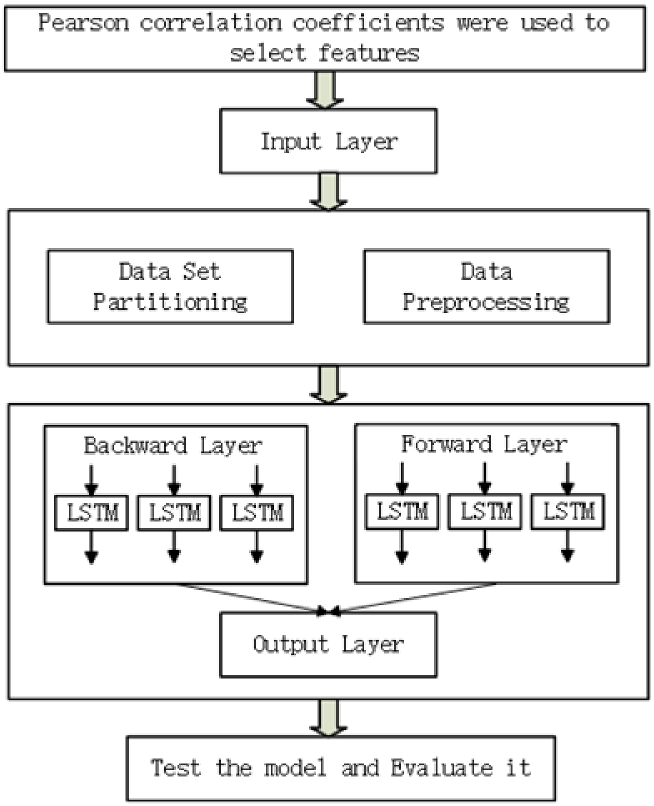

Figure 3: Structure of MI-BI-LSTM prediction model

4. Simulation experiment and result analysis

4.1 Data processing

Wind speed data from March 17, 2017 to April 18, 2017 collected by radar wind measuring device of a wind farm in China were used in this experiment, with a total of 4752 pieces of data. The wind speed sequence of a certain day at a height of 100 meters is shown in figure 4.

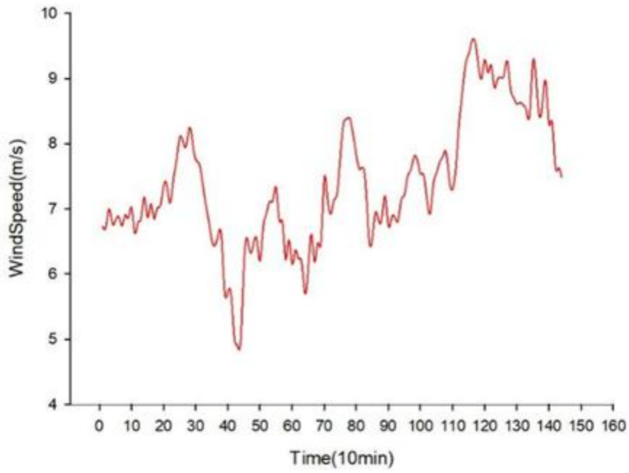

Figure 4: Daily wind speed series at a height of 100 m

The data set was normalized to remove the numerical difference between different input characteristic data and reduce the prediction error of the model. The specific normalization formula is as follows:

$$\tilde{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \qquad (14)$$

Where, x is the original data, $\tilde{x}$ is the normalized data, $x_{\max}$ and $x_{\min}$ are the maximum and minimum values of the original data.

4.2 Feature selection

Several parameters listed in Table 1 were selected as alternative input variables to analyze the Pearson correlation between alternative variables and predictive variables.

Table 1: 100.0mAvgWindSpeed related variables

	Min	Max	Mean	STD
Temperature	22	24	22.81	0.595
Humidity	90	96	92.7	1.881
Pressureh	1011	1015	1013.38	1.037
RainLevel	122.2	122.2	122.2	0
50.0mAvgWS	2.48	5.95	4.2644	0.90542
50.0mAvgWD	48.96	100.08	74.0925	14.71415
70.0mAvgWS	2.57	6.3	4.4649	0.95652
70.0mAvgWD	51.12	102.6	76.7175	14.74699
80.0mAvgWS	2.59	6.4	4.4919	0.95489
80.0mAvgWD	52.56	103.68	78.0675	14.68425
90.0mAvgWS	2.6	6.46	4.5228	0.95572
90.0mAvgWD	54	104.76	79.6375	14.67753
100.0mAvgWS	2.61	6.49	4.5523	0.95949

Pearson correlation coefficient [26] is used to measure the correlation between two variables X and Y, and its value is between -1 and 1. When the absolute value of correlation coefficient R is generally above 0.8, A and B are considered to be strongly correlated. Between 0.3 and 0.8 is considered to have a weak correlation; Below 0.3, there is no correlation. The predicted results are shown in Table 2. Wind speed at different heights has a strong correlation with the predicted variables, while ambient temperature, pressure and relative humidity have a weak correlation with the predicted variables. The final selected model input features include: Temperature, Pressure, 100.0mAvgWindSpeed etc.

Table 2: Pearson correlations between predictive variables and alternative correlated variables

Relevant variables	100.0mAvgWindSpeed
Temperature	-0.719**
Humidity	0.711**
Pressureh	0.423**
RainLevel	.b
50.0mAvgWindSpeed	0.988**
50.0mAvgWindDir	-0.233**
70.0mAvgWindSpeed	0.994**
70.0mAvgWindDir	-0.233**
80.0mAvgWindSpeed	0.997**
80.0mAvgWindDir	-0.236**
90.0mAvgWindSpeed	0.999**
90.0mAvgWindDir	-0.341**
100.0mAvgWindSpeed	1

Note: ** represents P less than 0.01, * represents P less than 0.05

4.3 Model performance evaluation index

RMSE, MAE and MAPE are the main parameters used to evaluate MI-BI-LSTM's prediction performance.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \qquad (15)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{y}_i - y_i\right| \qquad (16)$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \qquad (17)$$

4.4 Analysis of experimental results

The length of the historical input sequence was set as 18, that is, the wind speed of the next 10 minutes was predicted by the multidimensional wind speed characteristics of the first three hours. A batch size of 32 is used for training the prediction model, and the iteration parameters are 128. The model is optimized using the Adam algorithm, which takes the normalized mean square error as a loss function. In order to test the accuracy of MI-BI-LSTM wind speed prediction model, wind speed data from multiple dimensions is used to test the short-term wind speed forecasted by a model and compared with the actual wind speed forecast by the model, as shown in figure 5.

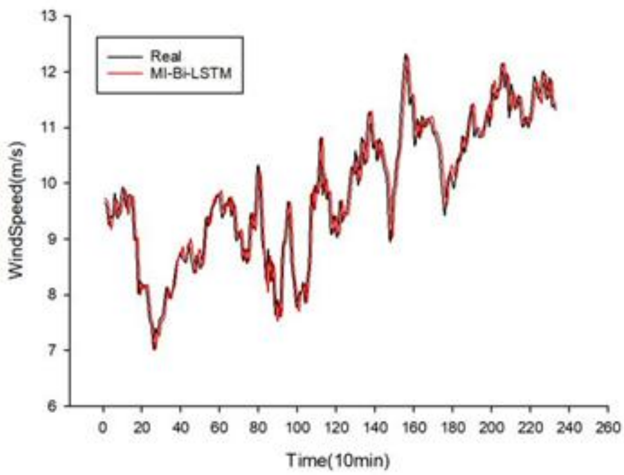

Figure 5: Comparison of results of MI-BI-LSTM wind speed prediction model

It can be seen from figure 5 that the prediction result of MI-BI-LSTM wind speed prediction model has a high degree of fitting with the real value. RMSE, MAE and MAPE of the prediction model are 0.3377, 0.2658 and 0.0274, respectively.

In this paper, RNN, LSTM and BI-LSTM models are established with the same data, and compared with MI-BI-LSTM model. The results are shown in Figure 7. It can be seen that the results of MI-BI-LSTM model are more consistent with the actual wind speed trend, and the errors are much smaller than those of the other three models. BI-LSTM model can retain the prediction results more effectively and has higher accuracy through LSTM gate structure optimization, but compared with MI-BI-LSTM model, the prediction results of MI-BI-LSTM model have smaller error.

It shows that the MI-BI-LSTM model, by extracting the forward and backward rules of time series, makes up for the deficiency of other models that only use one-sided information for prediction, and has better feasibility in actual data prediction. In order to clarify the specific error values of different prediction methods, different performance comparisons are shown in Table 3.

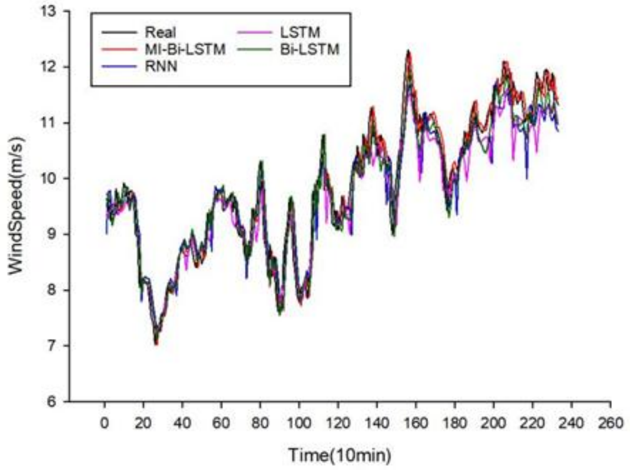

Figure 6: Prediction results of different models

Table 3: Model performance comparison

Model	RMSE	MAE	MAPE
RNN	0.4325	0.3413	0.0346
LSTM	0.3933	0.2992	0.0294
Bi-LSTM	0.3825	0.2921	0.0290
MI-Bi-LSTM	**0.3377**	**0.2658**	**0.0274**

According to the prediction results and evaluation index analysis in Table 3, the overall effect of the model based on LSTM and BI-LSTM network is better than that of RNN network, indicating that LSTM and BI-LSTM model have stronger prediction ability in extracting time series information based on RNN network. MAPE, RMSE and MAE based on MI-BI-LSTM wind speed prediction model are the smallest, with the highest fitting degree and the best prediction effect.

5. Conclusion

According to the current short-term wind speed prediction problem, a short-term wind speed prediction model based on MI-BI-LSTM was proposed. The hidden relationship about wind speed change was extracted from multi-input characteristic data, and the regular features were extracted from the wind speed data by using the forward and backward propagation characteristics of BI-LSTM network, and the wind speed size in the next 10 minutes was predicted. Compared with RNN, LSTM, BI-LSTM and other models, this method has better prediction performance in wind speed prediction. Experimental results verify the accuracy of network prediction performance and rationality of network results, providing a more scientific basis for wind speed prediction and large-scale grid connection of new energy power generation.

References

[1] Yaya Lei, Weisheng Wang, Yonghua Yin, et al. “Value analysis of wind power to power system operation” [J] Power grid technology, 2002, 26 (5): 5.

[2] Jian Wang, Gangui Yan, Wei Song, et al. “Overview of wind power prediction technology” [J] Journal of Northeast Electric Power University, 2011, 31 (3): 5.

[3] Hu Zhang. “Research and application of time series short-term prediction model” [D] Huazhong University of science and technology, 2013.

[4] Ak R, Fink O, Zio E . “Two Machine Learning Approaches for Short-Term Wind Speed Time-Series Prediction” [J]. IEEE Transactions on Neural Networks & Learning Systems, 2016, 27(8):1734-1747.

[5] Hochreiter, Sepp, Schmidhuber, et al. “Long short-term memory.” [J]. Neural Computation, 1997.

[6] Lijie Wang, Xiaozhong Liao, Yang Gao, et al. “Summary of research on Modeling and prediction of wind farm power generation” [J] Power system protection and control, 2009 (13): 4.

[7] “Research and application of time series short-term prediction model” [D] Huazhong University of science and technology, 2013.

[8] Xue Y, Chen N, Wang S, et al. “Review on Wind Speed Prediction Based on Spatial Correlatio”n [J]. Dianli Xitong Zidonghua/Automation of Electric Power Systems, 2017, 41(10):161-169.

[9] Chen M R, Zeng G Q, Lu K D, et al. “A Two-Layer Nonlinear Combination Method for Short-Term Wind Speed Prediction Based on ELM, ENN and LSTM” [J]. IEEE Internet of Things Journal, 2019:1-1.

[10] Ying Du, Jiping Lu, Qing Li, et al. “Short term wind speed prediction of wind farm based on least squares support vector machine” [J] Power grid technology, 2008, 32 (15): 5.

[11] Chunbo Xiu, Xiao Ren, Yanqing Li, et al. “Short term prediction method of wind speed series based on Kalman filter” [J] Journal of electrotechnics, 2014, 29 (2): 7.

[12] Jie Yan, Chengzhi Xu, Yongqian Liu, et al. “Short term wind power prediction method based on similar days of wind speed cloud model” [J] Power system automation, 2018, 42 (6): 7.

[13] Biwu Fang, Dichen Liu, Bo Wang, et al. “Short term wind speed prediction based on LSSVM optimized by wavelet transform and improved firefly algorithm” [J] Power system protection and control, 2016, 44 (8): 7.

[14] Yingying Su, Cuiying Li, Xiaofeng Wang, et al. “C-C and elm fast prediction method for short-term wind speed of wind farm” [J] Journal of power system and automation, 2019, 31 (7): 6

[15] “Anonymity Ultra short term wind speed prediction using spatial correlation” [J] Power system automation, 2017, 41 (12): 7.

[16] Mao Yang, Yi Zhou. “Ultra short term prediction of wind power considering the state of wind farm” [J] Chinese Journal of electrical engineering, 2019 (5): 10.

[17] Fu W, Wang K, Li C, et al. “Multi-step short-term wind speed forecasting approach based on multi-scale dominant ingredient chaotic analysis, improved hybrid GWO-SCA optimization and ELM” [J]. Energy Conversion and Management, 2019, 187:356-377.

[18] Bengio Y, Frasconi P, Simard P . "Problem of learning long-term dependencies in recurrent networks" [C]// Neural Networks, 1993. IEEE International Conference on. IEEE, 1993.
[19] Bing Li, Yan Zhang, Shi Liu. "Research on short-term wind speed prediction based on LSTM" [J] Computer simulation, 2018, 35 (11): 6.
[20] Mi X, Hui L, Li Y . "Wind speed prediction model using singular spectrum analysis, empirical mode decomposition and convolutional support vector machine - ScienceDirect" [J]. Energy Conversion and Management, 2019, 180:196-205.
[21] Zhang X, Trmal J, Povey D, et al. "Improving deep neural network acoustic models using generalized maxout networks" [C]// ICASSP 2014 - 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2014.
[22] Gu O T, Liu R, Yang H, et al. "Predict Atmosphere Electric Field Value with the LSTM Neural Network" [C]// 2017 International Conference on Computer Systems, Electronics and Control (ICCSEC). 2018.
[23] Jiapeng Zhu. "Prediction of protein RNA binding based on Bi LSTM and densenet" [D] Zhejiang University of technology.
[24] Yafei Wang, Jing Han, Huang Guo, et al. "Structural deformation prediction based on Bi LSTM" [J] Computer system application, 2021, 30 (11): 6.
[25] Aidong Xu, Yanwen Guo, Tao Wu, et al. "Prediction of household power consumption based on Bi LSTM" [J] Industrial control computer, 2020, 33 (4): 3.
[26] Yulei Zhang, Yan Dang, Ping'an He. "Quantitative analysis of biological kinship by Pearson correlation coefficient " [J] Computer engineering and application, 2005, 41 (33): 5.

Author Profile

Baolin Ma, Born in June 1997 in Guyuan, NingXia, he is a master's student in the school of control and computer engineering of North China Electric Power University. The main research direction is machine learning.