Journal of Computers Vol. 33 No. 4, August 2022, pp. 195-204
doi: 10.53106/199115992022083304016

Wind Power Prediction Based on Difference Method

Bao-Wei Zhang*, Hong-Bo Cui, Jiu-Xiang Song

College of Electrical Information Engineering, Zhengzhou University of Light Industry,
Zhengzhou 450002, Henan, PRC
{108200723, 981444028, 1454196320}@qq.com

Received 1 April 2022; Revised 19 April 2022; Accepted 20 April 2022

Abstract. The renewable wind power sources are difficult to be predicted in view of the fluctuating factors such as wind bearing, pressure, wind speed, and humidity of the surrounding atmosphere. An attempt is made in this paper to propose a difference method to build a neural network and a long short term memory (LSTM) model for wind power prediction. First, the correlation of each data is analyzed and then performing difference processing on the original data to solve the problem that the original data cannot be analyzed by probability distribution. The prediction is made by building the neural network and LSTM and feeding the original data and the difference-processed data into the neural network model respectively. Finally, the data are added for validation, and the raw data used include wind power data in Belgium from November 1, 2019 to November 30, 2019. The experimental results show that the LSTM prediction accuracy is improved by 178.67%, and is effective in predicting long-term wind power data with 216.06% accuracy improvement, the neural network prediction accuracy is improved by 154.07%, and the short-term wind power prediction accuracy is improved by 228%.

Keywords: wind power, volatility, difference method, LSTM, neural networks

1 Introduction

The unprecedented consumption of traditional energy sources not only resulted in global warming but posed serious environmental threats. Of late, several alternative clean and renewable energy sources, materials and technologies have been reported. A reliable wind power prediction system can predict the fluctuation of wind power in advance so as to ensure reliable measures, and improve the stability, safety and controllability of wind power generation in the power grid [1]. Wind energy has gradually become the third largest energy source after thermal power and hydroelectric power [2]. According to statistics, a record high, 71.67 million kilowatts of new grid-connected capacity was installed in 2020, and wind power generation was 466.5 billion KWH, an increase of about 15%. However, due to the highly random fluctuations and intermittent nature of the wind and the influence of the tail flow between units in the wind farms, the large-capacity wind turbines connected to the grid pose serious challenges to the security of the power system, power quality and the balance of power a record high, demand and supply [3]. Korprasertsak et al. have conceptualized statistical interpretation of multiple forecasting models for short-term prediction of wind power generation under uncertainty to study the fluctuation characteristics of wind power, improve the accuracy of real-time wind power prediction, and to overcome the adverse impact of wind power access on the grid [4].

Wind power prediction methods are generally classified into two types: statistical learning methods and physical models. The statistical learning approach uses measured data to make predictions of wind power by studying historical data from wind farms [5]. The physical model studies the progression of the weather and uses a mathematical model that respond to the evolution of the weather to predict the weather data based on boundary conditions. Neural networks are developed to handle nonlinear data to improve the accuracy of wind power prediction. However, the single neural network wind power prediction for short-term has low prediction accuracy and is prone to over learning [6]. An improved Elman neural network model and the BP neural network model are used to conduct a comparative experiment. The Elman neural network gives better experimental results than the single BP neural network whereas the wavelet decomposition improves the prediction accuracy [7]. To address the instability and volatility of wind power, support vector machine (SVM) and LSTM are used to conduct comparative experiments for testing the effect of wind speed on the prediction results, and the SVM method gives good results considering the wind speed [8]. A combined model was proposed to obtain the physical features using deep neural networks (DNN) and long and short-term memory neural networks (LSTM) time series for short-term wind power prediction to improve the accuracy and stability of wind power prediction [9]. The combined EMD-LSTM

* Corresponding Author

prediction model was used to correct the NWP wind speed; the EMD decomposed the wind power data series into data components while, the LSTM network model was used for accurate wind power prediction [10].

Wind power is influenced by environment and wind speed but the relevant data at first do not obey any distribution law. In view of this, the wind power data are first differenced, i.e., the next moment is differenced from the data at present moment to obtain a completely new set of data. The new data are imported into the LSTM network and neural network model to obtain the prediction results without differencing and the prediction results after differencing, respectively. Experiments show that the data processed by the difference method effectively improves the accuracy of wind power prediction by LSTM and neural network.

2 Model Building

This section uses both neural network and LSTM methods for wind power data prediction, the former is a non-time-series prediction method while the latter is a time-series prediction method. In the non-time-series forecasting method, the data are randomly discontinuous in order to ensure the discontinuity in time. While in time-series forecasting method, the data are continuous. A time steps of 15 minutes, 30 minutes, 1 hour, 3 hours, 5 hours, and 6 hours were selected for prediction, and the top 20, 50, and 200 data with the highest accuracy of the two methods were extracted. The prediction accuracy of the two methods and the extracted data are analyzed to understand the scope of application of each method and the basis of recommendation.

2.1 LSTM Model

Time series predictive analysis uses the characteristics of the time of an event in the past period to predict the characteristics of that event in the future. It is relatively a complex class of predictive modeling problem as compared to the regression analysis model. The time series models are dependent on the sequence of events, the same size values change the order of the input model to produce different results [11]. The structure of the LSTM model is shown in Fig. 1.

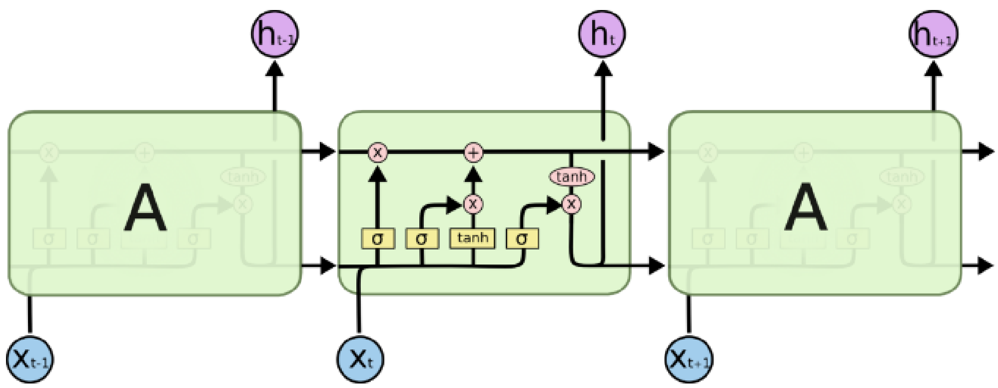

Fig. 1. LSTM network model

The LSTM is characterized by the addition of valve nodes for each layer in addition to the RNN structure. There are three types of valves: forget gate, input gate and output gate [12]. These valves can be opened or closed and are used to add the result determining whether the memory state (previous state of the network) of the model network has reached a threshold at the output of that layer and thus to the current calculation of that layer.

2.2 BP Neural Network Model

BP neural networks have been widely used for problems such as data prediction due to their outstanding advantages in last decade. The number of implicit layers in the BP neural network needs to be considered in the data prediction problem of quality in modeling process. In recent years, four-layer BP neural network is conventional in small sample prediction models due to its advantages of fast operation and high accuracy [13].

The standard BP neural network model consists of 3 parts, the outermost is the input layer, the middle layer has one or more hidden layers, and the last is the output layer of the network, which outputs the results of the operations. The process of BP neural network consists of two stages, the first stage is the forward propagation process of the input signals, from the input layer through the hidden layer, and finally reaches the output layer. The second stage is the back propagation of the error, from the output layer to the implied layer and finally to the input layer, adjusting in turn the weights and biases from the implied layer to the output layer, the weights and biases from the input layer to the implied layer [14].

A four-layer BP neural network model with double hidden layers for wind power prediction used in this paper is shown in Fig. 2.

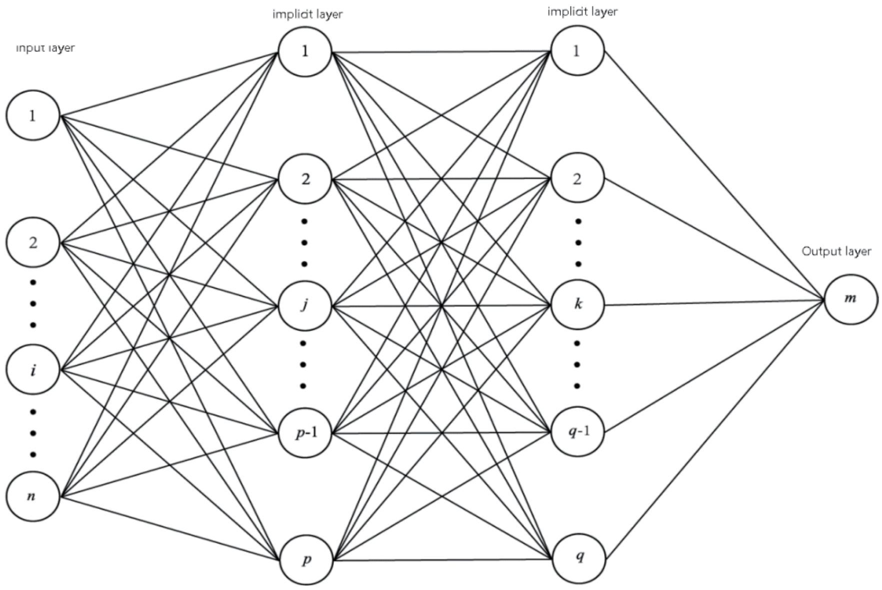

Fig. 2. Structure of four-layer BP neural network

This network used in this paper is for non-time-series prediction as the input data are randomly sampled, and the train_test_split function of the sklearn library is used for random cross-sampling, and 90% of the total data is taken as the training set while 10% as the test set. Among these, the dimension of each training set sample is based on the first 15 minutes, 30 minutes, 1 hour, 3 hours, 5 hours, 6 hours of the requirement to select the number of samples of 1, 2, 4, 12, 20, 24, whose corresponding output is the wind power at these moments, and the test set is also selected in the same way. The neural network used is a four-layer neural network with one input layer, two hidden layers, and one output layer with the number of neurons 32 in first two layers. The activation function of both hidden layers is RELU function, while the activation function of the output layer is Sigmoid function.

2.3 Difference Method

The difference method network model is constructed to handle the wind power time series data in a better way since it does not obey distribution law due to uncertainty, indirectness and wind volatility however, it does not exhibit any sudden surge or plunge.

2.3.1 Difference

First define an initial time dimension, and use the wind power data of the next time dimension and the current moment wind power data to make a difference to get a new data table. The difference processing is done not to find the fluctuation pattern of the data, but to use the new data to analyze what kind of distribution the original data obey, and then improve the accuracy of wind power prediction.

2.3.2 Differentiation

Let the collected wind power be x0, x1, x2, x3, …, xn, and the difference be Δx1, Δx2, Δx3, Δx4, …, Δxn, that is,

$\Delta x_1 = x_1 - x_0, \ldots, \Delta x_n = x_n - x_{n-1}$. Predicting wind power in different time dimensions requires doing different difference data. In order to explore the degree of influence of difference on wind power in different time spans, the data was collected at 15 minutes, 30 minutes, 1 hour, 3 hours, 5 hours, and 6 hours time dimension difference.

2.3.3 Inverse Differentiation

The collected wind power can be expressed as **xn=xn-1+Δxn** the difference $\sum_{n=1}^{n} \Delta \mathbf{x}_n$ is treated as a brand new data and saved in the difference table data. The difference data is used as input for the subsequent neural network model and LSTM model, and the obtained results are integrated to build a new output table. The differential power in the new output table is expressed as △ Yn, and then the differential power △ Yn is inverse differentialized to obtain the wind power prediction result Yn=Xn-1+ △ Yn.

2.3.4 Differentiated Neural Network Model

The difference table data from the inverse differencing is combined with the neural network model and LSTM model to construct the differencing network model. The difference table data are divided into different time dimensions and used as the input △ xn of the model. The selection of training and test sets is consistent with that of the LSTM model and the neural network model, with 90% of the data as the training set and 10% of the data as the test set. The model output values are then inverse-differentiated to obtain the new wind power data Yn. Fig. 3 shows the neural network model or LSTM model using a black box.

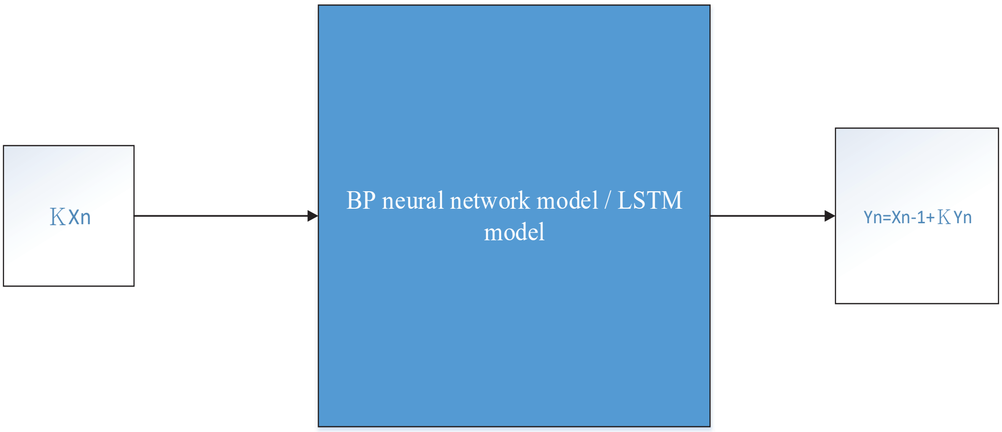

Fig. 3. Differentiated neural network model diagram

3 Model Building

The wind power data collected from November 1, 2019 to November 30, 2019 in Belgium were used as the original sample data, and the time interval for wind power data collection was 15 minutes. While, conducting data analysis the selection of time dimension also determines the prediction accuracy and the time dimensions selected are 15min, 30min, 1h, 3h, 5h, and 6h. The selection of multiple time dimensions can better analyze the improvement of wind power prediction accuracy after processing the data by the difference method.

3.1 Wind Power Data Pattern Analysis

Data analysis was first performed on the raw data, and preliminary judgments were made by the raw data histogram and probability distribution chart, which were verified using two tests. In order to better observe the overall wind power data and measure the relationship between wind power data and time, scatter plots and curve plots of wind power data were plotted, as shown below.

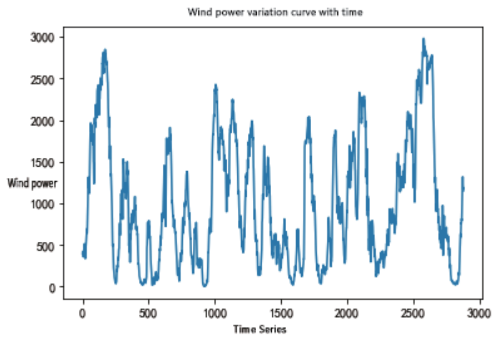

Fig. 4. Wind power curve

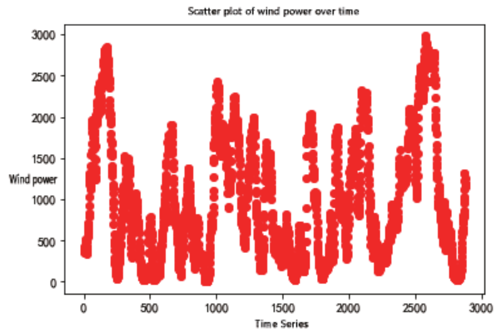

Fig. 5. Wind power scatter diagram

Note: Sine there are too many time interval series in the wind power database, the horizontal coordinates in Fig. 4, Fig. 5 cannot be displayed in their entirety, so they are expressed in the form of time series, where 0 means 0:0 on November 1, 2019, and every 15 minutes a point is recorded, the time series is added 1, and so on, to the last time point in the wind power database.

3.1.1 Data Verification Method

In order to verify whether the data conform to a normal distribution, a chi-square distribution or a t-distribution was carried out. Check the data, if it does not conform to the law of probability distribution, it means that the data is mainly affected by the randomness and volatility of wind power, and then the difference method is processed for such data.

The following two validation methods were used for the data:

The Shapiro-Wilk (S-W) test is a correlation-based algorithm. It is recommended by the national standard GB4882-85 to make the smallest type II error. According to formula (1), a correlation coefficient is obtained, and the closer it is to 1, the better the fitting of the data and the normal distribution [15].

$$W = \frac{(\sum_{i=1}^{n} a_i x_i)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}. \quad \textbf{(1)}$$

And is the average of the sample.

Kolmogorov-Smirnow (K-S) test is a non-parametric statistical test method for continuous distribution. This test is often used to compare whether a single sample conforms to a known distribution (compare the cumulative frequency distribution of the sample data with a specific theoretical distribution, and if the difference between the two is small, it is inferred that the sample is taken from a specific distribution cluster), the wo-sample K-S test compares the similarity of the cumulative distribution (continuous distribution) of the two datasets [16].

Most of the time, the test results obtained by these two methods are roughly the same, so that many people ignore the difference between the two methods: while analyzing small sample data with less than 50 rows, we tend to look at the results obtained by the S-W test. Normality test results; while analyzing small sample data with more than 50 rows, we tend to look at the normality test results obtained by the K-S test.

3.1.2 Wind Power Data Inspection

The hypothesis verification results are obtained by using the above two methods, as shown in Table 1.

Table 1. Validation results of normal distribution, t distribution and chi-square distribution (15 min)

Distribution	Method	V value	P value
Normal distribution	S-W Test	0.928957402706146	5.93653345757804E-35
Normal distribution	K-S Test	0.107432350831198	2.68549548177369E-29
t distribution	K-S Test	0.115625000000000	3.50588592959783E-17
chi-square distribution	K-S Test	0.09027777777777	1.24468109353966E-10

From Table 1, it is found that the P value is far less than 0.05, indicating that the original wind power data does not obey any of the three probability distributions. In order to better analyze the data and obtain the reason why the original wind power data does not obey the probability distribution, this paper verifies the obedience of the probability distribution of each day's data and selects 5 days with characteristics (1, 2, 9, 16, 25) for analysis, in order to briefly explain the compliance situation, the number of power data intervals taken is 20. The histogram and density curve are depicted in Fig. 6.

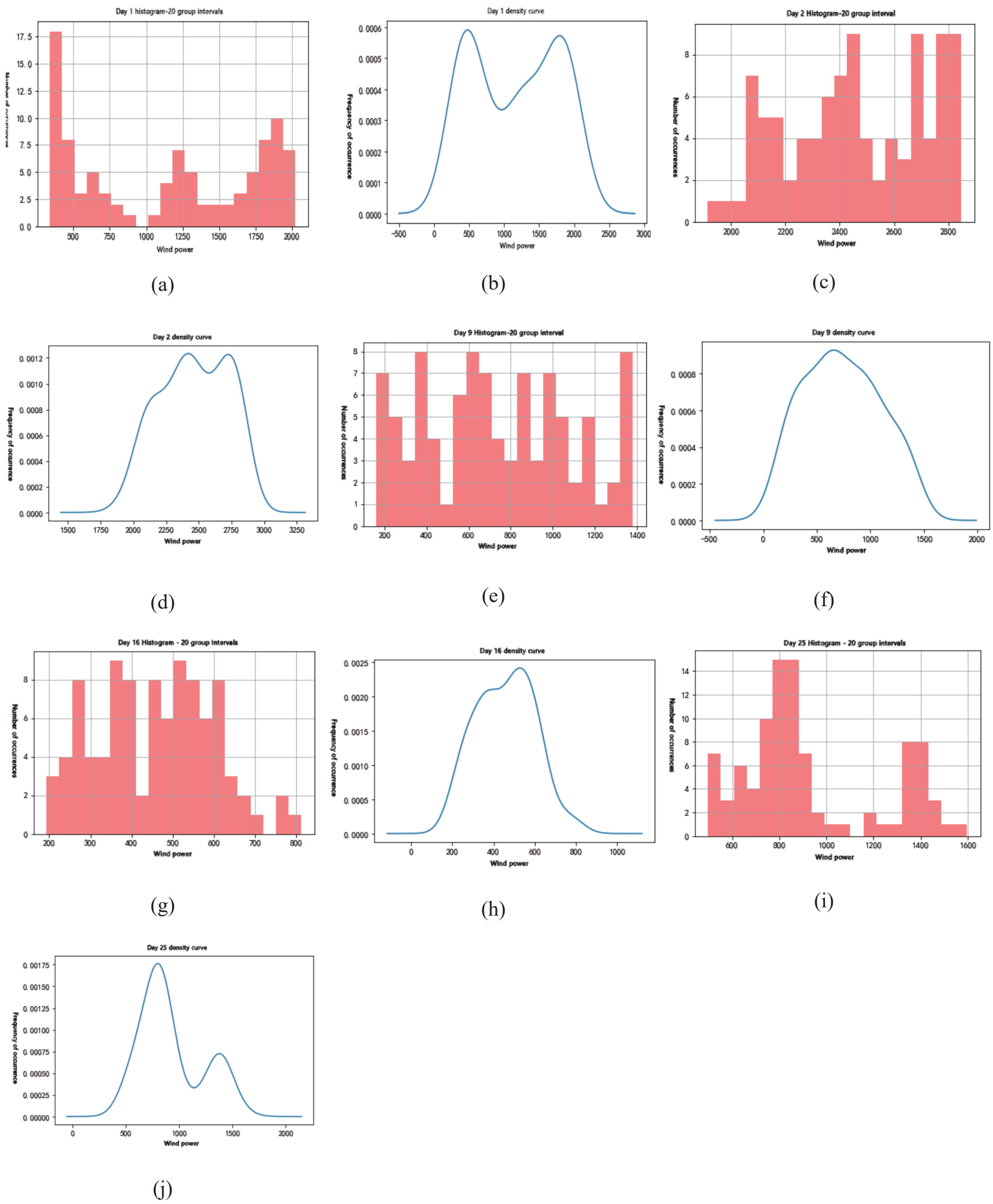

Fig. 6. Histogram and density curve of wind power data on days 1, 2, 9, 16, and 25

The validation results of the probability distribution of the histogram and the density curve of the five days data by using the K-S detection method are given below:

Table 2. Validation results of normal distribution, t distribution, and chi-square distribution

Time	Use Methodology	Gaussian distribution		t distribution		Chi-square distribution	
		V value	P value	V value	P value	V value	P value
2019/1/1	K-S Test	0.86914	0.00	0.23958	0.00789	0.98958	0.00
2019/1/2	K-S Test	0.94581	0.00060	0.15625	0.19230	0.10417	0.67764
2019/1/9	K-S Test	0.96043	0.00546	0.14583	0.25999	0.95833	0.00
2019/1/16	K-S Test	0.98014	0.15391	0.18750	0.06828	0.97917	0.00
2019/1/25	K-S Test	0.88221	0.00	0.28125	0.00095	0.15625	0.19229

It is observed from the above Table 2. that the distribution of the selected representative five-day data, in which the first day does not obey any distribution; the second day obeys the t distribution and the chi-square distribution, the ninth day obeys the t distribution, the 16th day obeys the Gaussian distribution and t distribution whereas the 25th day follows a chi-square distribution. By analyzing the data distribution of these five days, combined with the characteristics of wind power generation, it is found that the main reason why the wind power data does not obey the three probability distributions is attributed to the volatility and randomness of wind power.

3.1.3 Wind Power Data Differential Inspection

The difference method is adopted to improve the wind power data. The specific content of the difference method is to use the difference between the data at the next moment and current moment to obtain a new set of data. The S-W and K-S detection methods are used to verify the data processed by the difference method the validation results are presented in Table 3.

Table 3. Validation results of normal distribution, t distribution, and chi-square distribution (difference method)

Probability distribution	Method	V value	P value
Normal distribution	S-W Test	0.94258	$4.56*10^{-32}$
Normal distribution	K-S Test	0.47529	0
t distribution	K-S Test	0.03091	0.127661292
chi-square distribution	K-S Test	0.28899	$2.48*10^{-6}$

It is noticed from Table 3, that all P values of the t probability distribution function are greater than 0.05, which indicates that the wind power data improved by the difference method and obeys the t probability distribution. Thus, we find that this method enables the wind power data to obey the probability distribution: by using the wind power change as the processed data, the processed data obeys the t distribution, and its test statistic value is 0.03091.

3.2 Wind Power Forecast from Raw Data

In the process of verifying the effectiveness of the LSTM algorithm firstly, the training set is input into the LSTM for model training, and secondly, the test set is sent into the trained network model to obtain the evaluation results of the model. The LSTM model used in this paper is a single-layer bidirectional LSTM which includes an LSTM bidirectional layer, a hidden layer and an output layer wherein, the number of neurons in the LSTM layer is 64, the number of neurons in the hidden layer is 32, the activation function of the hidden layer is the SELU function, and the activation function of the output layer is the SELU function. Among these, the evaluation function selects mean square error (MSE) and root mean square error (RMSE). The formulae are shown in (2) and (3), and the evaluation results are shown in Table 4.

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(observeded_i - predicted_i)^2 \,. \tag{2}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(observeded_i - predicted_i)^2} \,. \tag{3}$$

Table 4. LSTM mean square error and root mean square error

Time interval	MSE	RMSE
15 minutes	3349.87	57.87804873
30 minutes	7688.72	87.68533733
1 hour	7442.09	86.26757552
3 hour	24071.6	155.1503941
5 hour	13676.5	116.9467353
6 hour	14245.4	119.3541417

In verification of the effectiveness of the neural network first, the training set is input into the neural network for model training and then the test set is sent into the trained neural network model to obtain the evaluation results of the model. The evaluation index function of the model is shown in formulae (2) and (3), and the evaluation results are presented in Table 5.

Table 5. Neural network output table

Time interval	MSE	RMSE
15 minutes	13080.9	114.372
30 minutes	4084.45	63.9097
1 hour	7069.78	84.082
3 hour	4923.75	70.1695
5 hour	4323.46	65.753
6 hour	5930.43	77.0093

3.3 Wind Power Prediction with Improved Difference Method

Considering that the original data does not obey any distribution, and the variation of the original data obeys the t distribution. Therefore, by using the difference of the original data as the input of the neural network or LSTM to improves the model.

The improved model gets the input of the model through the difference method, that is, the input is the difference between the data at the next moment and the current moment, and gives a new set of data. At the same time, considering the influence of time on the algorithm, the time series prediction method is used here, that is, 90% of all the data are taken as the training set in chronological order, and 10% are used as the test set. Finally, the validity of the model is verified by sequentially selecting the time periods of 15 minutes, 30 minutes, 1 hour, 3 hours, 5 hours and 6 hours in advance. The experimental results obtained are shown in Table 6.

The analysis shows that the low accuracy of wind power prediction is mainly due to the volatility and sudden change of wind power. The optimized neural network and LSTM have significantly improved the performance compared to that before optimization, and are very stable.

Table 6. Neural network and LSTM validity verification

	LSTM		Neural Networks	
Time interval	MSE	RMSE	MSE	RMSE
15 minutes	3285.03	57.31518	2589.72	50.88929
30 minutes	3022.23	54.97478	2546.95	50.46733
1 hour	3447.34	58.71408	2709.91	52.05678
3 hour	4730.89	68.78146	2682.85	51.79622
5 hour	3117.13	55.83129	2733.67	52.28455
6 hour	3062.11	55.33634	2756.95	52.50669

4 Conclusion

It can be seen from the prediction results of LSTM and the neural network on wind power data that LSTM is more suitable for short-term wind power forecasting, while the neural network is suitable for wind power forecasting in longer time dimension. The two models have their own strengths. The data processed by the difference method is used as an input to predict the result, and then compared with the previous wind power forecast data, the experiment shows that after the difference method data processing, the wind power prediction accuracy of the LSTM model and the neural network model has improved significantly. The difference method not only discovers the regularity of the original wind power data, but also improves the prediction accuracy of the wind power. After the improvement of the LSTM model, the prediction accuracy of wind power in the long-term dimension is greatly improved, such as the wind power prediction in the 6 hr time dimension. After the improvement of the neural network model, the prediction of wind power in the short time dimension is significantly improved, such as the wind power data in the 15-minute time dimension.

Acknowledgments

This work was supported in part by the Science and Technology Research Program of Henan Province, China (202102210181).

Conflict of Interest

The authors declare no conflict of interest.

References

[1] X. Yang, Y. Yang, Y. Liu, Z. Deng, A Reliability Assessment Approach for Electric Power Systems Considering Wind Power Uncertainty, IEEE Access 8(2020) 12467-12478.
[2] P. Sadorsky, Wind energy for sustainable development: Driving factors and future outlook, Journal of Cleaner Production 289(2021) 125779.
[3] A.N. de Paula, E.J. de Oliveira, L.W. Oliveira, C.A. Moraes, Reliability-constrained dynamic transmission expansion planning considering wind power generation, Electrical Engineering 102(4)(2020) 2583-2593.
[4] N. Korprasertsak, T. Leephakpreeda, Robust short-term prediction of wind power generation under uncertainty via statistical interpretation of multiple forecasting models, Energy 180(2019) 387-397.
[5] M. Arshad, B. O'Kelly, Global status of wind power generation: theory, practice, and challenges, International Journal of Green Energy 16(14)(2019) 1073-1090.
[6] J.D. Mao, X.J. Zhang, J. Li, Wind Power Forecasting Based on the BP Neural Network, in: Proc. the 2nd International Conference on Systems Engineering and Modeling (ICSEM 2013), 2013.
[7] X.L. Gong, Z.J. Hu, M.L. Zhang, H. Wang, Wind Power Forecasting Using Wavelet Decomposition and Elman Neural Network, Advanced Materials Research 608-609(2012) 628-632.
[8] E. Kang, N. Son, Support Vector Machine and Long Short-term Memory using Multivariate Models for Wind Power Forecasting, International Journal of Innovative Technology and Exploring Engineering 9(4)(2020) 2364-2369.
[9] B. Xiong, X. Meng, R. Wang, X. Wang, Z. Wang, Combined Model for Short-term Wind Power Prediction Based on Deep Neural Network and Long Short-Term Memory, Journal of Physics: Conference Series 1757(2021) 012095.
[10]F. Zhang, Z. Guo, X. Sun, J. Xi, Short-term wind power prediction based on EMD-LSTM combined model, IOP Conference Series Earth and Environmental Science 514(2020) 042003.

[11]A. Zeyer, P. Doetsch, P. Voigtlaender, R. Schlüter, H. Ney, A Comprehensive Study of Deep Bidirectional LSTM RNNs for Acoustic Modeling in Speech Recognition, in: Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2017.
[12]K.K. Chandriah, R.V. Naraganahalli, RNN / LSTM with modified Adam optimizer in deep learning approach for automobile spare parts demand forecasting, Multimedia Tools and Applications 80(17)(2021) 26145-26159.
[13]M.L. Zhang, Y.J. Zhang, X.L. He, Z.J. Gao, Adaptive PID Control and Its Application Based on a Double-Layer BP Neural Network, Processes 9(8)(2021) 1475.
[14]F. Wen, W. Pei, Q. Li, Z. Chu, W. Zhao, S. Wu, X. Zhang, C. Han, Insulation Monitoring of Dynamic Wireless Charging Network Based on BP Neural Network, World Electric Vehicle Journal 12(3)(2021) 129.
[15]W. Yamanappa, P.V. Sudeep, M.K. Sabu, J. Rajan, Non-Local Means Image Denoising Using Shapiro-Wilk Similarity Measure, IEEE Access 6(2018) 66914-66922.
[16]M.R.M. Akramin, A. Takahashi, M.N.M. Husnain, Z.L. Chuan, K-S Test for Crack Increment in Probabilistic Fracture Mechanics Analysis, IOP Conference Series Earth and Environmental Science 498(2020) 012035.