ARCHIVES OF ELECTRICAL ENGINEERING VOL. 69(2), pp. 271–286 (2020)

DOI 10.24425/aee.2020.133025

Ultra-short-term wind power prediction based on copula function and bivariate EMD decomposition algorithm

HAIQING LIU, WEIJIAN LIN, YUANCHENG LI

North China Electric Power University
China
e-mail: hqliu@ncepu.edu.cn, 2536973080@qq.com, yuancheng@ncepu.cn

(Received: 25.06.2019, revised: 04.11.2019)

Abstract: Against the background of increasing installed capacity of wind power in the power generation system, high-precision ultra-short-term wind power prediction is significant for safe and reliable operation of the power generation system. We present a method for ultra-short-term wind power prediction based on a copula function, bivariate empirical mode decomposition (BEMD) algorithm and gated recurrent unit (GRU) neural network. First we use the copula function to analyze the nonlinear correlation between wind power and external factors to extract the key factors influencing wind power generation. Then the joint data composed of the key factors and wind power are decomposed into a series of stationary subsequence data by a BEMD algorithm which can decompose the bivariate data jointly. Finally, the prediction model based on a GRU network uses the decomposed data as the input to predict the power output in the next four hours. The experimental results show that the proposed method can effectively improve the accuracy of ultra-short-term wind power prediction.

Key words: bivariate EMD decomposition, copula function, GRU network, meteorological factor, ultra-short-term wind power prediction

1. Introduction

Due to the serious energy crisis and environmental problems, renewable energy has gained worldwide attention for its large reserves, low carbon and renewable features [1, 2]. As a kind of very important renewable energy source, wind energy has developed vigorously in the past ten years [3]. The wind power capacity installed around the world has increased nearly 75-fold

in the past 20 years, from 75 GW in 1997 to about 564 GW in 2018. Between 2009 and 2013, wind power capacity doubled. In 2016, wind power accounted for 16% of renewable energy generation [2, 3].Wind energy has a wide range of application scenarios. However, since wind energy cannot be stored and it is intermittent, large-scale wind power grid connection will pose serious risks to the power grid, which will bring many problems to the power grid dispatching and frequency modulation [4]. Accurate and effective wind power forecasting is the key to realize the integration of large-scale wind power. It can effectively alleviate the problem of wind abandoning and power limiting, improve the grid dispatching, ensure the stable operation of wind turbines, cut down the operational cost, increase the revenue of wind farms and realize complementary power dispatching.

According to the requirements of the power generation system operation, the wind power prediction has four levels [5]: ultra-short-term, short-term, medium-term, and long-term. The predicted time range of the four levels are within four hours, four to twenty-four hours, one to seven days, and more than seven days respectively.

The previous research results show that there are many complex factors affecting the accuracy of wind power prediction [6]. Therefore, it is important to extract the key influencing factors which closely affect wind power generation. Currently, the most commonly used algorithms include correlation analysis [7], gray relational analysis [8], principal component analysis [9], regression analysis [10], autoregressive analysis [11]. Because the correlation analysis, gray relational analysis and principal component analysis are linear relationship analysis methods, they are not applicable to the analysis of the non-linear correlation between wind power and external impact factors. The regression analysis and autoregressive analysis will be restricted in some cases because the factors and expressions used in the regression analysis and autoregressive analysis are only estimates [10, 11]. In order to better analyze the non-linear correlation, we use the copula function which can effectively analyze the non-linear correlation between the wind speed and the external influence factors [5].

As the input of the prediction model, wind power data has the characteristics of randomness and volatility, which will also affect the accuracy of wind power prediction. In response to this problem, the data processing methods proposed by the related researches combine wavelet decomposition [12], empirical mode decomposition (EMD) [13], integrated empirical mode decomposition (EEMD) [14], complementary ensemble empirical mode decomposition (CEEMD) [15], κ-OCCO [16] and atomic sparse decomposition (ASD) [17]. These methods decompose wind power data with randomness and fluctuation into a series of relatively stable and periodic subsequences. In fact, there are many factors affecting wind power. Sometimes we need to input more than one kind of data into the prediction model, such as wind power and wind speed. Under this circumstance, the above-mentioned decomposition algorithm will not be able to use, because the above-mentioned method can only be used to decompose the data of a variable. Therefore, we use bivariate empirical mode decomposition (BEMD) to decompose the wind power data and the key influencing factors data jointly.

The methods of wind power prediction can be divided into a physical method [18], statistical method and machine learning method. Due to the consideration of various environmental physical factors, the physical method is difficult to model and has low prediction accuracy. The common statistical methods include a Bayesian model [19] and autoregressive model [20]. In recent years,

traditional machine learning has been applied to wind power prediction, methods that mainly include a support vector machine [21], radial basis function neural network [22], extreme learning machine [23] and Recurrent Neural Networks (RNN) [24]. A Gate Recurrent Unit (GRU) neural network is an improvement of an RNN. It is mainly used to study the time series problem, and can effectively solve the problems of gradient disappearance, gradient explosion and excessive model parameters. So we used the GRU neural network for wind power forecasting, a typical time series problem.

In order to improve the accuracy of ultra-short-term wind power prediction, we propose an ultra-short-term wind power prediction model. This model is based on a copula function, BEMD and GRU neural network. Our work mainly includes the following contributions:

1. We use a copula function which can effectively analyze the non-linear correlation to analyze the non-linear correlation between wind power and external factors, and extract the key factors affecting wind power.
2. In order to fully consider the key factors in data decomposition, we use a BEMD algorithm to decompose wind power data and the data of the key factors jointly.
3. We use a GRU neural network to construct our ultra-short-term wind power prediction model which can effectively reduce the number of parameters of the model and deeply explore the dependencies among data compared with an RNN and long short-term memory (LSTM).
4. The experiment shows that the proposed ultra-short-term wind power prediction model can effectively improve the prediction accuracy.

2. Methodology

In order to further improve the accuracy of ultra-short-term wind power prediction, we propose a new prediction method. The flow chart of the method presented in this paper is shown in the Fig. 1.

Our method mainly consists of three parts where:

1. A copula function is used to analyze the correlation between wind power and meteorological factors, and extract the key influencing factors with the greatest correlation with wind power.
2. Bivariate data composed of wind power and the key influencing factors are used as the input data of a BEMD algorithm, decomposed jointly to obtain a series of stable subsequence data.
3. We use a construct ultra-short-term wind power prediction model based on a GRU neural network.

Each sub-sequence decomposed in the second step corresponds to a GRU neural network prediction model, and each sub-sequence data is used as input to train the corresponding GRU neural network prediction model. Then we calculate predictive components of wind power corresponding to each subsequence by the trained GRU neural network prediction model.

Finally, all wind power prediction components are superimposed to get the final wind power prediction value. The next three chapters will give a detailed introduction to the above three parts.

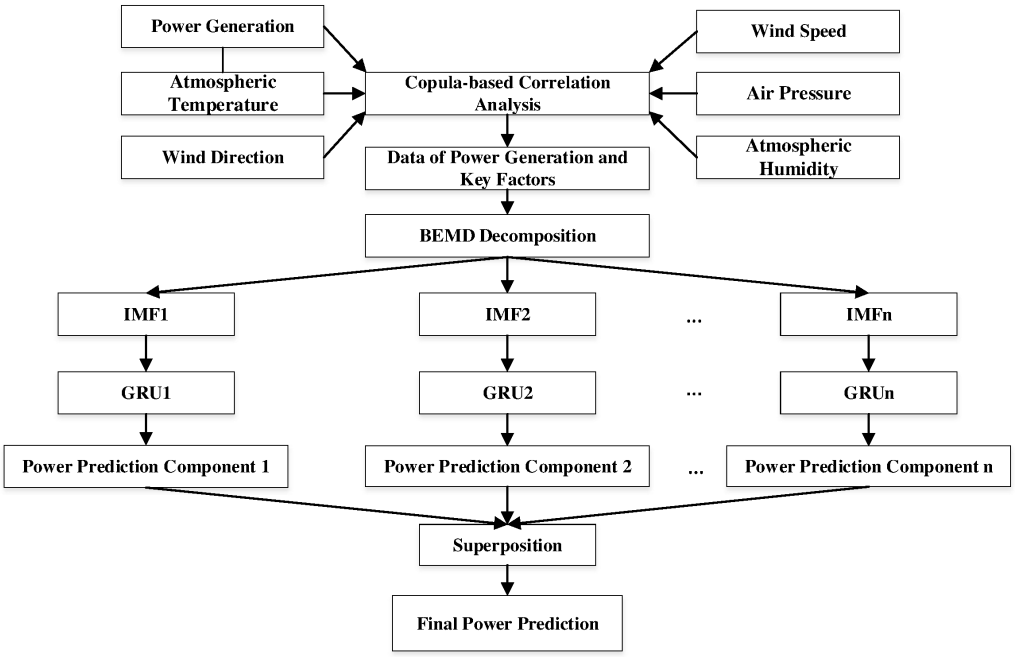


Fig. 1. Ultra-short-term wind power forecasting process

3. Correlation analysis

3.1. Principle of copula function

Copula theory points out that an N-dimensional joint distribution function can be decomposed into N marginal distribution functions and a copula function. The purpose of the copula function is to connect the joint distribution function of each random variable $X_1,\ X_2,\ \ldots, X_N$ with the marginal distribution function of each random variable [25] as shown in Equation (1).

$$F(x_1,\ x_2,\ \ldots,\ x_N) = C\left(F_{X_1}(x_1),\ F_{X_2}(x_2),\ \ldots,\ F_{X_N}(x_N)\right), \tag{1}$$

where: $F(x_1,\ x_2,\ \ldots,\ x_N)$ is the joint distribution function, $F_{X_1}(x_1), F_{X_2}(x_2), \ldots, F_{X_N}(x_N)$ is the marginal distribution function, $C(u_1,\ u_2, \ldots,\ u_n)$ is the copula function.

The commonly used binary copula functions are a binary normal copula, t-copula, the Gumbel copula, Clayton copula and Frank copula function [5], which can describe the linear and non-linear relationship among variables. Based on these copula functions, we use the Kendall rank correlation coefficient and the Spearman rank correlation coefficient to measure the correlation of random variables. These two coefficients describe the degree of consistency between random variables X and Y, their range of values is between -1 and 1, when the correlation coefficient takes a positive number, it means that the variables are positively correlated, and the negative number means that the variables are negatively related. The larger the absolute value of the correlation coefficient, the closer the correlation is.

Assume that the marginal distribution functions of the connected random variables X, Y are $F(x)$ and $G(y)$, respectively, and the corresponding copula function is $C(u, v)$. Then the Kendall rank correlation coefficient, the Spearman rank correlation coefficient and the copula function $C(u, v)$ have the following relationship in Equations from (2) to (6) [5].

$$\tau = 4 \int_0^1 \int_0^1 C(u, v)\, \mathrm{d}C(u, v) - 1, \tag{2}$$

$$\rho_s = 12 \int_0^1 \int_0^1 uv\, \mathrm{d}C(u, u) - 3 = 12 \int_0^1 \int_0^1 C(u, v)\, \mathrm{d}u\, \mathrm{d}v - 3, \tag{3}$$

$$U = F(x) \sim U(0, 1), \tag{4}$$

$$V = G(y) \sim V(0, 1), \tag{5}$$

$$\hat{C}(1 - u, 1 - v) = P(U > u, V > v) = 1 - u - v + C(u, v), \tag{6}$$

where: $\hat{C}(u, v) = u + v - 1 + C(1 - u, 1 - v)$ is called the survival copula function of X and Y.

3.2. Correlation analysis process

The correlation analysis steps are as follows [5, 26]:

1. Data pre-processing. Firstly, through the test of integrity and rationality, find out the missing and abnormal data. Secondly, for missing data points, we supplement them with the following steps. Step 1: calculate the slope between the known values above and below the missing value: $k = (p_n - p_m)/(n - m)$, p_n and p_m are the upper and lower known data points of missing data points, respectively. Step 2: calculate the corresponding missing value $a(i) = b_m + (i - m) \times k$. When the continuous lack of data points lasts more than eight hours, delete the missing data period. Finally, normalize the original data after supplementation and correction through min-max normalization as shown in Equation (7).

$$x_{\text{norm}} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}. \tag{7}$$

2. Determine the marginal distribution for each variable [5, 26]. The marginal probability density distribution function of each random variable is determined by the non-parametric kernel density estimation method based on the sample observation data of the random variable.
3. Select the appropriate copula function. After determining the marginal distribution of the random variables, select the appropriate copula function based on the shape of the binary frequency histogram. Maximum likelihood estimation is used here to determine the parameters of the copula [5, 26].
4. Calculation of the correlation coefficient. Calculate the Kendall and Spearman rank correlation coefficients according to Equations from (2) to (6).

4. BEMD decomposition algorithm

Empirical Mode Decomposition (EMD) proposed by Norden E. Huang [27] is widely used in the analysis of nonlinear and non-stationary signals. The essence of the method is to extract the Intrinsic Mode Function (IMF) from the original signal. Each IMF exhibits relative stationarity and local periodicity. So, the EMD is ideal for processing high frequency, non-linear wind power data [27, 28]. The EMD is based on the intuitive concept of "oscillation" which is naturally related to local extrema. Its basic idea is that "univariate signal = fast oscillation superimposed on slow oscillation". When the decomposed data represent a bivariate signal, the concept of oscillation is not applicable, and it is not clear how to define and interpret local extrema, so the EMD algorithm cannot decompose bivariate data. To solve this problem, a BEMD based on the concept of rotation is proposed. It is a two-dimensional extension of the usual concept of a univariate oscillation. The basic idea of the BEMD is to formalize the following idea: "a bivariate signal = superimposed on the fast rotation of slow rotation" [29].

For the BEMD, the IMF's criteria are as follows:

1. The zero point and pole of the projection in any direction are equal or differ by one.
2. The real part of the projection in any direction is locally symmetrical about the time axis.

The specific steps of the BEMD algorithm are as follows [30]:

Step 1: the bivariate signals $x(t)$are projected on the directions $\phi_k = 2k\pi/N$, $1 \leq k \leq N$ respectively. The real parts of the projection results are taken as shown in (8) [30]. Then the maximum and minimum points of these projections are obtained respectively.

$$P_{\phi_k}(t) = Re\left(e^{-i\phi_k} x(t)\right). \tag{8}$$

Step 2: firstly, each envelope curve is fitted in each direction by a cubic spline interpolation method and tangent $e_{1,\phi_k}(t)$ on the corresponding envelope curve at each time is obtained. And then the average value of the envelope can be calculated according to (9) [30]. The envelope of the bivariate signal is represented as a three-dimensional hose enclosing the original bivariate signal. Only the fixed points of the three-dimensional hose in the horizontal and vertical directions are considered here. So the extreme points can be respectively in the bottom, the top, the right side and the left side of the hose. For a given moment, the mean of the envelope takes the center of the four points.

$$m_1(t) = \frac{2}{N} \sum_k e_{1,\phi_k}(t). \tag{9}$$

Step 3: the difference between the original bivariate signal $x(t)$ and the envelope mean $m_1(t)$ is shown as (10) [30].

$$h_1(t) = x(t) - m_1(t). \tag{10}$$

In order to eliminate the modal aliasing, take $h_1(t)$ as the original signal and repeat the above three steps. As shown in (11) [30].

$$h_{1,1}(t) = h_1(t) - m_{1,1}(t). \tag{11}$$

Repeatedly screening k times to make $h_{1,k}(t)$ becomes an IMF component as shown in (12) [30].

$$h_{1,k}(t) = h_{1,k-1}(t) - m_{1,k}(t). \tag{12}$$

Let $d_1(t) = h_{1,k}(t)$, $d_1(t)$ is the first-order IMF of $x(t)$.

Step 4: the residual component of the original signal $r_1(t)$ can be obtained by subtracting the original bivariate signal $x(t)$ from the first-order IMF $d_1(t)$. Repetition of the above steps for the residual component yields IMFs of all orders and a the residual component that characterizes a certain trend of the original signal. Thus, the original bivariate signal can be expressed as the superposition of each IMF and residual component, as shown in (13) [30].

The stopping condition of the above decomposition process is any of the following:

1. When the IMF $d_n(t)$ or the residual component $r_n(t)$ is small enough;
2. When the residual component $r_n(t)$ becomes a single point bivariate function, the IMF cannot be decomposed from it.

$$x(t) = \sum_{i=1}^{n} d_i(t) + r_n(t). \tag{13}$$

5. Ultra-short-term wind power prediction model

5.1. Ultra-short-term wind power prediction

With the increasing installed capacity of wind power and the randomness of wind power, accurate wind power prediction is very important for the safe and stable operation of the power generation system. According to different time scales, wind power forecasting can be divided into different types. Ultra-short-term wind power forecasting is to predict the wind power output in the next four hours. Accurate ultra-short-term wind power forecasting can effectively strengthen power grid management, improve power consumption, promote wind power grid integration, realize mutual assistance and complementary dispatching of power.

5.2. GRU principle

As a variant of LSTM, there are only two gate structures in a GRU, an update gate and reset gate, respectively. The role of the update gate is to determine how much state information needs to be retained from past moments. The amount of the state information retained from the past moments is proportional to the value of the update gate and the reset gate is used to decide, whether to combine the current state information with the state information of the past moments. The larger the value of the reset gate, the less information is ignored. The structure of the GRU is shown in the Fig. 2 [31].

As shown in the structure diagram, x_t is the input of the GRU structure, h_t is the output of a hidden layer, and the related calculation of the unit structure is shown in (14) to (17) [31].

$$z_t = \sigma\left(W^{(z)}x_t + U^{(z)}h_{t-1}\right), \tag{14}$$

$$r_t = \sigma\left(W^{(r)}x_t + U^{(r)}h_{t-1}\right), \tag{15}$$

$$\widetilde{h}_t = \tanh\left(Wx_t + U(r_t * h_{t-1})\right), \tag{16}$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \widetilde{h}_t, \tag{17}$$

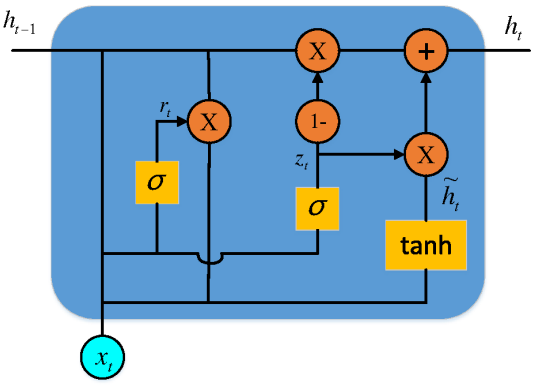


Fig. 2. Gated recurrent unit

where: z_t and r_t are the update gates and reset gates respectively; x_t is the input; h_{t-1} is the output of the upper hidden layer; $\widetilde{h}_t$ is the summary of x_t and h_{t-1}; σ is the Sigmoid function; tanh is the hyperbolic tangent function; $W^{(z)}$, $U^{(z)}$, $W^{(r)}$, $U^{(r)}$, W, U represent the training parameter matrix; $*$ is the product of the matrix.

5.3. Modeling process

The ultra-short-term wind power prediction model base on a GRU can be described by the following (18):

$$\begin{aligned}(P(t),\, P(t+1),\, \ldots,\, P(t+k)) = \\ = F(P(t-1),\, P(t-2),\, \ldots,\, P(t-n),\, m(t-1),\, m(t-2),\, \ldots,\, m(t-n)),\end{aligned} \tag{18}$$

where: $P(t)$, $P(t+1)$, ..., $P(t+k)$ represent the value of predicted power for the next four hours, $k = 15$; $W(t-1)$, $W(t-2)$, ..., $W(t-n)$ are the historical power value before the current time; $m(t-1)$, $m(t-2)$, ..., $m(t-n)$ represent the historical observation value of the meteorological factor filtered by the copula function; the parameter n is obtained by the ergodic experiment.

For the experiments of the paper, correlation analysis, BEMD and a GRU neural network prediction model are implemented in MATLAB R2016a. The implement environment for all of the calculations is a personal computer with an Intel i5-8250U CPU and 8 GB of RAM.

5.4. Prediction accuracy assessment

In order to evaluate the performance of the prediction model, we use root-mean-square percentage error (RMSPE) and mean absolute percentage error (MAPE) as the evaluation indicators for the performance of the prediction model. They estimate the total error and real-time bias respectively. The two indicators are calculated by Equations from (19) to (20) [5].

$$\text{RMSPE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\frac{(x_t - \hat{x}_t)^2}{xt}} \times 100\%, \tag{19}$$

$$\text{MAPE} = \frac{1}{N}\sum_{t=1}^{N}\frac{|x_t - \hat{x}_t|}{xt} \times 100\%, \tag{20}$$

where: x_t is the true value of wind power at time t; $\hat{x}_t$ is the power prediction value of the prediction model at time t; N is the number of data points in the test set.

6. Experimental analysis

6.1. Data

The data used in the three experiments bellow is collected from a wind farm in Jiangsu Province on the eastern coast of China. The data include wind speed and direction of wind turbine hub height, atmospheric temperature, atmospheric humidity, air pressure, and power generation. The time interval between data points is fifteen minutes, a total of 35040 data points from January 1 to December 31, 2013. The installed capacity of the wind farms is 49 MW.

6.2. Result

A. Correlation analysis between meteorological factors and wind power

In this part, the correlation analysis between the wind speed and the wind power is taken as an example to analyze. According to the correlation analysis step described above, the binary frequency histogram between wind speed and wind power generation (WPG) is shown in Fig. 3. According to Fig. 3, all the data are distributed on the diagonal, decreasing from the opposite ends of the diagonal to the middle. Correspondingly, the density function of the joint distribution will show the same characteristics. Therefore, according to these features expressed by the frequency histogram, the binary normal copula function and the t-copula function are selected here as a joint probability density function between wind speed and wind power.

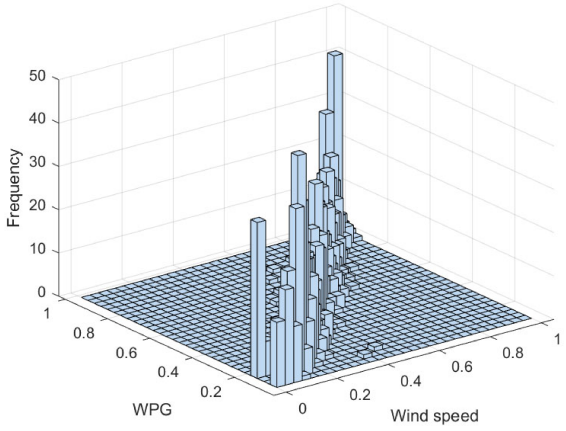
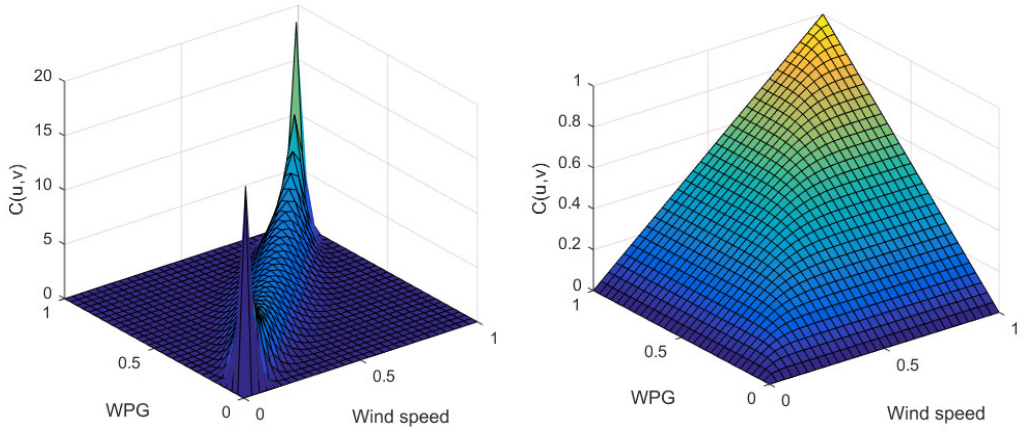


Fig. 3. Binary frequency histogram of wind speed and wind power

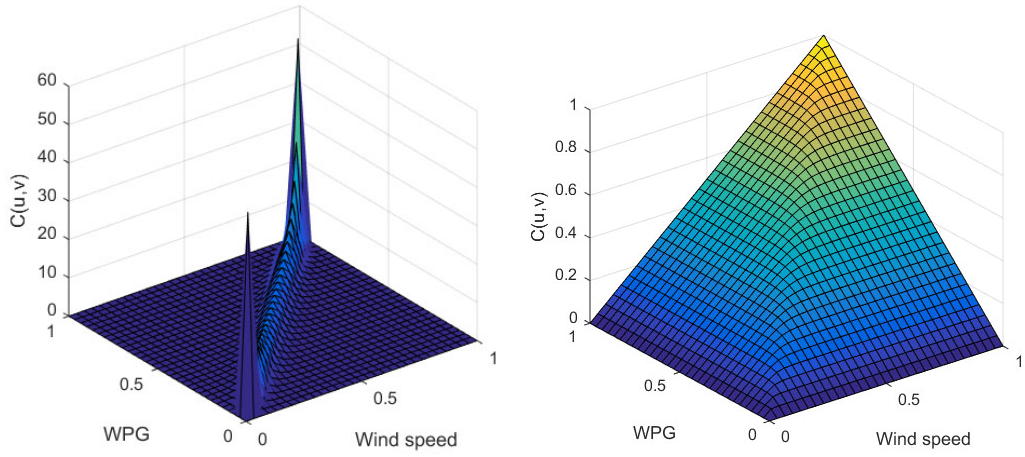
After determining the copula function, the joint distribution model of wind speed and wind power can be obtained. The experimental results of the density function and the distribution

function are shown in Fig. 4 and Fig. 5. As shown in Fig. 4(a) and Fig. 5(a), the wind speed and wind power distribution are mainly concentrated on a diagonal of 45°. Similar to the frequency histogram, the density function shows a peak at both ends of the diagonal and a thick tail in the middle. Wind speed and wind power have strong tail related features. When the wind speed and wind power is high or low, the correlation between the two is obvious. For a binary normal copula function, the calculation results of Kendall and Spearman rank correlation coefficients are 0.9233 and 0.9580, respectively. While the corresponding values of the binary t-copula function are 0.9884 and 0.9744, respectively. The experimental result shows a strong rank correlation between wind speed and wind power.



(a) (b)

Fig. 4. Joint distribution model based on binary normal copula function: density function (a); distribution function (b)



(a) (b)

Fig. 5. Joint distribution model based on the binary t-copula function: density function (a); distribution function (b)

Through the same method, we can obtain other results between different meteorological factors and wind power, which are shown in Table 1.

Table 1. Correlation analysis results between meteorological factors and wind power

Binary copula function	**Kendall**		**Spearman**	
	Normal copula	**t-copula**	**Normal copula**	**t-copula**
Wind speed	0.9233	0.9884	0.9580	0.9744
Wind direction	0.1479	0.1458	0.1217	0.1683
Pressure	−0.2344	−0.2239	−0.2015	−0.2778
Temperature	−0.2610	−0.2088	−0.2323	−0.2682
Humidity	0.3429	0.3049	0.3231	0.3534

B. Bivariate EMD decomposition of wind speed and wind power

According to the correlation analysis results above, the wind speed and wind power with the highest correlation are selected as input to the prediction model. The observation data of the wind speed and the wind power at the same time are combined into a plural form as an input of the BEMD. The wind power is taken as the real part while the wind speed is the imaginary part. Some of the original experimental data are shown in Fig. 6. From the graph, the non-stationary data of wind speed and wind power time series data can be observed. This feature has a great influence on the accuracy of prediction results. The original data is decomposed into 10 IMF components

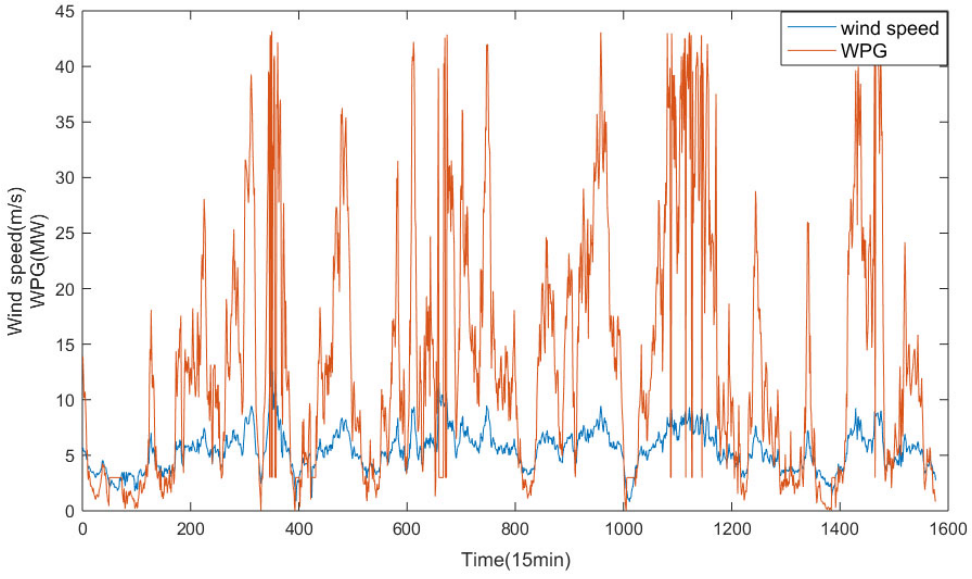


Fig. 6. Initial data of wind speed and wind power

by BEMD, where the last IMF component represents the residual component of the original data trend. Each IMF component is shown in Fig. 7. Comparing with the original data, it can be concluded that the IMF component fluctuation after BEMD decomposition gradually becomes

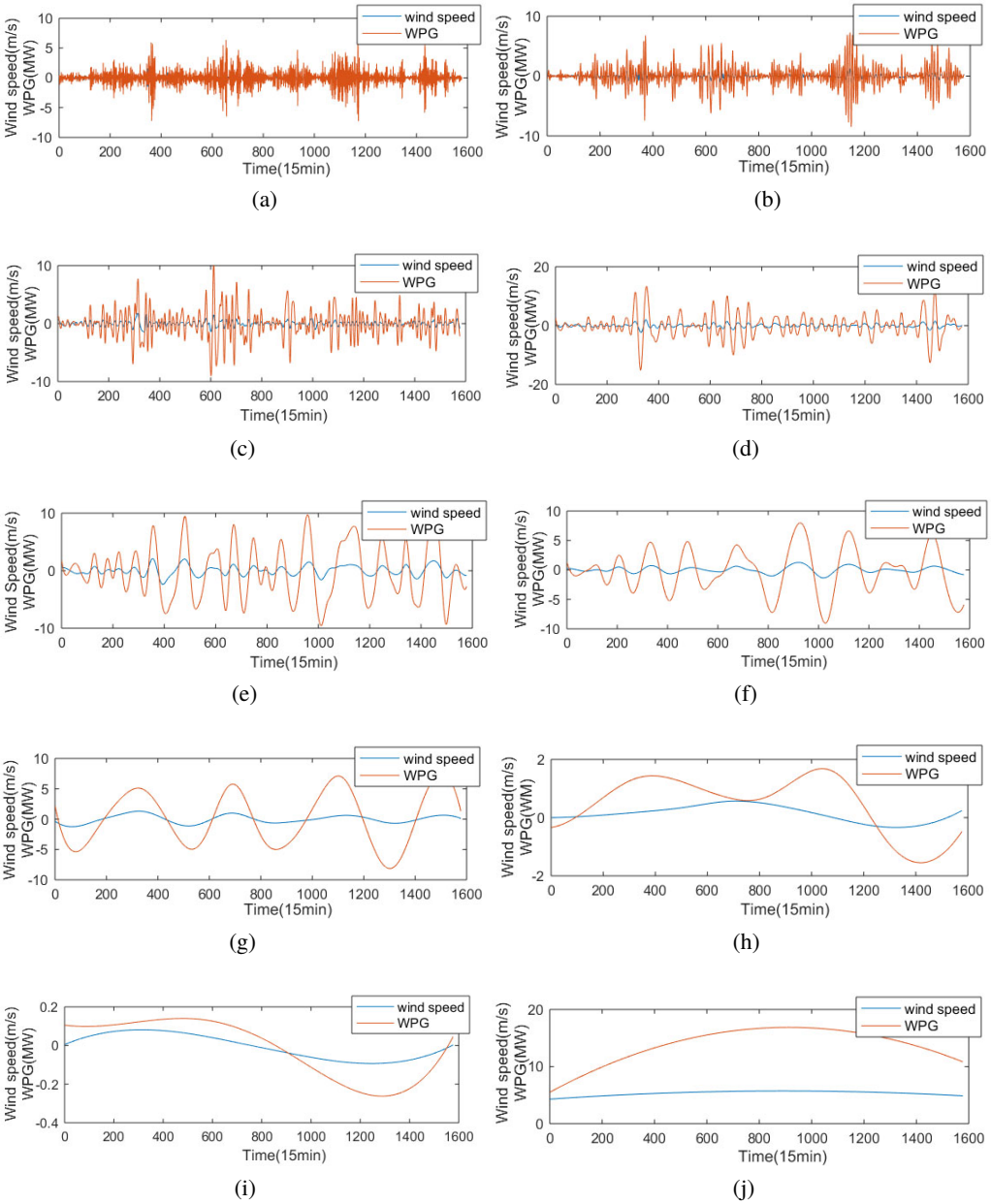


Fig. 7. BEMD decomposition results of wind speed and wind power: IMF1 (a); IMF2 (b); IMF3 (c); IMF4 (d); IMF5 (e); IMF6 (f); IMF7 (g); IMF8 (h); IMF9 (i); IMF10 (j)

smooth, and the later the IMF component, the smoother the fluctuation becomes. In addition, the decomposed data shows local relative periodicity and symmetry. Therefore, although the three components of IMF1, IMF2, and IMF3 fluctuate at a high frequency, the characteristics of local periodicity and symmetry make the prediction results of these three IMF components have good accuracy.

C. Analysis of ultra-short-term wind power prediction results

In order to verify the effectiveness of our method, we set up three groups of comparative experiments: Multilayer Perceptron (MLP) (power generation input) and GRU (power generation input), GRU (power generation input) and GRU (wind speed + power generation input), GRU (wind speed + power generation input) and GRU (wind speed + power generation input-BEMD). These three groups of comparative experiments verify respectively the validity of the correlation analysis based on a copula function, the data decomposition based on BEMD and a GRU neural network to improve the accuracy of ultra-short-term wind power prediction. The predicted results are shown in Fig. 8 and Table 2.

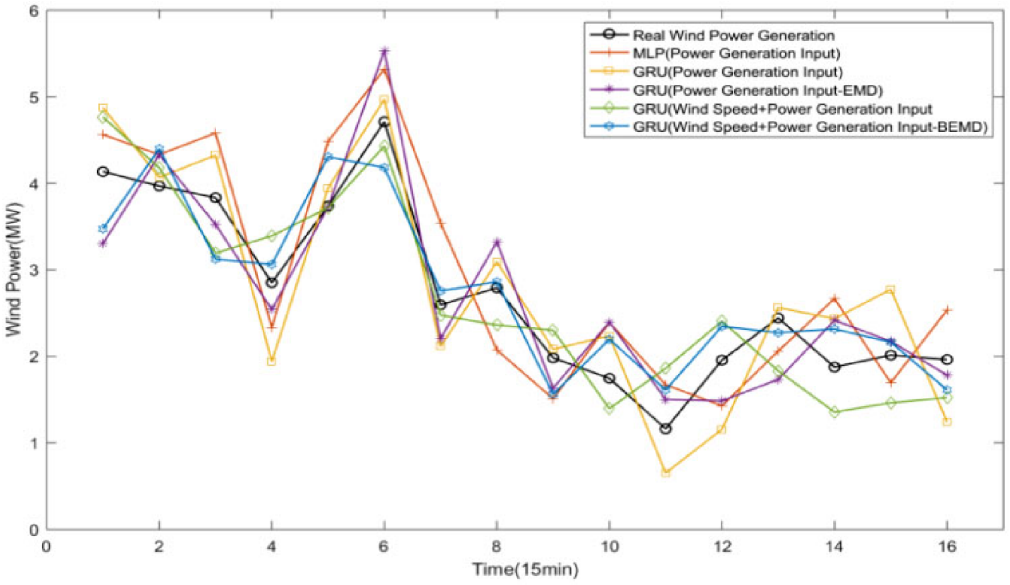


Fig. 8. Wind power prediction result for the next 4 hours

To further verify the validity of the proposed model, RMSPE is used to evaluate the overall error during the predicted period and MAPE is used to evaluate the real-time deviation. Compare the MLP (power generation input) and GRU (power generation input) experiments, the MAPE and RMSPE decrease by 3.37% and 2.91%, respectively. Compare GRU (power generation input) and GRU (wind speed + power generation input) experiments, the MAPE and RMSPE decrease by 1.75% and 4.29%, respectively. Compare GRU (wind speed + power generation input) and GRU(wind speed + power generation input-BEMD), the MAPE and RMSPE decrease by 3.46% and 5.29%, respectively.

Table 2. Performance comparison of different forecast models

Prediction model	MAPE	RMSPE
MLP (power generation input)	24.15%	38.73%
GRU (power generation input)	20.78%	35.82%
GRU (power generation input-EMD)	17.68%	30.28%
GRU (wind speed + power generation input)	19.03%	31.54%
GRU (wind speed + power generation input-BEMD)	15.57%	26.25%

7. Conclusions

We propose a method to predict ultra-short-term wind power based on a copula function, BEMD and GRU neural network. After theoretical analysis and setting up comparison experiments, the following conclusions were obtained.

1. Through the Kendall and Spearman rank correlation coefficients, the copula function can effectively analyze the nonlinear correlation between meteorological factors as well as wind power and extract key meteorological factors.
2. BEMD can effectively decompose bivariate data, and the accuracy of ultra-short-term wind power prediction based on the decomposition results has been improved.
3. The performance of a GRU neural network is better than a MLP neural network in dealing with the time series prediction of ultra-short-term wind power prediction. The above research results show that the proposed method for predicting ultra-short-term wind power is effective.

References

[1] Girouard N., Konialis E., Tam C., Taylor P., *OECD Green Growth Studies-Energy*, Preliminary Version, OECD (2011).

[2] Aliyu A.K., Modu B., Tan C.W., *A review of renewable energy development in Africa: A focus in South Africa, Egypt and Nigeria*, Renewable and Sustainable Energy Reviews, vol. 81, no. 2, pp. 2502–2518 (2018).

[3] https://irena.org/publications/2019/Mar/Renewable-Capacity-Statistics-2019, accessed March 2019.

[4] Waśkowicz B., *Statistical analysis and dimensioning of a wind farm energy storage system*, Archives of Electrical Engineering, vol. 66, no. 2, pp. 265–277 (2017).

[5] Han S., Qiao Y.H., Yan J., *Mid-to-long term wind and photovoltaic power generation prediction based on copula function and long short term memory network*, Applied Energy, vol. 239, pp. 181–191 (2019).

[6] Chen N., Qian Z., Nabney I.T., Meng X., *Wind power forecasts using Gaussian processes and numerical weather prediction*, IEEE Transactions on Power Systems, vol. 29, no. 2, pp. 656–665 (2014).

[7] Carvalho D., Rocha A., Gómez-Gesteira M., Santos C., *WRF wind simulation and wind energy production estimates forced by different reanalyses: comparison with observed data for Portugal*, Applied Energy, vol. 117, no. 3, pp. 116–126 (2014).

[8] Shi J., Ding Z., Lee W., Yang Y., Liu Y., Zhang M., *Hybrid Forecasting Model for Very-Short Term Wind Power Forecasting Based on Grey Relational Analysis and Wind Speed Distribution Features*, IEEE Transactions on Smart Grid, vol. 5, no. 1, pp. 521–526 (2014).

[9] He D., Ruiye Liu R., *Ultra-short-term wind power prediction using ANN ensemble based on PCA*, International Power Electronics and Motion Control Conference, Harbin, China, pp. 2108–2112 (2012).

[10] Gyu G.K., Jin H.C., So Y.P., Byeong G.B., Woo J.N., Hae L.C., *Prediction Model for PV Performance With Correlation Analysis of Environmental Variables*, IEEE Journal of Photovoltaics, vol. 9, no. 3, pp. 832–841 (2019).

[11] Jing P., Su Y., Jin X., Zhang C., *High-Order Temporal Correlation Model Learning for Time-Series Prediction*, IEEE Transactions on Cybernetics, vol. 49, no. 6, pp. 2385–2397 (2019).

[12] Sumit S., Aggarwal S.K., *Wind power forecasting using wavelet transforms and neural networks with tapped delay*, CSEE Journal of Power and Energy Systems, vol. 4, no. 2, pp. 197–209 (2018).

[13] Yang M., Chen X., Jian Du J., Cui Y., *Ultra-Short-Term Multistep Wind Power Prediction Based on Improved EMD and Reconstruction Method Using Run-Length Analysis*, IEEE Access, vol. 6, pp. 31908–31917 (2018).

[14] Su Y., Wang S., Xiao Z., Tan M., Wang M., *An Ultra-Short-Term Wind Power Forecasting Approach Based on Wind Speed Decomposition, Wind Direction and Elman Neural Networks*, 2018 2nd IEEE Conference on Energy Internet and Energy System Integration, Beijing, China, pp. 1–9 (2018).

[15] Yeh J., Shieh J., Huang N.E., *Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method*, Advances in Adaptive Data Analysis, vol. 2, no. 2, pp. 135–156 (2010).

[16] Wu J.L., Ji T.Y., Li M.S., Wu Q.H., *Multi-step wind power forecast based on similar segments extracted by mathematical morphology*, Power and Energy Engineering Conference, Hong Kong, China, pp. 1–6 (2015).

[17] Cui M., Peng X., Xia J., Sun Y., Wu Z., *Short term power forecasting of a wind farm based on atomic sparse decomposition theory*, IEEE International Conference on Power System Technology, Auckland, New Zealand, pp. 1–5 (2013).

[18] Chen N., Qian Z., Nabney I., Meng X., *Wind power forecasts using Gaussian processes and numerical weather prediction*, IEEE Trans. Power Syst., vol. 29, no. 2, pp. 656–665 (2014).

[19] Xie W., Zhang P., Chen R., Zhou Z., *A Nonparametric Bayesian framework for short-term wind power probabilistic forecast*, IEEE Trans. Power Syst., vol. 34, no. 1, pp. 371–379 (2019).

[20] Wan C., Wang J., Lin J., Song Y., Dong Z., *Nonparametric prediction intervals of wind power via linear programming*, IEEE Trans. Power Syst., vol. 33, no. 1, pp. 1074–1076 (2018).

[21] Zeng J.W., Qiao W., *Short-term wind power prediction using a wavelet support vector machine*, IEEE Transactions on Sustainable Energy, vol. 3, no. 2, pp. 255–264 (2012).

[22] Sideratos G., Hatziargyriou N., *Probabilistic wind power forecasting using radial basis function neural networks*, IEEE Trans. Power Syst., vol. 27, no. 4, pp. 1788–1796 (2012).

[23] Luo X., Sun J., Wang L., Wang W., Zhao W., Wu J., *Short-term wind speed forecasting via stacked extreme learning machine with generalized correntropy*, IEEE Trans. Ind. Inf., vol. 14, no. 11, pp. 4963–4971 (2018).

[24] Zaccheus O., *A 5-day wind speed and power forecasts using a layer recurrent neural network (LRNN)*, Sustainable Energy Technologies and Assessments, vol. 6, pp. 1–24 (2014).

[25] Xie Z.Q., Ji T.Y., Li M.S., Wu Q.H., *Quasi-Monte Carlo Based Probabilistic Optimal Power Flow Considering the Correlation of Wind Speeds Using Copula Function*, IEEE Transactions on Power Systems, vol. 33, no. 2, pp. 2239–2247 (2018).

[26] Xie Z., *Statistical Analysis and Application of MATLAB: An Analysis of 40 Cases*, Beihang University Press (2010).

[27] Huang N.E., Shen Z., *The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis*, Proceedings: Mathematical, Physical and Engineering Sciences, vol. 454, pp. 903–995 (1998).

[28] An X., Jiang D., Zhao M., Liu C., *Short-term prediction of wind power using EMD and chaotic theory*, Communications in Nonlinear Science and Numerical Simulation, vol. 17, no. 2, pp. 1036–1042 (2012).

[29] Rilling G., Flandrin P., Goncalves P., *Bivariate Empirical Mode Decomposition*, IEEE Signal Processing Letters, vol. 14, no. 12, pp. 936–939 (2007).

[30] Jiang T., *Subband-based EMD Signal Decomposition Algorithm*, PhD Thesis, School of Electronic Engineering, Xidian University, Xi'an (2013).

[31] Li W.X., Logenthiran T., Woo W.L., *Multi-GRU prediction system for electricity generation's planning and operation*, IET Generation, Transmission and Distribution, vol. 13, no. 9, pp. 1630–1637 (2019).