Wind speed time series forecasting method in wind farm based on EEMD

Yu Dongyang¹*, Sun Fengchang¹, Deng Xiaochuan¹, Wang Zheng¹, Wu Jiahua¹, Zhang Zhichao¹
¹Beijing Branch-State Grid Information and Telecommunication Group, Beijing, China
*Corresponding author’s e-mail: yudongyang@sgitg.sgcc.com.cn

Abstract. The article described the relevant knowledge of wind speed prediction and the verification method of the forecast. First, the characteristics of wind speed are general analyzed. Through the form of wind, speed frequency of wind, shear speed and turbulence characteristics of wind, it was found that the wind speed had instability and non-linear characteristics. By analysing the basic principles of the three methods included the time series, neural network method and empirical mode decomposition, the wind speed time series prediction method was proposed. On the basis of this method, it was applied to a raw wind factory speed time series and simulated reality by Matlab. Compared the predictive conclusion with original data, it gained a higher accuracy and verified the efficiency of the method.

1. Introduction
Wind speed prediction is to study or use a series of systematic mathematical methods to deal with the past and future wind speed under the condition of considering all the influencing factors. It is used to determine the average wind speed that can meet the accuracy requirements at a certain time or period in the future. Research methods at home and abroad are mainly divided into physical methods and statistical methods. In this paper, time series method, EEMD and neural network method are combined in wind speed prediction. First, the EEMD algorithm is used to decompose the original wind speed time series, and a series of eigenmode components with relatively stable changes can be obtained. Then the neural network method is used to predict each component one by one, and the final wind speed prediction value can be obtained by adding the predicted values of each component.

2. Time series method
Time series refers to the series formed by arranging the values of a certain phenomenon and a certain statistical index at different times in chronological order. The so-called time series analysis is to study all the data, obtain the law of their long-term changes, understand the dynamic system to be studied by revealing the law, forecast and control future events, and improve the level of business decision-making[1].

The development of time series model has experienced a process from linear to nonlinear. Linear time series models include AR model, MA model, ARMA model, ARIMA model, etc. Nonlinear models mainly include threshold autoregressive model, exponential autoregressive model, GARCH model and so on. Since the time series model has a poor smoothing effect when dealing with wind speed data, this paper chooses the EEMD method to decompose wind speed data.
3. EMD and EEMD

3.1. EMD

Empirical Mode Decomposition (EMD) is a data analysis and processing method proposed by N. E. Huang et al. It is a new method for nonlinear and non-stationary signal processing. This method is intuitive, direct, posterior and adaptive, because the basis function is decomposed by the signal itself[2].

The essence of empirical mode decomposition (EMD) is to stabilize the non-stationary signal, and the result is to decompose the fluctuation or trend of different scales step by step. Thus, a series of data components with different characteristic scales are obtained, which are called intrinsic mode function. When EMD is used to decompose the data signal, the obtained IMF must satisfy the following two conditions[3]:

- For the whole time series, the number of extreme points must be equal to the number of crossing zeros, or the difference between them is 1.
- At any point, the average value of the maximum and minimum envelope is 0.

For the wind speed series \( \{x(t)\} \), empirical mode decomposition can be carried out by the following method:

- All the local extremum points of \( x(t) \) are determined, and all the maximum and minimum points of \( x(t) \) are connected by cubic spline interpolation to form the upper envelope \( x_{\text{max}}(t) \) and lower envelope \( x_{\text{min}}(t) \). The mean value \( m(t) \) of the two envelope lines is calculated, and then the difference \( h(t) \) between \( x(t) \) and the mean value is obtained.

\[
m(t) = \frac{x_{\text{max}}(t) + x_{\text{min}}(t)}{2} \tag{1}
\]

\[
h(t) = x(t) - m(t) \tag{2}
\]

- Judge whether \( h(t) \) meets the conditions of IMF, if so, let \( h(t) \) be the first IMF, record \( c_1(t) = h(t) \), and the difference \( r_1(t) \) between the original signal and the IMF is calculated; If not, repeat the above process several times until the new \( h(t) \) satisfies the IMF condition.

- Take \( r(t) \) as the signal to be decomposed, repeat steps (1) and (2). When the remaining signal is displayed as a monotone function, the decomposition process stops. Then the original sequence can be expressed as:

\[
x(t) = \sum c_i(t) + r_n(t) \tag{3}
\]

Compared with many signal processing methods, empirical mode decomposition has certain advantages. However, due to the discontinuity of the initial signal, the phenomenon of mode aliasing sometimes occurs in the decomposition process (that is, one IMF contains a wide range of totally different scale signals, or different IMF components contain similar scale signals). It will affect the actual effect of decomposition and reduce the prediction accuracy. The causes of mode aliasing can be divided into two types: The first one is caused by the discontinuity of the signal, which can be caused by the insufficient number of extremum points or uneven distribution; The second type is the problem of mode aliasing caused by the failure to separate some frequency signals correctly[4].

3.2. EEMD

In order to reduce the adverse effects of modal aliasing and improve the accuracy of EMD decomposition, N. E. Huang proposed a method, which is to add noise auxiliary analysis in the decomposition process, that is the EEMD method[5]. The core idea of this method is to believe that each observed data incorporates actual time series information and varying degrees of noise, and its
The overall mean value will be close to the real time series. Therefore, in order to extract the actual signal of the data, a white noise sequence is added to the original sequence. The amplitude of the white noise sequence must be limited. The combination of signal and noise is used as a signal to be decomposed, and the uniform distribution of the white noise spectrum is used. When the signal is loaded on a white noise background with uniform distribution throughout the time-frequency space, the signals of different time scales will be automatically distributed to the appropriate reference scale, and due to the characteristics of zero mean noise, they are respectively decomposed by EMD, and then the mean value of the corresponding IMF component is regarded as the true component.

The specific decomposition steps of EEMD are as follows:
- The white noise sequence is added to the target data sequence.
- The signal with white noise is decomposed into multiple IMFs according to the above EMD method.
- Repeat the above steps n times, but the white noise sequence added each time is different from each other; Finally, the mean value of IMF obtained by decomposition is taken as its final result.

The size of white noise in ensemble empirical mode decomposition algorithm needs to be determined according to the experience of the experimenters, which is lack of adaptability and reliability for different signals. If the amplitude of the added white noise sequence is too small, it has no good effect to solve the problem of mode aliasing. If the amplitude is too large, it will cause some interference to the decomposition process and affect the final decomposition result. Since the EEMD method is more sensitive to auxiliary noise, the amplitude of the added auxiliary white noise sequence is usually relatively small. In general, the standard deviation of the added auxiliary white noise is 0.2 times of the signal standard deviation.

4. Neural network method
Artificial neural network (ANN) is a highly integrated interdisciplinary subject. Its research and development involves many fields of computer science and mathematical science. It has achieved remarkable results in applications in many fields such as signal processing, pattern recognition, target tracking, robot control, expert systems, combinatorial optimization, and network management. Artificial neural network is a simplified model to simulate human brain neural process. Because of its excellent ability in self-learning ability, nonlinear fitting and generalization ability, it has been widely used in the prediction of wind power, and achieved better prediction accuracy than other methods in a certain range. Artificial neural networks include back propagation artificial neural network (BP-ANN) and radial basis function (RBF).

The dynamic process of artificial neural network is divided into two stages: learning stage and working stage. In the learning phase, the selected learning samples are sequentially added to the network with the initial weight randomly set in the form of (input, output) sample pairs. When the actual output generated by the sample input through the network is different from the expected sample output, according to the calculation deviation, the weights and structure of the artificial neural network are adjusted according to a certain algorithm until the actual output of the network is exactly the same as or close to the expected output. Since the weights adjusted by one (input, output) sample pair may not meet the requirements of other (input, output) sample pairs, it is necessary to input all the learning sample pairs repeatedly for learning. The learning process is also called the slow process in the whole dynamic process. Since then, Ann has entered the working stage. At this time, the connection right is fixed and has reached a certain stable state. The neural network calculates the corresponding output mode according to the input mode. This process is also called fast process.

5. Example analysis
In this paper, the wind speed data of a wind farm is taken as the test sample to verify the proposed prediction algorithm. In this experiment, the standard of the original data is sampling every hour, and the data of 744 consecutive sampling points are selected according to the time sequence. The last 24
sampling points are taken as the test samples, and the latest 720 points are used to predict the wind speed value at the next moment. The original wind speed time series is shown in Figure 1.

- **EMD decomposition**
  EMD is used to decompose the original wind speed time series, and six IMF are obtained. The decomposition results are shown in Figure 2. Compared with the original wind speed time series, these components change more smoothly.
- **EEMD decomposition**

  The original wind speed data are decomposed by ensemble empirical mode decomposition (EMD), and white noise with equal length and normal distribution is added to the sequence for many times. A total of 10 IMF are generated. The decomposition results are shown in Figure 3. By comparison, it can be found that the number of IMF decomposed by EEMD is more than that decomposed by EMD. These components are not only more stable than the original wind speed time series, but also more stable than those decomposed by empirical mode, which greatly weakens the adverse effect of mode aliasing and improves the prediction accuracy.

- **Prediction by neural network method**

  The first 720 sample points of wind speed time series are taken as the training samples of neural network method, and the 10 decomposed components are predicted one by one. Finally, all the predicted values are accumulated to get the prediction results of the next time. Figure 4 shows the comparison chart of all predicted values and actual values. It can be seen that the two curves basically overlap, indicating that the prediction effect of this method is better.
6. Conclusion
Accurate wind speed prediction is an important foundation to ensure the safe and stable operation of power systems after large-scale wind power integration. Due to the non-stationary and non-linear characteristics of wind speed time series, it is difficult to guarantee the accuracy of multi-step prediction. Therefore, this paper proposes a wind speed time series forecasting method based on EEMD and neural network method. EEMD can effectively solve the modal aliasing phenomenon of the original EMD decomposition. The IMF obtained by the EEMD decomposition is more stable than the EMD decomposition result, which simplifies the interference and coupling between different feature information and improves the prediction accuracy. The neural network method is used to predict each component one by one, and then the prediction results are superimposed to get the final prediction value. The prediction accuracy is high, which is an effective prediction method.

References
[1] Zhan, S.S., Chen, R.W. (2015) Development and Application of Time Series Analysis Technology. J. Higher Education Online, 824: 153-154.
[2] Chu, F.L., Peng, Z.K., Feng, Z.P, et al. (2009) Modern Signal Processing Methods in Mechanical Fault Diagnosis. Science Press, Beijing.
[3] Wu, Z.H., Huang, N.E. (2003) A study of the characteristics of white noise using the empirical mode decomposition method. J. Process of the Royal Society of London Series A, 460(2046):1597-1611.
[4] Meng, F.L. (2012) Research on Theories and Applications of Ensemble Empirical Mode Decomposition. D. Zhenjiang: Jiangsu University of science and Technology.
[5] Wu, Z., Huang, N.E. (2005) Ensemble empirical mode decomposition: a noise-assisted data analysis method. R. Calverton: Center for Ocean-Land-Atmosphere Studies.
[6] Yi, F. (2005) Research on Neural Network Prediction. D. Chengdu: Southwest Jiaotong University.