

Short-Term Wind Power Prediction Based on Empirical Mode Decomposition and Extreme Learning Machine

Jiajia Wu^{1, a}, Changliang Liu^{2, b}

¹ School of Control and Computer Engineering, North China Electric Power University, Baoding, 071003, China

² State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (NCEPU), Beijing, 100085, China

^aemail: wu_jia_jia@126.com, ^bemail: changliang_liu@163.com

Keywords: Wind Power; Power Prediction; Empirical Mode Decomposition; Extreme Learning Machine

Abstract. Wind power prediction is important for the power system with plenty of wind power. This paper studies the method combined with empirical mode decomposition and extreme learning machine for short-term wind power prediction. The empirical mode decomposition method is utilized to decompose the signal of wind power into sequences with different characteristic scale. The extreme learning machine method is used to model and predict each sequence. Eventually, the prediction results of each sequence are added to obtain the final wind-power prediction results. The simulation result shows that the proposed method in this study improves the prediction accuracy of wind power prediction.

Introduction

Wind-energy generation is a clean and renewable energy. Given significant advantages of wind-energy generation, more and more attention has been drawn around the world [1]. Wind-energy generation is inherently intermittent and random, which leads to strong fluctuation of wind-energy generation. Therefore, large-scale wind-power networks will inevitably bring the difficulty of controlling the power grid and enormous challenges of safe and stable operation [2]. Accurate prediction of wind-energy generation will enable power system dispatching department to make timely adjustments to scheduling based on wind power changes, which can effectively reduce the adverse effects of wind power fluctuations on the grid, ensure the quality of the grid, significantly reduce the grid spinning reserve capacity, and effectively reduce the cost of power system operation[3].

Nowadays, there are vast numbers of domestic and foreign scholars working on wind power prediction. Traditional wind power prediction methods include regression analysis [4], time series [5], Kalman filtering method [6] etc. But the randomness and nonlinear of wind power constrains the application of traditional prediction methods. Currently, artificial neural networks (ANN) [7] and support vector machine (SVM) [8] and other intelligent algorithms are widely used in short-term wind power prediction. Extreme Learning Machine (ELM) is a new feed-forward neural network method [9] proposed by Hang et al 2006, which has good training speed and generalization ability, and so on. ELM overcomes the disadvantages of traditional neural networks. ELM is widely applied in classification and regression [9], the time series prediction [10], wind power interval prediction [11], economic dispatch [12] and so on.

This paper presents a short-term wind power prediction model based on EMD and nuclear extreme learning machine. For non-stationary time series of wind power, first of all, time series of wind power will be broken down into a series of data sequences with different characteristics scales using empirical mode decomposition (EMD); then ELM models each data series of IMF and predicts each data series; eventually predictions for each data series are added to get the final value of original wind power prediction. The simulation based on the data of a real wind farm show that the proposed short-term wind power prediction models predict faster with higher accuracy and good

generalization performance.

Empirical Mode Decomposition

EMD is an adaptive signal decomposition method, which is widely used to analyze nonlinear and non-stationary signals. In this study, EMD will be applied to analyze time series of wind power. EMD is a smooth signal processing, which will decompose the fluctuation at different scales and trends present in the signal under stepwise, producing a series of data sequences with different characteristics scales called intrinsic mode function (IMF). The decomposed IMF components comprise local characteristics of different time scales of the original signal. The procedure of EMD decomposition in wind power time series is shown as follows:

(1) Make sure all maxima and minima of signals $x(t)$, forming the upper envelope $e_1(t)$ and the lower envelope $e_2(t)$ to calculate the average of the upper and lower envelope m :

$$m = \frac{1}{2}(e_1(t) + e_2(t)) \quad (1)$$

(2) The difference of data $x(t)$ and the mean value of upper and lower envelope m is expressed as follows

$$h_1 = x(t) - m \quad (2)$$

h_1 is seen as the new $x(t)$. Repeat the above step, until h_i satisfies the IMF conditions. The first component of IMF chosen from the original wind power series is denoted as c_1 , $c_1 = h_i$;

(3) To obtain the remaining components, separate c_1 from $x(t)$:

$$r_1 = x(t) - c_1 \quad (3)$$

r_1 is regarded as the new signal $x(t)$. Repeat the above step. Obtain the remaining components c_2 , $c_3 \dots$, sequentially, until the n-order remaining component becomes monotone function. Then do not decompose IMF components. So

$$x(t) = \sum_{i=1}^n c_i(t) + r_n(t) \quad (4)$$

Extreme Learning Machine

Extreme learning machine is a new type of feedforward neural networks. The weights connecting the hidden layer and the input layer and the thresholds of hidden layer neurons are randomly given and kept the same in the training process in this method, then the output weights is calculated by regularization principle, the neural network is still able to approximate any continuous system. For N arbitrary distinct samples $\{(x_i, t_i)\}_{i=1}^N$, where $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in R^n$ is input data and $t_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T \in R^m$ is target output value, standard single hidden layer feedforward networks with K hidden nodes and activation function $g(x)$ are mathematically modeled as

$$o_j = \sum_{i=1}^K \beta_i g_i(x_j) = \sum_{i=1}^K \beta_i g(\omega_i \cdot x_j + b_i), \quad j = 1, \dots, N \quad (5)$$

where $\omega_i = [\omega_{i1}, \omega_{i2}, \dots, \omega_{in}]^T$ is the weight vector connecting the i th hidden node and the input nodes, $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$ is the weight vector connecting the i th hidden node and the output nodes, b_i is the threshold of the i th hidden node, o_j is the output of network, $g(\omega_i \cdot x_j + b_i)$ is the activation function of the i th hidden node, generally the Sigmoid function.

When training begins, ω_i and b_i are randomly initialized and kept the same in the training process, only the output weights β_i need to be determined through training. That standard single hidden layer feedforward networks with K hidden nodes with activation function $g(x)$ can

approximate these N samples with zero error means that $\sum_{i=1}^N \|o_i - t_i\| = 0$, i.e., there exist β_i , ω_i and b_i such that

$$\sum_{i=1}^K \beta_i g(\omega_i \cdot x_j + b_i) = t_j, \quad j = 1, \dots, N \quad (6)$$

The above N equations can be written compactly as

$$H\beta = T \quad (7)$$

Where

$$H = \begin{bmatrix} g(\omega_1 \cdot x_1 + b_1) & \cdots & g(\omega_K \cdot x_1 + b_K) \\ \vdots & \vdots & \vdots \\ g(\omega_1 \cdot x_N + b_1) & \cdots & g(\omega_K \cdot x_N + b_K) \end{bmatrix}_{N \times K} \quad \beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_K^T \end{bmatrix}_{K \times m} \quad T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times m} \quad (8)$$

H is called the hidden layer output matrix of the neural network; T is desired output vector.

Traditionally, train a single hidden layer feedforward networks is simply equivalent to find a least-squares solution $\hat{\beta}$ of the linear system $H\beta = T$:

$$\|H(\omega_1, \dots, \omega_K, b_1, \dots, b_K)\hat{\beta} - T\| = \min_{\beta} \|H(\omega_1, \dots, \omega_K, b_1, \dots, b_K)\beta - T\| \quad (9)$$

where $\hat{\beta} = H^\dagger T$ is the smallest norm least-squares solution of the linear system $H\beta = T$; H^\dagger is the Moore–Penrose generalized inverse of the hidden layer output matrix H . The training process of extreme learning machine is a simple linear regression process. It effectively overcome the problem of traditional neural networks which will easily fall into local minimums, and the training speed and generalization ability are also greatly improved.

Wind Power Prediction Model based on EMD-ELM

Time series of wind power has obvious nonlinearity, non-stationary and randomness, so larger error will generate when using traditional prediction methods. Given EMD technology has a good non-stationary capability for data processing, this paper presents a combination method for short-term wind power prediction based on EMD-ELM. Prediction scheme is shown in Figure 1.

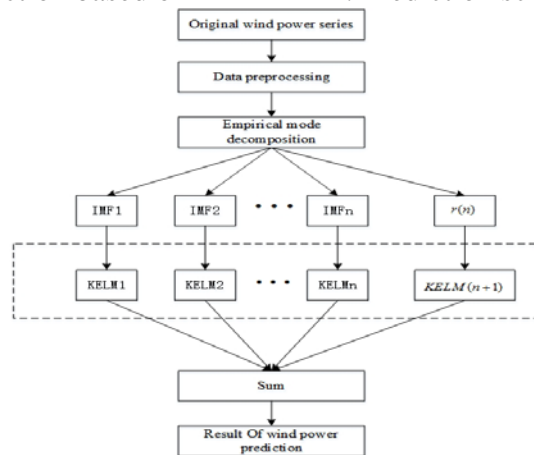


Fig.1. Flowchart of EMD-ELM prediction

(1) Utilize EMD method to decompose original wind power time series to obtain each component $c_i(t)$ and the remaining components $r_n(t)$ of IMF.

(2) ELM prediction models are established for each component $c_i(t)$ and the remaining components $r_n(t)$ of IMF. Wind speed, the sine and cosine value of wind direction are regarded as input for wind prediction model to obtain the predicted value of each component sequence IMF.

(3) The predictive value of each IMF component is superimposed to obtain a final wind power

prediction value.

(4) Compare with the actual data of wind power. Obtain prediction error indicators and error analysis.

Example analysis

To verify the feasibility of wind power combination prediction model presented in this paper, wind power data for a wind farm on October 2014 is used for simulation. The input sample is sampling a point every 15 minutes. Take seven days' data, i.e., 672 sampling points are as the experimental samples. Among them, the former 480 sampling points are for the model input sample and the latter 192 sampling points are for the test samples. Original wind power time series are shown in Figure 2.

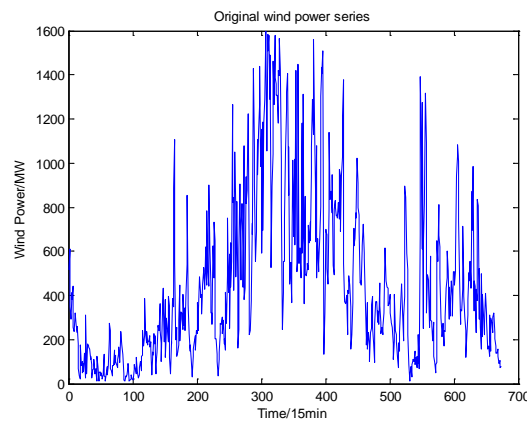


Fig.2. Original wind power series

First, the sample data (including the original wind power, wind speed, wind direction) is normalized, i.e. the original data is linearly transformed into interval $[0, 1]$. The wind power time series EMD are decomposed to six IMF components and a residual component, as shown in Figure 3. According to Figure 3, it shows that IMF1 frequently fluctuates and the value change is not great, which illustrates that a high frequency portion of the signal power of wind can be seen as a random noise. Therefore, IMF1 is not predicted. IMF1 and the rest of prediction values are added directly to get the final wind power prediction value; IMF5 and IMF6 denote wind power low-frequency part of the signal. The oscillation period is relatively large. The remaining portion Res represents the whole trend of wind power original signal. The mean value is closer to the average of the original wind power signal. Based on the above analysis, establish ELM prediction model for each of the IMF (except IMF1) component signal and the residual signal Res. Then predict each IMF (except IMF1). Finally, each IMF (except IMF1) component, the predictive value of the residual signal Res, and IMF1 component are superimposed to give the final result of wind power prediction.

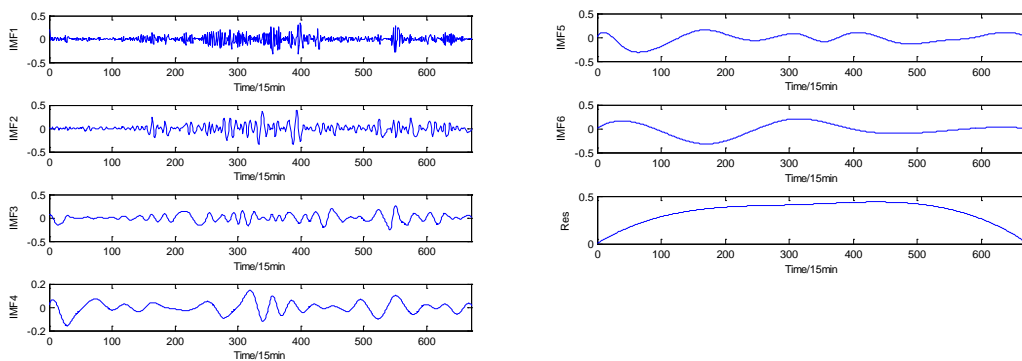


Fig.3. EMD results of wind power data

In this paper, predict the wind power 16 steps in advance. Because the sampling frequency is sampled once every 15min, this result is the prediction of wind power with 4 hours in advance. Because the test sampling points is 192 points, 12 times for multi-step prediction are needed, i.e., 12

times wind power prediction with 4 hours in advance. The detailed content is to complete a 16-step prediction, then the tested wind power sequence is put into the prediction model for training. Subsequently continue to the next 16-step prediction. Repeat 12 times to obtain the final wind power forecast results. To verify the effectiveness of the proposed method, establish EMD-ELM, ELM and EMD-BP prediction model of the input data, respectively. Compare the results with the actual wind power data. The prediction results are shown in Figure 4.

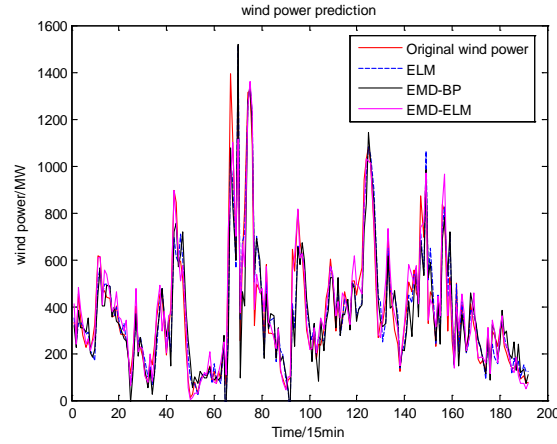


Fig.4. Actual value of wind power and prediction results ahead 4 hours of each prediction model

Select the root mean square error E_{nmse} , mean absolute error E_{nmae} , and mean percentage error E_{mape} to quantitatively evaluate the accuracy of the model prediction results by the following:

$$E_{nmse} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2} \quad (10)$$

$$E_{nmae} = \frac{1}{N} \sum_{i=1}^N |x_i - \hat{x}_i| \quad (11)$$

$$E_{mape} = \frac{1}{N} \sum_{i=1}^N \left| \frac{x_i - \hat{x}_i}{x_i} \right| \times 100\% \quad (12)$$

where x_i is actual value; \hat{x}_i is model prediction value; N is the number of the test sample. Table 1 lists performance of various errors.

TABLE 1. Error Comparison of Different Wind Power Prediction Methods

Prediction methods	Root mean square error	Mean absolute error	Mean absolute percentage error /%
EMD-ELM	11.83	10.09	20.88
ELM	14.74	11.28	24.98
EMD-BP	16.12	13.45	27.57

From Figure 4, we can see EMD-ELM wind power prediction and actual values agree well in advance of 2 hours. As shown in Table 1, compared EMD-ELM with direct ELM method, the accuracy of prediction is improved, which illustrates that decomposing wind power wind power by EMD method can improve prediction accuracy to some extent. What's more, compared the error results of EMD-ELM method and EMD-BP method, we can see the prediction accuracy of EMD-ELM is improved, which shows that the proposed prediction method can better identify trends and patterns inherent in the data and the prediction curve can follow the actual power curve to achieve a high accuracy of short-term wind power prediction.

Conclusion

Due to non-stationary and randomness in wind power series, empirical mode decomposition

combined with extreme learning machine is utilized for short-term wind power prediction. EMD method can gradually decompose real trend or fluctuation amount with the same scale in wind power time series, which can better reduce wind power signal in unsteady behavior. Extreme learning machine can better map out the non-linear relationship between the signal and reduce the influence of nonlinear wind power data for wind power prediction results. The simulation results show that short-term wind power prediction method based on empirical mode decomposition and extreme learning machine can follow the change rule of wind power and effectively improve the prediction accuracy of short-term wind power.

Acknowledgement

This project was supported by State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources. Thanks a lot.

References

- [1] Yin M, Wang C S, Ge X B. Review of Economic-Technical Assessment of Wind Power Integration[J]. Proceedings of the Chinese Society of Universities for Electric Power System and its Automation, 2010, 5: 019.
- [2] Yang X, Xiao Y, Chen S. Wind speed and generated power forecasting in wind farm[J]. Proceedings-Chinese Society of Electrical Engineering, 2005, 25(11): 1.
- [3] Zhang L, Ye T, Xin Y, et al. Problems and Measures of Power Grid Accommodating Large Scale Wind Power [J]. Proceedings of the CSEE, 2010, 25(002).
- [4] J Conejo A J, Plazas M A, Espinola R , et al. Day-ahead electricity price forecasting using the wavelet transform and ARIMA models[J]. IEEE Transactions on Power Systems, 2005, 20(2): 1035-1042.
- [5] Ding M, Zhang L, Wu Y. Wind speed forecast model for wind farms based on time series analysis [J]. Electric Power Automation Equipment, 2005, 8: 32-34.
- [6] Bossanyi E A. Short-term wind prediction using Kalman filters[J]. Wind Engineering, 1985, 9(1): 1-8.
- [7] Wang L, Dong L, Liao X, et al. Short-term power prediction of a wind farm based on wavelet analysis[J]. Proceedings of the CSEE, 2009, 28: 006.
- [8] Lin Y E, Peng L I U. Combined model based on EMD-SVM for short-term wind power prediction [J]. Proceedings of the CSEE, 2011, 31(31): 102-108.
- [9] Huang G B, Wang D H, Lan Y. Extreme learning machines: a survey[J]. International Journal of Machine Learning and Cybernetics, 2011, 2(2): 107-122.
- [10] Xin-Ying W, Min H. Multivariate chaotic time series prediction based on extreme learning machine[J]. 2012.
- [11] GUAN Wenyuan, LIU Yuqi, XIAO Yunqi, Prediction Intervals Forecasts of Wind Power based on PSO-KELM [J], Proceedings of the CSEE, 2015, 35:146.
- [12] Hongming Yang, JunYi, Junhua Zhao, et al. Extreme learning machine based genetic algorithm and its application in power system economic dispatch [J]. Neurocomputing, 2013(102): 154-162.