Wind Speed Forecasting Based on EEMD and ARIMA

Yu Min, Wang Bin, Zhang Liang-li and Chen Xi
School of Information Science and Engineering
Wuhan University of Science and Technology
Wuhan, China
E-mail: yufeng3378@163.com

***Abstract*-This paper proposes a prediction model based on Emsemble Empirical Mode Decomposition (EEMD) and Autoregression Integrated Moving Average (ARIMA) model for the characteristics of the wind speed as the nonlinear and non-stationary sequence. Firstly, the wind speed time series is decomposed into a number of Intrinsic Mode Functions (IMFS) and one residual series which are smoother than the original sequence using EEMD .Then the ARIMA model is applied to forecast the IMF and residue series. Finally, the prediction result of the wind speed is obtained by summing the predicted results of each IMF and residue component. The results gained in this paper show that the prediction accuracy of EEMD-ARIMA model is higher than that of EMD - ARIMA model and ARMA model.**

Keywords-Ensemble Empirical Mode Decomposition (EEMD); Autoregression Integrated Moving Average (AIRMA); wind speed

I. INTRODUCTION

With the support of national policy, the wind power generation technology has a rapid development, and the installed wind-power capacity of China ranked first in the world in 2010 for the first time. But with the development of the wind power installed capacity, it brings a big impact to the power grid, the quality problem is more complicated, the difficulty of the power grid dispatching is increased. Accurate prediction of wind speed can improve the utilization of wind energy, reduce the impact of wind power on the power grid [1], will provide effective basis for the power grid scheduling, and improve the reliability of power grid operation. Wind speed forecasting plays a more and more important role in the operation of the grid-tied wind power system, so wind speed forecasting has been the key technology and hot research of the grid-tied wind power system.

At present, the forecasting methods of wind speed mainly include Support Vector Machine Model[2,3],Neural Network Model[4],Chaotic Time Series forecasting[5,6],Time Series Model[7],and so on. These methods have their own applicable occasions. Support Vector Machine is applicable for fast forecasting of small samples, we can gain global optimum, but the selection of some parameters can directly affect the performance of algorithm and effect of forecasting, so the Support Vector Machine's selection of parameters lacks a valid structured method; It's easy for Neural Network algorithm to produce mistakes in local minims so that the forecasting results may be non-stationary; The method of chaotic time series forecasting needs series to be chaotic series; Although Time Series Model needs a little information in modeling, and it is widely applied, it needs original series to be stationary series. Because of the nonlinear and non-stationary characteristics of wind speed, the wind speed series is needed to be stabilized and linearized at first.

The ways of data stabilization process mainly include Wavelet Transformation[8,9] and Empirical Mode Decomposition[10](EMD), but it is easy for Wavelet Transformation and Empirical Mode Decomposition to produce frequency band mixing and mode mixing, components lose real physical meaning, reduce the stability of component, and influence the efficiency of prediction. Empirical Mode Decomposition[11](EEMD) is the improvement of EMD, it is an adaptive method to represent a nonlinear and non-stationary signal as the superposition of signal components stabilized and linearized with a noise-assisted analysis technique.

According to the non-stationary property of wind speed, this paper presents a model of wind speed forecasting, which is based on EEMD and ARIMA. Firstly, the wind speed series is decomposed by EEMD ,and gain a number of relative stationary sub-sequences adaptively; then the sub-series forecasting results are obtained using ARIMA; Finally, we can gain the forecasting value of wind speed by summing the forecasting results of the sub-sequences .

II. METHODOLOGIES

A. EMD and EEMD

Empirical mode decomposition (EMD) is a nonlinear signal processing method proposed by Huang et al. It can decompose the signal adaptively, not only for stationary signals, but also for non-stationary signals. Wind speed time series has very strong volatility, which can be regarded as the superposition of different frequency components. The wind speed components can be decomposed into a group of functions by EMD; these functions are called Intrinsic Mode Function (IMF). EMD decomposition of the wind speed time series can be expressed as follows

$$x(t)=\sum_{i=1}^{n}c_i(t)+r(t) \qquad (1)$$

The project is supported by National Nature Science Foundation of China (51407131, 51308426, 61473338).

Where $r(t)$ is remainder term, $c_i(t)$ is the components of IMF, its frequency ranks from high to low.

Because wind speed is intermittent, its extreme points are unevenly distributed. The procedure of EMD mainly depends on the selection of extreme points, if the extreme points are distributed unevenly; the mode mixing phenomenon is prone to occur. EEMD uses white noise's statistical property which is distributed evenly, and adds it to original signals; this can change the characteristics of the original signals' extreme points. After decomposition the signal in different scales is continuous, and improve mode mixing phenomenon. And because of the characteristics of zero mean noise, the noise can be offset by the average after multiple averaging; the result of integration can be used as the final result.

Concrete steps of EEMD are as follows:

(1) The white noise sequence is added to the original sequence;

(2) Gain sub-sequence by EMD ;

(3) Adding different white noise to the original sequence, repeat steps (2), after taking every IMF component's mean, the mean values are IMFs of EEMD.

B. ARIMA Model

The basic models of time series proposed by Box and Jenkins include Auto-regression Model, Moving Average Model, Auto-regression Moving Average Model, Auto-regression Integrated Moving Average Model; ARMA is a relatively mature model which contains AR Model, MA Model and ARMA Model. ARMA (p, q) Model is expressed as follows:

$$\begin{aligned} y_t &= \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} \\ &\quad + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \cdots - \theta_q \varepsilon_{t-q} \end{aligned} \tag{2}$$

Where y_t is forecasting value of time, t, $\phi_i, \theta_j (i = 1, 2, \cdots, p; \quad j = 1, 2, \cdots, q)$ are model parameters, ε_t is the random process of white noise whose mean value is zero, p and q are orders of the model.

ARIMA Model is the extension of ARMA Model, ARMA Model is usually used to dispose stable time series, if the series are non-stationary, it can be transformed into a stationary time series using the d'th difference process, d is usually zero, one or two. Then the series after difference is modeled by ARMA. The whole above process is called ARIMA. Using ARIMA to predict the process is as follows:

(1) Stationary identification of sequence. Check series' stationary with test methods of ADF root of unity.

(2) Series' stationary processing. If data series are unstable, we need to conduct difference processing until the data after disposed meet stationary condition. The order of difference is d when time series are stable.

(3) The establishment of the model. If partial auto-correlation function of stable series is clipped, auto-correlation function is trailing, we can assert that AR Model is suitable to the series; If partial auto-correlation function of stable series is trailing, auto-correlation function is clipped, we can assert that MA Model is suitable to the series; If partial auto-correlation function and auto-correlation function of stable series are both trailing, ARMA Model is suitable to the series. We can determine the order p and q by using AIC principle.

(4) The estimation of parameters, we need check whether it has statistical significance.

(5) Conduct hypothesis testing, we need to judge whether residual sequence of the model is white noise.

(6) Conduct forecasting analysis by using models that has been checked to be qualified.

As for the non-stationary property of wind speed, this paper selected ARIMA Model to establish forecasting model to wind speed sub-sequence decomposed by EEMD.

B. The modeling Process

Because wind speed time series is strongly fluctuant, wind speed series can be viewed as a sum of sub-sequence of single frequency components. This paper firstly decompose original wind speed into IMF whose frequency ranks from high to low by using EEMD, sub-sequence is stable relatively to original wind speed, If sub-sequence is unstable, we need to conduct difference to it, then conduct modeling forecasting to sub-sequence by ARMA, finally, we can gain the forecasting value of wind speed by summing the forecasting results of the sub-sequence. The process of the wind speed forecasting modeling based on EEMD and ARIMA is denoted in figure 1.modeling processes are as follows:

(1) Gain IMF components and residue using EEMD

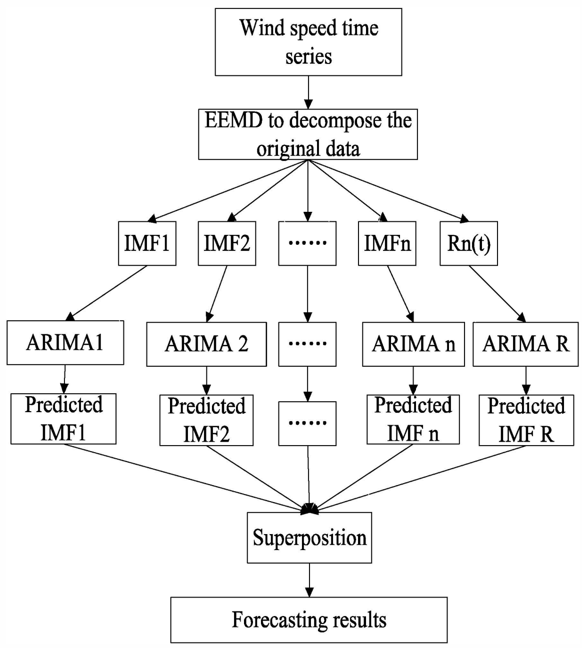

Fig.1. Flowchart of wind speed forecasting based on EEMD-ARIMA.

method.

(2) Each IMF component and residue is modeled by ARIMA as a forecasting tool, and obtains prediction result of each component.

(3) Obtain the final forecasting result of wind speed by summing every component forecasting value.

III. EXAMPLE ANALYSIS

A. Sample Selection and Processing

In this paper, the data were from some wind power farm from 1998-09-14-17:00 to 1998-10-22-05:00, chosen 901 data that updated once every one hour.

The forecasting model was built by the first 801 data, and the latter 100 data were used as the model test data, and the predicted values were compared with the 100 values. Firstly the original data was decomposed into 8 IMFs and residual component using EEMD, the decomposition result was given in figure3, and the original wind speed data was given in figure 2. From figure 3 we can see that the IMF component is more stationary compared with the original sequence in figure 2, and it is more easy to use ARIMA to model.

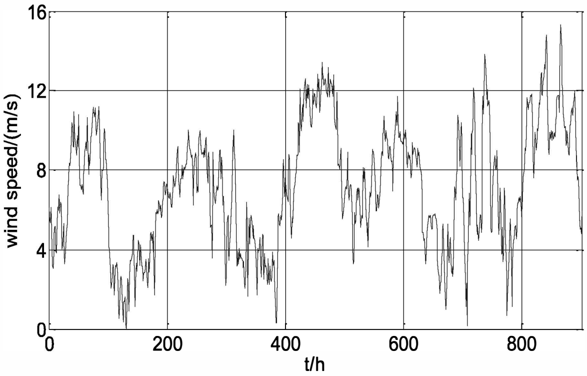

Fig.2. Original wind speed

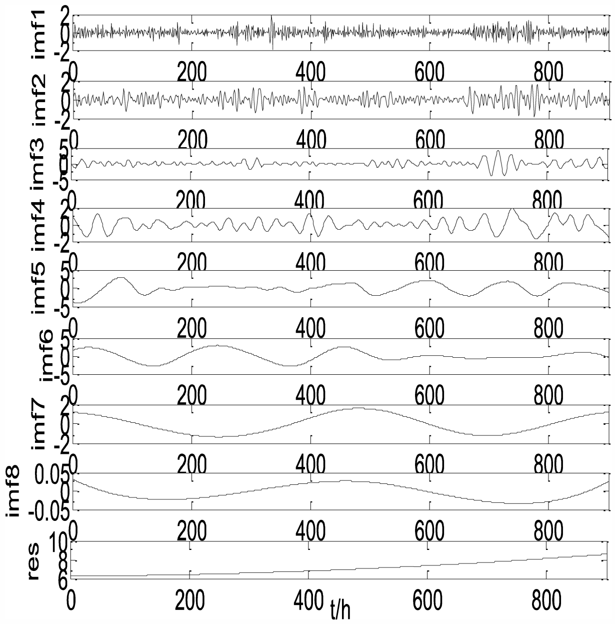

Fig.3. EEMD results of wind speed signal

In order to assess forecasting model presented in this paper properly, Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE) and Root Mean Square Error (RMAE) are selected to evaluate. The three criteria are smaller, the prediction effect is better.

B. The Analysis of Forecasting Results

Forecasting results using ARIMA, EMD-ARIMA and EEMD-ARIMA are denoted in figure 4-6.Forecasting errors is given in table 1.

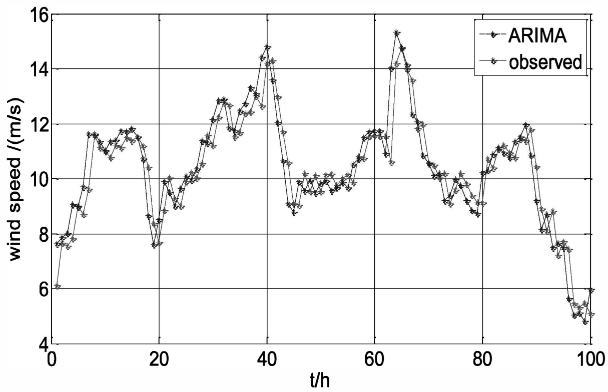

Fig.4. Wind speed forecasting of the ARIMA

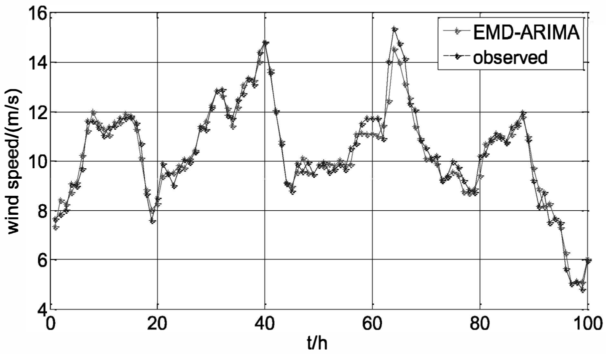

Fig.5. Wind speed forecasting of the EMD-ARIMA

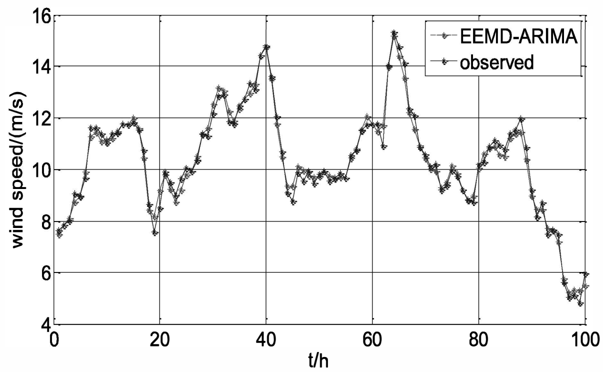

Fig.6. Wind speed forecasting of the EEMD-ARIMA

TABLE I. ERROR INDEXES OF THREE MODELS

Model for	Evaluation Criteria of Forecasting		
	MAPE	***MAE***	***RMSE***
ARIMA	0.0618	0.6066	0.8013
EMD-ARIMA	0.0280	0.3036	0.4014
EEMD-ARIMA	0.0208	0.1911	0.2676

From Table 1 we can see that the error indicators of EEMD-ARIMA and EMD-ARIMA are closer, the indexes decreased greatly compared with ARIMA mode, and every criteria of EEMD-ARIMA is smallest, showing that the forecasting effect is best by EEMD-ATIMA model, which also proves the validity of the model presented in this paper.

The prediction of wind speed using EEMD-ARMA method is relatively accurate as a whole, but there are big errors in some points. The three models have relatively good forecasting effects in stable times of wind variation, but the effects are bad in times of violent wind variation, such as wind speed changes quickly in the points from 46 to 54, we can see from figure 4-6, the fitting degree of forecasting results and observed results is worse than the 5 points from 40 to 44 whose wind speed changes smoothly. RMSE of the two time periods of EEMD-ARIMA forecasting model was calculated, RMSE of 5 points from 40 to 44 is 0.2388, which is lower than the overall RMSE 0.2676, while the RMSE of the 9 points from 40 to 44 is 0.3574, higher than the overall RMSE 0.2676.That is, the faster the speed of wind speed changes, the worse the prediction effect is, and the slower the wind speed fluctuate, the more accurate. The prediction effect is.

IV. Discussions and Conclusions

As wind speed is nonlinear and unstable, it is hard to forecast wind .According to the feature of wind speed, the EEMD-ARIMA hybrid model was proposed in this paper, and the following conclusions were obtained from study of wind speed of actual wind farm:

(1)The forecasting accuracy of EMD-ARIMA and EEMD-ARIMA are higher than single ARIMA methods.

(2)EEMD can decompose nonlinear and unstable wind speed series into a group of stable components, and it is more suitable to handle wind speed compared to EMD. EEMD is more powerful than EMD in anti-mode mixing, which can ensure each IMF and residue have a clear physical meaning, the IMFs are more stationary, each IMF component and trend term forecasted by ARIMA model have better prediction effect.

(3) Three evaluation measures (MAPE, MAE and RMSE) are adopted to evaluate the effect of three models, the results obtained in this paper show that EEMD-ARIMA model can effectively improve the prediction accuracy for wind speed forecasting.

Acknowledgment

The project is supported by National Nature Science Foundation of China (51407131, 51308426, 61473338).The authors gratefully acknowledge the support of those foundation management organizations.

References

[1] L-Y. Zhang, T-L Ye, Y-Z Xin, F.Hang, G-F. Fan, "Problems and measure of power grid accommodating large scale wind power," Proceedings of the CSEE, vol. 30, pp. 1-9, 2010.

[2] H.Wang, Z-J. Hu, Z. Chen, M-L. Zhang, J-B. He, C.Li, " A hybrid model for short-term wind speed forecasting based on Ensemble Empirical Mode Decomposition and Least Squares Support Vector Machines," Transactions of China Electro technical Society, vol. 29,pp.237-245,2014.

[3] X-Y. Yang, B-J. Sun, X-F. Zhang, L-X, Li, "Short-term wind speed forecasting based on Support vector machine with similar data," Proceedings of the CSEE, vol.32, pp.35-41, 2012.

[4] E.Cadenas,W.Rivera , "Wind speed forecasting in three different regions of Mexico using a hybrid ARIMA-ANN model," Renewable Energy, vol.35, pp. 2732-2738, 2010.

[5] Y-Y. Zhang, J-P. Lu, Y-Y. Meng, H.Yan, H. Li, "Wind power shot-time forecasting based on Empirical Mode Decomposition and chaotic phase space reconstruction," Automation of Electric Power Systems, vol. 36, pp.24-28, 2012.

[6] H-Y. Luo, T-Q. Liu, X-Y. Li, "Chaotic forecasting method of short-term wind speed in windfarm," Power System Technology, vol. 33, pp. 67-71, 2009.

[7] C. Peiyuan, T.Pedersen , J B.Bak , " ARIMA-based time series model of stochastic wind power generation," IEEE Transactions on Power System, vol. 25, pp. 667-676, 2010.

[8] K. Xu, J-W. Xu, X-J. Ban, "Forecasting of some non-stationary time series based on wavelet decomposition,"Acta Electronica Sinica, vol. 29,pp. 1-3, 2001.

[9] H. Zhang, R-E. Ren, "Non-stationary time series forecasting based on wavelet decomposition and remnant GM (1, 1)-AR," Systems Engineering-Theory &Practice, vol.30, pp.1016-1020, 2010.

[10] N E. Huang, Z. Shen, S R. Long, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," Proc. of the Royal Society of London, vol. A454, pp. 903-995, 1998.

[11] Z-H. Wu, N E. Huang, "Ensemble empirical mode decomposition; a noise-assisted data analysis method," Advances in Adaptive Data Analysis, vol. 1, pp.1-41, 2009