A New Method for Wind Speed Forecasting Based on Empirical Mode Decomposition and Improved Persistence Approach

Chengchen Sun
College of Energy and Electrical Engineering, Hohai University,
Nanjing, 211100, China
sunchengchen@126.com

Yue Yuan
College of Energy and Electrical Engineering, Hohai University,
Nanjing, 211100, China
yyuan@hhu.edu.cn

Qiang Li
Jiangsu Electric Power Company Research Institute,
Nanjing, 211103, China
ee.liqiang@gmail.com

***Abstract*-Wind speed forecasting plays an important role in sizing the capacity of the energy storage system and guaranteeing the security and stability of power system. In order to forecast wind speeds more accurately, a hybrid forecasting method based on empirical mode decomposition (EMD) and an improved persistence approach has been proposed in this paper. Employing the EMD technique to decompose the measured wind speeds into many intrinsic mode function (IMF) components and a residue, which represent the original signal in both high-frequency and low-frequency signals. Meanwhile each IMF is analyzed and predicted using Moving Average method (high-frequency signals) and Persistence Approach (low-frequency signals), so does the residue. The sum of the predictive value for each decomposed component is the forecasted data. A set of measured wind speed data from a given wind farm locating at Jiangsu Province in China were modeled using the proposed method and the forecasted results were compared to the measured wind speeds as well as those predicted with other traditional methods. The results indicate that the forecasting precision can be improved with the developed model.**

Keywords-Wind Speed Forecasting, Empirical Mode Decomposition, Persistence Approach

I. INTRODUCTION

For the sake of protecting the environment, wind is supposed to be one of the best substitutes for coal or other primary energy. With the growth of wind power in power system, many more new challenges are brought to traditional power system, the reliability and stability, for example, have been challenged due to the randomness and intermittency of the wind. The large-scale integration of wind power in the grid may surely cause problems, including power quality, stability, allocation of the power storage system and power dispatching[1]-[2]. Sizing of the storage system is also limited by the output of the wind turbine and reliable operation of the power system. Accurate and reliable forecasting of the wind speeds and wind production is a key part to solve problems above. Besides, in order to use the wind in an affordable way, wind forecasting is essential.

At present the main methods for wind speed forecasting focus on persistence approach, which assumes that the forecasting value of the wind speeds is the last measured one[3], Time-series Analysis[4], Kalman Filters[5], Artificial Neural Networks[6], Fuzzy Logic[7], etc. Reference [8] presented a novel method for wind speed forecasting based on relating the predictive value to their corresponding historical value in previous years within the same period. Reference [9] proposed a new model based on local and spatial relations of the wind speed so as to improve the efficiency of short and long range forecasting ranging from minutes to several hours ahead. Methods that use the relationships of wind speeds among several sites have been proposed.

Due to the wind speed is sensitive to temperature, terrain, pressure and other factors, the data of the wind speeds are nonlinear and fluctuating. However, preprocessing the wind speed data and studying the internal characteristic information

This work is supported by National Science and Technology Support Program "Study on the integration technology of large scale battery energy storage system" *(2011BAA07B07)*

may improve the forecasting precision.

Considering the efficiency of the prediction of short-term wind speeds and enhancing the forecast precision, a hybrid forecasting method based on Empirical Mode Decomposition (EMD) and Improved Persistence Approach (IPA) is presented in this paper. Empirical Mode Decomposition (EMD) technique has been verified to be more practical than other signal processing methods in many areas[10]. To some extent, EMD is an adaptive wavelet decomposition strategy. However, EMD doesn't require choosing the wavelet base. It decomposes non-stationary signals into some smooth and stationary intrinsic mode function (IMF) with different frequencies. The local features of original speed series are highlighted in the intrinsic mode functions so that it is more obvious to observe the cycle, random and trend parts of the original sequence[11]. The hybrid forecasting method employs EMD to decompose the original wind speed data into a residue and many intrinsic mode function (IMF) components, which can be divided into high-frequency signals and low-frequency signals. Afterwards, the high-frequency parts are forecasted using Moving Average method, while the low-frequency parts are predicted using traditional Persistence Approach. The final results can be obtained from both the prediction of high and low frequency components. The proposed method(Empirical Mode Decomposition-Improved Persistence Approach, EMD-IPA) was applied to the modeling and forecasting of a set of wind speed data from a given wind farm in Jiangsu Province, China. The results demonstrate the validity and practicability of the novel method

II. EMPIRICAL MODE DECOMPOSITION

The theory of intrinsic mode function (IMF) has been pioneered by Norden E. Huang et al when they did research on nonlinear problems and Hilbert Transform. At the same time, they proposed a method for signal decomposition, namely Empirical Mode Decomposition (EMD). It not only makes the signal decomposition unique but also has good local characteristics both in time domain and frequency domain.

The nature of the EMD technique is to identify the intrinsic oscillatory modes by their characteristics time scales in the data empirically, and then decompose the data accordingly[12]. The data are then categorized into many IMF components, which are those functions that satisfy two constraints: (1) in the whole data set, the number of extrema and that of zero crossing must either equal or differ at most by one; and (2) at any point, the mean value of the envelope defined by the local maxima and that defined by the local minima is zero.

For a given signal X(t), we should first let r(t)=X(t): if r(t) is a monotonous function or the value of the difference between its upper and lower envelopes is smaller than the given threshold value, the algorithm stops; otherwise, go to the next step. Then we let h(t)=r(t), if h(t) is an IMF, then C(t) is equal to h(t), otherwise, we calculate the low frequency component of h(t), which is designated as m(t), we wouldn't stop this cycle until h(t) satisfy the terminal conditions.

We let $r_k(t) = r_{k-1}(t) - C(t)$ and repeat the process above. Thus, we can get a series of C(t).

Here, X(t) is the original time data, r(t) is the residue component and C(t) is the IMF. The details of the process will be explained in the following sections.

III. IMPROVED PERSISTENCE APPROACH

At present, there are still many drawbacks when use the main methods to forecast the wind speeds. Time-series method cannot guarantee the precision of the predictive results when the wind speed fluctuates fiercely. Kalman Filters method takes the wind speeds as state variables, this method builds the state space model and forecasts the wind speeds. However, it is hard to estimate the statistical properties of noise which is an essential element in this method. The algorithm used in Artificial Neural Networks is easy to fall into local minimum and difficult to obtain the optimal solution. Fuzzy Logic method is always used in conjunction with other methods. Compared with the above methods, Persistence Approach is easy to understand and implement. This method takes the measured values as the predictive values of the next sampling points. It turns out to be precise in short-term wind speed forecasting. However, as the traditional Persistence Approach requires, the original signal is supposed to have a greater similarity between two sampling points, which is unrealistic in high-frequency parts of wind speeds. The proposed method, Improved Persistence Approach (IPA), can solve this problem.

The Improved Persistence Approach is composed with traditional Persistence Approach and Moving Average method.

A. Traditional Persistence Approach

The original wind speed is separated into several parts. Due to the speed of the low-frequency parts changes softly and the measured wind speed is similar to that of sampling points nearby, persistence approach can guarantee the prediction accurate. It considers that the latest measured value is the predictive value of the next sampling point[13]:

$$v'_{t+k} = v_t (k = 1,2,3...) \quad (1)$$

v'_{t+k} is the predictive value at the time $t+k$; v_t is the measured value at the time t .

B. Moving Average Method

In terms of the high-frequency signals, Moving Average method is effective to forecast the wind speeds. This method is the moving average of several historical data. It can smooth the forecasting results when there are dramatic changes among sampling points[14].

$$v'_{t+k} = \frac{v_t + v_{t-1} + ... + v_{t-n}}{n+1}$$

$$(k = 1,2,3,...; n = 1,2,3,...) \quad (2)$$

v_{t-n} is the measured values at the time $t-n$ .In this paper, the windowsize of MA is supposed to be a constant.

The sum of all the predictive values is the forecasting results.

IV. THE MODEL FOR WIND SPEED FORECASTING BASED ON EMD-IPA

As is shown in Fig.1, a novel model of wind speed forecasting based on EMD-IPA was developed. The EMD unit was used to decompose the original wind speed data into a residue and some IMF components; C_i is the obtained IMF and r_n is the residue; MA is the Moving Average modeling unit; PA is the Persistence Approach modeling unit.

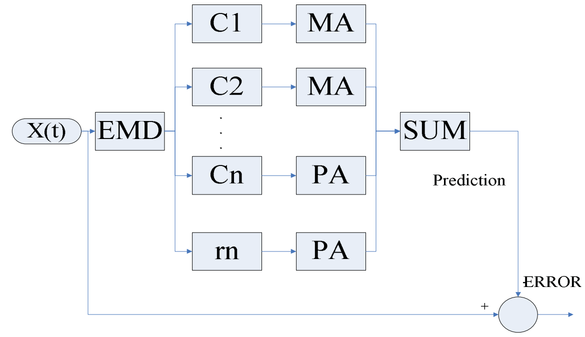

Fig.1. The EMD-IPA model

For a given wind speed data X(t) ,the process of decomposition is as follows[15]-[16]:

(1) Identify all extrema of X(t), all the extrema can be connected by a cubic spline line.

(2) The upper and lower envelope can be obtained then, denoted by u(t) and v(t), and the mean of the upper and lower envelope is denoted by m(t);

$$m(t) = \frac{u(t) + v(t)}{2} \quad (3)$$

(3) Subtract from the signal to obtain h(t);

$$h(t) = X(t) - m(t) \quad (4)$$

(4) Whether h(t) satisfy the constraints? Yes: go to (5); No: h(t) is treated as the new data and put back to (1).

$$h_k(t) = h_{k-1}(t) - m_{k-1}(t) \quad (5)$$

(5) Let c=h(t), then the first IMF c(t) and the residual of the signal r(t) can be obtained.

$$r(t) = X(t) - c(t) \quad (6)$$

(6) After extracting an IMF, this same IMF is subtracted from the signal X(t). The residual is treated as the new data. Then do the steps (2)-(5). The cycle can stop by the predetermined criteria.

$$r_n(t) = r_{n-1}(t) - c_n(t) \quad (7)$$

(7) The last r(t) which cannot be extracted is the residue. It can be either a monotonic function or a single cycle. The original signal can be represented in the following equation:

$$X(t) = \sum_{i=1}^{n} c_i(t) + r(t) \quad (8)$$

The above screening process may repeat many times. To make sure that the IMF components retain enough physical characteristics of amplitude and frequency, a criterion for the process to stop is needed. A standard deviation, SD, is chosen as a threshold value to satisfy this. The SD is computed from the two consecutive sifting results as:

$$\mathrm{SD}=\left\{\sum_{t=0}^{T}\left[\frac{\left|\left[h_{1(k-1)}(t)-h_{1k}(t)\right]\right|^{2}}{h_{1(k-1)}^{2}(t)}\right]\right\}^{1/2} \tag{9}$$

A typical value of SD should be lower than 0.1.Through the screening process, the data, X(t), is finally decomposed into several IMFs and a residue r(t), i.e.

$$X(t)=\sum_{i=1}^{n}C_i+r_n \tag{10}$$

V. FORECASTING RESULTS AND DISCUSSION

This paper takes a set of wind speed data, which is collected from a certain wind farm located at Jiangsu Province in China, as an example. The sampling interval is 10 minutes. EMD-IPA model was built up on the basis of every 10 minutes mean wind speed samples, thus guaranteeing the precision of the prediction of wind speeds at future time-step.

In experiments, we regard the first two days (288 data) values as sampling points, while the following one day (144 data) values are used as checking data for validating the EMD-IPA method we proposed.

The original wind speed data were shown in Fig.2. Obviously, the wind speed time series are fluctuated and non-stationary.

According to the terminal constraints introduced in above equation (9), we can obtain all of the IMF and a residue, the process as shown from Fig.3 to Fig.9.

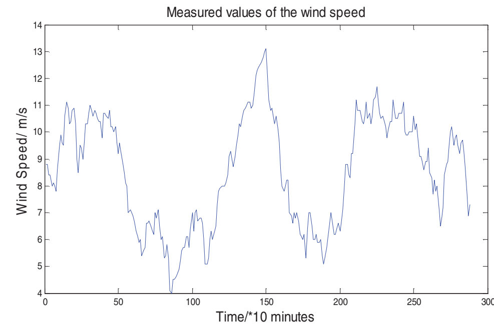

Fig.2. The original wind speeds

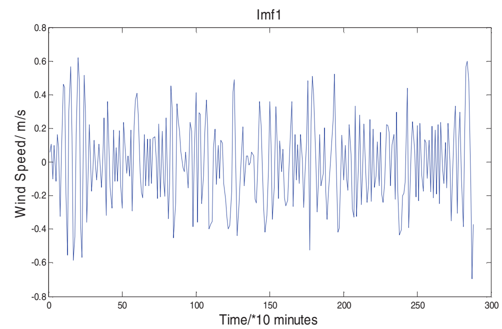

Fig.3. Wind speeds of IMF1

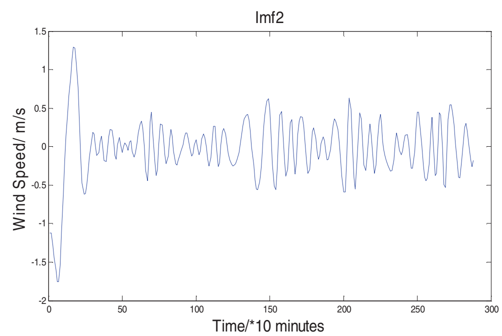

Fig.4. Wind speeds of IMF2

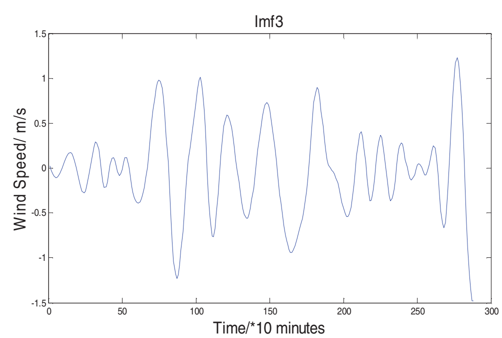

Fig.5. Wind speeds of IMF3

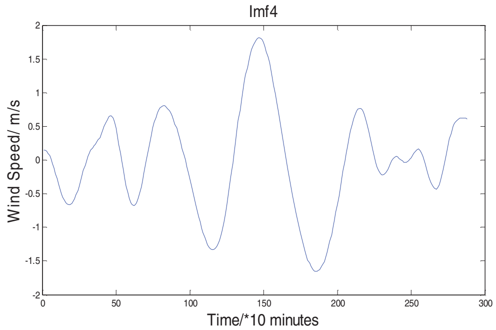

Fig.6. Wind speeds of IMF4

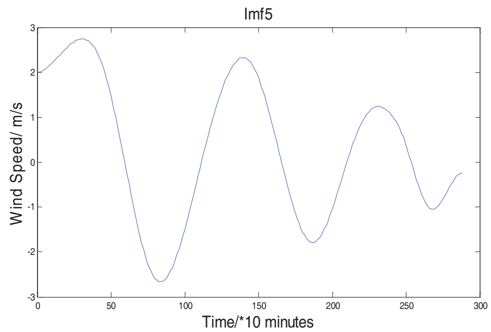

Fig.7. Wind speeds of IMF5

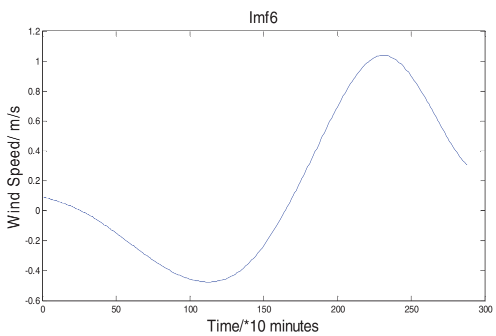

Fig.8. Wind speeds of IMF6

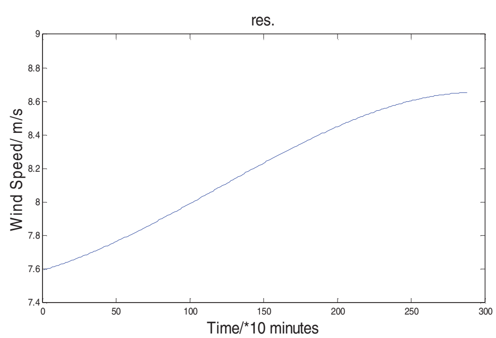

Fig.9. Residue

Before forecasting each IMF and residue, we should divide them into high and low frequency parts. Fig.10 demonstrates the frequency of the wind speeds of IMF4, while Fig.11 illustrates that of IMF5.

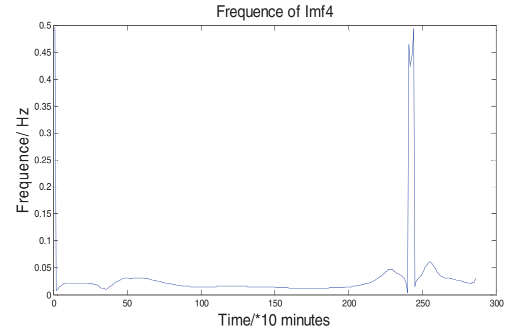

Fig.10. Frequency of IMF4

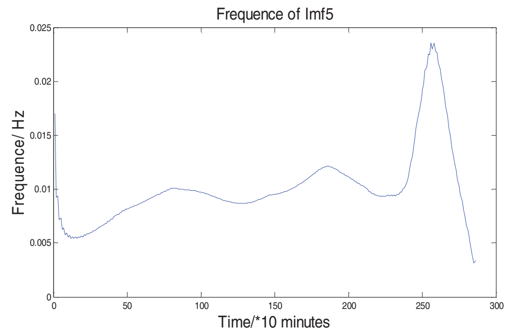

Fig.11. Frequency of IMF5

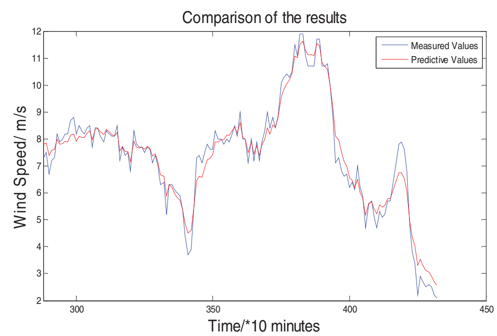

Fig.12. Comparison of the results

Clearly, the frequency of the first four components is high, while the frequency of the other components is low. Based on the data updated every ten minutes, we may predict the wind speeds of each component, respectively. The forecasting results from each IMF and residue are summarized as the future wind speed, which was compared to the measured set in Fig.12. It is obvious that the trends of the forecasting result can match that of the real one. Due to the characteristic of the Moving Average method, the tracking of the extrema still needs to be improved. Tab.1 listed the Relative Mean Square Error(RMSE) and Mean Absolute Percentage Error(MAPE) employing the traditional persistence approach, time-series analysis together with the proposed method.

TABLE.1. THE COMPARISON OF THE FORECASTING ERROR

Error	Persistence	ARMA	EMD-IPA
RMSE	1.5467	0.9472	0.3785
MAPE	20.37%	12.15%	8.14%

Compared to the other two traditional methods, EMD-IPA reduced almost a 12% or 4% MAPE. Also the RMSE was greatly decreased. It can be seen that the proposed method has higher precision.

VI. CONCLUSIONS

In this paper, a novel approach for wind speed forecasting employing EMD-IPA is proposed. The key problem of the proposed method is to preprocess the wind speed data and divide them into several stationary components which can be distinguished by their different frequencies. As a result, the coupling between all the characteristics information in data are weakened to some extent. Further, the corresponding IPA model for each divided component is easier to build. After each wind speed data are predicted in built IPA model, the forecasting results are summarized as the wind speed forecasting results. Ten-minute ahead forecasting is used in this paper. The rigor of the arithmetic has been testified by example above. The interrelationship between the precision accuracy and the sampling period is analyzed in this paper. The proposed method is used to forecast a set of wind speed data collected from Jiangsu Province in China and the results are compared with the real data. The results indicate the effectiveness of the developed method in promoting the prediction precision.

VII. REFERENCES

[1] Yang Xiu-yuan, Xiao Yang, Chen Shu-yong, “Wind speed and generated power forecasting in wind farm,” Proceedings of the CSEE, vol. 25, no. 11, pp. 1-5, 2005.

[2] Zhou Shang-xi, Wang Hai-chao, Chen Shou-sun, “Analysis on operation value of wind power resources,” Power system Technology, vol. 26, no. 5, pp. 100-105, 2002.

[3] Alexiadis M, Dokopoulos P, Sahsamanoglou H, et al. Short-term Forecasting of Wind Speed and Related Electrical Power[J]. Solar Energy, 1998, 63(1):61-68.

[4] Torres J L, Garcia A, Blas M De, et al. Forecast of hourly average wind speed with arma models in Navarre(spain)[J].Solar Energy, 2005, 79: 65-77.

[5] Bossanyi E A. Short-term wind prediction using kalman filters[J]. Wind Engineering, 1985, 9(1): 1-8.

[6] Kariniotakis G, Stavrakis G, Nogaret E. Wind power forecasting using advanced neural network models[J]. IEEE Trans on Energy Conversion, 1996, 11(4): 762-767.

[7] Damousis I G. Dokopoulos P. A fuzzy expert system for the forecasting of wind speed and power generation in wind farms[A]. 22nd IEEE Power Engineering Society International Conference[C], 2001, 5(20-24): 63-69

[8] T.H.M.El-Fouly, E.F.El-Saadany, M.M.A.Salama, “One day ahead prediction of wind speed using annual trends,” IEEE, pp. 1-7, 2006.

[9] I. G. Damousis, M. C. Alexiadis, J. B. Theocharis, and P. S. Dokopoulosy, “A fuzzy expert system for the forecasting of wind speed and power generation in wind farms using spatial correlation,” IEEE Trans. on Energy Conversion, vol. 19, no. 6, pp. 352-361, 2004.

[10] Huang N E, Shen Z, Long S, Wu M, Shih H, Zheng Q. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series. In: Proceedings of the Royal Society of London Series A. London, UK: 1998. 903-995.

[11] Ran Li, Yue Wang. Short-term wind speed forecasting for wind farm based on empirical mode decomposition. Electrical Machines and Systems, 2008. ICEMS 2008. International Conference on. Page(s): 2521-2525.

[12] Liu Xing-Jie, Mi Zeng-Qiang, Bai Lu, Wu Tao. A novel approach for wind speed forecasting based on EMD and time-series analysis. Power and Energy Engineering Conference, 2009. APPEEC 2009, Asia-Pacific. Page(s): 1-4.

[13] Kizhner S,Blank K,Flatley T,Huang N.E,Petrick D,Hestnes P, “On certain theoretical development underlying the Hilbert-Huang transform,”Aerospace Conference,2006 IEEE, pp.1-14, 2006.

[14] Han Shuang, Yang Yongping, Liu Yongqian. Application study of the methods in wind speed prediction[J]. Journal of North China Electric Power University, 2005, 35(3): 57-61.

[15] Rato. R. T, Ortigueira. M. D, Batista. A. G. 2008. On the HHT, its problems, and some solutions[J]. Mechanical Systems and Signal Processing, 22: 1374-1394.

[16] Wen-Xian Yang. 2008. Interpretation of mechanical signals using an improved Hilbert-Huang transform[J]. Mechanical Systems and Signal Processing, 22: 1061-1071.