A Novel Approach for Wind Speed Forecasting Based on EMD and Time-Series Analysis

Liu Xing-Jie[1], Mi Zeng-Qiang[1]
Dept. of Electrical Engineering
North China Electric Power University
Baoding, China
lxj5085@163.com

Bai Lu[2]
Dept. of production management
Dalate Power Plant of North Union Power Company
Baotou, China

Wu Tao[3]
Dept. of power system
North China Electric Power Research Institute
Beijing, China

***Abstract*—Wind speed forecasting is significant to the security and stability of electric power system. Aiming to forecast wind speed more efficiently, a hybrid forecasting method based on empirical mode decomposition(EMD) and time-series analysis has been presented in this paper. Employing the EMD technique to decompose the original data into a residue and many intrinsic mode function(IMF) components, which represent the oscillation modes embedded in the data. Afterwards each IMF is modeled and forecasted using time-series analysis, so does the residue. The forecasting value for each decomposed component is summarized as that for the original data. A set of wind speed data from a given wind farm were modeled using the proposed method and the forecasted data were compared to those of measured wind speed as well as those calculated with other conventional methods. The results obtained indicate that the building model is simple and the forecasting precision has been greatly improved using the proposed method.**

Keywords-wind speed; forecasting; EMD; time-series analysis;

I. INTRODUCTION

Wind energy is the fastest growing technology in the range of alternative power generation sources. With the growth of wind power in electric power system, however, the reliability and stability of power system have been challenged owing to the uncertain and variable nature of wind energy. As a crucial tool to support operators, traders and decision makers, wind power forecasting is attracting more and more interests. A proved effective way for wind power forecasting is short-term wind speed forecasting together with power curve of wind turbine[1].

At present the main approaches for wind speed forecasting focus on persistence approach[2], time-series analysis[1,3], kalman filters[4], neural network[5], fuzzy logic[6], etc. And some of them have also been combined to forecast wind speed[7]. Most of the above mentioned approaches target the original wind speed data. Due to wind speed is susceptible to temperature, pressure and other factors, the wind speed data are nonlinear and non-stationary. Through preprocessing the wind speed data and extracting the inner characteristic information, the complexity of forecasting model will be reduced and the forecasting precision is expected to get improved.

Empirical mode decomposition (EMD) technique has been testified to be more effective than other signal processing methods in many areas[8]. Hereby, we intend to use EMD ahead to deal with the original wind speed data. On the other hand, time-series analysis is a conventional wind speed forecasting approach and has been widely used due to its lower requirements for the data and lower forecasting costs[9]. So a novel approach to forecast wind speed using EMD in conjunction with time-series analysis is proposed in this paper. This approach employs EMD to decompose the original wind speed data into a residue and many intrinsic mode function(IMF) components, which are modeled and forecasted using time-series analysis, respectively. These forecasting results are summarized as the future wind speed. The introduced method was applied to the modeling and forecasting of a set of wind speed data from a given wind farm. The results demonstrate the validity and practicability of the proposed method.

II. EMD

The essence 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. The data are then divided into many IMF components, which are those functions that satisfy two conditions: (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.

A systematic way to extract IMF, designated as sifting process, is described as follows:

(1) let $r(t)=X(t)$;

(2) if $r(t)$ is a monotonous function or the difference between its amplitude is lower than the given threshold value, the algorithm stops; otherwise, go to step(3);

(3) let $h(t)=r(t)$;

(4) if $h(t)$ is an IMF, then go to step(7);

(5) calculate the low frequency component of $h(t)$, which is designated as $m(t)$;

(6) let $h(t)=h(t)-m(t)$, and go to step(4);

(7) let $C(t)=h(t)$;

(8) let $r(t)=r(t)-C(t)$, and go to step(2).

Here, $X(t)$ is the original time data, $r(t)$ is the residue component and $C(t)$ is the IMF.

978-1-4244-2487-0/09/$25.00 ©2009 IEEE

The above sifting procedure may be repeated many times. To guarantee that the IMF components retain enough physical sense of amplitude and frequency modulations, a criterion for the sifting process to stop is needed to be determined. A standard deviation, SD, is defined as a threshold value to accomplish this. The SD is computed from the two consecutive sifting results as

$$S_d = \sum_{t=0}^{T} \frac{\left|h_{1(k-1)}(t) - h_{1k}(t)\right|^2}{h_{1k}^2(t)} \qquad (1)$$

A typical value for SD can be set between 0.2 and 0.3.

Through the sifting process, the data, X(*t*), is finally decomposed into n-IMFs and a residue r_n, i.e.

$$X(t) = \sum_{i=1}^{n} C_i + r_n \qquad (2)$$

III. TIME-SERIES ANALYSIS

Time series is a popular tool for the analysis and forecasting of a large number of data. This method has been used for many applications such as economic forecasting, budgetary analysis and inventory studies. When employing this method, the fitting of time series models can be an ambitious undertaking. There are many methods of model fitting including AR, MA, ARMA, ARIMA, etc. The expression that governs ARMA model is given from

$$\varphi(B)x_t = \theta(B)\alpha_t \qquad (3)$$

where $\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_n B^n$,

$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_m B^m$, $\{x_t\}$ (t=1,2,⋯N) is the time series, N is the number of time series, and $\{\alpha_t\}$ represents a normally distributed white noise, with zero mean and variance σ_a^2, where φ_i (i=1,2,…,n) is the autoregressive parameter, θ_j (j=1,2,…,m) is the moving average parameter, n and m is the order of the model, and B is the delay operator.

In the ARMA model, the parameters of φ_i and θ_j are estimated using a certain algorithm. Prior to this, it is required to identify the specific number of ARMA parameters to be estimated, that is the order of the ARMA model. Parameters estimation and order identification are very significant to the forecasting performance of time series.

IV. THE MODEL FOR WIND SPEED FORECASTING BASED ON EMD AND TIME-SERIES ANALYSIS

A. Model architecture

A new model of wind speed forecasting based on EMD and ARMA was built, which is illustrated in Fig.1. The EMD unit is 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; D_i is the time-series modeling unit; $ARMA_i$ is the built time-series model of IMF or residue for forecasting.

B. Algorithm presentation

A set of wind speed data, X(*t*), *t*=1,2,…,N, are firstly sifted using EMD and decomposed into many IMF components with different scale and a residue (the detailed procedure can refer to part II). Then these divided components are separately analyzed using time series and corresponding ARMA model is built.

- parameters estimation[10]

The parameter φ_i is first estimated from the following Yule-Walker equation set:

$$R_A = R_B \varphi \qquad (4)$$

in which $R_A = [R_{m+1} \quad R_{m+2} \quad \cdots \quad R_{m+n}]^T$,

$R_B = [BR_A \quad B^2R_A \quad \cdots \quad B^nR_A]$, $\varphi = [\varphi_1 \quad \varphi_2 \quad \cdots \quad \varphi_n]^T$,

where R_k is the autocorrelation function values of X(*t*).

A new time series $\{y_t\}$ could be formed through the expression

$$y_t = \varphi(B)x_t \qquad (5)$$

Through the estimations of autoregressive function of y_t and the MA(m) model fitted from y_t, a nonlinear equations set can be obtained:

$$R_{y,k} = \sigma_a^2 \sum_{j=0}^{m} \theta_j \theta_{j+k} \qquad (\theta_0 = -1, j+k \le m, k = 0,1,\cdots,m) \qquad (6)$$

Here θ_j and σ_a^2 are as unknown variables, which can be solved by Gauss-Seidel algorithm.

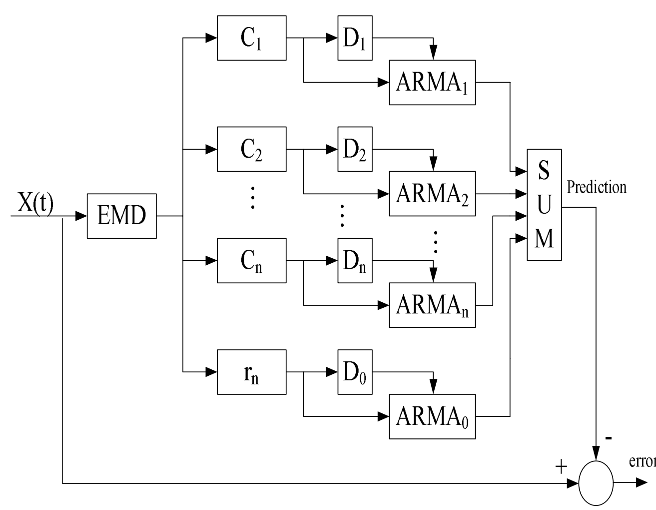

Figure 1. The EMD-ARMA model

- order identification

The ARMA(n,n-1) model introduced by Pandit-Wu is adopted in this algorithm[11]. Akaike information criterion(AIC) function is chosen to check the model's validity, which is defined as

$$AIC(n) = N \ln \sigma_a^2 + 2n - 1 \quad (7)$$

The model order n begins to increase from 2. The correlative parameters are estimated according to (4)-(6) and the AIC value is calculated from (7). When the AIC value reaches the minimum for the first time, the corresponding n is right the most appropriate model order.

Now the residue and IMF components can be orderly forecasted through the built ARMA models. For example, the forecasting of the k^{th} IMF is carried out through the expression:

$$C_k(t+\tau) = \sum_{i=1}^{n} \varphi_{ki} C_k(t-i) + \sum_{j=0}^{m} \theta_{kj} \alpha_k(t-j) \quad (8)$$

here τ is the forecasting horizon. Ultimately, these forecasted results corresponded to all IMF and residue components are summarized as the future wind speed, i.e.

$$X(t+\tau) = \sum_{i=1}^{n} C_i(t+\tau) + r_n(t+\tau) \quad (9)$$

V. FORECASTING RESULTS AND DISCUSSION

A set of wind speed data, which were collected from a certain wind farm in North China, were chosen to demonstrate the performances of the proposed method in this paper. There were 800 samplings in this time-series and the last 200 samplings were regarded as tested set.

The original wind speed data were shown in Fig.2. Clearly, the wind speed time series are oscillating and non-stationary. Fig.3 displayed the obtained IMF and residue decomposed from the original wind speed data using EMD. It can be seen that the original complicated time series were decomposed into 11 IMF components (C_i, i=1,2…,11) , which exhibited stable and regular variation. This means that the interruption and coupling between different characteristic information embedded in the original data have been weakened to an extent. Thus the forecasting model becomes easier to build.

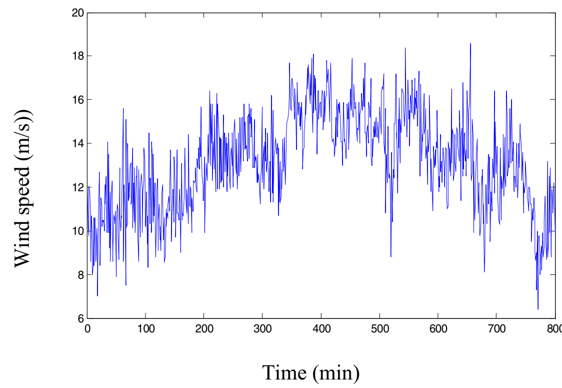

Figure 2. The original wind speed

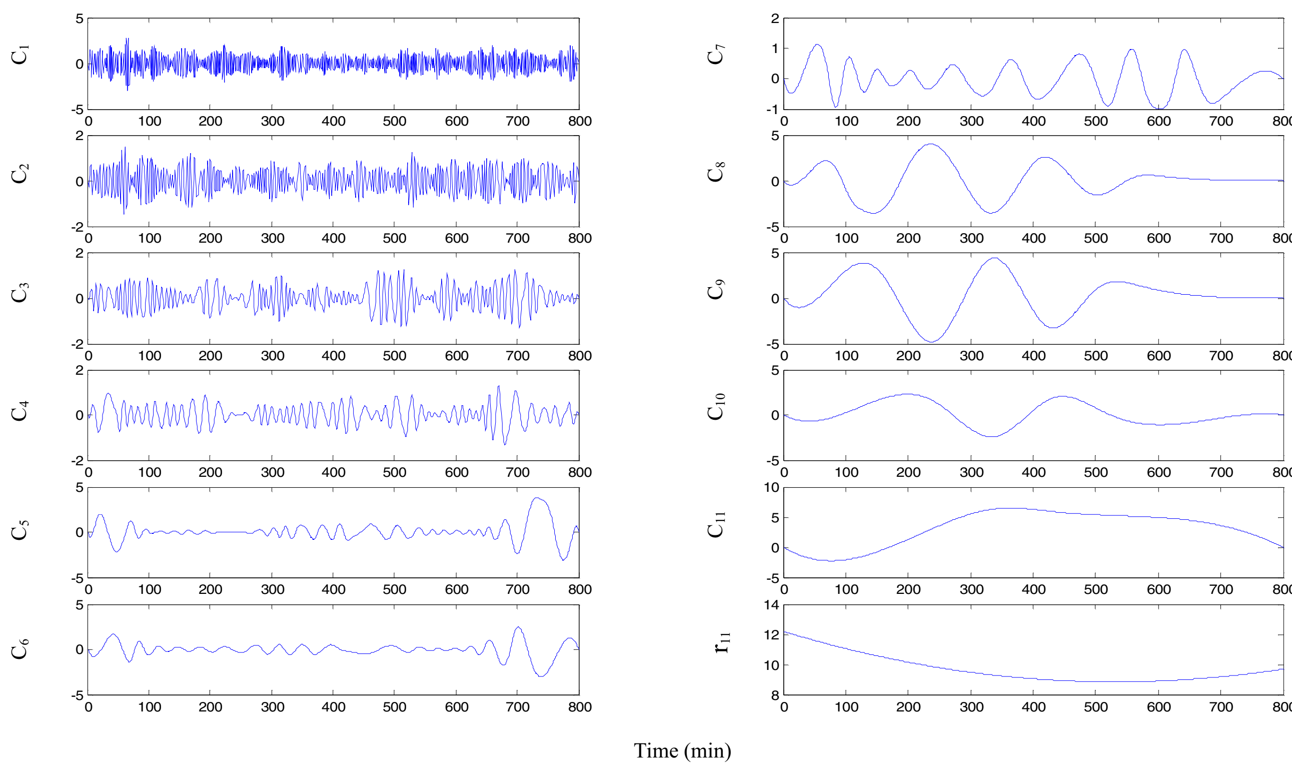

Figure 3. The results of EMD decomposition

Tab.1 listed the model parameters and the AIC values corresponding to C_4. Apparently, when n equaled to 4, the AIC value reached the first minimum. According to the principle of order identification, the ARMA(4,3) model was recognized as the most appropriate one. In this way, the ARMA model for each C_i was successively built.

TABLE I. THE CALCULATIVE RESULTS OF ARMA MODELS PARAMETERS OF C_4

Model Orders	Parameters φ_i	Parameters θ_j	AIC Value
(2,1)	1.8445; -0.9666	0.8291	-7.2092
(3,2)	2.6691;-2.5347;0.8418	-0.9937;-0.4476	-8.4961
(4,3)	3.4580;-4.7161;2.9846; -0.7415	-0.2573;0.1916; 0.1528	-8.5146
(5,4)	2.9500;-2.8907;0.4190; 0.9226;-0.4203	-0.7831;0.1045; 0.3559;0.1903	-8.5085

The forecasting result of C_i was summarized as the future wind speed, which was compared to the tested set in Fig.4. It is obvious that the forecasting result and the real one met very well. Tab.2 listed the relative mean square error(RMSE) and relative mean error(RME) employing the persistence approach, time-series analysis together with the proposed method. In contrast to the other two methods, a 6% or 4% RME was reduced using the proposed method. Also the RMSE was greatly decreased. It can be seen that the proposed method has higher precision.

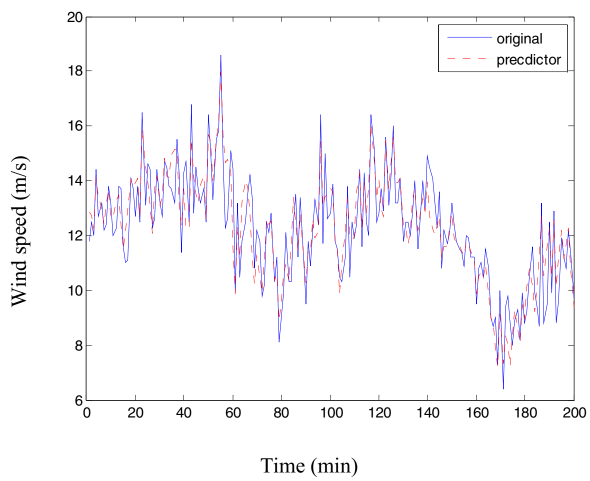

Figure 4. The original wind speed and the forecasting results

TABLE II. THE FORECASTING ERROR

Error	Persistence	ARMA	EMD-ARMA
RMSE	1.8521	1.4653	0.8872
RME	0.1226	0.0962	0.0569

VI. CONCLUSIONS

A novel approach for wind speed forecasting employing EMD and time-series analysis is proposed in this paper. The core of the proposed method is to preprocess the wind speed data and decompose them into more stationary and regular components(IMF or residue) using EMD technique. As a result, the interruption and coupling between all the characteristics information in data are weakened to an extent. Further the corresponding ARMA model for each divided component is easier to build. After the IMF components and residue are forecasted in built ARMA model, the forecasting values are summarized as the wind speed forecasting result. A set of wind speed data collected from a certain wind farm in North China were tested using the proposed method. The validity and practicality of the method are proved through comparing forecasting results with the real data and those calculated using other forecasting methods.

REFERENCES

[1] Yang Xiuyuan, Xiao Yang, Chen Shuyong. Wind speed and generated power forecasting in wind farm [J]. Proceddings of the CSEE,2005,25(11):1-5.

[2] 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.

[3] 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.

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

[5] 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.

[6] 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.

[7] Liu Yongqian, Han Shuang, Yang Yongping, et al. Study on combined prediction of three hours in advance for wind power generation[J]. ACA ENERGIAE SOLARIS SINICA, 2007,Vol.28, No.8, 839-843.

[8] 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.

[9] Ding Ming, Zhang Lijun, Wu Yichun. Wind speed forecast model for wind farms based on time series analysis[J]. 2005, 25(8):32-34.

[10] Yang Shuzi, Wu Ya, Wang Zhifan, et al. Time Series Analysis and Its Engineering Application[M]. Huazhong University of Science and Technology Press,2003.

[11] Yang Weiqin, Gu lan. Time Series Analysis and Dynamical Modeling[M].Beijing University of Technology Press,1988.