Short-Term Wind Speed Forecasting for Wind Farm Based on Empirical Mode Decomposition

Ran Li, Yue Wang
Department of Electrical Engineering, North China Electric Power University, Baoding, 071003, China

***Abstract*-As an important renewable energy form, wind power obtains rapid development recently. More advanced accurate and reliable techniques for wind speed forecasting are required. It can reduce the disadvantageous impact to the power system. According to the outstanding feature of EMD algorithm, this paper presents a new technique for wind speed forecasting based on Empirical Mode Decomposition (EMD) and ARMA. EMD is a new method for analyzing nonlinear and non-stationary signal. It is an adaptive wavelet decomposition strategy. We make full use of the characteristic of the EMD and the ARMA in the EMD-ARMA model. Actual wind speed data are used to test the approach. It concludes that the EMD-ARMA model is an effective method in wind speed forecasting.**

***Key words*-wind speed forecasting; EMD; ARMA**

I. Introduction

With the lack of general energy source, it is provided with very important meanings to solve energy lack by using renewable energy source as substituted energy source in the future. Wind generated power is one of the renewable energy sources. It is paid more and more attention nowadays [1].

Wind power brings new challenges to traditional power system due to the randomness and intermittency. The large-scale integration of wind power in the grid may cause some problems, including power quality, stability, allocation of spinning reserve and power dispatching [2]-[3]. Wind penetration is limited by the requirement for secure and reliable operation of the power system. Accurate and reliable forecasting of the wind speed and wind production is an effective way of increasing wind penetration. At the same time, it can reduce penalties in a spot market coming from over or underestimation of the production. It will be an effective tool for optimizing operating costs and improving reliability.

Wind power is a function of wind speed. So, forecasting of wind power is generally derived from forecasts of speed. Then time series of speed are transformed to series of wind power using manufacturers' curves [4]. Wind speed has a certain periodicity and randomness by itself, so it is difficult to construct the forecasting model. At present, the prediction accuracy can not meet the requirement in spot control.

At present, there are some wind speed forecasting models. Not only historical data but also other factors including meteorological (orography, roughness, obstacles) and wind turbines technical information (hub height, power curve, power penetration) is used in the forecasting model. It is usually the best option for longer term prediction.

There are some models built by wind speed series only. The simplest method is the persistent method, which assumes that the forecasting value of the wind speed is the last measured one. Time series model is developed based on historical values. Reference [4] present a new technique for wind speed forecasting based on relating the forecasted value to their corresponding historical value in previous years within the same time period. Reference [5] proposed a 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. Some other time series analysis techniques are the ARMA model [6], Artificial Neural Networks [7]. These models were used for forecasting average values for the wind speed for one step ahead. They required a large set of historical data.

In order to predict the short-term wind speed effective and level up the forecast precision, a hybrid forecasting method based on EMD and ARMA is presented in this paper. EMD is a new method for processing the nonlinear and non-stationary signals. EMD is an adaptive wavelet decomposition strategy. It has multi-resolution of the Wavelet Decomposition. EMD doesn't require choosing the wavelet base. EMD can decompose non-stationary signals into some smooth and stationary intrinsic mode function (IMF) with different frequency in the different scale space by the shifting process. The local features of original speed series are prominent in the intrinsic mode functions so that it is more obvious to observe the cycle, random and trend parts of the original sequence. After sequence stable processing of the signal, the coupling of characteristic information was reduced.

In this paper, we attempt to combine the EMD and ARMA. Firstly, the history wind speed was decomposed into IMF and residua. Secondly, build the model for IMF and residua. Finally, obtain the final results by Superposition. We deal with short-term wind speed forecasting for a wind farm. At last, hybrid method is used to obtain the forecasting results of original wind speed.

II. EMD-ARMA Model

A. Basic Theory of EMD

The concept of IMF has been pioneered by Norden E. Huang et al. at the end of 1990s. At the same time, they proposed the decomposition method for signal, namely EMD. At recently, scholars have deeply researched in this field [8]-[9]. The EMD method gets extensive application in the power system

[10]-[11].

An IMF satisfies the following two properties [12]:

(a) the number of extrema and the number of zero crossing in a whole sampled data set must either equal or differ at most by one; (b) At any point ,the mean value of the envelope defined by the local maxima and the envelop defined by the local minima is zero.

For a given signal $X(t)$,the process of the EMD is as follows:

1) Identify all extrema of X(t), all the local extrema can be connected by a cubic spline line. Then the upper and lower envelope can be obtained. The upper and lower envelope is denoted by $u(t)$ and $v(t)$.

2) The mean of the upper and lower envelope is denoted by $m(t)$: $m(t)=(u(t)+v(t))/2$. Subtract from the signal to obtain $h(t)= X(t)-m(t)$, $h(t)$ should be an IMF. It should satisfy all the requirements of IMF. If $h(t)$ dose not meet the definition of the IMF , $h(t)$ is treated as the new data and subjected to the same shifting process as described above until an IMF is obtained. The process can be expressed as follows:

$$\begin{gathered} m_1(t)=[u_1(t)+v_1(t)]/2 \\ h_2(t)=h_1(t)-m_1(t) \\ \cdots \\ m_{k-1}(t)=[u_{k-1}(t)+v_{k-1}(t)]/2 \\ h_k(t)=h_{k-1}(t)-m_{k-1}(t) \end{gathered} \tag{1}$$

Then the first IMF $C_1(t)$ and the residual of the signal $r_1(t)$ can be obtained. The relationship between $C_1(t)$ and $r_1(t)$ is :

$$r_1(t)=X(t)-C_1(t) \tag{2}$$

3) After extracting an IMF, this same IMF is subtracted from the signal. The residual is treated as the new data. Then do the step 2. The sifting process can stop by the predetermined criteria. This process can be expressed as follows:

$$\begin{gathered} r_2(t)=r_1(t)-C_2(t) \\ \cdots \\ r_n(t)=r_{n-1}(t)-C_n(t) \end{gathered} \tag{3}$$

Finally, the original signal can be represented in the following equation:

$$X(t)=\sum_{i=1}^{n} C_i(t)+r(t) \tag{4}$$

$r(t)$ is the residue. It can be either a monotonic function or a single cycle.

B. Basic Theory of Time Series

Time series is a set of observation that is arranged chronologically in time. $\{X_t\}$ is stationary time series. The ARMA model can be defined as follows:

$$X_t-\varphi_1 X_{t-1}-\cdots-\varphi_p X_{t-p}=a_t-\theta_1 a_{t-1}-\cdots-\theta_q a_{t-q} \tag{5}$$

Formula (5) was denoted by ARMA (p, q). If p=0, then equation becomes an AR model of order q, is denoted by MA (q). If q=0, the equation becomes an MA model of order p, is denoted by AR (p). If $p=q=0$, the model is $X_t=a_t$. In this ARMA (p, q) model, a_t is whit noise.

Model building of the ARMA (p, q) is important for improving the prediction accuracy. Throughout the following of this section, the model building method will be given.

1) Model identification.

According to the characteristics of the partial correlation function, we can build the ARMA (p, q) model. Time series was undertook correlation analysis. Calculate the sample autocorrelation function and the sample partial correlation function. We can determine the model by truncation and tailing preliminarily (see Table I).

TABLE I
THE PROPERTIES OF ARMA (p, q) MODEL

	MA(q)	AR(p)	ARMA(p, q)
ACF	qth setp truncating	trailing	trailing
PACF	trailing	pth step truncating	trailing

2) Judging model's order.

Akaike's Information Criterion (AIC) can be used to define p and q .It used the maximum likelihood functions to judge model's order.

AIC can be defined as follows:

$$\mathrm{AIC}(n,m)=\ln \hat{\sigma}_a^2+2(n+m+1)/N \tag{6}$$

For time series, let P be the upper bound for p. let Q be the lower bound for q. Calculate the AIC for every (n,m) , $0\le n\le P$, $0\le q\le Q$. If the equation (7) is satisfied, then the order of the ARMA model is (p, q).

$$AIC(p,q)=\min_{0\le n,m\le L} AIC(n,m) \tag{7}$$

$\hat{\sigma}^2$, refers to the maximum likelihood estimate value for this series.

3) Model parameter estimation

Moment estimation, maximum likelihood estimation or least square estimation can be used for calculating the parameters: $(\hat{\varphi}_1,\cdots,\hat{\varphi}_p)$, $(\hat{\theta}_1,\cdots,\hat{\theta}_q)$.

4) Model test

Calculate the residual $\hat{\varepsilon}_t$ .If the model passes the white noise test, the model built is reasonable and it can be used for forecasting. Otherwise, we should improve the model.

III. EMD-ARMA MODELs

In the paper, this method firstly adopts EMD to decompose the historical data, then use ARMA (p, q) model to forecast each IMF. The concrete steps are as follows (Fig.1):

1) The history actual historical wind speed data is decomposed into a series of stationary intrinsic mode function (IMF) in different scale space via EMD sifting procedure.

2) According to all intrinsic mode functions and the residual, ARMA (p, q) models were built. Then do the wind speed forecasting by these ARMA (p, q) models.

3) Superpose the results forecasted by ARMA (p, q) model. We can obtain the finally predictive value.

4) Use the method proposed in [2] to mend the value forecasted by ARMA (p, q) model. According to persistence method, we know that y_t is highly correlated with y_{t-1} .So, we can mend the value by the formula (8).

$$\hat{y}_t = (\hat{y}_t + y_{t-1})/2 \tag{8}$$

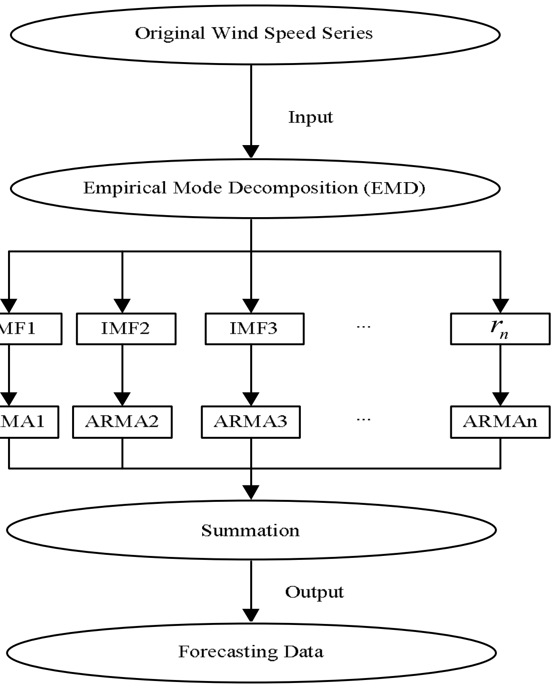

Fig.1. Flow chart of EMD-ARMA method

IV. EXAMPLE AND ANALYSIS OF RESULTS

This paper takes real data of a certain wind farm as an example. The real data's sampling interval is 1h. EMD-ARMA model was established on the basis of consecutive hourly mean wind speed samples, thus allowing the prediction of wind speeds at future time-step.

In experiments, we use the first 600 data values for model building, while the following 30 data values are used as checking data for validating the EMD-ARMA model we proposed.

Firstly, do successive iteration (1), until the series satisfy the IMF properties. But it is often difficult to satisfy the second property. So, we should determine a terminal condition to end the iteration.

This paper adopts the terminal conditions proposed by Rilling et al. as in [13-14]. They introduce a new criterion based on two thresholds θ_1 and θ_2, aimed at guaranteeing globally small fluctuations in the mean while taking into account locally large excursions.

Let $m(t)$ is the mean of the upper and lower envelope $u(t)$、$v(t)$, $m(t) = (u(t) + v(t))/2$. Let $a(t)$ be the mean of the difference value of the upper and lower envelope.

$$a(t) = (u(t) - v(t))/2 \tag{9}$$

$$\sigma(t) = m(t)/a(t) \tag{10}$$

Then the terminal conditions are two:

(a) Namely, sifting is iterated until $\sigma(t) < \theta_1$ for some prescribed fraction $(1-\alpha)$ of the total duration. We can express the condition as follows:

$$\#\{t \in D \mid \sigma(t) < \theta_1\} / \#\{t \in D\} \geq 1 - \alpha \tag{11}$$

In the equation (11), D is sustained scope of the signal. #A is the number of the element in A.

(b) At any time, $\sigma(t) < \theta_2$, let $\theta_2 = 10\theta_1$, namely $\theta_2 = 0.5$.

According to the terminal conditions introduced in above paragraphs, we can obtain the first IMF, the process as shown in Fig.2.

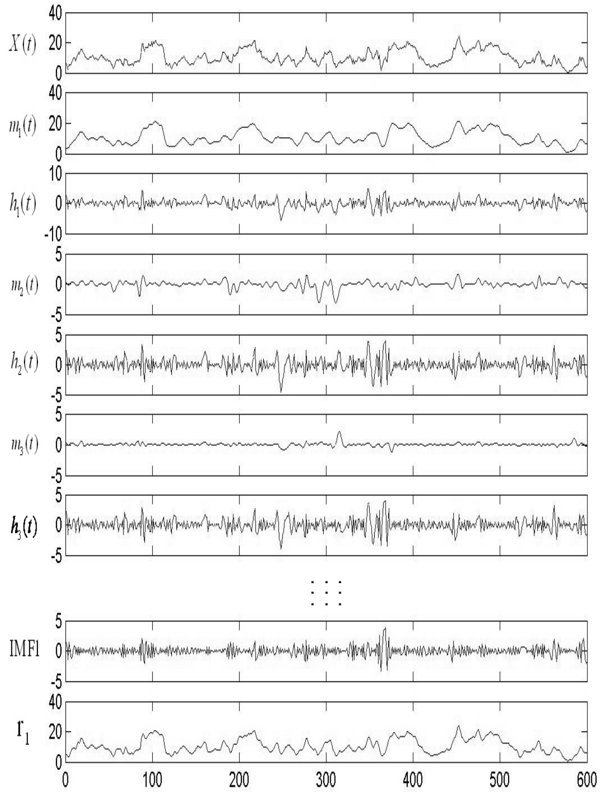

Fig.2. Process of IMF1 in EMD

Then we can obtain each IMF and the residue, the process is shown as in Fig.3.

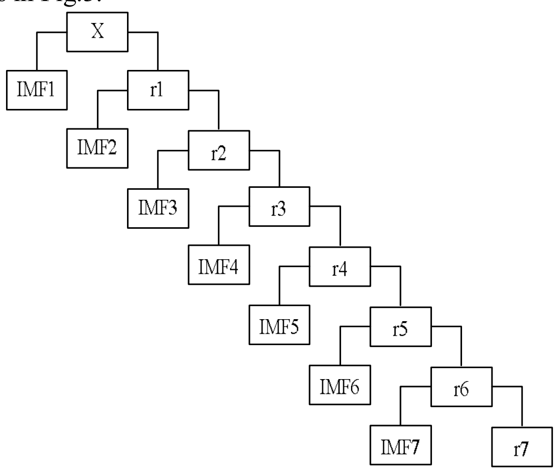

Fig.3. Flow chart of IMF

So, we can obtain seven IMF and the residue. The final result is shown as Fig.4.

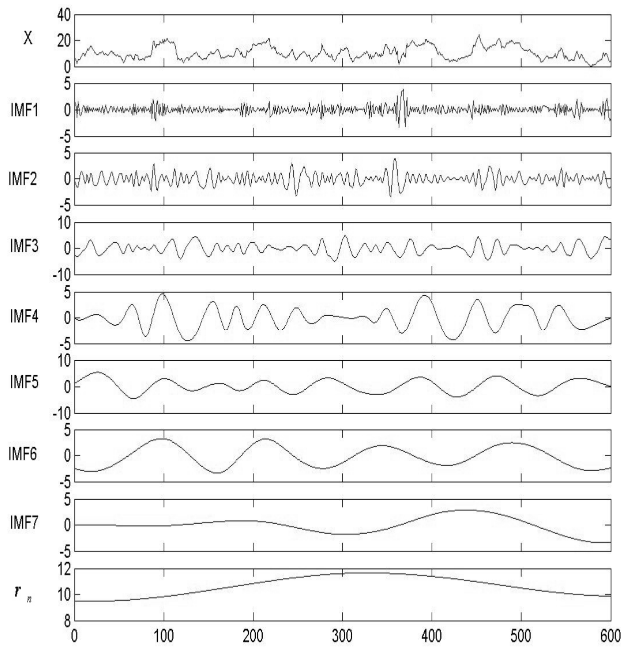

Fig.4. Wind speed series and decomposed time series by EMD

The corresponding iteration number k of each IMF is shown in Table 2.

TABLE II
THE ITERATION STEP OF EACH IMF

IMF	IMF1	IMF2	IMF3	IMF4
k	24	14	13	18
IMF	IMF5	IMF6	IMF7	
k	10	9	5	

Compared with original wind speed series, decomposed series IMF_i and r_n are approximate stationary series. Build model for IMF_i and r_n.

We proposed the results of IMF_3:

$$\begin{aligned} x_{3,t} &= 2.613x_{3,t-1} - 2.386x_{3,t-2} + 0.752x_{3,t-3} \\ &+\varepsilon_t + 1.105\varepsilon_{t-1} + 0.4984\varepsilon_{t-2} \end{aligned} \tag{12}$$

Then the forecasting model of IMF_3 is as follows:

$$\begin{aligned} x_{3,t+h} &= 2.613x_{3,t+h-1} - 2.386x_{3,t+h-2} + 0.752x_{3,t+h-3} \\ &+\varepsilon_{t+h} + 1.105\varepsilon_{t+h-1} + 0.4984\varepsilon_{t+h-2} \end{aligned} \tag{13}$$

Let h be the prediction step .The more prediction steps, the more error. Let h=1.

The forecasting values of IMF_n and r_n are $I\hat{M}F_n$ and $\hat{r}_n$ respectively, then the final forecasting result is:

$$\hat{X} = I\hat{M}F_1 + I\hat{M}F_2 + \cdots + I\hat{M}F_n + \hat{r}_n \tag{14}$$

Prediction results and error curve are just as follows (Fig.5, Fig.6):

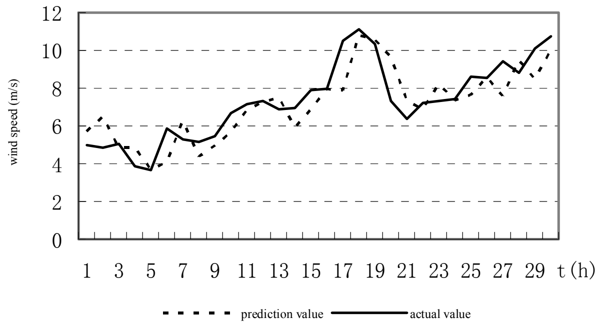

Fig.5. Comparison of wind velocity predicted with actual data in 1h-ahead forecasting

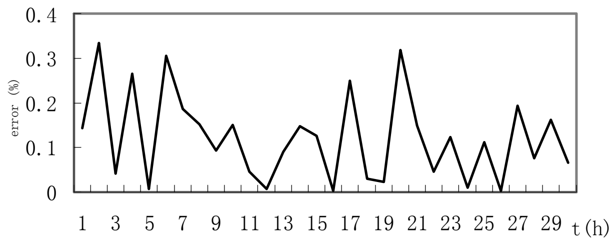

Fig.6. Error curve of 1h-ahead forecasting

In this study, an index, namely mean absolute percent error (MAPE), is used as forecasting precision measure. The index is represented as follows:

$$E_{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right| \tag{15}$$

Y_i is the actual time series value at period i, $\hat{Y}_i$ is the forecasting time series value at period i, and N is the number of forecasting results.

From the statistical analysis of the forecasting result, we know that the MAPE is 12.15%. At the same time, we do forecasting by single ARMA model. The MAPE is 14.55%. The error curve is shown in Fig.7.

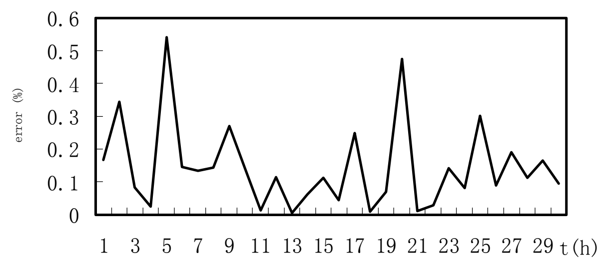

Fig.7. Error curve of 1h-ahead forecasting by single ARMA model

Then we take the real data whose sampling interval is 10 min as an example. In this experiment, we also use the first 600 data values for the training, while the following 30 data values are used as checking data (Fig.8, Fig.9). The MAPE is 3.62%.

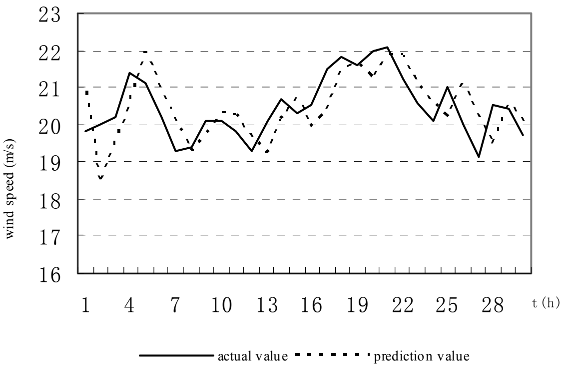

Fig.8. Comparison of wind velocity predicted with actual data in 10min-ahead forecasting

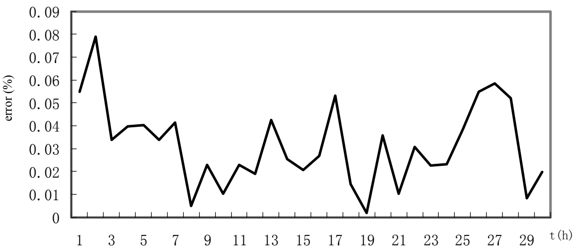

Fig.9. error curve of 10min-ahead forecasting

We can conclude that the prediction accuracy is associated to the prediction period. If the prediction period is short, wind speed varies slowly, prediction error was relatively small, and vice versa. Analyzing the error curve, we can conclude that prediction accuracy is influenced greatly by the regularity of the wind speed. EMD is a good data processing method for time series modeling. The EMD-ARMA model is an effective method of short-term wind speed forecasting.

V. CONCLUSION

In this paper, we proposed a method based on the nonlinear and non-stationary wind speed series. Wind speed sequence is decomposed into several approximate stationary IMF and the residual. EMD-ARMA model was built. We do one-hour ahead and ten-minute-ahead wind speed forecasting. The validity of this arithmetic has been testified by example. We analyze the relationship between the prediction accuracy and the prediction period.

The formation of wind is impacted by the objective law of atmospheric motion. Combination of the objective law and the weather forecasting is an effective way of improving the prediction period and prediction accuracy. Now, the short-time wind speed forecasting researched is the average wind speed forecasting. The terrain and the landform in the wind farm were not considered. Owing to the significance of these two points, data exchanging between power department and meteorological department by internet is an important way in grid connected wind power. Building refined wind speed forecasting model based on weather forecasting will provide guarantee for secure, stable and economic operation of the power system.

REFERENCES

[1] Lei Ya-zhou, “Studies on wind farm integration into power system,” *Automation of electric power system*, vol. 27, no. 8, pp. 84-89, 2003.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] Ding Ming，Zhang Li-jun, Wu Yi-chun, “Wind speed forecast model for wind farms based on time series analysis,” *Electric power automation equipment*, vol.25, no.8, pp.32-34, 2005.

[7] Kariniotakis G, Stavrakakis G, Nogaret E, “Wind power forecasting using advanced neural network models,” *IEEE Trans. on Energy Conversion*, vol.11, no.4, pp.762-767, 1996.

[8] 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.

[9] Haiyong Zhang, Qiang Gai. “Research on properties of empirical mode decomposition method,” *Proceeding of the 6th World Congress on Intelligent Contontrol and Automation*, pp.10001-10004, Jun. 2006.

[10] Z.Lu,J.S.Smith,Q.H.Wu,J.Fitch, “Empirical mode decomposition for power quality monitoring,” *2005 IEEE/PES Transmission and Distribution Conference & Exhibition*, pp.1-5, 2005.

[11] Li Tian-yun, Zhao Yan, Li Nan, “Apply empirical mode decomposition based Hilbert transform to power system transient signal analysis,” *Automation of electric power systems*, vol.29, no.4, pp.49-51, 2005.

[12] Zhu Zhi-hui, Sun Yun-lian, Ji Yu, “Short-term load forecasting based on empirical mode decomposition and least square support vector machine,” *Relay*, vol.35, no.8, pp.38-40, 2007.

[13] Zheng Tian-xiang, Yang Li-hua. “Discussion and improvement on empirical mode decomposition algorithm,” *ACTA scientiarum naturalium universitatis sunyatseni*, vol.46, no.1, pp.1-6, 2007.

[14] RILLING G，FLANDRIN P，GONCALVES P , “On empirical mode decomposition and its algorithms,” I*EEE-EURASIP Workshop on Nonlinear Signal and Image Processing , Grado (I)*, 2003.