TELKOMNIKA, Vol. 11, No. 12, December 2013, pp. 7007~7013
e-ISSN: 2087-278X

A Novel Strategy for Wind Speed Prediction in Wind Farm

Guang Yang[1], Ziqiang Hu[*2], Xinrong Liu[3]
[1]Electrical & Electionic Experment Training Center, Yantai University, Yantai, 264005, China
[2]School of Opto-electronic Information Science and Technology, Yantai University, Yantai, 264005, china
[3]Department of computer, Shandong Institute of Business and Technology, Yantai, 264005, China
*Corresponding author, e-mail: ygyantai@163.com

Abstract

The empirical mode decomposition (EMD) is well known for predicting wind speed.However, but the joint application of relevance vector machine (RVM) and empirical mode decomposition in wind speed forecasting is seldom found in the field. This paper proposes a relevance vector machine model based on empirical mode decomposition to predict the wind speed. Before the wind speed forecasting with RVM, EMD algorithm is used to decompose wind speed signal in order to weaken the disadvantageous influences of nonlinearity and uncertainty in wind spped. By the decomposition process, a series of intrinsic mode functions (IMFs) are generated. To each IMF, RVM algorithm is used to construct the model and carry on the forecast espectively. The final predicted result is obtained by the superposition of all prediction results. By the simulation experiment, the comparison of several algorithms is shown. The results showed that EMD-RVM model is effective, and has better forecasting precision

Keywords: *relevance vector machine, wind speed forecast, empirical mode decomposition, application, wind farm*

1. Introduction

Along with the increase of wind power investment, the wind power capacity that is connected to the power system is increasing rapidly [1]. However, Due to the intermittent and uncertain characteristics of the wind, the safe and stable operation of power system is affected severely. The accurate prediction of wind speed is helpful for the reliable and high-quality operation of power system, and reduces the operating costs of wind power generation [2]. The wind speed prediction can also let people to predict the output power of generator in advance, and can effectively reduce the impact of wind power fluctuations on the power grid.

Nowadays, wind speed forecasting methods usually are divided into two categories, namely one sort of method is based on the weather prediction, and another is based on historical data [3-4]. With the aid of numerical weather prediction, prediction time can reach 24h, 48h, 72h, or even longer. The physical method is based on the weather prediction, and in fact, it is not easy to acquire faultless weather prediction data of to establish the prediction model. In china, because of the lack of weather forecast information for wind farm, the forecast methods based on the historical data is mostly used. These methods include the method of time series analysis, persistence method, the method of Kalman filter, Fuzzy logical method, method of artificial neural network and the method of support vector machine (SVM) [5-6].

By pattern recognition and parameters estimation, the method of time series analysis builds its mathematical model. But it has a low forecasting precision to the low order model, and it is difficult to fulfill the parameters estimation for the higher order model.

In persistence method, the historical data is used to predict wind speed, and this method has the large prediction error.

The method of Kalman filter builds state space model of wind speed based on the state variable, but this method dose not run efficiently when the noise statistical characteristics are not given.

In Fuzzy logical method, a fuzzy linear model is set to approximate the nonlinear variation of wind speed. But the fuzzy forecasting has some limitation in predicting wind speed because it has weak learning ability and the relating theory need to be perfected.

Received March 31, 2013; Revised July 23, 2013; Accepted August 18, 2013

There is a good self-learning ability and adaptability for method of artificial neural network, but the accuracy of the prediction is low. That because some drawbacks such as the slowing convergence speed, easily falling into local minimum, and so on.

SVM is a machine learning algorithm based on statistics theory [7-8]. By means of the structural risk minimization principle, SVM can solve the small-sample problems well. It has well generalization ability and can not be trapped in local minimum. But the kernel function need to satisfy Mercers theorem and the parameters of kernel function are difficult to determine [9].

The above methods have their own characteristics, but also have their own defects. Research shows that it has large forecast error when the single forecasting method above is used to forecast the wind speed. Therefore, the single forecast method can not meet the requirements. Recent studies have indicated that the prediction accuracy using mixed method is better than that using a single method [10].

In order to predict wind sppeed more effectively and accurately, this paper presents a wind speed forecast model based on hybrid algorithm. The RVM and EMD are jointed to forecast wind speed. EMD is an algorithm which can decompose a nonlinear and non-stationary signal into a group of stable data series, and it is an effective way for wind speed forecasting. Compared with SVM, RVM has more advantages. RVM is based on Bayesian estimation theory. RVM algorithm can achieve the same recognition performance as SVM, while the computational complexity is reduced greatly.

By the process of EMD, the non-stationary wind speed signal is decomposed into some sequences of stable data at different scales. To each data sequence, the RVM model is built respectively. Each data sequence is forecasted by RVM algorithm, and the final forecast result is obtained by the superposition of all predictive value.

2. Research Method

2.1. Principle of EMD

Huang is a Chinese American, and he put forward the Empirical mode decomposition (EMD) in 1999. As the wavelet transforms (WTs), EMD is used to analysis the nonlinear and non-stationary signals. While keeping the multiresolution feature of wavelet transform, EMD avoids the difficulty in the choice of the wavelet bases. So, EMD has its superior in dealing with the complex non-stationary and nonlinear signals, it is a kind of practical signal self-adapted decomposition method [11].

A signal can be decomposed into a set of intrinsic mode functions (IMFs) by EMD, and each IMF should meet all the following conditions:

1) The difference between the zero crossing number and the extreme point number is not greater than 1.

2) At any point of IMF, the envelope mean is zero.

The decomposition process of EMD is described below:

First, all the extreme points of the original signal *(x(t))* should be found out, then by using cubic spline function, the minimum envelope and the maximum envelope are worked out. The envelope mean of original signal is calculated by the formula:

$$m_1(t) = \frac{e_+(t) + e_-(t)}{2} \tag{1}$$

Where $e_+(t)$ and $e_-(t)$ is the maximum envelope and the minimum envelope respectively and m1(t) means the envelope mean of original signal.

By subtracting the $m_1(t)$ from the original signal, a new signal $(h^1_1(t))$ is obtained and the formula is shown as below:

$$h_1^1(t) = x(t) - m_1(t) \tag{2}$$

Where *x(t)* is the original signal.

Generally, $h^1_1(t)$ dose not satisfy the criteria of IMF, and $h^1_1(t)$ is assigned to *x(t)*. Then recalculate the $m_1(t)$. Repeating the above process for several times until an IMF can be acquired. k is used to express the repeat times , the first IMF is shown as follows:

$$c_1(t) = h_1^k(t) \quad (3)$$

Where, $c_1(t)$ means the first IMF.
Subtracting the $c_1(t)$ from the initial signal$(x(t))$, we obtain a new signal $(r_1(t))$:

$$r_1(t) = x(t) - c_1(t) \quad (4)$$

Then, we assign $r_1(t)$ to $x(t)$. From the new $x(t)$, the second IMF $(c_2(t))$ is obtained by repeating the above process, and the new $x(t)$ is decomposed. So repeatedly, we can work out all IMFs and the residue.

The above process should go to end when the residue became monotonous function or a constant. The final expression is shown as follows:

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t) \quad (5)$$

Where $c_i(t)$ represents the ith IMF; n represents the number of IMF; $r_n(t)$ represents the residue.

Actually, the mean of envelope is not zero. The following formula is used to judge if the second condition of IMF is satisfied:

$$\frac{\sum [h_n^{k-1}(t) - h_n^k(t)]^2}{\sum [h_n^{k-1}(t)]^2} \leq \varepsilon \quad (6)$$

Here, ε represents threshold value, $0.2 \leq \varepsilon \leq 0.3$.
When the formula (6) is valid, the second condition of IMF is deemed to be met.

2.2. Principle of RVM

Though the SVM algorithm has been successfully applied to wind speed prediction [12], defects still exist in some respects. There are too many number of support vector in SVM, and this can lead to a long training time and over-fitting. The kernel function of SVM need to be meet Mercers theorem, and the penalty factors are required to be set in advance [13]. The different parameters can bring the predicted results great difference, and improper parameters can lead to an error result. Due to the above disadvantages, SVM has not been effective applied in practice.

Based on Bayesian estimation theory, the RVM algorithm can overcome the above defects of SVM. Due to the structure based on the probability and the serious sparse character, RVM not only improves the prediction accuracy and running time is shorter. The number of relevant vectors that are used to train is less than that in SVM. In addition, RVM dose not need to meet the Mercer conditions, and there is a more widely range to the choice of kernel function.

Similar to the SVM, the output of RVM is expressed as follow formula:

$$y(X, w) = \sum_{i=1}^{N} w_i K(X, X_i) + w_0 \quad (7)$$

Where X means the n-dimensional input vector; X_i represents the given training data; w means the weight parameter vector, and $w=(w_0, w_1, \ldots w_N)$.

Assuming that each target value of independent sample has error, and the probability formula is used to express the objective function of RVM. The target value is given as below:

$$t_i = y(x_i; w) + \varepsilon_i \quad (8)$$

t_i means the target value; ε_i means the error which meets Gaussian distribution with mean zero and variance σ^2.

The likelihood function of all samples is given as follows:

$$p(t \mid w,\sigma^2) = (2\pi\sigma^2)^{-\frac{N}{2}} \exp(-\frac{\|t-\Phi w\|^2}{2\sigma^2}) \tag{9}$$

Here $\Phi=[\varphi(x_1),\varphi(x_2),\ldots,\varphi(x_N)]^T$, $\varphi(x_i)=[1,k(x_i,x_1),\ldots.k(x_i,x_N)]^T$, and $k(\cdot)$ is kernel function; $t=(t_1,\ldots,t_N)^T$.

When the maximum likelihood estimation is used to evaluate w and σ^2 directly, it will lead to a severe over-fitting result. So, we define w as obeying Gaussian distribution with mean zero:

$$p(w \mid \alpha) = \prod_{i=0}^{N} N(w_i \mid 0,\alpha_i^{-1}) = \prod_{i=0}^{N} \frac{\alpha_i}{\sqrt{2\pi}} \exp(-\frac{\alpha_i w_i^2}{2}) \tag{10}$$

Here $\alpha=[\alpha_0, \alpha_1,\ldots\alpha_N]^T$.

According to Bayesian theory, the posterior probability distribution of unknown parameter is expressed as below:

$$p(w \mid t,\alpha,\sigma^2) = (2\pi)^{-\frac{N+1}{2}} |\Sigma|^{-\frac{1}{2}} \exp(-(w-\mu)^T \Sigma^{-1}(w-\mu)) \tag{11}$$

Σ means posterior covariance matrix; $A=diag(\alpha_0, \alpha_1,\ldots\alpha_N)$;$\mu=\sigma^{-2}\Sigma\Phi^T t$; $\Sigma=(\sigma^{-2}\Phi^T\Phi+A^{)-1}$.
The α and σ^2 can be finally worked out by using maximum likelihood method:

$$\alpha_i^{new} = \frac{\gamma_i}{\mu_i^2} \tag{12}$$

$$(\sigma^2)^{new} = \frac{\|t-\Phi\mu\|^2}{N-\sum_{i=0}^{N}\gamma_i} \tag{13}$$

Where N means the number of sample; $\gamma_i=1-\alpha_i\Sigma_{i,i}$; $\Sigma_{i,i}$ means the diagonal element of Σ.
For the given input, the output probability is subject to Gaussian distribution:

$$p(t^* \mid t,\alpha_{MP},\sigma_{MP}^2) = N(t^* \mid y^*,\sigma^*) \tag{14}$$

Here σ^2_{MP} means the optimal value of σ^2, and α_{MP} means the optimal value of α; t^* means the output predictive value; y^* is the mean of the predictive value and it is given as below:

$$y^* = \Phi(x^*)\mu \tag{15}$$

2.3. EMD-RVM Model

Due to the nonlinearity and nonstationarity of wind speed, the single forecasting method above can't complete accurate predictions. EMD is good at dealing with non-stationary signal, and wind speed is forecasted by joining EMD and RVM in this paper. We call the hybrid model EMD-RVM. In EMD-RVM, the wind speed signal is decomposed into many IMF components by EMD algorithm at first, and then, the RVM model is built for each IMF component respectively. To the RVM model, the kernel function need to be determined and the predicted value of each IMF component can be work out. By summing the predicted value of each IMF component, the final predicted value can be obtained. The specific process is given in Figure 1.

3. Experimental Results and Analysis

3.1. Decomposition of Wind Speed Signal

The wind speed signal that is used as the experimental data come from the wind farm of Horqin Left Wing Middle Banner (Inner Mongolia of china) on January of 2005. Every one hour, a group of wind speed time sequence is sampled. The first 600 data points are used to establish model. The hourly wind speeds of the next day, namely, the next 24 data points are forecasted. The decomposition process of EMD is carried out with matlab, and the flowchart of EMD decomposition is illustrated in Figure 2. The original wind speed series and the decomposed data component are given by Figure 3.

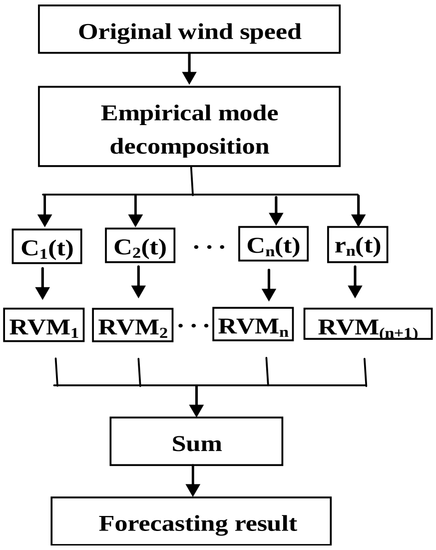

Figure 1. EMD-RVM Prediction

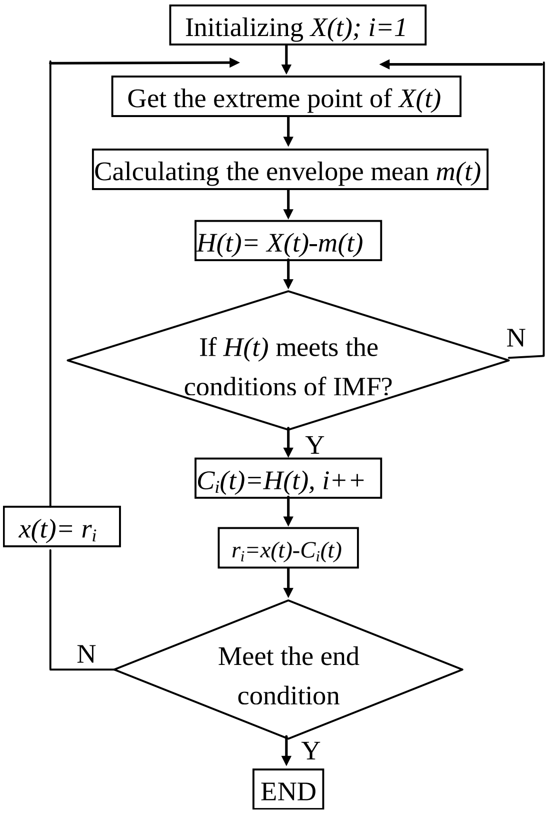

Figure 2. EMD decomposition

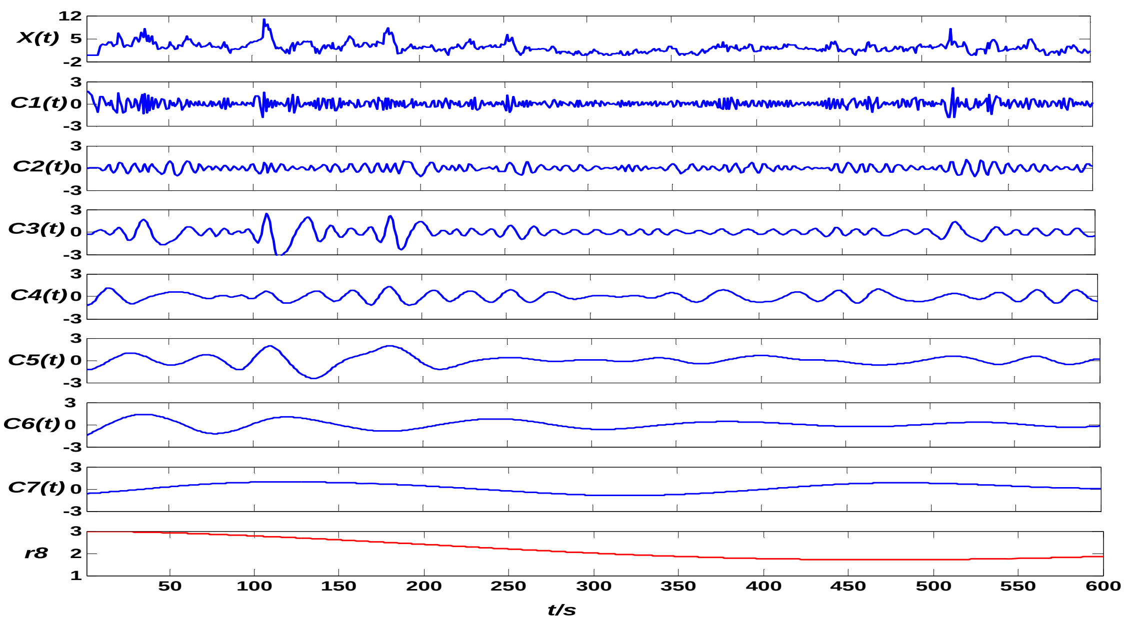

Figure 3. EMD Decomposition Results

X(t) is the original wind speed series; C1(t)- C7(t) are the IMF components after EMD decomposition and r8 is the residue. From Figure 3 we can see that the nonlinearity and uncertainty of original wind spped are reduced effectively, and C3(t)- C7(t) have preferable regularity.So, the more accurate forecasting results about C1(t) - C7(t) and r8 can be acquired.

3.2. RVM Forecast Model

The RVM models about C1(t) - C7(t) and r_8 are established respectively with the aid of matlab. At first, the kernel functions of RVM and their parameters must be determined. As the kernel function of RVM need not to meet Mercers theorem, theoretically, the kernel function can be chosen arbitrarily. But the kernel function is often used include radial basis function (RBF), linear kernel function and polynomial kernel. RBF has a good learning ability and broad scope of application. In this paper, RBF is chosen as the kernel function of all RVM models. RBF is expressed as follow formula:

$$k(v, v_i) = \exp(-\frac{\| v - v_i \|^2}{\delta^2}) \tag{16}$$

v means the input vector; *vi* means the ith dimensional vector of *v*; δ^2 means the kernel function’s width.

The δ^2 value in BRF has a great influence on the precision of prediction. In this paper genetic algorithm (GA) is used to calculate the optimal value of δ. The Mean absolute percentage error (MAPE) is elected to the fitness function of genetic algorithm:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{| y_i - y_i^* |}{y_i} \times 100 \tag{17}$$

Where y_i* and y_i means the predicted value and actual value respectively; *n* means the number of sample data.

The final predicted value is worked out by adding all predicted value together. To illustrate the effect of EMD-RVM, the predicted value by RVM and BP Neural network algorithm are also calculated. In Figure 4, the comparison curves for the BP predicted value, the RVM predicted value, EMD-RVM predicted value, and the actual value in the next 24 hours are given. As can be shown from Figure 4, all the BP, the RVM and EMD-RVM can forecast the wind speed, but the cure of EMD-RVM can better reflect the actual value. The reason is that the EMD has been used to the predicted model, and the forecasting process of non-stationary signal is become to the forecasting process of some relatively stable signals. That makes the predicted results more accurate.

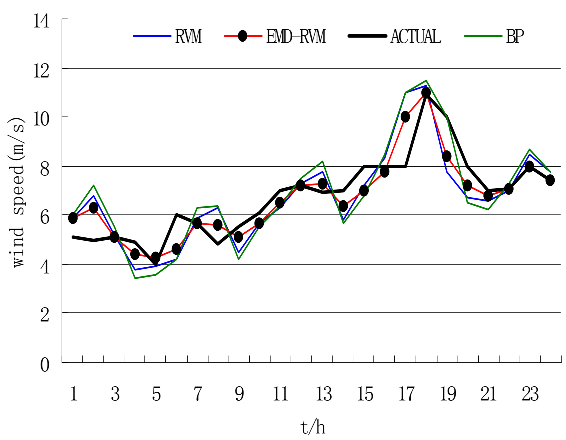

Figure 4. Comparison Curve of Predicting Outcomes

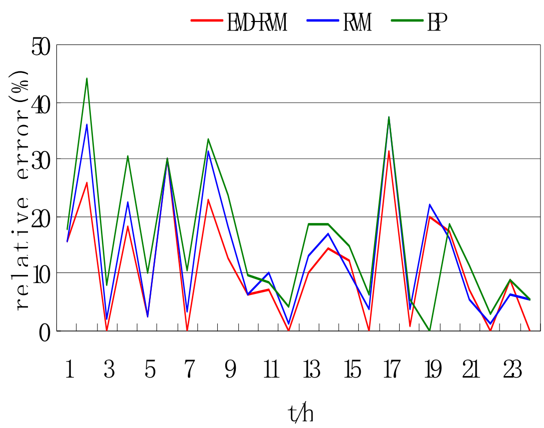

Figure 5. Comparison Curve of Relative Error

Because of the non-stationary and the nonlinearity of wind power, the predicted result has errors [14]. In Figure 5, the comparison curves of relative error are shown. From the comparison curve we can see that EMD-RVM has a smaller error.

The relative error is calculated by the follow formula:

$$E = \frac{|y_i - y_i^*|}{y_i} \times 100$$

where E is the of i point's relative error.

The MAPE is anther performance parameter used to measure the forcasting error of each arithmetic. The MAPE of BP is14.6; the MAPE of RVM is 12.4, and the MAPE of EMD-RVM is 9.72. Therefore, we may come to a conclusion that EMD-RVM has better prediction accuracy than the other arithmetic.

4. Conclusion

In this paper, a wind speed prediction model that joints EMD and RVM is proposed. The non-stationary wind speed signal is decomposed by EMD and the RVM model is set to fulfill the forecasting process. The results of experiment show that the method is feasible and the forecasting model has better prediction accuracy.

Acknowledgment

Project support by the Shandong Province Higher Educational Science and Technology Program (No. J10LG64); the Yantai Science and Technology Development Project (No. 2011057) and the Yantai University Youth Fund Project (No. sx11z9).

References

[1] Wang Xiao lan, Li Hui. *Effective wind speed forecasting in annual prediction of output power for wind farm*. Proceedings of the CSEE. 2010; 30(8): 117-122.
[2] Yang Xiu yuan, Xiao Yang, Chen Shu yong. *Wind speed and generated power forecasting in wind farm*. Proceedings of the CSEE. 2005; 25(11): 1-5.
[3] Ding Ming, Zhang Lijun, Wu Yichun. Wind speed forecast model for wind farms based on time series analysis. *Electric Power Automation Equipment*. 2005; 25(8): 32-34.
[4] James WT, Patrick E, Sharry M. Wind power density forecasting using ensemble predictions and times series models. *IEEE Transactions on Energy Conversion*. 2009; 24, (3): 775-782.
[5] Harikrishna D, Srikanth NV. Dynamic stability enhancement of power systems using neural-network controlled static compensator. *TELKOMNIKA Indonesia Journal of Electrical Engineering*. 2012; 10(1): 9-16.
[6] Huang Jianzhao, Xie Jian, Li Hongcai, Tian Gui, Chen Xiaobo. Self-adaptive decomposition level de-noising method based on wavelet transform. *TEIKOMNIKA Indonesia Journal of Electrical Engineering.* 2012; 10(5): 1015-1020.
[7] Damousis IG, Alexiadis MC, Theocharis JB, et al. A Fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation. *IEEE Transactions on Energy Conversion*. 2004; 19(21): 352-361.
[8] Feng Shuanglei, Wang Weisheng, Liu Chun, Dai Huizhu. *Study on the Physical Approach to Wind Power Prediction.* Proceedings of the CSEE. 2010; 30(2): 1-6.
[9] Li Gang, Wang Gui Long, Xue Hui Feng. GA Optimizing Method to Kernel Function Parameters of RVM. *Control Engineering of China*. 2010; 17(3): 335-337.
[10] Ye Lin, Liu Peng. Combined Model Based on EMD-SVM for Short-term Wind Power Prediction. *Proceedings of the CSEE*. 2011; 31(31): 102-108.
[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*. 2005; 29(4): 49-51.
[12] Zhang Hua, Zeng Jie. Wind Speed Forecasting Model Study Based on Support Vector Machine. *Acta Energiae Solaris Sinica*. 2012; 31(7): 928-932.
[13] Sun Guo Qiang, Wei Zhinong, Zhai Weixing. Short Term Wind Speed Forecasting Based on RVM and ARMA Error Correcting. *Transactions of China Electrotechnical Society*. 2012; 27 (8): 187-193.
[14] Pan Difu, Liu Hui, Li Yanfei. *Optimization algorithm of short-term multi-step wind speed forecast*. Proceedings of the CSEE. 2008; 28(26): 87-91.