**Lei DONG[1], Shuang GAO[1], Xiaozhong LIAO[1], Yang GAO[2]**

Beijing Institute of Technology (1), Shenyang Institute of Engineering (2)

# Short-Term Wind Power Forecasting with Combined Prediction Based on Chaotic Analysis

***Abstract**. With the integration of wind energy into electricity grids, it is becoming increasingly important to obtain accurate wind power forecasts. In this paper, models for short-term wind power prediction in large wind farms are discussed. The analysis of modeling with low dimensions nonlinear dynamics indicates that wind power time series have chaotic characteristics and wind power can be predicted in the short-term. The wind power prediction models are built with phase space reconstruction method and the combination model with different embedding dimensions is tested.*

***Streszczenie.** Opisano metodę krótkoterminowego prognozowania mocy elektrowni wiatrowych bazującą na chaotycznym charakterze wiatru w krótkich odstępach czasu. (**Krótkoterminowa prognoza mocy elektrowni wiatrowych bazująca na teorii chaosu**).*

**Keywords:** Wind Power Generation, Short Term, Chaotic Characteristic, Phase Space Reconstruction.
**Słowa kluczowe:** sieci elektrowni wiatrowych, prognozowanie.

## Introduction

Wind power generation has rapidly developed over the last decade and global installed wind power continues to grow at around 28% per year [1]. By the year 2020 wind power generation will supply about 12% of the total world electricity demands [2]. Because of the wind power fluctuations in direction and speed, wind power generation integrated in electrical power system may cause several problems, these problems include; power quality, stability and especially power dispatching. The problems increase as the penetration of wind energy increases [3-6].

Prediction of wind power, along with load forecasting, permits scheduling the connection or disconnection of wind turbines or conventional generators, thus achieving low spinning reserve and optimal operating cost. The horizon of prediction may vary: long-term horizon, one year; short-term horizon, 36hours, 10 hours, and 1hour.

Different horizons correspond to different objectives. There are two main state-of-art approaches, one based on physical deterministic modelling [3,7-9] and a second one based on statistical or time series modelling [2,4-6]. Physical models are usually used for long-term prediction and physical consideration has a great effect on the accuracy. Since wind power is a function of wind speed, predictions of wind power are generally derived from predictions of wind speed [4,5]. The statistical models are direct transformations of the input variables to wind power. Advanced statistical analysis methods are ARMA [10], Kalman Filters [11], artificial neural network (ANN) [12-14], fuzzy logic [15], support vector machine [16] and so on.

This paper is organized as follows. Firstly, the method of dynamical phase space reconstruction is given. The chaos ANN prediction model is presented. Then, the linear combination and ANN combination prediction method are presented. The experimental results of comparing the algorithm proposed in this paper with other algorithms are also presented. Finally, our work of this paper is summarized in the last section.

## Reconstruction of dynamical phase space

The wind power generation system is a complex dynamical system. The reconstruction of dynamics from the time series of wind power generation should be done before discussing the chaotic behavior of time series of wind power generation, which is built based on Takens' embedding theorem. This theorem provides the mathematical basis of the dynamic reconstruction problem. It states that model reconstruction of a nonlinear dynamical system, using just one observation of the system, should succeed to a certain extent and the reconstruction is independent of which signal component is used [17].

With a chaotic time series, consider a $D$-dimensional compact manifold $M$, Takens' embedding theorem says with $m \geqslant 2D+1$, the map $\Phi: M \rightarrow R^m$ defines a corresponding trajectory, the trajectory of the reconstructed space is diffeomorphism with original dynamical system, where m is the embedding dimension [17].

Typically, a wind power time series $x_1, x_2, \ldots, x_{n-1}, x_n$ is a series of record values of observations of a dynamical system at regular, discrete time intervals ($\Delta t$). Assume the embedding dimension is $m$, and the time delay is $\tau = k\Delta t$, then the reconstructed phase space becomes

$$Y(i) = [x(i), x(i+\tau), x(i+2\tau), \ldots, x(i+(m-1)\tau)], \quad i = 1,2,\ldots,N, \quad N = n-(m-1)\tau \tag{1}$$

Embedding dimension $m$ and time delay $\tau$ are two important parameters to perform the dynamic reconstruction. They will impact the aspect of the optimal prediction of a reconstructed dynamical model. Embedding dimension $m$ is the component number of each point in the reconstructed phase space, and it can be the input number of ANN prediction model. Time delay $\tau$ is the time interval between two nearby components of each point in the reconstructed phase space and it can be the inputs time interval of ANN prediction model, i.e. the input values of ANN represent a phase point in the reconstruct phase space.

A nonlinear approach is used to measure time delay $\tau$ called the mutual information $I$. As far as the quality of prediction with neural networks is concerned, it is best to choose the value of $\tau$ somewhere in between 1 and the first local minimum of the mutual information function $I(\tau)$ [18].

The mutual information of a time series as follows

$$\begin{aligned} I(Q,S) &= H(Q) - H(Q \mid S) = H(Q) + H(S) - H(S,Q) \\ &= -\sum_i P(q_i)\log P(q_i) - \sum_j P(s_j)\log P(s_j) + \\ &\sum_i\sum_j P(s_j,q_i)\log P(s_j,q_i) \end{aligned} \tag{2}$$

where: $Q$, $S$ – two discrete variables, $H(*)$ – information entropy, the more $H(*)$is, the stronger uncertainty is. $H(S,Q)$ – united information entropy, $H(Q|S)$ – conditional entropy, which indicates the uncertainty of $Q$ in case that $S$ is known. $P(q_i)$ – appearance probability of $q_i$ event, $P(q_i, s_j)$ – simultaneous appearance probability.

The value of $I(Q,S)$ indicates the relevance level between $Q$ and $S$. The smaller $I(Q,S)$ is, the weaker the relevance between $Q$ and $S$.

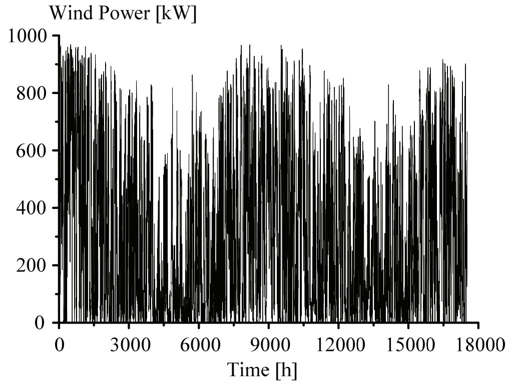

Fig.1. Measured wind power time series from Fujin wind farm (2006. 1. 1 to 2007. 12. 31)

Figure 1 shows original wind power time series from 2006. 1. 1 to 2007. 12. 31, which are from the wind farm located in the Fujin China. The wind power time series from 2006. 1. 1 to 2006. 12. 31 in Fujin Wind Farm are used for training the prediction model and the wind power time series from 2007. 1. 1 to 2007. 12. 31 used to test the prediction results.

Assume $S=x(k)$ and $Q=x(k+\tau)$, the $I(Q,S)$ is shown in Figure 2. When the time delay $\tau$ changes from 1 to 50, the first local minimum of the mutual information function $I(\tau)$ is 12, so we choose 12 as time delay of the wind power time series.

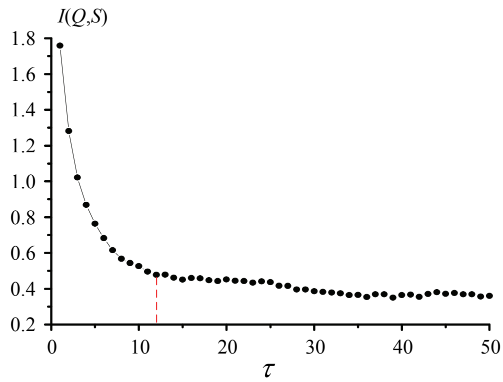

Fig.2. Mutual information versus time delay

Paper [15] presents a method to determine the minimum embedding dimension. Consider a wind power time series $x_1$, $x_2$, …, $x_{n-1}$, $x_n$. The time delay has been determined above, so the reconstructed phase space is shown as equation (1). Define

(3) $$a(i,m)=\frac{\| Y_{m+1}(i)-Y_{m+1}(n(i,m)) \|}{\| Y_m(i)-Y_m(n(i,m)) \|}$$

(4) $$E(m)=\frac{1}{N-m\tau}\sum_{i=1}^{N-m\tau} a(i,m)$$

(5) $$E1(m)=\frac{E(m+1)}{E(m)}$$

where: $i=1,2,\cdots,N-m\tau$. $Y_{m+1}(n(i,m))$ is the nearest neighbour of $Y_{m+1}(i)$, and $Y_m(n(i,m))$ is the nearest neighbour of $Y_m(i)$. $E(m)$ is a middle function, which can simplify the process to access embedding dimension $m$. If $m$ is qualified as an embedding dimension by the embedding theorem, then any two points which stay close in the m-dimensional reconstructed space will be still close in the ($m$+1)-dimensional reconstructed space [15].

Suppose the time series comes from an attractor, $E1(m)$ will stop changing when m is greater than some value m0. Then the $m_0$+1 is the minimum embedding dimension. To investigate the wind power time series of Figure 1, $E1(m)$ is calculated and is shown in Figure 3. When m is greater than 9 the $E1(m)$ stops changing and then the embedding dimension of wind power generation time series in Figure 1 is 10.

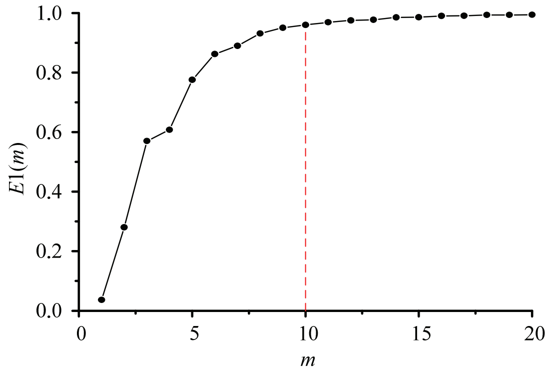

Fig.3. The values $E1(m)$ for wind power time series

It has to be confirmed that the wind power generation system is chaotic before applying chaos theory to predict the wind power, to satisfy this need, chaotic characteristics of wind power must be determined. In practice, it can be defined that a bounded deterministic dynamical system with at least one positive Lyapunov exponent is a chaotic system [17]. The deterministic and bounded properties are obvious of the wind power generation system in terms of finite installation power capacity and finite attractor dimension. Then the criterion of at least one positive Lyapunov exponent plays a key role in the characterization of a process as chaotic, because the Lyapunov exponents not only show qualitatively the sensitive dependence on initial conditions, but also give a quantitative measure of the average rate of separation or attraction of nearby trajectories on the attractor.

An algorithm developed by Wolf et al [19], which implements the theory in a very simple and direct fashion, can be used to calculate the largest Lyapunov exponent. As mentioned above, the initial point of the reconstructed phase space of wind power time series in equation (1) is $Y(t_0)$. Assume the nearest neighbour of the initial point is $Y_0(t_0)$. Let $L_0$ denote the Euclidean distance between them. Next, we have to iterate both points forward for a fixed evolution time, which should be of the same order of magnitude as the time delay $\tau$, and finally they evolve to two points, $Y(t_1)$ and $Y_0(t_1)$, along with their trajectories separately. The distance of the two new points is $L_0^{'}=|Y(t_1)-Y_0(t_1)|>\varepsilon$, where $\varepsilon$ >0. Then another nearest neighbour $Y_1(t_1)$ of the point $Y(t_1)$ should be found, which has the minimum distance between the pair of points, $L_1=|Y(t_1)-Y_1(t_1)|<\varepsilon$. This procedure is repeated until the initial point $Y(t_0)$ reaches the end of the time series, shown in Figure 4 [19]. Finally, the largest Lyapunov exponent is calculated according to the equation

(6) $$\sigma_{\max}=\frac{1}{t_M-t_0}\sum_{i=0}^{M}\ln\frac{L'_i}{L_i}$$

where: $M$ – the total number of replacement steps. $\sigma_{\max}$ indicates the average separation rate of nearby trajectories in the reconstructed phase space, and it determines the future prediction time.

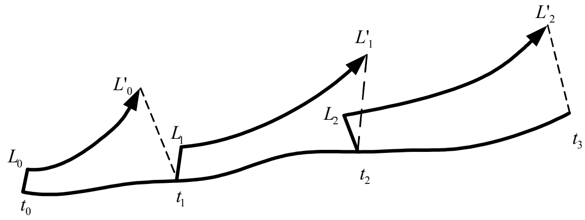

Fig.4. Wolf method for largest Lyapunov exponent calculation

By using equation (6), the largest Lyapunov exponent for the attractor presented in Figure 1 can be calculated. The largest Lyapunov exponent converges very well to $\sigma_{max}$=0.0638. This is a firm proof for the chaotic behavior of the wind power generation system.

ANN model is suitable for wind power prediction but the ANN model is difficult to be configured. There are 3 steps for ANN model design, 1) identify the appropriate input parameters, 2) set up appropriate network structure, and 3) design a training algorithm that yields the best training performance. Shuhui Li, et al [7] presents an ANN model with 4 input data, which are wind speed and direction from two meteorological towers. Rohrig, K, et al [20] presents another ANN model with 8 input data in terms of 3 wind speed values, 3 direction values, and 2 values related to the time series. Unfortunately, these patterns are limited by the meteorological towers and cannot give out the hub height wind speed for each turbine. This issue can be considered with the application of the chaos-based ANN model. An ANN serves as the reconstructed phase space model of the wind power generation system. The input data are the $Y(i)=[x(i), x(i+\tau), x(i+2\tau),\ldots, x(i+(m-1)\tau)]$, $i=1,2,\ldots, N$, $N=n-(m-1)\tau$), and the number of input parameters is the embedding dimension $m$. The predicted value of the time series point $i+(m-1)\tau)+1$, i. e. $x(i+(m-1)\tau+1)$, is set as the output of the chaos-based ANN model. The nonlinear function $\hat{\Phi}: M \to R^{D_E}$ will be obtained by training the ANN. This ANN model is diffeomorphism with original dynamical system of wind power generation system, and has locally (i.e. short-term) predictable ability.

The configuration of ANN for wind power prediction shows in the Figure 5 [14]. The prediction accuracy of the models with monthly normalized mean absolute error (*NMAE*) and normalized root mean squared error (*NRMSE*) are expressed in equations (7)–(8).

(7) $$NMAE = \frac{1}{P_{cap}} \times \frac{1}{N} \sum_{i=1}^{N} | x'(i) - x(i) | \times 100\%$$

(8) $$NRMSE = \frac{1}{P_{cap}} \times \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x'(i) - x(i))^2} \times 100\%$$

where: $x(i)$ – actual value, $x'(i)$ – predicted value, $N$ – the number of prediction sample, $P_{cap}$ – the rated capacity of the wind turbine.

The chaos ANN model presented in this paper performs well, *NMAE* of 2007 is 6.17%, and *NRMSE* of 2007 is 9.22%.

Persistence is a simple method, which considers that the wind power production remains the same in all look-ahead times [15]. The same data as the chaos method are used. The *NMAE* of 2007 is 5.96%, and the *NRMSE* of 2007 is 9.35%.

From comparisons, although the pure chaos-based ANN model does not play well as the persistence model from the *NMAE*, the *NRMSE* is smaller than persistence model. And the prediction accuracy can be improved by building a combined prediction model on the basis of chaos-based ANN model.

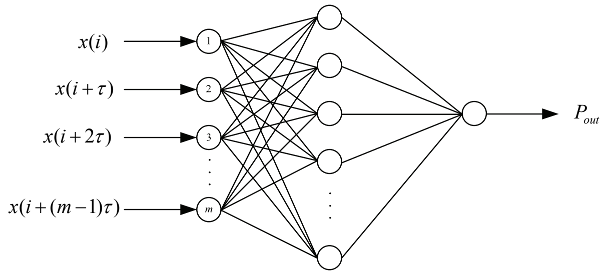

Fig.5. Multi-layer neural network for wind power generation prediction

**Combined prediction model**

Wind power prediction of large wind farm using only a single model brings the risk of high and costly errors, particularly in the case of extreme events where individual models can go wrong. The choice of the embedding dimension of reconstructed dynamical model will be not accurate because the inflexion of $E1(m)$ curve is very difficult to be determined. In order to diminish the effect of reconstructed parameters of prediction of chaotic system, a combined model for wind power prediction based on multi-dimension embedding is proposed.

We introduce the linear prediction model first. $x(t)$ ($t=1,2,\ldots, n$)is the measured wind power generation at time $t$. Assume there are $p$ different embedding dimensions around the determined one in the above-mentioned. And the predicted values of respective ANN models with different embedding dimensions are $\hat{x}_1(t), \hat{x}_2(t), \cdots, \hat{x}_p(t)$ . The linear combination model can be written as

(9) $$\hat{x}(t) = l_1 \hat{x}_1(t) + l_2 \hat{x}_2(t) + \cdots + l_p \hat{x}_p(t)$$

where: $L = (l_1, l_2, \cdots, l_p)^T$ – the weights of each model and

(10) $$l_1 + l_2 + \cdots + l_p = 1$$

The prediction error of the model $i$ at time $t$ is

(11) $$e_{it} = x(t) - \hat{x}_i(t)\, , i = 1,2,\cdots, p\, ,\ t = 1,2,\cdots, n$$

Then the error matrix shows as

(12) $$E = [(e_{it})_{p \times n}][(e_{it})_{p \times n}]^T$$

The quadratic sum of error of the linear combination model is

(13) $$J = \sum_{t=1}^{n} [x(t) - \hat{x}(t)]^2 = \sum_{t=1}^{n} (\sum_{i=1}^{p} l_i e_{it})^2 = L^T EL$$

The minimum of the quadratic sum of error is selected as objective function, then combination weights can be obtained by calculating the optimal function (14)

(14) $$\min J = L^T EL \qquad s.t. R^T L = 1, R = (1,1,\cdots,1)^T$$

According to the linear combination model, the weights are fixed, so it can not be suitable for changing weather situation. Choosing a combination of the best weights for the specific weather situation, leads to a significantly improved wind power prediction.

If the predicted $\hat{x}(t)$ satisfied

(15) $$\hat{x}(t) = \phi(\hat{x}_1(t), \hat{x}_2(t), ..., \hat{x}_p(t))$$

where: $\phi$ – a nonlinear function. This combination model is nonlinear prediction model i.e. the combination weights are changing. Because the ANN can approximate any nonlinear function, the ANN combination model will be a good choice for wind power prediction. The ANN combination model can also regulate the weights of the different models after training of the ANN, so more history information can be used for future prediction.

**Experimental result**

The wind power time series from 2006. 1. 1 to 2006. 12. 31 in Fujin Wind Farm are used for training the prediction model and the wind power time series from 2007. 1. 1 to 2007. 12. 31 used to test the prediction results.

Three models based on phase space reconstruction are built. The dimensions are $m_1$=9, $m_2$=10, $m_3$=11, respectively. By using quadprog function in MATLAB optimization toolbox, the optimal combination weights of the three prediction models are obtained by calculating the optimal function (14), $w_1$=0.1866, $w_2$=0.5667, $w_3$=0.267. The $NMAE$ of linear combination model of 2007 is 6.16%, while the values for the individual predictions of 2007 are 6.21%, 6.17% and 6.20%. And the $NRMSE$ of linear combination model in 2007 is 9.21% which is smaller than persistence model.

To find an optimal combination of different phase space reconstruction model with $m_1$=9, $m_2$=10, $m_3$=11, the ANN is used for regulating weights of each model, which can synthesize information and fuse prediction deviation in different embedding dimensions, resulting in the prediction accuracy is improved. The prediction results of the three embedding dimensions are used as three inputs of ANN and the actual wind power is the output. Then the weights of the three prediction methods in ANN combination model can be obtained by self-learning. The node number of hidden layer is determined by cut-and-try method, and the network structure of ANN combination model is 3-2-1. The $NMAE$ of ANN combination model in 2007 is 6.10% and $NRMSE$ is 9.21%, which is also smaller than persistence model.

The $NRMSE$ of different prediction models are shown in Figure 6. From the comparisons in Figure 6, although the errors are very similar, the $NRMSE$ of persistence model are the biggest for all months of 2007. The linear combination model and ANN combination model have good performance, and they perform a little better than the pure ANN model.

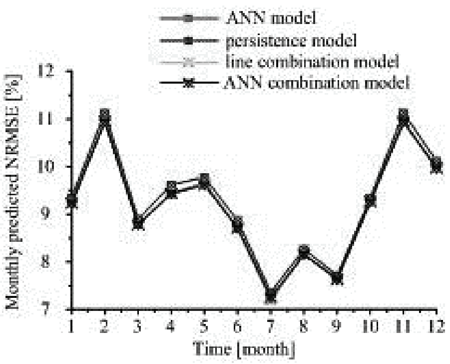


Fig.6. $NRMSE$ of different prediction models

The errors of different models of 2007 are shown in Table 1. From the comparisons in Table 1, it can be concluded that: (1) When the single embedding dimension is used, m=10 is the best and the weight in linear combination model is the largest; (2) The combination model is better than ANN model with the single embedding dimension; (3) The $NRMSE$ of combination models are smaller than persistence model.

Table 1. Errors of Different Models.

| Models | $NMAE$ [%] | $NRMSE$ [%] |
|---|---|---|
| ANN model ($m$=9) | 6.21 | 9.25 |
| ANN model ($m$=10) | 6.17 | 9.22 |
| ANN model ($m$=11) | 6.20 | 9.24 |
| Linear combination model | 6.16 | 9.21 |
| ANN combination model | 6.10 | 9.21 |
| Persistence model | 5.96 | 9.35 |

**Conclusions**

The time series of wind power generation are analyzed in this paper. The wind power system indicates typical chaotic characteristics. According to the chaotic behaviors of wind power generation, the ANN model based on phase space reconstruction is used for wind power prediction. The design method of ANN model is discussed in this paper.

Prediction accuracy is affected by different embedding dimension which is very difficult to choose in phase space reconstruction. In order to diminish the effect of reconstructed parameters of prediction of chaotic system, a combined model for wind power prediction based on multi-dimension embedding is proposed. The linear combination model and the ANN combination model show a little better performance of wind power prediction when compared with the pure ANN model. The $NRMSE$ of combination models are smaller than persistence model.

*This study was supported by China Power Investment Corporation science & technology project (2009064ZDGKJX).*

REFERENCES

[1] Analysis and Prospects of Global Renewable Energy Development Situation in 2008 [Online].Available:http://www.chinapower.com.cn
[2] El-Fouly THM, El-Saadany EF, Salama MMA. One day ahead prediction of wind speed using annual trends. Proceedings of the Power Engineering Society General Meeting 2006
[3] Ferreira LM. Evaluation of Short-Term Wind Predictability. IEEE Transactions on Energy Conversion, 7 (1992) 409-417
[4] Alexiadis MC, Dokopoulos PS, Sahsamanoglou HS. Wind speed and power forecasting based on spatial correlation models. IEEE Transactions on Energy Conversion, 14 (1999) 836-842
[5] Ioannis G, Minas C, John B. A Fuzzy Model for Wind Speed Prediction and Power Generation in Wind Parks Using Spatial Correlation. IEEE Transactions on Energy Conversion, 19 (2004), 352-361
[6] Li S. Wind power prediction using recurrent multilayer perceptron neural networks. Power Engineering Society General Meeting, IEEE Conference, 4 (2003), 2325-2330
[7] Li S, Wunsch DC, O'Hair EA, Giesselmann MG. Using neural networks to estimate wind turbine power generation. IEEE Transactions on Energy Conversion, 16 (2001), 276-282
[8] Watson SJ, Landberg L, Halliday JA. Application of wind speed forecasting to the integration of wind energy into a large scale power system. IEE Proceedings-Generation, Transmission and Distribution, 141 (1994), 357-362
[9] Pison P, Kariniotakis GN. Wind Power Forecasting using Fuzzy Neural Networks Enhanced with On-line Prediction Risk Assessment. IEEE Bologna PowerTech Conference, Bologna, Italy, 2003
[10] Yang X, Yang X, Chen S. Wind speed and generated power forecasting in wind farm. Proceedings of the CSEE, 25 (2005), 1-5
[11] Bossanyi EA. Short-term wind prediction using Kalman filters. Wind Engineering, 9 (1985), 1-8.

[12] Ruddy B. Very short-term wind power forecasting with neural networks and adaptive Bayesian learning. Renewable Energy, 36 (2011), 1118-1124
[13] Flores P, Tapia A, Tapia G. Application of a control algorithm for wind speed prediction and active power generation. Renewable Energy, 30 (2005), 523-536
[14] Li G, Shi J, Zhou J. Bayesian adaptive combination of short-term wind speed forecasts from neural network models. Renewable Energy, 36 (2011), 352-359
[15] Sideratos G, Hatziargyriou N. Using radial basis neural networks to estimate wind power production. In Proc. Power Eng. Soc. General Meeting, Tampa, FL, 2007
[16] Du Y, Lu J, Li Q. Short-term wind speed forecasting of wind farm based on least square-support vector machine. Power System Technology, 32 (2008), 62-66
[17] Haykin S, Li X. Detection of signals in chaos. Proceedings of the IEEE, 83 (1995), 95-122
[18] Rape R, Fefer D, Drnovsek J. Time series prediction with neural networks: a case study of two examples. Instrumentation and Measurement Technology Conference, 1994. IMTC/94. Conference proceedings. 10th Anniversary. Advanced Technologies in I & M., IEEE, 1 (1994), 145–148
[19] Wolf A, Swift JB, Swinney HL, Vastano JA. Determining Lyapunov exponents from a time series. Physica D, 16 (1985) 285–317
[20] Rohrig K, Lange B. Application of wind power prediction tools for power system operations. Power Engineering Society General Meeting 2006, 5

***Authors**: Lei Dong, Beijing Institute of Technology, E-mail: pemc.bit@163.com; Shuang Gao, Beijing Institute of Technology, E-mail: gao19820808@163.com; Xiaozhong Liao, Beijing Institute of Technology, E-mail: liaoxiaozhong@bit.edu.cn; Yang Gao, Shenyang Institute of Engineering, E-mail: motor.com@163.com.*