The Short-term Wind Power Forecast Based on Phase-space Reconstruction and Neural Networks

Jin-hua Zhang[1], Yong-qian Liu[2], De Tian[2]
[1,2] State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (North China Electric Power University)
Beijing, China
[1]School of power, North China University of Water Resources and Electric Power
Zhengzhou. China
zhangjh@ncwu.edu.cn, yqliu@ncepu.edu.cn

Journal of Digital Information Management

ABSTRACT: *The short-term forecast of the wind power of a wind farm is of great significance for the security and stability of a grid-connected generation system. An accurate forecast may reduce the spinning reserve of a grid while providing reliable references for operation dispatch of a wind farm. In order to improve the accuracy of short-term forecasts, introducing the phase-space reconstruction technique of the chaos theory, this paper was established forecasting models by reconstructing the historical time-series data of the wind power of a single unit based on the dynamical properties of chaos sequences, choosing the best delay time with the mutual information method, determining the best combination of embedding dimensions with the Cao algorithm, as well as utilizing the Elman recurrent neural network and others like the BP. As comparative case analysis shows, the forecasts of Elman model are more accurate than that of the others, exhibiting a positive prospect of utilizing this combined model of phase-space reconstruction and neural network in wind power forecasting of a single unit.*

Categories and Subject Descriptors:
C.1.3 [Network Architecture Styles]: Neural nets **F.1 [Computation by Abstract Devices]:** Self-modifying machines

General Terms:
Neural networks, Embedding techniques

Keywords: Phase-Space Reconstruction, Embedding Dimensions, Delay time, The short-term Wind Power forecasts, Elman Neural Network, BP Neural Network

Received: 19 August 2012, Revised 1 October 2012, Accepted 8 October 2012

1. Introduction

The wind speed, with its random values, brings flexibility to the power output of wind turbines, which, if connected to a power grid in large quantities, may present serious challenges to the security, stability and quality of the grid. Short-term forecasts of the wind speed and the generated power is an effective way to alleviate the negative impacts of wind power on grid and raise the installment ratio of wind power generators in high-voltage grids as well as offering reliable references for optimization of wind farm operation dispatch.

Generally, forecasts of generated power of wind farms can be categorized into long-term forecasts, mid-term forecasts, short-term forecasts and ultra-short-term forecasts. The short-term forecast, the forecast of the generated power 24 to 72 hours prior to its happening, aims at facilitating the operation dispatch of a grid and ensuring the power quality [1-3]. The current approaches to forecasts of generated power of a wind farm mainly include physical methods, duration, time series, the Kalman filter method, neural networks, fuzzy logic, spacial correlation and their combinations. The Chaos, a reciprocating motion generated by the system of deterministic kinetics with features of built-in randomness, impossibility of long-term prediction and high sensitivity to its initial value, is one of the major objects of non-linear dynamical studies. The non-linear time series can be extracted from the chaotic system and the internal laws of the series can be reflected by the phase-space reconstructed. Although the long-term prediction of the chaotic time series is rather difficult, the accuracy of the short term forecast is high.

Statistics and calculations out of long-term researches show that the time series of wind speed or wind power fall into the same chaotic time series class [9- 11]. By establishing models with the actual historical data of a large wind power plant in China, this paper confirms their chaotic properties, based on which neural networks of Elman and BP are established with the approach of phase-space reconstruction for the short-term forecast of single units of wind farms. Cases prove that, to some degree, the Elaman model helps to improve the accuracy of the forecast.

2. The Phase-space Reconstruction of the Chaotic Dynamical System

The phase-space reconstruction, also known as the dynamical reconstruction of the system, is the approach to extracting the whole original system hidden in some variable time series through a series of concrete algorithms. For time series $x(n)$ and $n = 1, 2, 3, ... N$, if the dimensions embedded m and the delay time τ can be properly chosen and determined, then the phase-space $X(i) = \{x(i), x(i+\tau), ..., x(i + (m-1)\tau)\}$, $i = 1, 2, 3, ... M$, can be reconstructed. According to the phase-space theory brought forward by Packard and Takens, if all the dynamical information needed to determine the status of the system are contained in the time series of any variable of the system, when the time series of a single variable are embedded into a new coordinate system, the attractors can be restored in that dimension embedded as long as the number of the dimension embedded is large enough, which means the status track obtained keeps the primary features of the original dimension track. In the reconstruction of phase-space, the choice of delay time τ and dimensions embedded m affects directly the quality of the reconstruction, and thus the accuracy of the forecast [4].

2.1 The Choice of Delay Time τ

The delay time chosen must be the best one in order to make the phase space reconstructed a good showcase of the dynamical features of the system. In the wind in the forecast actual Motor Assembly state the best delay time is to refactor the system to participate in a non-related, as far as possible to point to the same as in the embedded space is maintained between each component of the dynamical of relationships, and that the information on the attractor as much as possible. The delay time too small, the track of the phase-space may squash to the same position, making the information revelation not remarkable, causing redundant errors. The delay time too large, the changes of dynamical morphs at a certain moment will be too dramatic, making simple geometric patterns complex, distorting the dynamical signals, causing uncorrelated errors. The current approaches to the best delay time include mainly the autocorrelation functions and the mutual functions.

The approach of autocorrelation functions, mainly used to extract the linear correlation of time series, is mature in determining the delay time. Essentially, it is a linear concept suitable to decide the linear correlation. The chaotic system, however, is a non-linear system. Therefore, Fraser and Swinney suggested using the mutual functions to determine the non-linear correlation of a system. The informatics provides a measurement for the non-linear correlation within and between the time series.

For a time sequence $\{x_i, i = 1, 2, 3, ... n\}$ with the given delay time τ, the load time sequence will become $\{x_{i+\tau}, i = 1, 2, 3, ... n\}$. The probability of the appearance of x_k in sequence $\{x_i, i = 1, 2, 3, ... n\}$ is $P(x_k)$. The probability of the appearance of $x_{k+\tau}$ in sequence $\{x_{i+\tau}, i = 1, 2, 3, ... n\}$ is $P(x_{k+\tau})$. The joint probability of the appearance of both x_k and $x_{k+\tau}$ in the two sequences is $P(x_k, x_{k+\tau})$, wherein probability $P(x_k)$ and $P(x_{k+\tau})$ can be obtained via the frequency of their appearance in corresponding time sequence and the joint probability $P(x_k, x_{k+\tau})$ can be obtained by counting the corresponding checkers on plane $(x_k, x_{k+\tau})$. Then the mutual functions will be:

$$I(\tau) = \sum_{k=1}^{N} P(x_k, x_{k+\tau}) \ln \frac{P(x_k, x_{k+\tau})}{P(x_k)\, p(x_{k+\tau})} \quad (1)$$

The value of the delay time τ is determined by the first minimum value of the mutual function.

2.2 The Determination of the Embedding Dimensions

Selection of embedding dimensions generally involves the G-P method of saturated correlation dimension, the false nearest neighbor algorithm and the Cao algorithm.

The false nearest neighbor (FNN) algorithm, once regarded as one of the most effective ways, was later found to have defects such as sensitivity to noise in the signal, fluctuation, instead of tedious variation, of the number of false nearest points subject to the impact of noise and fluctuations. In addition, this method needs two parameters set manually in practice, which brings it a strong subjectivity. The calculation accuracy of the G-P method was easily subject to influences of data length, noise and other factors. An improved FNN algorithm (the Cao method) brought forward by Cao Liangyue aims mainly at overcoming the shortcoming of threshold settinging in the FNN algorithm. This method has the following advantages: the calculation needs only one parameter, the delay time τ; the embedding dimension can be obtained with a relatively small amount of data; the capability of distinguishing chaos time series and random time series; the suitability for analysis of high-dimensional time series; the relatively high efficiency in calculation.

For a time sequence $\{x_i, i = 1, 2, 3, ... N\}$, where n is the length of the sequence, Reconstruct the d-dimension and $(d + l)$ dimension phase-space. X_j^{d+1} is a phase point with ordinal number i, $X_j^d(j = 1, 2, 3, ... k\}$ is the neareast neighbor point of X_j^d, $\| \bullet \|$ is the Euclidean distance, calculate $a(i, d) = \frac{\| x_i^{(d+1)} - x_j^{(d+1)} \|}{\| x_i^{(d)} - x_j^{(d)} \|}$, $E(d)$ is the mean value of $a(i, d)$, then $E(d) = \frac{1}{n - d\tau} \sum_{i=1}^{n-d\tau} a(i, d)$, thus the change $E_1(d) = \frac{E(d+1)}{E(d)}$ from m-dimension to $(m+1)$ dimension.

If a wind power time sequence is a chaotic time sequence, then $E_1(d)$ will go towards saturation with the increase of *d*. Therefore, a time sequence can be determined as to whether it is a chaotic time sequence according to the fact that whether $E_1(d)$ will go towards saturation with the increase of *d*. Here $(d+1)$ is the best embedding dimension wanted. The degree of fluctuation of $E_2(m)$ can reflect the degree of randomness and determinacy of the signals. The bigger the amplitude of the noise, the smaller will the fluctuation be, and the stronger the randomness of the sequence. For a time sequence with noise pollutions, the G-P method was not satisfactory. Considering the fact that the wind power short-term load is inevitably affected by noise pollution (sudden changes of wind speed, errors in the data of SCADA records, etc.), this paper uses the Cao method to select the best embedding dimension.

3. The Neural Network

3.1 Structures of the Neural Networks of BP and Elman

The BP (Back Propagation) network brought forward in 1986 by the science team headed by Rumelhart and McCelland, is a multi-layer feed-forward network trained by error back propagation algorithm and currently one of the most widely used neural network models. By continuously adjusting the weight values and threshold values of the network via the back-propagation, BP neural network minimizes the square sum of error. Figure 1 is a topology of a BP neural network models. The Elman neural network brought forward by Jeffrey L Elman in 1990 is a recurrent neural network which has a better computing power than a feed-forward neural network. In addition to units of the input layer, the hidden layer and the output layer, a basic Elman neural network has a special access layer, also known as the context unit, or the state layer. The input layer unit only serves as a signal conveyor. The output layer plays a role of linear weighted sum. The hidden layer unit may involve linear or non-linear functions. The unit of access layer receives feedback signals from the hidden layer and memorizes the output value of the neurons in hidden layer at a former moment. The output of the neuron of the access layer, after delay and storage, is input again into the hidden layer. This makes it sensitive to historical data, and increases the network's ability to deal with dynamic information, which is helpful to the modeling of the dynamic process. The structure of a standard Elman network is shown in Figure 2.

The mathematical model of the Elman neural network is as follows:

$$\begin{cases} x(k) = f(w^1 x_c(k) + w^2 u(k-1)) \\ x_c(k) = a\, x_c(k-1) + x(k-1) \\ y(k) = g(w^3 x(k)) \end{cases} \tag{2}$$

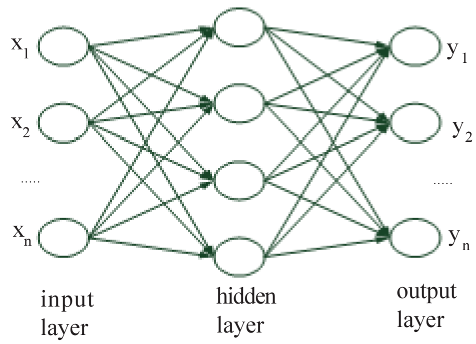


Figure 1. BP neural network model

Wherein w^1 and w^2 and w^3 are the respective link weight matrix from the structural unit to the hidden layer, from the input layer to the hidden layer and from the hidden layer to the output layer. Function f and function g are the non-linear vector functions composed of the excitation functions of the output layer unit and the hidden layer. $x_c(k)$ stands for the output of the access layer unit and the hidden layer unit. $y(k)$ stands for the output of the output unit. $0 \leq \alpha < 1$ is the auto-link feedback gain factor.

3.2 Characteristics of the Elman Neural Network

Currently, most of the forecasts and research of generated power of wind are based on a static feed-forward neural network of the BP algorithm. To identify a dynamic system using a static feed-forward network actually turns dynamic time modeling into static space modeling, while selecting the order of the model structure, especially when the increase of the system orders or unknown orders, the dramatic expanded network structure slows down the convergence rate of network learning as well as causing excessive network input knots, training difficulties, sensitivity to exterior noises, etc. In contrast, the Elman neural network is able to reflect more vividly and directly the dynamic characteristics of the system, representing the direction of development of neural network modeling, identification and control. The Elman neural network is a typical dynamic recurrent neural network. Based on the BP network infrastructure, it gives the network the function of mapping dynamic characteristics by storing internal state, which allows the system to adapt to time variant characteristics. The Elman neural network can learn either the space-domain model or the time-domain mode, and is also able to give non-linear and dynamic properties to trained networks while avoiding the shortcomings of traditional neural network, such as not being able to change in real time the structure of a model or lacking adaptability to sudden changes in the future. Compared with feed-forward neural networks (like BP neural network), the Elman neural network has links of feed-forward and feed-back, overcoming the shortcomings of feed-forward neural networks such as low speed in net convergence, frequent subjection to local minimum, incapability of dynamic learning, etc.

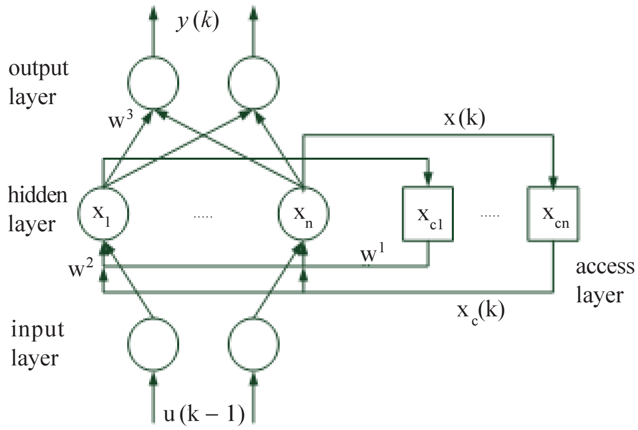


Figure 2. Elman neural network model

4. The Forecasting Model of Wind Power

This paper establishes forecasting model by combining phase-space reconstruction technology with neural networks of Elman, BP and RBF. The steps are as follows: Collect the historical data of a state variable (such as amount of the wind turbine electricity generation) of the forecast system and sort them out as required by time sequences. Set the sorted time sequence to be: $\{x(t), t = 1, 2, 3, \ldots N\}$

Determine the best delay time and the embedding dimension m;

Use the best delay time and the embedding dimension to carry out the phase-space reconstruction of the original time series to obtain the new phase-space vector:

$Y(t) = (x(t), x(t+\tau), x(t+2\tau), \ldots, x(t+(m-1)\tau))$, wherein $t = 1, 2, 3, \ldots N-(m-1)\tau$;

(4) Build the Elman neural network prediction model. The neural network input dimension equals the embedding dimension m. The time before each of the input data differs with a point of τ. Take $(x(t), x(t+\tau), x(t+2\tau), \ldots, x(t+(m-1)\tau))$ as the input of neural network. The middle layer is a single layer. The amount of the neurons in the middle layer can be determined by trial calculation or $2n+1$ (nstands for the number of neurons of the input layer). The output layer contains a neuron and its output is forecast value of the time point to be predicted, or $x(t+(m-1)\tau)$);

(5) Train the network until it meets the requirements;

(6) Select test samples. If they meet the requirements, go to step (7) to forecast. If a large error appears, return to step (5) for re-training, or step (4) to re-design the network structure;

(7) Select the time point of the forecast and apply the established model to forecast.

5. Case Analysis

5.1 Forecast Evaluation Indicators

The most frequently used wind power forecast error indicators are root mean square error and percentage error.

Root mean square error is defined as follows:

$$e_{RMSE} = \sqrt{\frac{\sum_{t=1}^{N}(P_t' - P_t)^2}{N}} \tag{3}$$

e_{RMSE} stands respectively for the root mean square error of the power forecast of a wind farm. P_t' and P_t stand respectively for the predicted value and the acutual value of the power. N stands for the number of predicted values.

5.2 Wind Power Forecast Experiment

Select the former 4641 data out of the 4921 active data collected from a generating unit in a wind farm from 0:00, Feb. 1st to 23:00, March 31 as training samples, and the latter 280 data as test samples to predict the wind power generated in 48 hours. Figure 3 presents the time series of wind power during the period, from which it can be easily found that the wind power sequence shows a strong non-linear with no apparent pattern of changes.

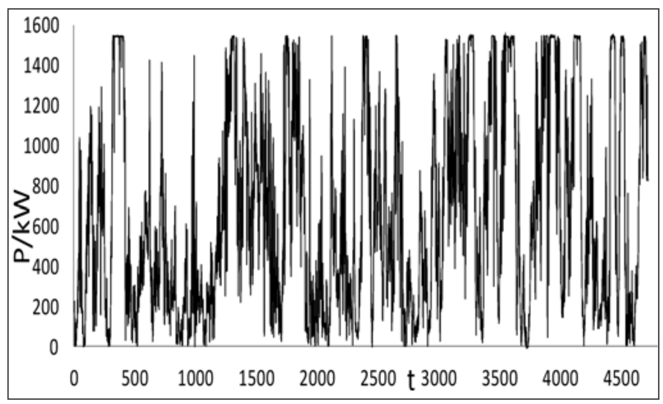


Figure 3. The wind power time series

The calculating result of the best delay time of the wind power time series of this period using the mutual functions is shown in Table 1. As it can be seen in Figure 4, the obtained best delay time is $\tau = 9$.

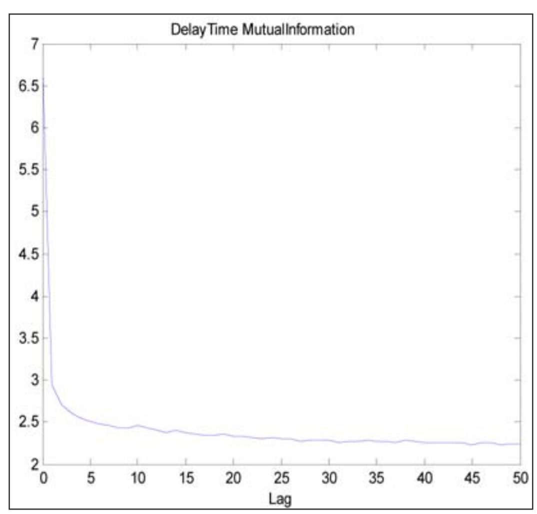


Figure 4. Compute delay time using mutual function

τ	Autocorrelation	τ	Autocorrelation
0	6.6868	10	2.4603
1	2.9372	11	2.4267
2	2.7145	12	2.4042
3	2.6089	13	2.3744
4	2.5451	14	2.3962
5	2.502	15	2.3792
6	2.4796	16	2.355
7	2.4607	17	2.3474
8	2.4361	18	2.3392
9	2.4339	19	2.3516

Table.1 Rresults of Mutualinformation Method

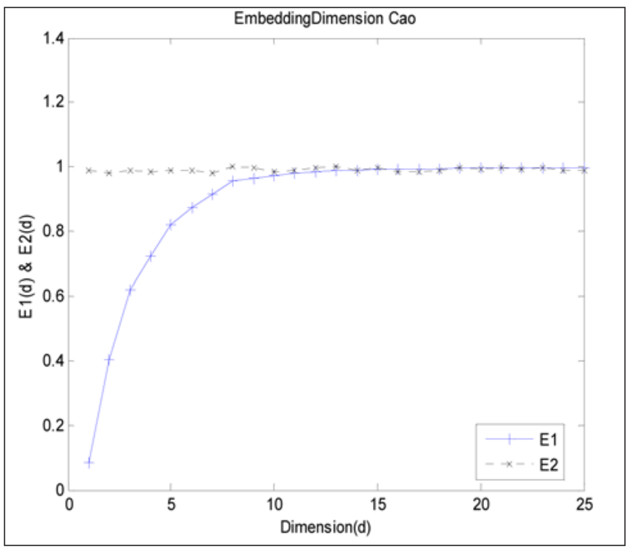


Figure 5. Compute optimal embedding information using Cao method

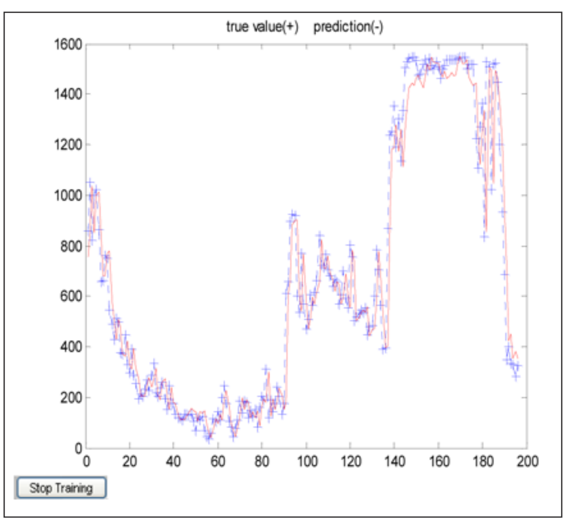


Figure 6. Power prediction curve of the chaos-Elman neural network

For the same time series data, the best embedding dimension obtained by using the Cao method is shown in Figure 5. The value of the embedding dimension of time series can be obtained as the value of $E_1(m)$ goes towards 1 with the increase of embedding dimensions. As can be seen in the figure: the calculation result of the optimal embedding dimension is 10, i.e. $M = 10$.

Based on the analysis above, the number of input nodes of the neural network is determined as the value of the embedding dimension, i.e., 10. Conduct the single-step forecast and predict the power within the future 48 hours, or 192 points, and the forecast results are shown in Figure 6. With a line graph, Figure 7 shows more vividly the comparison of the values obtained via the four forecasting methods. As can be seen in Figure 7, the error of the prediction based on chaos-Elma neural network is relatively small while its forecast accuracy and stability is remarkably better than those based on chaos-BP.

Prediction algorithms	Root mean square error/ %
Chaos-Elman	17. 62
Chaos-BP	18. 37

Table 2. Comparison of the root mean square error of various chaos-based predictions

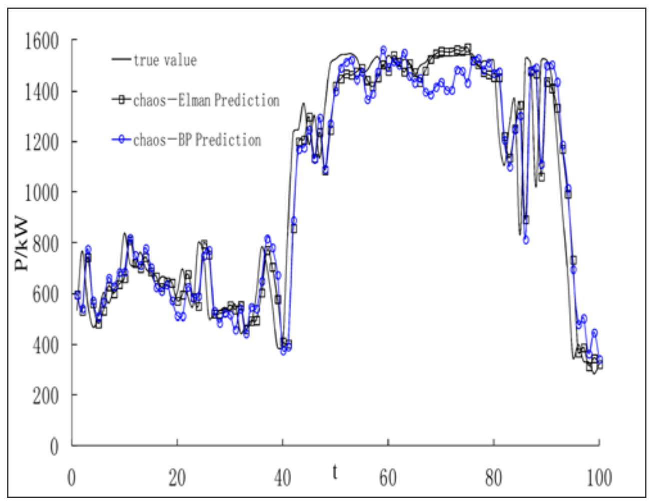


Figure 7. Comparisons of calculation results of various prediction algorithms

5.3 Analysis of Lab Conclusions

Concluded here are a number of phenomena and problems are found in trainings of various networks and prediction models and their combinations, which may serve as basis for future researches.

Comparison of the training results between chaos-Elman and chaos-BP shows that, under the circumstance of the same amount of data, the same training time intervals and training frequencies, the network models of chaos-BP can go into the training of higher-precision earlier, but the time taken for training and prediction is at leat 5 times longer than that of chaos-Elman. On occasions of abrupt changes (sudden increase/decrease in wind speed/ power), the relative prediction error of the chaos-Elman network is much smaller than that of chaos-BP. This also indicates that compared with feed-forward neural networks, the Elman neural network has more adaptability to abrupt changes as well as a stronger capability in dynamic learning.

6. Conclusions

(1) As neural network technology is still in development and new models and algorithms are being brought forward, there is much to explore in the selection of the most effective model and algorithm.

(2) In the system of prediction, for the strong randomness and non-linear of data, the size of the deviation in the selection of the prediction reference point will directly affect the predicted results. The selection of prediction reference point in this paper is based on the Euclidean distance. The prediction accuracy is relatively high with low embedding dimension factors but not ideal with high embedding dimension factors. In future studies, therefore, approaches capable of presenting better the correlation of the evolution tracks of phase points could be considered.

References:

[1] Yongqian, Liu., Shuang, Han., Yongsheng, Hu. (2007). Review on Short-term Wind Power Prediction, *Modern Electricity,* 24 (5) 6-11.

[2] Junfeng, Li., Pengting, Shi., Gaohu. (2010). Report on China Wind Power Development 2010. Haikou: Hainan Press, 6-10.

[3] Hassan, G. (2006). Forecasting short term wind farm production [EB/O], [2006-06-01]. http://www.garradhassan .com.

[4] Jinhu, Lv . (2002). Analysis and Application of Chaos Time Series . Wuhan: Wuhan University Press.

[5] Zhiling, Yang., Yongqian, Liu. (2011). Short-Term Wind Power Prediction With Particle Swarm Optimization. *Grid Technology*, (05).

[6] Shuang, Huan. (2008). A Study Shor-term Wind Power Forecasting Approaches. Beijing: North China Electric Power University.

[7] Bernhard, L., Kurt, R., Bernhard, E. et al. (2006). Wind power prediction in Germany—recent advances and future challenges. European Wind Energy Conference, Athens.

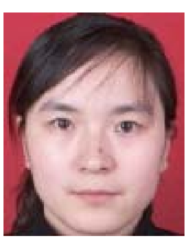
[8] Hatziargyriou, N., Contaxis, G., Matos, M. et al. (2001). More care-advice for secure operation of isolated power systems with increased renewable energy penetration and storage. European Wind Energy Conference, Copenhagen.

[9] Lei, Dong., Shuang, Gao., Xiaozhong, Liao. (2007). etc. Chaos Characteristic Analysis on The Time Series of Wind Power Generation Capacity, *Solar Energy*, 28 (11) 1290 - 1294.

[10] Tao, Lv., Wei, Tang., Li, Suo. (2010). Prediction of short-term wind speed in wind farm based on chaotic phase space reconstruction theory. *Power System Protection and Control*, 38 (21) 113-117.

[11] Jia, Tao., Hong, Zhang., Guorong, Zhu. (2011). etc. Wind Power Prediction Based on Technology of Advanced Phase Space Reconstruction. *Electrical Engineering Journal*, China, 31 (28) 9-14.

Author Biography



Jin-hua Zhang was born in July 1980 in Xun County, Henan Province. In July 2008, she received a master's degree in water resources and hydropower engineering in North China Institute of Water Resources and Hydroelectric Power. From September 2010 to the present, she is reading for her doctorate in renewable energy and clean energy in North China Electric Power University. Her research direction is wind power generation.

At present she works in North China Institute of Water Resources and Hydroelectric Power and her main research directions include the optimal dispatch of wind farm output and the wind power prediction technology.

Jin-hua Zhang, is a doctoral candidate and also involved in teaching.