Short-term wind speed forecasting method based on wavelet packet decomposition and improved Elman neural network

Ruili Ye, Zhizhong Guo, Ruiye Liu
Department of Electrical Engineering
Harbin Institute of Technology
Harbin, China
yeruili_hit@163.com

Jiannan Liu
State Grid AC Engineering Construction Company
Beijing, China;
18911267942@163.com

***Abstract*—Accurate prediction of wind speed is of great significance to the operation and maintenance of wind farms, the optimal scheduling of turbines and the safe and stable operation of power grids. This paper puts forward a new method for short-term wind speed forecasting based on the wavelet packet decomposition(WPD) theory and an improved Elman neural network(ENN), and the concrete application steps of the method is given.WPD theory is firstly adopted to decompose wind speed data into several wavelet spaces, and according to the correlation, the optimal decomposition tree will be persisted and random data will be rejected. Then a new particle swarm optimization (PSO) training algorithm with disturbance is proposed to improve the training speed of neural networks and deal with the drawback of PSO's easily falling into local optimal solution. Finally, ENNs with different structures are established and used to find the laws of wind speed in different frequency bands, prediction results are hence received. The correctness and effectiveness of the proposed method is verified by wind speed data of a wind farm in south China.**

Keywords—wind power, wind farm, short-term wind speed forecasting, wavelet packet decomposition(WPD), Elman neural network(ENN), particle swarm optimization(PSO)

I. INTRODUCTION

For the past few years, wind power industry has developed vigorously worldwide. According to the latest statistical report[1] of the Global Wind Energy Council (GWEC) released in Feb. 2nd 2016, 2015 set a new record of more than 63.01 GW installed in a single year, bringing the global total up to 432.42 GW. Given the high environment sensitivity and low predictability, large-scale integration of wind power accompanies great influence on power system and big challenges to grid dispatching operation[2-3]. It is of great significance to increase wind power forecast accuracy so as to ensure the safe operation of power system[4-5].

The accuracy of wind power prediction may be affected by a variety of factors, among which the precision of wind speed forecasting is a key one. Wind speed forecasting not only plays a decisive effect on wind power prediction, but also is of great significance to safe and stable operation of wind farms and the grids[6]. When predicting wind power of a wind farm, a common and effective method is to forecast wind speed of the wind farm firstly, and then find the predicted value of wind power according to wind power curve[7-8].

Wind speed forecasting methods can be divided into two types: methods based on physical models and methods based on statistical models. The former needs to establish physical models of wind farms together with the surrounding environments and the terrain, roughness and other information of wind farms will all be taken in consideration when conducting prediction via physical equations. The latter methods, however, usually carry on statistical analysis on historical data and measured data of wind farms and then identify the inherent laws and mapping relationship between system inputs and wind speed to conduct wind speed prediction.

In short-term and ultra short-term forecasting, statistical models get more applications as they can automatically match the location and environment of wind farms. Relevant researches mainly focus on time series[9-11], Kalman filtering[10-13], support vector machine[14-15], wavelet analysis[16-17] and artificial neural networks[18-19], etc.

In carrying out wind speed forecasting based on statistical models, many scholars usually process data firstly in order to effectively improve the prediction accuracy and help find the laws of wind speed[20].For example, Luo et al.[12] found the the basic parameters of wind speed by time series method, and took them as model inputs. ZHANG et al.[21] made use of the empirical mode decomposition theory to decompose wind speed data into a trend component and a random component. And the authors of [22-23] proposed to conduct pre-forecasting firstly and then modify the pre-forecasting values so as to improve the prediction precision.

Wavelet packet decomposition(WPD)[24] is a more sophisticated decomposition method than wavelet analysis and empirical mode decomposition, as it can decompose both the high frequency parts and low frequency parts of a complex signal into different frequency bands and adaptively select the appropriate frequency band to match the signal frequency spectrum according to the features of the signal and analytical requirements. Therefore, in order to effectively improve the accuracy of wind speed data to be input, this paper adopts WPD theory to decompose wind speed data into several wavelet spaces and scale spaces, and

978-1-5090-1970-0/16/$31.00 ©2016 IEEE PMAPS 2016

persists the optimal decomposition tree according to relevance.

Elman neural network(ENN)[25] is a recurrent neural network, and can approximate any function within a limited period of time.ENN has a faster training speed than BP network, and has a better dynamic memory than the traditional static feed-forward neural network (FNN). Therefore this paper chooses to adopt ENN model to conduct the prediction. In order to further improve the training speed of neural networks and deal with the drawback of PSO easily falling into local optimal solution, this paper puts forward a new PSO training algorithm with disturbance, and then proposes an improved ENN model, namely the ENN model based on the PSO training algorithm with disturbance, to predict the wind speed.

II. WAVELET PACKET DECOMPOSITION THEORY

WPD is generated and gets developed on the basis of wavelet transformation. Wavelet transformation is a kind of localized time-frequency analysis method with changeable time and frequency windows, and is very suitable for processing non-stationary transient signals. In the decomposition process, low-frequency signals will be decomposed, while high-frequency signals will not, which makes the frequency resolution of the method decrease with the increase of frequency. Thus wavelet transformation has a poor performance in processing slowly-changed signals.

Both low-frequency and high-frequency signals will be decomposed when adopting WPD, and according to the characteristics of the signal and analysis requirements, WPD can help adaptively select the appropriate frequency band to match the signal spectrum. Hence, WPD has a wider application in dealing with complex slowly-changed signals, as it can help to grasp the detail characteristics of the signals and improve the temporal resolution of the signals.

The basic principle of WPD is as follows:

Assume $\{V_j; j \in Z\}$ is a series of closed sub-spaces of space $L^2(R)$. If $\{V_j; j \in Z\}$ generates an orthonormal multiresolution analysis of $L^2(R)$ with function $\varphi(t)$, then there exists a series of low-pass filter coefficients $\{h_n; n \in Z\}$ to meet expression (1).

$$\varphi(t)=\sqrt{2}\sum_{n\in Z}h_n\varphi(2t-n) \tag{1}$$

Set $g_n=(-1)^{n-1}\overline{h}_{1-n}$, $n \in Z$, where g_n are high-pass filter coefficients. Construct a orthogonal wavelet function shown as expression (2).

$$\psi(t)=\sqrt{2}\sum_{n\in Z}g_n\varphi(2t-n) \tag{2}$$

For $\forall j \in Z$, the two function families $\{2^{\frac{j}{2}}\psi(2^j t-n); n \in Z\}$ and $\{2^{\frac{j}{2}}\varphi(2^j t-n); n \in Z\}$ constructed by function $\varphi(t)$ and function $\psi(t)$ are normal orthogonal systems in space $L^2(R)$. Based on the two normal orthogonal systems, we can obtain two function space series $W_j = Closespan\left\{2^{\frac{j}{2}}\psi(2^j t-n); n \in Z\right\}$ and $V_j = Closespan\left\{2^{\frac{j}{2}}\varphi(2^j t-n); n \in Z\right\}$.

To improve the high frequency resolution of multi-resolution analysis, W_j and V_j need further decomposition, namely to express W_j and V_j using spaces N_j^m.

Set $\mu_0(t)=\varphi(t), \mu_1(t)=\psi(t)$,

$\mathrm{H}_0(\omega)=\frac{1}{\sqrt{2}}\sum_{n\in Z}h_n e^{-in\omega}$, $\mathrm{H}_1(\omega)=\frac{1}{\sqrt{2}}\sum_{n\in Z}g_n e^{-in\omega}$,

according to scale equation (1) and signature equation (2), the Fourier transformation of $\mu_0(t)$ and $\mu_1(t)$ can be given by expressions (3) and (4).

$$\mathrm{N}_0(\omega)=\mathrm{H}_0\left(\frac{\omega}{2}\right)\mathrm{N}_0\left(\frac{\omega}{2}\right) \tag{3}$$

$$\mathrm{N}_1(\omega)=\mathrm{H}_1\left(\frac{\omega}{2}\right)\mathrm{N}_0\left(\frac{\omega}{2}\right) \tag{4}$$

The wavelet packets determined by scale equation $\varphi(t)$ can be defined as the function series $\{\mu_m(t); m=0,1,2,...\}$,

$$\mu_{2m}(t)=\sqrt{2}\sum_{n\in Z}h_n\mu_m(2t-n) \tag{5}$$

$$\mu_{2m+1}(t)=\sqrt{2}\sum_{n\in Z}g_n\mu_m(2t-n) \tag{6}$$

Obviously, for any non-negative integer m ,the Fourier transformation $\mathrm{N}_{2m+l}(\omega)$ of $\mu_{2m+l}(t)$ can be written as follows:

$$\mathrm{N}_{2m+l}(\omega)=\mathrm{H}_l\left(\frac{\omega}{2}\right)\mathrm{N}_m\left(\frac{\omega}{2}\right),\quad l=0,1 \tag{7}$$

For $\forall m \in N$, expression (9) gives the binary form of m .

$$m=\sum_{l=0}^{+\infty}\varepsilon_l\times 2^l \tag{8}$$

Where $\varepsilon_l \in \{0,1\}, l \geq 0$. Thus the Fourier transformation $\mathrm{N}_m(\omega)$ of $\mu_m(t)$ can be expressed as follows:

$$\mathrm{N}_m(\omega)=\prod_{l=0}^{+\infty}\mathrm{H}_{\varepsilon_1}\left(2^{-(l+1)}\omega\right) \tag{9}$$

Fig.1 gives the schematic diagram[18] of WPD with three layers. For wind speed forecast, S is the original wind speed.

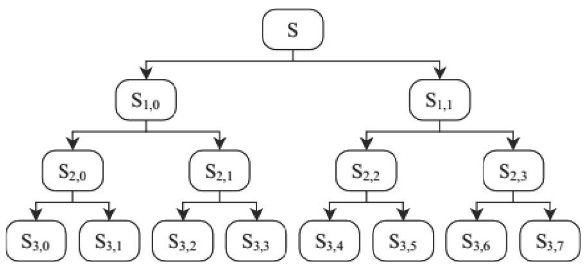


Fig. 1. Schematic diagram of WPD with three layers

From Fig.1, we can find that WPD is essentially obtained by the further decomposition of the high frequency part of wavelet decomposition.And the final decomposition result is that signal S will be mapped to 2^i wavelet packet sub-spaces, where i is the number of decomposition layers.

III. Improved Elman neural network model based on Particle Swarm Optimization training algorithm with disturbance

A. Elman neural network principle

ENN is a typical dynamic neural network and is based on the mechanism of FNN. By storing internal status of the neural network, ENN gets the function of mapping dynamic features and helps the system to obtain the ability of adapting to time-varying characteristics.

Fig.2 shows the structure of ENN,which usually has four layers, namely the input layer, the hidden layer,the correlation layer and the output layer. The connections among the input layer, the hidden layer and the output layer of ENN are similar with that of FNN.The correlation layer is used to record the output values of the hidden layer of the previous time and form an internal feedback connection together with the hidden layer, which makes ENN has better dynamic memory property than FNN.ENN can be used to approximate any function in a limited time and is more suitable to establish the forecast model of time series.

The space expression of ENN can be expressed as follows:

$$\begin{cases} p(k) = g(w_3 x(k)) \\ x(k) = f(w_1 y(k) + w_2 u(k-1)) \\ y(k) = x(k-1) \end{cases} \tag{10}$$

Where, $g(\cdot)$ is the transfer function of input layer. $f(\cdot)$ is the transfer function of hidden layer. k is the training times. u is the input vector, p is the output vector, x is the output vector of hidden layer neurons, y is the feedback state vector. Besides, w_1, w_2, w_3 are separately the weight coefficients of correlation layer to hidden layer,input layer to hidden layer,and hidden layer to output layer.

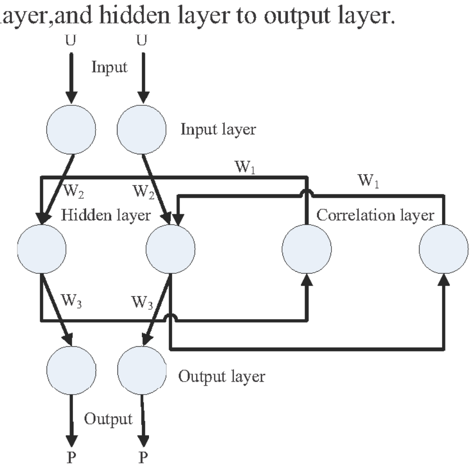


Fig. 2. Structure of ENN

B. Particle swarm optimization training algorithm with disturbance

The common training algorithms of ENN usually include the back-propagation algorithm (BP algorithm), PSO algorithm, etc. In this paper, the PSO algorithm is chosen to train ENN.

PSO algorithm is an evolutionary intelligence optimization algorithm and its mathematical description is as follows[26]: Suppose in a n-dimensional space, $X = \{x_1, x_2, \cdots, x_m\}$ are m particles and they are used to search the best solution.The position and the speed of the i-th particle can be expressed as $x_i = \{x_{i,1}, x_{i,2}, \cdots, x_{i,n}\}^{\mathrm{T}}$ and $v_i = \{v_{i,1}, v_{i,2}, \cdots, v_{i,n}\}^{\mathrm{T}}$. $p_i = \{p_{i,1}, p_{i,2}, \cdots, p_{i,n}\}^{\mathrm{T}}$ expresses the best individual particle position expression and $p_g = \{p_{g,1}, p_{g,2}, \cdots, p_{g,n}\}^{\mathrm{T}}$ expresses the best global position. Thus ,the speed and position of the t-th generation particles can be updated according to equation (10) and (11).

$$v_{i,d}^{(t+1)} = v_{i,d}^{(t)} + c_1 r_1 (p_{i,d}^{(t)} - x_{i,d}^{(t)}) + c_2 r_2 (p_{g,d}^{(t)} - x_{i,d}^{(t)}) \tag{11}$$

$$x_{i,d}^{(t+1)} = x_{i,d}^{(t)} + v_{i,d}^{(t+1)} \tag{12}$$

where $d = 1, 2, \cdots, n$, $i = 1, 2, \cdots, m$, r_1, r_2 are two random numbers uniformly distributed between [0, 1] and c_1 , c_2 are acceleration constants.

From expression (10) and (11), it can be found that if the global optimal particle goes into the local optimal solution,the other particles will be high likely attracted into the local optimal solution. Therefore, a disturbance is introduced to a dimension d of the global optimal solution to help the global optimal particle jump out of the local optimal solution. First we should randomly select a disturbance dimension d and then introduce the disturbance according to expressions (12) and (13).

$$p_{g,d}^{(t)} = p_{g,d}^{(t)} + (x_d^{\max} - x_d^{\min}) \cdot Gaussian(0, \sigma^2) \tag{13}$$

$$\sigma = \sigma_{\max} - (\sigma_{\max} - \sigma_{\min}) \frac{t}{G} \tag{14}$$

Where G is the total iteration number. $x_d^{\max}$ and $x_d^{\min}$ are the maximum and minimum values of all generation particles occurred in disturbance dimension d .

At the beginning of the proposed algorithm, the speed of particles is high and the algorithm has a stronger global searching ability. In that case the second item in expression (12), namely the disturbance item, is obviously smaller than the first item in expression (12), the influence of the disturbance item to the searching ability of the algorithm is very small and can be ignored. However , in the middle or at the end of the calculation process, the speed of the particles gets smaller and smaller because of the convergence of the particles, the existence of the disturbance item in expression (12) helps to ensure the particle speed not to fall to zero, which can keep the local searching going, and provide the possibility for particles to jump out of the local optimal solution.

IV. Short-term wind speed forecasting method based on wavelet packet decomposition and improved Elman neural network

This paper puts forward a short-term wind speed forecasting method based on WPD theory and improved ENN algorithm(WPD-ENN method). When adopting the method, WPD is firstly used to pre-process the given wind speed data, then PSO training algorithm with disturbance is used to accelerate the training of ENN, finally, the pre-processed results of the given wind speed data will be input to the ENN to forecast wind speed. The concrete steps of WPD-ENN method mainly includes:

1) Decompose the given wind speed data by WPD theory and obtain the optimal decomposition tree.The parts with smaller correlation coefficient can be considered as random fluctuations of wind speed, similar to signal noises, and can be ignored in wind speed forecasting. After removing the random fluctuations, we obtained the wind speed data in the selected frequency band, and they can be used as the training samples of ENN and the historical data for wind speed forecasting.

2) Build different ENNs for each frequency band of wind speed data, and determine the structures of the ENNs and the neuron numbers of different layers.

3) Adopt the proposed PSO training algorithm with disturbance to train the ENNs.

4) Input the selected wind speed data in different frequency bands obtained in step 1) into the trained ENNs to predict wind speed and output the results.

V. Analysis of Examples

Wind speed data of a wind farm in south China are adopted in this paper to verify the validity and efficiency of the proposed method. There are totally 33 wind turbines (600kW) in the wind farm, with the total installed capacity of 19.8MW. Since 1998 when the wind farm was integrated to the grid, meteorological parameters and output power values of the wind farm have been taken sample every 10 minutes.

A. wind speed forecast based on WPD-ENN method

1) Application of wavelet packet decomposition

First, adopt WPD to decompose the historical wind speed data. Figure 3 shows the best WPD tree of the given data, and each space is corresponding to one frequency band. [1,1], [3,0], [3,1], [3,2], [3,3] are the five best WPD spaces, and the wind speed data corresponding to the above five frequency bands are the data that can be used as neural network training samples or the historical data for wind speed forecasting. Fig 4 gives the decomposed wind speed waves in different frequency bands.

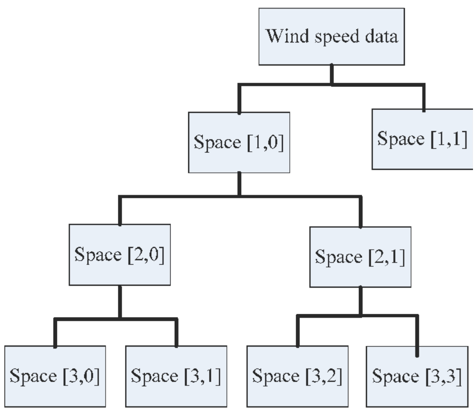


Fig. 3. Optimal WPD decomposition tree of the original wind speed data of the wind farm

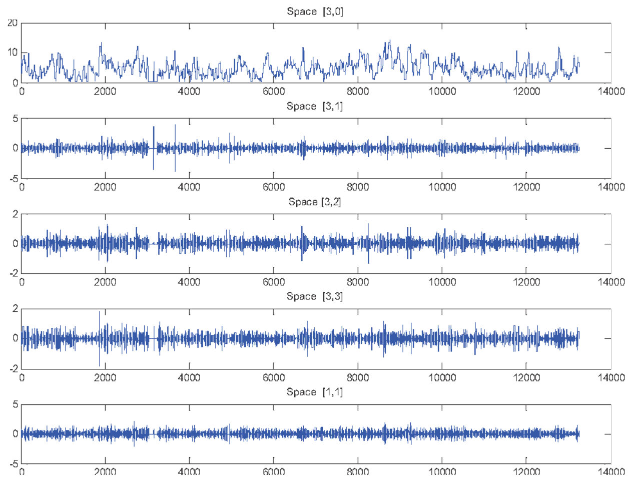


Fig. 4. Wind speed waves in different frequency bands

2) build different ENNs for each frequency bands

Assume that all the neural networks are single hidden layer design. TABLE I gives the number of neurons in input layers, hidden layers and output layers of the established ENNs.

TABLE I. Corresponding Structures of ENNs According to the Wind Speed in Different Frequency Bands

best WPD space	number of neurons in different layers
[3,0]	24-20-1
[3,1]	12-20-1
[3,2]	6-20-1
[3,3]	3-20-1
[1,1]	2-20-1

3) Adopt the proposed PSO algorithm with disturbance to train the ENNs

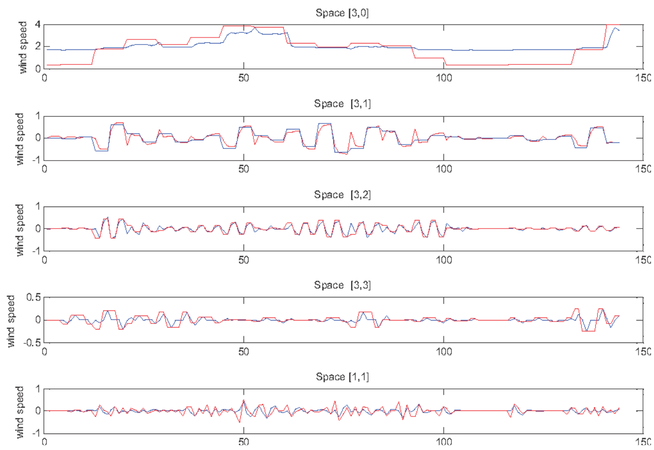
Set the original parameters of the proposed PSO algorithm with disturbance as TABLE II. Using the algorithm to adjust and train the relevant parameters of ENN. When the error objective function reaches the threshold, stop the process and save the deviation and the corresponding network parameters and weights.

TABLE II. Initial Parameters of New PSO with Disturbances

Parameter	value
number of particles (m)	25
iteration number(G)	2000
acceleration constant(c_1)	2.0
acceleration constants(c_2)	2.0

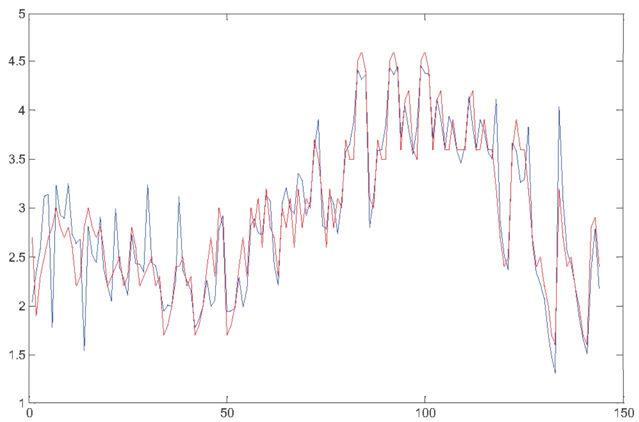
4) Conduct wind speed forecasting

Input the selected wind speed data corresponding to the five best WPD spaces in Fig.4 into the trained ENNs to predict wind speed. Take April 1st, 2011 as the typical prediction day. Fig.5 shows the predicted wind speed curves calculated by the proposed WPD-ENN method corresponding to the five best WPD spaces and the curve of the composite predicted wind speed curve of the five spaces is shown in Fig.6. The real wind speed curve is also given as a comparison in the two figures.



(blue—predicted value; red—real value)

Fig. 5. Curves of predicted value and real value of wind speed corresponding to the five best WPD spaces



(blue—predicted value; red—real value)

Fig. 6. Curves of composite predicted value and real value of wind speed

B. Analysis of forecast error

Wind speed data from March 3rd to April 19th in 2011 of the wind farm are adopted in this paper to further measure the prediction effect. The proposed WPD-ENN method as well as ARMA and traditional ENN are applied to forecast wind speed and the mean absolute prediction error (MAE), which can be calculated by expression (14), is used to evaluate the prediction results.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{|P_{mi} - P_{pi}|}{C_i}\right) \tag{15}$$

TABLE III gives the comparison of MAE results of different methods and Fig.6 is the statistical chart of the forecast error of WPD-ENN method.

TABLE III. MAES OBTAINED BY DIFFERENT PREDICTION METHODS

forecasting method	MAE（m/s）
ARMA method	1.2719
Traditional ENN method	0.8221
WPD-ENN method	0.6181

From Tab.III, Fig.5 and Fig.6, it can be found that the proposed method of this paper is has an ideal effect in predicting wind speed, the predicted and measured value curves have a high degree of coincidence and can properly reflect the law of wind. The MAE of WPD-ENN method is 0.6181m/s，which is obviously better than that of ARMA and traditional ENN methods.

VI. CONCLUSIONS

In this paper,WPD theory is introduced to process the original wind speed data. WPD theory can help reasonably eliminate random fluctuations in wind speed data to improve the accuracy and highlight valuable details of the raw data. The example results show that the selected data processed by WPD theory reflect the laws of wind speed much better than the original data.

It is of better pertinence to apply ENNs with different structures to find the laws of wind speed in different frequency bands in this paper.Besides, the use of the PSO algorithm with disturbance in the training of ENN can indeed accelerate the training process and can effectively solve the problem of the optimal particle trapped in the local optimal solution.

This paper puts forward an short-term wind speed forecasting method based on WPD theory and improved ENN algorithm and the concrete steps are also presented.The example results show that the predicted values of the proposed approach are close to the real values，and the MAE of the approach is significantly smaller than traditional ENN and ARMA methods, which verifies the correctness and effectiveness of the proposed method.

REFERENCES

[1] Global Wind Power Statistics 2015.Available online: http://120.52.72.30/www.gwec.net/c3pr90ntcsf0/ wp-content/uploads/vip/GWEC-PRstats-2015_LR_corrected.pdf/ (accessed on 10 Feb 2016).

[2] A. COSTA, A. CRESPO and J. NAVARRO, “A review on the young history of the wind power short-term prediction,” Renewable and Sustainable Energy Reviews, vol. 12, pp. 1725-1744, 2008.

[3] XUE Yusheng, CAI Bin, JAMES G, et al. Primary energy congestion of power systems[J]. Journal of Modern Power Systems and Clean Energy, 2014, 2(1): 39-49.

[4] Y. Ruili, G. Zhizhong and L. Ruiye, “A Method for Designing Optimal Energy Storage System Based on Analysis of Wind Power Forecast Error,” Automation of Electric Power Systems, vol. 38, pp. 28-34, 2014.

[5] ZHANG Chunyu, DING Yi, NORDENTOFT NC, et al. FLECH: a Danish market solution for DSO congestion management through DER flexibility services[J]. Journal of Modern Power Systems and Clean Energy, 2014, 2(2): 126-133.

[6] LI Yuanzheng, LI Mengshi, WU Qinghua. Optimal reactive power dispatch with wind power integrated using group search optimizer with intraspecific competition and lévywalk[J]. Journal of Modern Power Systems and Clean Energy, 2014, 2(4): 308-318.

[7] DING Yi, CHENG Lin, ZHANG Yonghong, et al. Operational reliability evaluation of restructured power systems with wind power penetration utilizing reliability network equivalent and time-sequential simulation approaches[J]. Journal of Modern Power Systems and Clean Energy, 2014, 2(4): 329-340.

[8] ZHANG Hong, ZHAO Dongmei, GU Chenghong, et al. Economic optimization of smart distribution networks considering real-time pricing[J]. Journal of Modern Power Systems and Clean Energy, 2014, 2(4): 350-356.

[9] B. BROWN, W. KATZ R and H. MURPHY, "Time series models to simulate and forecast wind speed and wind power," Journal of Climate and Applied Meteorology, vol. 23, pp. 1184-1195, 1984.

[10] P. Difu, L. Hui, and L. Yanfei, "Optimization algorithm of short-term multi-step wind speed forecast," Proceedings of the CSEE, vol. 28, pp. 87-91, 2008.

[11] P. Difu, L. Hui, and L. Yanfei, "A wind speed forecasting optimization model for wind farms based on time series analysis and Kalman filter algorithm," Power System Technology, vol. 32, pp. 82-86, 2008.

[12] L. Wen and W. Lina, "Short-term wind speed forecasting for wind farm," Transactions of China Electrotechnical Society, vol. 26, pp. 68-74, 2011.

[13] X. Chunbo, R. Xiao, and L. Yanqing, "Short-Term Prediction Method of Wind Speed Series Based on Kalman Filtering Fusion," Transactions of China Electrotechnical Society, vol. 02, pp. 253-259, 2014.

[14] M.A. Mohandes, T.O. Halawani, and S. Rehman, "Support Vector Machines for wind speed prediction," Renewable Energy, vol. 29, pp. 939–947, 2004.

[15] Ming YANG,You LIN,Simeng ZHU,Xueshan HAN,Hongtao WANG.Multi-dimensional scenario forecast for generation of multiple wind farms[J]. Journal of Modern Power Systems and Clean Energy, 2015, 3(3):361-370.

[16] W. Dongfeng, W. Fuqiang, and N. Chenglin, "Analysis of wavelet decomposition for multi-step prediction of short-term wind speed," Power System Protection and Control, vol. 08, pp. 82-89, 2014.

[17] S. CATALAOJP, H. POUSINHO and M. MENDS, "Hybrid wavelet-PSO-ANFIS approach for short-term wind power forecasting in Portugal," IEEE Trans on Sustainable Energy, vol. 2, pp. 50-59, 2011.

[18] K. METHAPRAYOON and Y. TANAPONG, "An integration of ANN wind power estimation into unit commitment considering the forecasting uncertainty," IEEE Trans on Industry Applications, vol. 43, pp. 1441-1448, 2007.

[19] M. Meiqin, Z. Songlin, and S. Jianhui, "Short-term windpower forecast based on ridgelet neural network," Automation of Electric Power Systems, vol. 35, pp. 70-74, 2011.

[20] M. FOLEY, P. LEAHY and A. MARVUGLIA, "Current methods and advances in forecasting of wind power generation," Renewable Energy, vol. 37, pp. 1-8, 2012.

[21] Z. Yiyang, L. Jiping, and M. Yangyang, "Wind power short-term forecasting based on empirical mode decomposition and chaotic phase space reconstruction," Automation of Electric Power Systems, vol. 36, pp. 24-28, 2012.

[22] Y. Tiejiang, C. Qin, and L. Yiyan, "Short-term wind power output forecasting model for economic dispatch of power system incorporating large-scale wind farm," Proceedings of the CSEE, vol. 30, pp. 23-27, 2010.

[23] W. Songyan, and Y. Jilai, "Joint conditions probability forecast method for wind speed and wind power," Proceedings of the CSEE, vol. 31, pp. 7-15, 2011.

[24] H. Xiaojuan, C. Yueyan, and Z. Hao, "Application of hybrid energy storage technology based on wavelet packet decomposition in smoothing the fluctuation of wind power," .Proceeding of the CSEE, vol. 33, pp. 8-13, 2013.

[25] Y. Dongyang, S. Yifa, and J. Mingjie, "Short-term wind speed forecasting using Elman neural network based on rough set theory and principal components analysis," Power System Protection and Control, vol. 42, pp. 46-52, 2014.

[26] L. Yanmin. Research of Partical Swarm Opitimization and its appplication," Shandong Normal University, 2011.