Sensors & Transducers, Vol. 161, Issue 12, December 2013, pp. 241-246

Sensors & Transducers

http://www.sensorsportal.com

Study on Wind Power Forecasting Based on ISODATA

[1] Gu Bo, [2] An Chao, [1] Ren Yan, [1] Li Yanpin

[1] School of electric power, North China University of Water Conservancy and Electric Power, Zhengzhou, Henan, 450011, China

[2] Yellow River Engineering Consulting Co., Ltd, Zhengzhou, Henan, 450003, China

Tel.: 13592597463

E-mail:gb1982@ncwu.edu.cn

Received: 17 August 2013 /Accepted: 22 November 2013 /Published: 30 December 2013

Abstract: The short-term prediction of wind power generation is of great significance to the security and stability of grid-connected wind power system. the principle and calculation steps of iterative self-organizing data analysis (ISODATA) is presented, which is used for clustering numerical weather prediction data (NWP) in history, in order to gather weather prediction data with numerical similarity to a class. According to the numerical weather prediction data, a BP neural network model has been constructed, to train the BP neural network and to predict wind power, with both clustered and raw data. The predicted results show that, the predictive accuracy of the BP neural network model which employs the clustering algorithm of ISODATA is better than that of itself. *Copyright © 2013 IFSA.*

Keywords: Wind power prediction, Numerical Weather Prediction (NWP), Iterative Self-organizing Data Analysis (ISODATA), BP neural network.

1. Introduction

Energy is the material foundation to support the progress of human civilization, and is also an integral part for the development of modern society. In China, in the process of modernization and common prosperity, energy resources are an extremely strategic issue. In-depth developing renewable energy is key strategic measures to promote clean and multi-energy development and thus fostering emerging industries. It is also urgently required to protect our environment, to respond to climate change and to realize sustainable development. Through unswerving efforts in developing renewable energy sources, China makes great endeavor to increase non-fossil fuel shares in primary energy consumption to 11.4 % and installed generating capacity 30 %, respectively, by the end of the 12th Five-Year Plan [1].

Wind power as the most developing value of renewable energy in the world, has been widely developed and utilized. However, due to random characteristics of the wind speed, wind power has a lot of volatility and uncertainty, which exerts an impact on stable operations on power system, thus limiting wind power integration. According to the latest statistic report of the National Energy Bureau, installed capacity of wind power accounted for about 30 % of the total energy in Inner Mongolia, the share of wind turbine capacity connected to power grid, however, is less than 2 % of that. In order to improve the development and utilization of wind power, there is a need to further improve the accuracy of wind power prediction. J. P. S. Catalão *et al.* have combined neural network, wavelet analysis and particle swarm optimization algorithm to build a wind power hybrid prediction model, of which the accuracy is higher than single prediction algorithm. But the amount of calculation and calculation time of hybrid prediction model are greater than the latter [2-5]. Hans Bludszuweit *et al.* employs various and

improving statistics method to predict wind power and to analysis and correct the prediction error. The prediction structure shows that the prediction method based on statistics has better prediction accuracy and is more practical-concerned [6-9]. Maria Grazia De Giorgi *et al.* have studied on wind power prediction error and its influence on economics of power system, and proposed corresponding solutions [10-12]. M. Jafarian *et al.* have improved fuzzy neural network and clustering analysis, and the improved algorithm has been used to carry on forecasting wind power. The prediction result of such improved algorithm is better than before [13-15].

Firstly, this paper presents the principle and calculation steps of ISODATA, which is used for clustering the historical NWP data and numerical similarity of NWP data are gathered for a class automatically. Secondly, a BP neural network model has been constructed according to the NWP data. Clustering data and original data are used, respectively to train the BP neural network and to predict the wind power. Compared and analyzed the prediction results between BP neural network with ISODATA clustering algorithm and single BP neural network, it is concluded that the prediction accuracy of BP neural network with ISODATA is higher than BP neural network used alone.

2. Steps of ISODATA Clustering Algorithm

Before introducing the ISODATA algorithm, the sample data of NWP are needed to be processed, and daily numerical weather prediction data were selected as a sample. The assumption that there is a sample set composed of N samples currently $\{y_1, y_2, \ldots, y_N\}$. The steps of ISODATA algorithm are as follows.

Step 1:

Set control parameters, K- the number of clusters expected; θ_N - the minimum number of samples in a cluster; θ_S - standard deviation parameters; θ_C - consolidation parameters; L - the maximum couple of clusters allowed to merge in each iteration; I - the allowed number of iterative. The initial number of classes C and the initial cluster center m_i =1, 2, …, c are given.

Step 2:

According to the Equation (1) to classify the sample set.

$$if\ \left\|y-m_j\right\|<\left\|y-m_i\right\|\ \ i=1,2,\cdots c, i\neq j,\ then\ y=\Gamma_j \qquad (1)$$

In Equation (1), Γ_j is the j-th class, the centre is m_j. If the number of samples of any one Γ_j has $\theta_j < \theta_N$, deletes Γ_j, and makes c=c-1.

Step 3:

To update the entire mean vector according to the Equation (2), in Equation (2) N_j is the number of samples of the j-th class.

$$m_j=\frac{1}{N_j}\sum\nolimits_{y\in\Gamma_j} y\ \ ,\ \ \mathrm{j}=1,2,\cdots,c \qquad (2)$$

The average distance of each sample in Γ_j from the center m_j is calculated according to the Equation (3).

$$\overline{\delta_j}=\frac{1}{N_j}\sum\nolimits_{y\in\Gamma_j}\left\|y-m_j\right\|\quad ,j=1,2,\cdots,c \qquad (3)$$

According to the Equation (4), the average distance of all samples away from its corresponding cluster center can be calculated.

$$\overline{\delta_j}=\frac{1}{N}\sum_{j=1}^{c}N_j\overline{\delta_j} \qquad (4)$$

Step 4:

If this is the last iteration (determined by iteration parameters I), then make θ_C=0, turn to Step 8.

If $c \le K/2$, then turn to Step 5.

If it is an even iteration, or if it is $c \ge K/2$, then turn to Step 8.

Otherwise, continue down.

Step 5:

For each one cluster j, using Equation (5) to calculate the standard deviation σ_j= $[\sigma_{j1}, \sigma_{j2}, \ldots, \sigma_{jd}]^T$

$$\sigma_{ji}=\sqrt{\frac{1}{N_j}\sum\nolimits_{y_k\in\Gamma_j}(y_{ki}-m_{ji})^2} \qquad (5)$$

In the Equation (5), y_{ki} is the i-th component of the sample of the k-th, y_k belongs to Γ_j, m_{ji} is the i-th component of the j-th cluster center; σ_{ji} is the i-th component of the j-th standard deviation, d is the dimension of the sample.

Step 6:

Each cluster has a maximum standard deviation of the component σ_{jmax}, j=1, 2, …, c.

Step 7:

If anyone σ_{jmax}, j=1, 2, …, c, there is $\sigma_{jmax} > \theta_S$, and $\overline{\delta_j} > \overline{\delta}$, $N_j > 2(\theta_N+1)$, or $c \le K/2$, Γ_j is split into two classes, its center correspondingly becomes m_j^+ and m_j^-, then delete the original m_j, and make c=c+1. m_j^+ and m_j^- are calculated as follows.

1) Given a value of k, make 0<k≤1;

2) Make γ_j=$k\sigma_j$, or γ_j=k[0,0, …, σ_{jmax}, …,0]^T;

3) $m_j^+ = m_j + \gamma_j$, $m_j^- = m_j - \gamma_j$.

The value of k should be chosen such that the distance of each sample to m_j^+ and m_j^- is different, but all the samples of Γ_j should remain in those two new collections.

Step 8:

For the entire cluster center, calculate the distance between every two classes.

$$\delta ij = \left\| m_i - m_j \right\| ,\ m_i = 1,2,\cdots,c-1,\ m_j = i+1, i+2,\cdots,c \qquad (6)$$

Comparing δ_{ij} and θ_C, if the δ_{ij} is less than θ_C, then sorted in a order of increasing distances $\delta_{i1j1}<\delta_{i2j2}<\ldots<\delta_{iljl}$, where l is given in Step 1, it is the maximum couples of cluster allowed to merge in each iteration.

Step 9:

Starting from the smallest δ_{i1j1}, for each δ_{iljl}, corresponding to two class centers for m_{il} and m_{jl}. If in the same iteration, m_{il} and m_{jl} don't be merged into one, then merges m_{il} and m_{jl} as Equation (7).

$$m_l = \frac{1}{N_{il} + N_{jl}} [N_{il} \cdot m_{il} + N_{jl} \cdot m_{jl}] \qquad (7)$$

After the merge processing is carried out according to Equation (7), the number of classes is c=c-1. Since each iteration, a class center only can be combined once, and so the actual merging couple is always less than or equal to j.

Step 10:

If this is the last iteration, the program is terminated. Alternatively, when initial parameters need to be changed, the algorithm returns to Step 1 and continues; otherwise, the execution of the algorithm turns to Step 2. The iteration counter increases 1.

3. Classification Calculation

The NWP data obtained from the meteorological department, including wind speed, wind direction, temperature, humidity and air pressure, etc. Wind turbines could be automatically against the wind, thus, the information of wind direction for wind power prediction system is taken out of consideration. In NWP data, because wind speed and pressure data are different in the order of magnitude, before the distance calculated, they should be normalized.

The normalized data were classified and calculated using ISODATA method, the initial parameter settings: K - the number of classes is 20; θ_N - the minimum number of samples in a class is 15; θ_N - standard deviation parameters is 0.08; θ_N - consolidation parameters is 0.02; L - each iteration allowed to merge the maximum couple of clusters is 2; I - the allowed iterative number is 500. Algorithm converges after 32 iterations; the final samples were divided into 11 classes. The center coordinates of the classes are shown in Table 1, and each row represents as a class of center coordinates.

Fig. 1 is the clustering result of samples; no crossover phenomenon can be seen from the figure, between various samples, indicating that the sample has been correctly assigned to various classes.

Table 1. ISODATA automatically clustering results.

Classes	C1	C2	C3	C4
Class 1	0.4878	0.9899	0.6340	0.6434
Class 2	0.3748	0.9910	0.7011	0.5077
Class 3	0.5788	0.9897	0.4868	0.9043
Class 4	0.2580	0.9883	0.8140	0.3324
Class 5	0.4446	0.9890	0.7314	0.2864
Class 6	0.2761	0.9902	0.5385	0.5749
Class 7	0.3439	0.9914	0.5044	0.7294
Class 8	0.4806	0.9915	0.5740	0.4244
Class 9	0.3608	0.9892	0.4099	0.9014
Class 10	0.2146	0.9905	0.6362	0.3376
Class 11	0.5515	0.9940	0.4497	0.5934

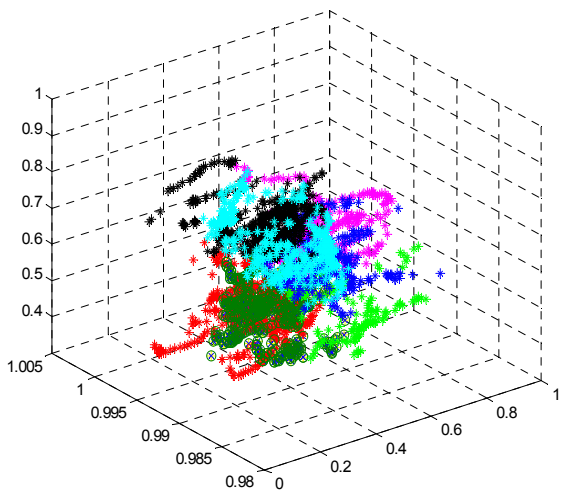

Fig. 1. Samples clustering results.

4. The BP Neural Network Design

The BP network is also called error back propagation algorithm. A typical BP network is composed of input layer, hidden layer and output layer, and a whole link between each layer. The learning process of BP network is composed of mode forward propagation, error back propagation, memory training and learning convergence.

4.1. Mode Forward Propagation

Mode forward propagation begins with the input layer of the network, and each unit of input layer corresponding to the individual elements of the input pattern vector. Make input pattern vector Xz=($x1$, $x2$, …, xn) and (z=1, 2…, m), where z is the pairs of learning mode, n is the number of input layer unit. Corresponding to the input pattern, a desired output vector is Yz= ($y1$, $y2$ …, yq), where q is the number of output layer unit. According to the

calculation principle of BP network, the input of hidden layer can be expressed as follow:

$$S_j = \sum_{i=1}^{n} w_{ij} a_i - \theta_j \tag{8}$$

In the Equation (8), S_j is the input value of hidden layer; θ_j is the threshold of hidden layer, j=1, 2, …, p; p is the number of hidden layer unit; w_{ij} is the connection weight between input layer and hidden layer; a_i is the i-th neuron of the input layer, i=1, 2, …, n.

In order to simulate the nonlinear characteristics of biological neurons, calculate the output of the hidden layer unit with selecting that S_j is the independent variable of S function. The mathematical expression of S function is

$$b_j = f(s_j) = \frac{1}{1+e^{-s_j}} \tag{9}$$

In the Equation (9), b_j is the output value of j-th neural unit in hidden layer.

The threshold value θ_j of unit output has been set for simulating the threshold potential of biological neural, it is constantly being modified. The inputs of each unit in output layer follow as.

$$L_t = \sum_{j=1}^{p} v_{jt} - \gamma_t \tag{10}$$

$$C_t = f(L_t) \tag{11}$$

In the Equation (10), (11), v_{jt} is the connection weight between the neural j in hidden layer and the neural t in output layer; γ_t is the threshold value of the output layer neuron, t is number of output layer neurons, t=1, 2, …, q; f is the S function; L_t is the input value of the output layer neuron.

4.2. Error Back Propagation

The first step of the error back propagation is to calculate the error. The process of error back propagation is the process of passing the error d_j of output layer to the error e_j of hidden layer. The calibration error of output layer can be expressed as follow.

$$d_t^k = (y_t^k - C_t^k) f'(L_t) \tag{12}$$

In the Equation (12), $(y_t^k - c_t^k)$ is the absolute error between expected output and actual output of the network, k=1, 2, …, m; $f'(L_t)$ is the amount of deviation adjustment based on the actual response of each unit.

To complete the error passed to the hidden layer, the calibration error of each unit in hidden layer needs to calculate. The calibration error of hidden layer can be expressed as:

$$e_j^k = \sum_{t=1}^{q} v_{jt} d_t^k f'(s_j) \tag{13}$$

Physical meaning of the formula (13) and (12) are similar. But the calibration error of each intermediate unit in the hidden layer comes from the transformation of q units' calibration error in output layer.

After the calibration error d_t^k and e_j^k obtained, adjust the connection weights between output layer and hidden layer, hidden layer and input layer, and the output threshold value of each unit in reverse direction. The adjustment formula can be expressed as:

$$\Delta v_{jt} = \partial d_t^k b_j^k \tag{14}$$

$$\Delta \gamma_t = \partial d_t^k \tag{15}$$

$$\Delta w_{ij} = \beta e_j^k a_i^k \tag{16}$$

$$\Delta \theta_j = \beta e_j^k \tag{17}$$

In the formula (14) to (17), ∂, β are the learning rate, $0<\partial<1$, $0<\beta<1$.

According to the principle of BP neural network above, combined with the numerical weather forecast data and power data of a wind farm, a BP neural network model has been built. The neural network model is constituted by 3-layer, namely an input layer, a hidden layer and output layer. The input data of the input layer is mainly numerical weather prediction data, which includes wind speed, wind direction, temperature, humidity and pressure, so the neuron number of input layer in neural network is 4. According to the construction experience of neural network, the number of neurons in the hidden layer is approximately twice as much as that in the input layer. In this article, the number of neuron in the hidden layer is 9. The output in the output layer is mainly wind power, so the number of neurons in the output layer is 1. Fig. 2 illustrates a neural network structure designed in accordance with the paper required.

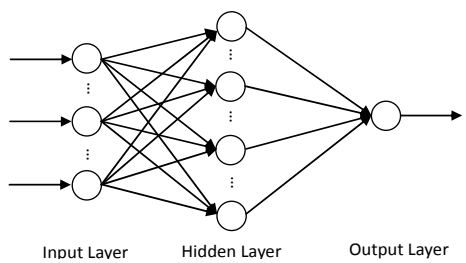

Fig. 2. Neural network structure.

5. Conclusion Network Training and Data Analysis

The content of the principle of ISODATA clustering algorithm and BP neural network model is introduced previously. This section employs numerical weather prediction data to complete the model training of BP neural network and to implement it to wind power prediction.

Fig. 3 is the training processes of BP neural network, where the solid line represents training processes with the help of clustering sample. The dot line indicates training processes of the original sample. Known from the figure, the overall error produced by the neural network which is clustered, is smaller than that derived from original samples.

Fig. 4 is a comparison between predictive and real power of samples. It can be seen that, in some local area, compared to the real power, prediction error of samples is larger. Fig. 5 compares the prediction power results of the clustering samples and the real power. Seen from the figure, the predictive value can accurately reflect the actual variation of wind power, and there appears to be no larger deviation area between predicted and real wind power.

Through the contrast analysis among Fig. 4 and Fig. 5, the predictive results of neural network model trained by clustering samples, are better than that by original data.

In Table 2, CBP represents the clustered BP, BP represents normal BP. From Table 2, it is can be known that the CBP MAE and RMSE are almost equal BP's when prediction time is 12 h, the CBP MAE and RMSE are smaller than BP's, so the conclusion can be got that the prediction result of CBP is better than normal BP when prediction time is 24 h, 48 h, 72 h.

Table 2. The error of the prediction result comparison.

Prediction Time	CBP MAE (%)	CBP RMSE (%)	BP MAE (%)	BP RMSE (%)
12 h	6.34	7.91	6.33	8.02
24 h	7.16	8.26	8.81	12.07
48 h	10.54	13.57	13.67	16.93
72 h	9.98	13.32	13.03	16.33

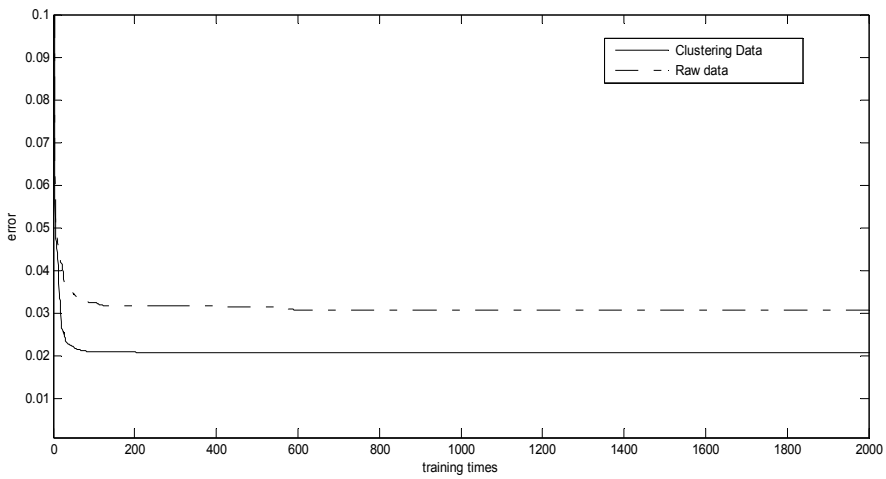

Fig. 3. The training process of the neural network.

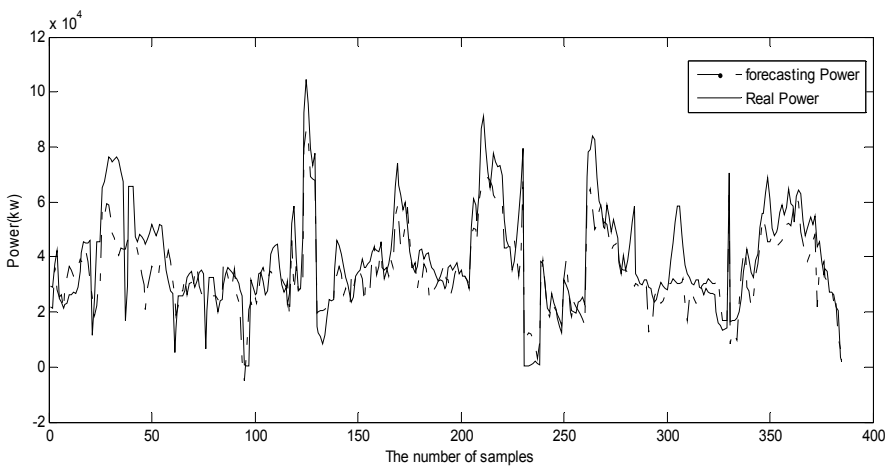

Fig. 4. Original sample forecast results.

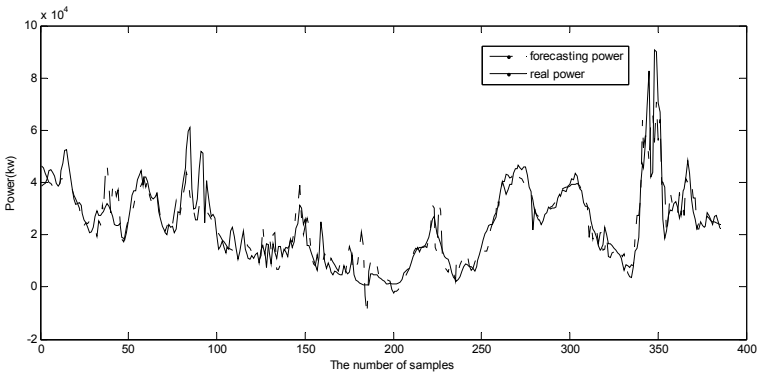

Fig. 5. The number of samples.

6. Conclusions

A wind power prediction technology which combined the ISODATA clustering algorithm and the BP neural network was proposed. The specific work is as follows.

1) The steps of ISODATA clustering algorithm are given, which have been used to cluster the NWP data, gathering those data with similar value to a class;

2) The BP neural network model has been built based on the numerical weather prediction data, and then training such with the original data, and finally the trained network was carried out to predict wind power;

3) The BP neural network has been trained by applying clustering data and then using the trained network to forecast wind power. Compared with the single BP neural network, the results of BP neural network with clustering algorithm predicts better.

References

[1]. News Office of the State Council issued on 24, China's energy policy (2012) the White Paper 2012, *Xinhua News Agency*, Beijing, 2012.

[2]. J. P. S. Catalão, H. M. I. Pousinho, Hybrid Wavelet-PSO-ANFIS Approach for Short-Term Wind Power Forecasting in Portugal, *IEEE Transactions on Sustainable Energy*, 2, 1, 2011, pp. 50-58.

[3]. Zhen-Hai Guo, Jie Wu, Hai-Yan Lu, *et al.*, A case study on a hybrid wind speed forecasting method using BP neural network, *Knowledge-Based Systems*, 24, 2011, pp. 1048-1056.

[4]. Xueli An, Dongxiang Jiang, Chao Liu, *et al.*, Wind farm power prediction based on wavelet decomposition and chaotic time series, *Expert Systems with Applications*, 38, 2011, pp. 11280-11285.

[5]. Erasmo Cadenas, Wilfrido Rivera, Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA-ANN model, *Renewable Energy*, 35, 2010, pp. 2732-2738.

[6]. Hans Bludszuweit and José Antonio Domínguez-Navarro, A Probabilistic Method for Energy Storage Sizing Based on Wind Power Forecast Uncertainty, *IEEE Transactions on Power System*, 26, 3, 2011, pp. 1651-1658.

[7]. Hui Liu, Hong-Qi Tian, Chao Chen, *et al.*, A hybrid statistical method to predict wind speed and wind power, *Renewable Energy*, 35, 2010, pp. 1857-1861.

[8]. Maria Grazia De Giorgi, Antonio Ficarella, Marco Tarantino, Assessment of the benefits of numerical weather predictions in wind power forecasting based on statistical methods, *Energy*, 26, 2011, pp. 3968-3978.

[9]. Gong Li, Jing Shi, Application of Bayesian model averaging in modeling long-term wind speed distributions, *Renewable Energy*, 35, 2010, pp. 1192-1202.

[10]. Maria Grazia De Giorgi, Antonio Ficarella, Marco Tarantino, Error analysis of short term wind power prediction models, *Applied Energy*, 88, 2011, pp. 1298-1311.

[11]. U. Focken, M. Lange, K. Mnnich, *et al.*, Short-term prediction of the aggregated power output of wind farms-a statistical analysis of the reduction of the prediction error by spatial smoothing effects, *Wind Energy*, 90, 3, 2002, pp. 231–246.

[12]. A. Fabbri, T. Gomez San Roman, J. Rivier Abbad, Assessment of the cost associated with wind generation prediction errors in a liberalized electricity market, *IEEE Transactions on Power System*, 20, 3, 2005, pp. 1440-1446.

[13]. M. Jafarian, A. M. Ranjbar, Fuzzy modeling techniques and artificial neural networks to estimate annual energy output of a wind turbine, *Renewable Energy*, 35, 2010, pp. 2008-2014.

[14]. Andrew Kusiak, Wenyan Li, Short-term prediction of wind power with a clustering approach, *Renewable Energy*, 25, 2010, pp. 2362-2369.

[15]. Yi-Long Tu, Tsang-Jung Chang, Cheng-I Hsieh, *et al.*, Artificial neural networks in the estimation of monthly capacity factors of WECS in Taiwan, *Energy Conversion and Management*, 51, 2011, pp. 2938-2946.