Sensors & Transducers, Vol. 176, Issue 8, August 2014, pp. 237-243

Sensors & Transducers

http://www.sensorsportal.com

Study on Parameters Modeling of Wind Turbines Using SCADA Data

Yonglong YAN, Jian LI, Peng SUN, Xiaomeng ZHANG
State Key Laboratory of Power Equipment & System Security and New Technology
College of Electrical Engineering
Chongqing University, Chongqing, 400030, China
Tel.: +86 23 65106880, fax: +86 23 65102442
E-mail: lijian@cqu.edu.cn

Received: 18 March 2014 /Accepted: 31 July 2014 /Published: 31 August 2014

Abstract: Taking the advantage of the current massive monitoring data from Supervisory Control and Data Acquisition (SCADA) system of wind farm, it is of important significance for anomaly detection, early warning and fault diagnosis to build the data model of state parameters of wind turbines (WTs). The operational conditions and the relationships between the state parameters of wind turbines are complex. It is difficult to establish the model of state parameter accurately, and the modeling method of state parameters of wind turbines considering parameter selection is proposed. Firstly, by analyzing the characteristic of SCADA data, a reasonable range of data and monitoring parameters are chosen. Secondly, neural network algorithm is adapted, and the selection method of input parameters in the model is presented. Generator bearing temperature and cooling air temperature are regarded as target parameters, and the two models are built and input parameters of the models are selected, respectively. Finally, the parameter selection method in this paper and the method using genetic algorithm-partial least square (GA-PLS) are analyzed comparatively, and the results show that the proposed methods are correct and effective. Furthermore, the modeling of two parameters illustrate that the method in this paper can applied to other state parameters of wind turbines. *Copyright © 2014 IFSA Publishing, S. L.*

Keywords: Wind turbines, SCADA data, Parameter selection, Data modeling, Neural network.

1. Introduction

With the rapid development of wind power, the Supervisory Control and Data Acquisition (SCADA) system has been installed in almost all wind farms to improve the reliability of wind turbines and to minimize operation and maintenance costs [1-5]. Thus, studies on condition monitoring and fault diagnosis by applying SCADA data and establishing the models of state parameters have drawn increasing attention among researchers both at home and abroad, gradually becoming a hotspot in the field of condition monitoring for wind power [6-8].

SCADA collects information on all of the important subassemblies of wind turbines in wind farms [9], including various parameters that can amount to 100 or more [10]. Reducing the dimensions of data and establishing models with the most closely related parameters is a premise in simplifying models and ensuring prediction accuracy. Although parameters selection has been studied by some researchers who used several algorithms to establish the parameter prediction model, the process of parameter selection has not been discussed clearly and selected parameters have not been validated [11-12]. In some studies, the input parameters of the established parameter prediction model have been validated, and importance degrees of input parameters were compared. However, the significance of importance degrees of input parameters in simplifying models has not been further studied, and the existing model has not been adjusted [13].

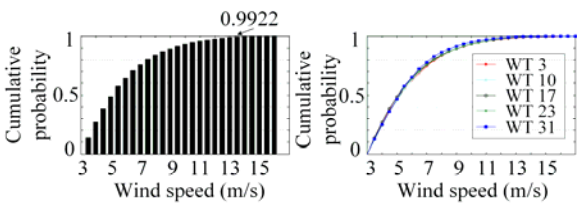
These parameters have been modeled using the regression method and neural networks reported in [14], and the two generated models have been compared. However, the effect of input parameters on the result has been ignored, especially in neural network modeling. Although some of the state parameters of wind turbines have been modeled in [11-15], the influence of the mutual relationships of these parameters on model output are neglected because the input parameters were selected according to the experience of experts and that in the field. Therefore, the accurate modeling of parameters for wind turbines as constructed based on parameter selection has academic value and application potential.

With regard to the aforementioned problems, a method of parameter selection for wind turbines is introduced in this study after studying the SCADA data. In this method, state parameters are modeled according to these data. In part 2 of this paper, SCADA data are selected, and the adopted SCADA monitoring parameters are introduced; in part 3, the temperature parameters of generator bearing B and of cooling air are explained; and in part 4, the methods are verified and analyzed, followed by the conclusions.

2. Collecting SCADA Data in a Wind Farm

The wind turbines of a wind farm have a horizontal shaft and an active yawing and pitch control system. The rated power is 1.5 MW; cut-in wind speed is 3 m/s; rated wind speed is 12 m/s; and cut-out wind speed is 25 m/s.

The wind farm houses 31 WTs that are similar. SCADA monitoring data were obtained at random from WT 3, 10, 17, 23, and 31. Fig. 1(a) shows the probability of the accumulative distribution of the wind speed of WT 17 after wind speed is analyzed using the Bin method [16].



(a) Distribution of wind speed of WT 17

(b) Distribution of wind speed of several WTs

Fig. 1. Wind speed distribution of WT.

When the speed is 13.5 m/s, the accumulative probability is 99.22 %. Thus, the monitoring data constitutes less than 0.8 % of the probability when the speed is above 13.5 m/s. This wind speed analysis can determine accurate monitoring data for further study.

SCADA records the operating parameters, such as nacelle vibration parameters, the phase voltage and current of the power grid, and the temperatures of the gearbox, generator, mainshaft bearing, wind speed, and environment. All of these parameters are presented in Table 1.

Table 1. Monitoring parameters of SCADA system.

Monitoring units	Assessment indices	Modeling or not
Nacelle	Nacelle vibration along X direction X_1	No
	Nacelle vibration along Y direction X_2	No
	Temp. of base cabinet X_3	Yes
	Temp. of top cabinet X_4	Yes
Gearbox	Temp. of input shaft X_5	Yes
	Temp. of output shaft X_6	Yes
	Temp. of inlet oil X_7	Yes
	Temp. of oil X_8	Yes
Bearing of main shaft	Temp. of rotor-side bearing X_9	Yes
	Temp. of gearbox-side bearing X_{10}	Yes
Generator	Temp. of winding X_{11}	Yes
	Temp. of cooling air X_{12}	Yes
	Temp. of bearing A X_{13}	Yes
	Temp. of bearing B X_{14}	Yes
Grid factors	Phase voltage X_{15}	No
	Phase current X_{16}	No
	Active power X_{17}	Yes
	Reactive power X_{18}	No
Environment factors	Wind speed X_{19}	No
	Temp. of ambient X_{20}	No

SCADA data may be problematic given that data selection is based on field experience. Issues include low sampling rate, inaccurate data, and deficiencies in parameter recording. In further study, the operating parameters that accurately reflect the condition of the wind turbines are chosen, such as the parameter of generator bearing temperature. This parameter is closely related to generator operation. The parameter of cooling air temperature also directly influences the heat dissipated by the generator. Therefore, parameters models of generator bearing and cooling air temperatures are established.

3. Establishing the Parameters Models

Artificial neural networks, which have a nonlinear mapping ability that does not require precise mathematical models and which exhibit good ability to learn useful knowledge from input and output data, have been widely used to model nonlinear systems. The back propagation (BP) neural network algorithm is the most widely used neural network in this field [17-19]. Using neural networks to predict parameters has achieved good results [12]; thus, the BP neural network was used in the present study to establish the normal state models of the parameters in WTs.

The BP neural network adopted three layers of network structure, with 19 nodes in the input layer and 1 node in the output layer. The 14 nodes in the hidden layer were determined by an empirical equation, i.e., Eq. (1). And the hidden layer of all models has 14 nodes for comparative analysis in this paper.

$$l = \sqrt{m+n} + a\,, \tag{1}$$

where l is the nodes of a hidden layer; m is the nodes of an input layer; n is the nodes of an output layer; a is the constant in the range of [1, 10].

A total of 30000 data under normal operating conditions were selected from the SCADA system of WT 17 for a year, including 27000 training data and 3000 testing data. The system of the wind turbines is complicated and non-linear; thus, any change in SCADA data may provide useful information. As a result, all of the data used in this study are unprocessed.

3.1. Parameter Modeling of Temperature of Generator Bearing B

3.1.1. The Initial Parameter Model

The modeling process is described as follows.

1) The optimal normal state model was obtained based on the training data described earlier. The model structure was 19:14:1.

2) The model was tested with the test data, as shown in Eq. (2). The average error was 0.0105 and the root-mean-square error (RMSE) was 1.7656.

3) The input parameters of the test data were successively increased and decreased by 5 %. The model was tested. The average error and RMSE were recorded, and arranged in descending order of RMSEs, as shown in Table 2.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(r_i - \hat{r}_i\right)^2}{n}}\,, \tag{2}$$

where $RMSE$ is the root-mean-square error; r_i is the raw data; $\hat{r}_i$ is the predicted value.

Table 2. Influence of the parameters on the temperature of generator bearing B.

No.	Parameters		RMSE
1	Phase voltage X_{15}	+5%	2.5363
		-5%	2.3748
2	Temp. of output shaft X_6	+5%	2.2950
		-5%	2.5846
3	Temp. of oil X_8	+5%	2.4915
		-5%	2.2599
4	Temp. of bearing A X_{13}	+5%	2.2858
		-5%	2.2690
5	Temp. of input shaft X_5	+5%	2.2428
		-5%	2.1590
6	Temp. of top cabinet X_4	+5%	2.0841
		-5%	2.2222
7	Temp. of winding X_{11}	+5%	2.1577
		-5%	2.1073
8	Temp. of gearbox-side bearing X_{10}	+5%	2.1546
		-5%	2.0624
9	Temp. of rotor-side bearing X_9	+5%	2.0581
		-5%	2.1309
10	Temp. of cooling air X_{12}	+5%	1.9859
		-5%	1.9977
11	Temp. of base cabinet X_3	+5%	1.9135
		-5%	1.8659
12	Temp. of inlet oil X_7	+5%	1.8679
		-5%	1.9051
13	Active power X_{17}	+5%	1.8538
		-5%	1.8698
14	Temp. of ambient X_{20}	+5%	1.8416
		-5%	1.8607
15	Phase current X_{16}	+5%	1.8572
		-5%	1.8403
16	Wind speed X_{19}	+5%	1.8219
		-5%	1.8177
17	Reactive power X_{18}	+5%	1.8193
		-5%	1.8190
18	Nacelle vibration along X direction X_1	+5%	1.8192
		-5%	1.8190
19	Nacelle vibration along Y direction X_2	+5%	1.8192
		-5%	1.8189

3.1.2. Selecting the Input Parameters of the Model

According to the order of the parameters shown in Table 2, the first model structure was 1:14:1, and the input parameter was the number 1 parameter. The second model structure was 2:14:1, and the input parameters were the first two parameters. By analogy, the *n*th model structure is *n*:14:1, and the input parameters are the first *n* parameters. Each model had 14 hidden nodes, was trained 30 times with the same training data, and tested with the same test data. The average RMSE at 30 times was obtained, as shown in Fig. 2.

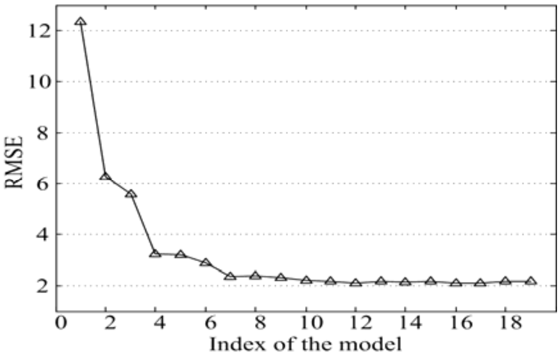


Fig. 2. The change of RMSE versus the increase of input parameters of the model.

As shown in Fig. 2, the RMSEs dropped sharply from the 1st to the 7th model, and the RMSE difference between the 7th and 19th models was smaller. The RMSE changed slightly after the 7th model and exhibited certain fluctuations, thus indicating that the subsequent parameter information was found in the first seven input parameters. According to the figure, the normal state model established based on the first seven input parameters exhibits good accuracy. Thus, the model built from input parameters that include phase voltage, as well as the temperatures of the output shaft, gearbox oil, generator bearing A, input shaft, top cabinet, and generator winding reflects the actual changing situation of the temperature in generator bearing B, thus illustrating the validity of modeling based on the parameter selection method according to the order set in Table 2.

3.2. Parameter Modeling of Temperature of Generator Cooling Air

The parameter of cooling air temperature is modeled using the method applied to model the temperature of generator bearing B. The RMSE of the model are calculated, and the influence of each parameter on cooling air temperature is determined as shown in Table 3. The parameter models are established according to the order in Table 3, and the RMSEs of the models are obtained. Fig. 3 shows that the 6th model is highly accurate. Therefore, the parameter model whose input parameters include the temperature of the input shaft of the gearbox, of the inlet oil of the gearbox, of the output shaft of the gearbox, of gearbox oil, and of generator bearings A and B can effectively reflect the actual changes in the cooling air temperature of generators.

3.3. Parameter Modeling

After the influences of various parameters on the target parameters were considered, the normal state model was established by the following steps (Fig. 4).

Table 3. Influence of the parameters on temperature of generator cooling air.

No.	Parameters		RMSE
1	Temp. of input shaft X_5	+5%	5.3504
		-5%	3.8300
2	Temp. of inlet oil X_7	+5%	3.9634
		-5%	4.0442
3	Temp. of output shaft X_6	+5%	3.2942
		-5%	4.6651
4	Temp. of oil X_8	+5%	4.0026
		-5%	3.4799
5	Temp. of bearing A X_{13}	+5%	3.0371
		-5%	3.2676
6	Temp. of bearing B X_{14}	+5%	2.9222
		-5%	2.8212
7	Phase voltage X_{15}	+5%	2.8991
		-5%	2.7293
8	Temp. of winding X_{11}	+5%	2.8561
		-5%	2.7145
9	Temp. of gearbox-side bearing X_{10}	+5%	2.5705
		-5%	2.5691
10	Temp. of rotor-side bearing X_9	+5%	2.5567
		-5%	2.5645
11	Temp. of top cabinet X_4	+5%	2.5182
		-5%	2.5237
12	Temp. of ambient X_{20}	+5%	2.4501
		-5%	2.4525
13	Phase current X_{16}	+5%	2.4466
		-5%	2.4441
14	Active power X_{17}	+5%	2.4449
		-5%	2.4452
15	Temp. of base cabinet X_3	+5%	2.4421
		-5%	2.4421
16	Wind speed X_{19}	+5%	2.4311
		-5%	2.4411
17	Reactive power X_{18}	+5%	2.4357
		-5%	2.4342
18	Nacelle vibration along X direction X_1	+5%	2.4357
		-5%	2.4342
19	Nacelle vibration along Y direction X_2	+5%	2.4356
		-5%	2.4341

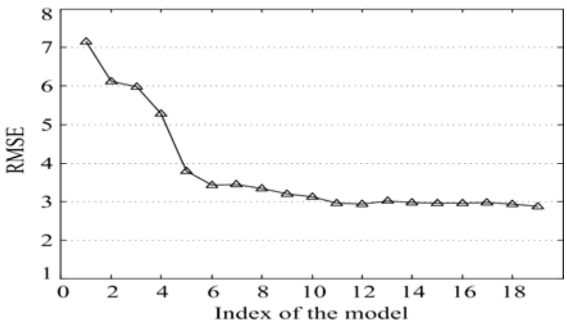


Fig. 3. The change of RMSE versus the increase of input parameters of the model.

Step 1: The first three parameters with higher RMSEs were considered as the input parameters. The number of hidden nodes was determined with Eq. (1). The BP neural network model, with a model structure of 3:14:1, was built. The optimal normal state model was established with the previous training data. Proceed to step 3.

Step 2: The parameter with the highest RMSE in the remaining parameters in Table 2 was added. The normal state model based on the BP neural network was established.

Step 3: The newly built model was tested with the test data. The RMSE was compared with that obtained from 19 initially built input parameters model. When the difference is greater than the threshold, return to the step 2. When it is smaller than the threshold, then the model is the final normal state model.

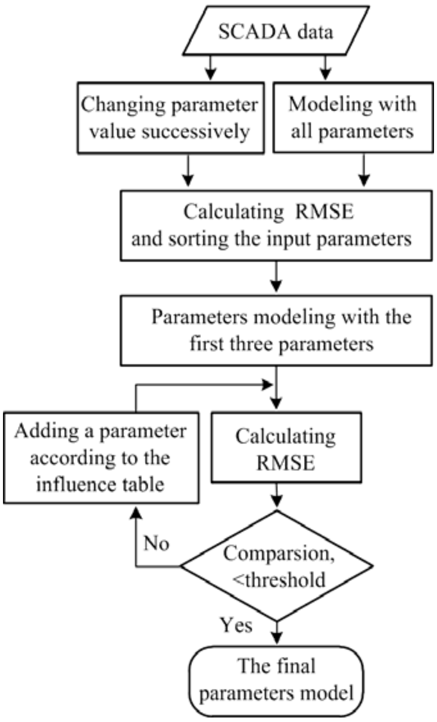


Fig. 4. Flowchart of parameters modeling.

4. Verification and Analysis

To validate the method of parameter selection in this study, SCADA data from the wind farm are used to verify the parameter models of the temperature of generator bearing B and of the cooling air temperature of the generator.

Parameter selection is vital to the establishment of the models because it can simplify them by removing spare parameters. It can also ensure model accuracy, enhance the interpretability of the model, and clarify the effects of parameters on one another. Seven parameters are selected as inputs for the temperature model of generator bearing B through genetic algorithms and partial least squares in [20]. These parameters are temperatures of generator winding, of generator bearing A, of the output shaft of the gearbox, of generator cooling air, phase current, active power, temperature of rotor-side bearing. The models are validated using SCADA data from January 14, 2012 to February 15, 2012, and the results are depicted in Fig. 5. The models are constructed using the method in [20] and the method proposed in this study. Fig. 5(a) and 5(b) indicate that residues mainly distribute at temperatures ranging from −6 °C to +5 °C, whereas the residue from the model generated using the method in [20] mainly distributes at temperatures ranging from −8 °C to +5 °C. The difference of the two models is clear from Figure 5. The calculated RMSEs are 2.0959 and 3.1504, as displayed in Fig. 5 (a) and 5 (b), thereby implying that the model developed with the proposed method is more accurate.

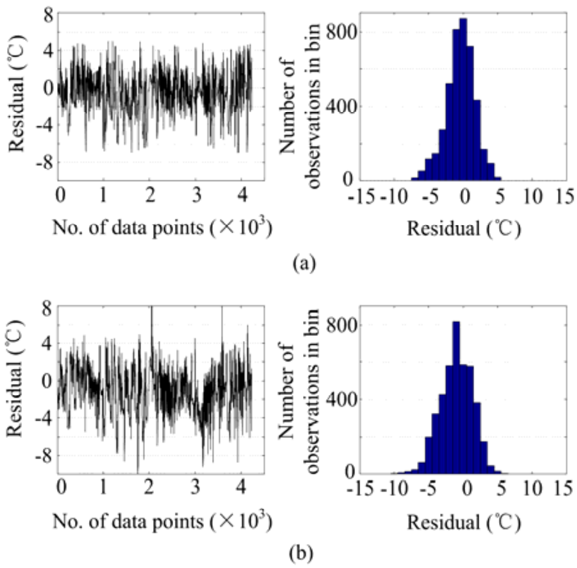


Fig. 5. (a) Residual and distribution of parameter model for temperature of generator bearing B using the proposed method in this paper and (b) the method in [20].

The parameter model of the cooling air temperature of the generator reveals the generalizability of the proposed method. Using the method in [20], six parameters are selected as inputs for this model, including phase current, active power, and the temperatures of generator bearings A and B, of generator winding, and of the output shaft of the gearbox. The structure of the model of the back-propagation neural network is expressed in the ratio 6:14:1. The two models are evaluated by SCADA data from February 5th, 2012 to March 7th, 2012, and the test results are exhibited in Fig. 6. Figs. 6(a) and 6(b) indicate that residues mainly distribute at temperatures ranging from −8 °C to +10 °C, whereas the residue from the model generated using the method in [20] mainly distributes at temperatures ranging from 12 °C to +10 °C. Thus, the residual distribution of the proposed model is more concentrated. The calculated RMSEs are 3.5246 and 4.3969, as displayed in Figures 6(a) and 6(b), thereby implying that the model developed with the proposed method is more accurate.

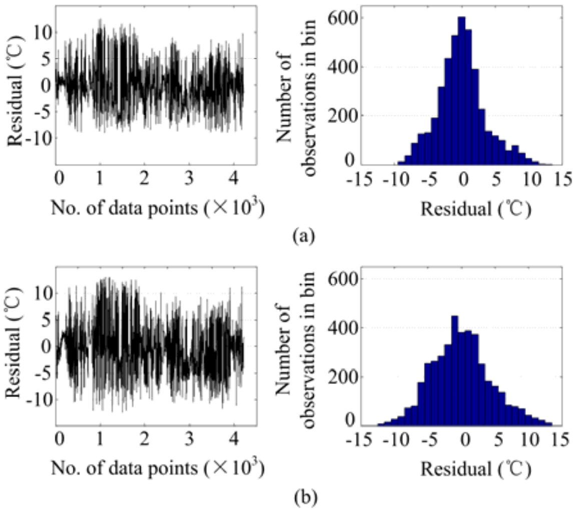


Fig. 6. (a) Residual and distribution of parameter model for temperature of generator cooling air using the proposed method in this paper and (b) the method in [20].

5. Conclusions

To evaluate abnormal parameters, a method to model the parameters of wind turbines is developed using SCADA data. The conditions of wind turbines can thus be assessed, and faults can be diagnosed. This diagnosis can improve the reliability of wind turbines in operation and reduce operation and maintenance costs. In full consideration of the characteristics of the SCADA data obtained from a wind farm, strong representative data are chosen for this study. An artificial neural network algorithm is adopted, and models of temperature of generator bearing and cooling air are built. To simplify the structure and accuracy of the models, a method of parameter selection is proposed. This method effectively reduces the dimensionality of the input parameters and selects the parameters that strongly affect the target parameters. Aside from the proposed method, another technique reported in previous literature was used to select the parameters for this study. The corresponding parameter model of the temperature of generator bearing B is thus established. The proposed method is validated through comparison and analysis; to verify its generality, it is utilized to develop the parameter model of the cooling air temperature of the generator.

Acknowledgements

The authors acknowledge the funding of the 973 Program (No. 2012CB215205). The Funds for Innovative Research Groups of China (51021005), and the Programme of Introducing Talents of Discipline to Universities, China (B08036) are also appreciated.

References

[1]. The World Wind Energy Association, *World Wind Energy Report,* 2011.

[2]. J. K. Kaldellis, D. Zafirakis, The wind energy revolution: A short review of a long history, *Renewable Energy*, Vol. 36, Issue 7, 2011, pp. 1887-1901.

[3]. W. X. Yang, P. J. Tavner, C. J. Crabtree, M. Wilkinson, Cost-Effective condition monitoring for wind turbines, *IEEE Transactions on Industrial Electronics*, Vol. 57, 2010, pp. 263-271.

[4]. D. McMillan, G. W. Ault, Condition monitoring benefit for onshore wind turbines: sensitivity to operational parameters, *IET Renewable Power Generation*, Vol. 2, 2008, pp. 60-72.

[5]. J. Nilsson, L. Bertling, Maintenance management of wind power systems using condition monitoring systems—Life cycle cost analysis for two case studies, *IEEE Transactions on Energy Conversion*, Vol. 22, 2007, pp. 223-229.

[6]. Z. Hameed, Y. S. Hong, Y. M. Cho, S. H. Ahn, C. K. Song, Condition monitoring and fault detection of wind turbines and related algorithms: A review, *Renewable and Sustainable Energy Reviews*, Vol. 13, 2009, pp. 1-39.

[7]. B. Lu, Y. Y. Li, X. Wu, Z. Z. Yang, A review of recent advances in wind turbine condition monitoring and fault diagnosis, *IEEE Conference on Power Electronics and Machines in Wind Applications,* Lincoln, NE, June, 2009,pp. 1-7.

[8]. W. X. Yang, R. Court, J. S. Jiang, Wind turbine condition monitoring by the approach of SCADA data analysis, *Renewable Energy*, Vol. 53, 2013, pp. 365-376.

[9]. IEC61400-12-1 Wind turbines – Part 25-1: Communications for monitoring and control of wind power plants – Overall description of principles and models, *IEC International Standard,* 2006.

[10]. A. Kusiak, A. Verma, A Data-Driven Approach for Monitoring Blade Pitch Faults in Wind Turbines, *IEEE Transactions on Sustainable Energy*, Vol. 2, 2011, pp. 87-96.

[11]. A. Kusiak, A. Verma, Analyzing bearing faults in wind turbines: A data-mining approach, *Renewable Energy*, Vol. 48, 2012, pp. 110-116.

[12]. A. Kusiak, W. Y. Li, Virtual models for prediction of wind turbine parameters, *IEEE Transactions on Energy Conversion*, Vol. 25, 2010, pp. 245-252.

[13]. A. Zaher, S. D. J. McArthur, D. G. Infield, Y. Patel, Online wind turbine fault detection through automated SCADA data analysis, *Wind Energy*, Vol. 12, 2009, pp. 574-593.

[14]. M. Schlechtingen, I. F. Santos, Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection, *Mechanical Systems and Signal Processing*, Vol. 25, 2011, pp. 1849-1875.

[15]. M. C. Garcia, M. A. Sanz-Bobi, J. Pico, SIMAP: Intelligent System for Predictive Maintenance. Application to the health condition monitoring of a wind turbine gearbox, *Computers in Industry*, Vol. 57, 2006, pp. 552-568.

[16]. IEC61400-12-1 Wind turbines – Part 12-1: Power performance measurements of electricity producing wind turbines, *IEC International Standard,* 2005.

[17]. Z. Z. Shi, Neural Networks, *Higher Education Press*, Beijing, 2009.

[18]. M. Nisha, S. P. Srinivas, Soft Computing Techniques for Employee Evaluation: Designing Framework of Artificial Neural Network for Employee Evaluation, *Journal of Computational Intelligence and Electronic Systems*, Vol. 1, 2012, pp. 94-98.

[19]. J. H. Wang, K. Liu, Feature-Based Fuzzy Neural Network Approach for Intrusion Data Classification, *Journal of Computational Intelligence and Electronic Systems*, Vol. 1, 2012, pp. 99-103.

[20]. M. Ferrand, B. Huquet, S. Barbey, F. Barillet, F. Faucon, H. Larroque, O. Leray, J. M. Trommenschlager, M. Brochard, Determination of fatty acid profile in cow's milk using mid-infrared spectrometry: Interest of applying a variable selection by genetic algorithms before a PLS regression, *Chemometrics and Intelligent Laboratory Systems*, Vol. 106, 2011, pp. 183-189.