Condition monitoring of wind turbines based on extreme learning machine

Peng Qian, Xiandong Ma, Yifei Wang

Engineering Department
Lancaster University
Lancaster, UK LA1 4YR
p.qian@lancaster.ac.uk

Abstract—nowadays, wind turbines have been widely installed in many areas, especially in remote locations on land or offshore. Routine inspection and maintenance of wind turbines has become a challenge in order to improve reliability and reduce the energy of cost; thus adopting an efficient condition monitoring approach of wind turbines is desirable. This paper adopts extreme learning machine (ELM) algorithms to achieve condition monitoring of wind turbines based on a model-based condition monitoring approach. Compared with the traditional gradient-based training algorithm widely used in the single-hidden layer feed forward neural network, ELM can randomly choose the input weights and hidden biases and need not be tuned in the training process. Therefore, ELM algorithm can dramatically reduce learning time. Models are identified using supervisory control and data acquisition (SCADA) data acquired from an operational wind farm, which contains data of the temperature of gearbox oil sump, gearbox oil exchange and generator winding. The results show that the proposed method can efficiently identify faults of wind turbines.

Keywords-component; Wind turbines, condition monitoring, SCADA data, model-based approach, artificial neural network, extreme learning machine.

I. INTRODUCTION

Nowadays, wind power has been considered as one of most sustainable and eco-friendly energy sources. Wind turbines are widely installed in different places, especially in remote locations on land or offshore, because in these locations the wind resource is stronger and more reliable, and visual and noise impacts can be reduced [1]. However, routine inspection and maintenance of wind turbines has become a challenge in these remote areas. The maintenance cost for wind turbines usually accounts for a considerable proportion of income [2-3]. Consequently, it is essential to develop efficient condition monitoring techniques for wind turbines, providing information about the past and current conditions of the turbines and enabling optimal scheduling of maintenance activities. Supervisory control and data acquisition (SCADA) data have been considered to be the commonly used monitoring data system that are applied to monitor and control devices in many industrial applications [4], such as space flight and aviation, transportation, biological medicine, and power energy sectors. These systems can generate monitoring data that helps to build models of a process operating under different conditions.

The schematic diagram of the model-based condition monitoring approach is illustrated in Fig. 1 where the data generated from condition monitoring system or equivalents are used as inputs using models to predict the output signals of a physical process. Then, actual output signals are compared with the signals that are predicted by the model for given input signals. Differences between actual output signals and the model predicting signals could be caused by changes in the process, possibly due to the occurrence of faults [5]. The residual signal can be an important indicator to provide an early warning of the impending component failure.
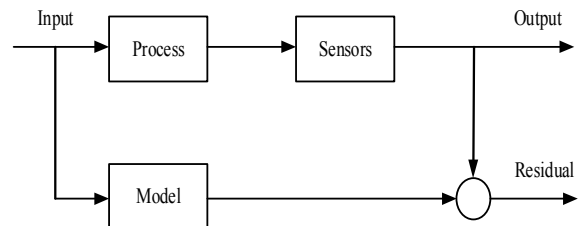

Fig. 1 Schematic diagram of model-based condition monitoring system [5]

A number of supervised learning methods have been applied in model-based condition monitoring system of wind turbines, such as artificial neural networks (ANNS) [6]. However, ANNS have suffered a drawback in the real-time implementation [7], which limits their engineering applications. Extreme learning machine (ELM) is considered in this paper due to its extremely fast learning speed [8-9]. The paper is organised as follows. The principle of ELM is described in Section 2; SCADA data including the temperatures of gearbox oil exchange, gearbox oil sump and generator winding are then employed to verify the effectiveness of the proposed method and the results are presented in Section 3. Conclusions and suggestions for future work are given in Section 4.

II. EXTREME LEARNING MACHINE

A single-hidden layer feed forward neural network (SLFN) has been widely used in many fields such as mode recognition and state prediction, because of its efficient learning skills [10-12]. However, SLFN always selects a gradient-based neural network algorithm as its training algorithm. The traditional gradient-based training algorithms have some disadvantages such as trapping at local minima, the overtraining, and the high computing burdens, which causes longer training time of the SLFN during the learning process [7]. As a relatively new

technique, the ELM can avoid this drawback [8-9]. Compared with the traditional gradient-based training algorithms, ELM randomly chooses the input weights and hidden biases and needs not be tuned in the training process. Thus, ELM algorithm features an extremely faster learning speed than most popular learning algorithms such as back-propagation, and thus dramatically reduce learning time. Furthermore, if the chosen activation function is infinitely differentiable, the ELM can identify distinct samples exactly with zero error under the condition of equal number of hidden-layer neurons and distinct samples processed in the ANN.

The schematic diagram of a single-hidden layer feed forward neural network is shown in Fig.2, which is consisted of an input layer, a hidden layer and an output layer. It assumes that the input layer and the hidden layer have n and L neurons, respectively, while the output layer has m neurons.
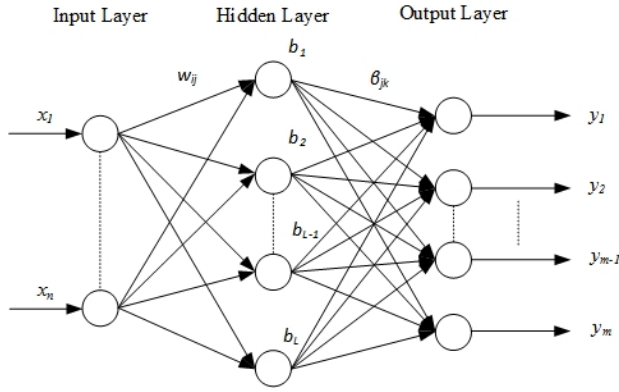

Fig. 2 Schematic diagram of single-hidden layer feed forward neural network (SLFN)

The coefficients ω_{ij} represent the input weight matrix between the input layer and the hidden layer; β_{jk} denotes the output weight matrix connecting the hidden layer and the output layer; b represents the bias of the hidden layer matrix. These parameters can be defined

$$\omega = \begin{bmatrix} \omega_{11} & \omega_{12} & \cdots & \omega_{1n} \\ \omega_{21} & \omega_{22} & \cdots & \omega_{2n} \\ \vdots & \vdots & & \vdots \\ \omega_{L1} & \omega_{L2} & \cdots & \omega_{Ln} \end{bmatrix}_{L \times n} \tag{1}$$

$$\beta = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2m} \\ \vdots & \vdots & & \vdots \\ \beta_{L1} & \beta_{L2} & \cdots & \beta_{Lm} \end{bmatrix}_{L \times m} \tag{2}$$

$$b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_L \end{bmatrix}_{L \times 1} \tag{3}$$

Given datasets with Q distinct training samples, X and Y are input matrix and output matrix, respectively.

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1Q} \\ x_{21} & x_{22} & \cdots & x_{2Q} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nQ} \end{bmatrix}_{n \times Q} \tag{4}$$

$$Y = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1Q} \\ y_{21} & y_{22} & \cdots & y_{2Q} \\ \vdots & \vdots & & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mQ} \end{bmatrix}_{m \times Q} \tag{5}$$

If there exists an ELM with L neurons in the hidden layer and an activation function $g(.)$ can approximate the Q samples with zero error, this ELM can be represented by

$$T = \begin{bmatrix} \sum_{i=1}^{L} \beta_{i1} g \left(\omega_i x_j + b_i \right) \\ \sum_{i=1}^{L} \beta_{i2} g \left(\omega_i x_j + b_i \right) \\ \vdots \\ \sum_{i=1}^{L} \beta_{im} g \left(\omega_i x_j + b_i \right) \end{bmatrix}_{m \times 1} \begin{pmatrix} j = 1, 2, \cdots, Q \\ i = 1, 2, \cdots, L \end{pmatrix} \tag{6}$$

$$\omega_i = \begin{bmatrix} \omega_{i1} & \omega_{i2} & \cdots & \omega_{in} \end{bmatrix}$$

$$x_j = \begin{bmatrix} x_{1j} & x_{2j} & \cdots & x_{nj} \end{bmatrix}^T$$

For simplicity, eq. (6) can be compactly described as,

$$H \beta = T' \tag{7}$$

where T' is the transpose matrix of T and H is the output matrix of the hidden layer of the ELM. The matrix H can be expressed as,

$$H = \begin{bmatrix} g(\omega_1 x_1 + b_1) & g(\omega_2 x_1 + b_2) & \cdots & g(\omega_L x_1 + b_L) \\ g(\omega_1 x_2 + b_1) & g(\omega_2 x_2 + b_2) & \cdots & g(\omega_L x_2 + b_L) \\ \vdots & \vdots & & \vdots \\ g(\omega_1 x_Q + b_1) & g(\omega_2 x_Q + b_2) & \cdots & g(\omega_L x_Q + b_L) \end{bmatrix}_{Q \times L} \tag{8}$$

When the input weight matrix ω and the hidden layer bias matrix b are initialized, the hidden layer output matrix H can be uniquely determined. The output weight matrix β can be calculate as follow,

$$\min_{\beta} \| H \beta - T' \| \tag{9}$$

During this process, the input weight matrix ω and the hidden layer bias matrix b do not need to be changed and the solution can be expressed,

$$\hat{\beta} = H^+ T' \tag{10}$$

This process is equivalent to finding a unique smallest norm least-squares solution of the linear system in eq. (7). The matrix H^+ is the Moore-Penrose generalized inverse of the hidden layer output matrix H, which can be derived through the singular value decomposition (SVD) method.

The Fig. 3 demonstrates the flowchart of ELM algorithm. There are four steps to implement the ELM algorithm, including (i) design of the SLFN structure, (ii) random choice of the input weights ω and hidden biases b, (iii) acquisition of the initial hidden layer output matrix H and the output weights β, and finally (iv) improvement and updating of the hidden layer output matrix H and output weights β.

Fig. 3 The flowchart of extreme learning machine (ELM)

III. RESULTS AND ANALYSIS

Supervisory control and data acquisition (SCADA) system is an industrial automation control system, which has been widely used in many industries, including oil and gas, metallurgy, manufacturing, railway system, transportation, energy, and power systems. Modern SCADA systems work relying on multiple hardware and software elements and IT technologies to monitor, gather, and process data. In power systems, SCADA is a mature technology and it is used for data collection, device control, parameter adjustment, signal alarm generation and other functions.
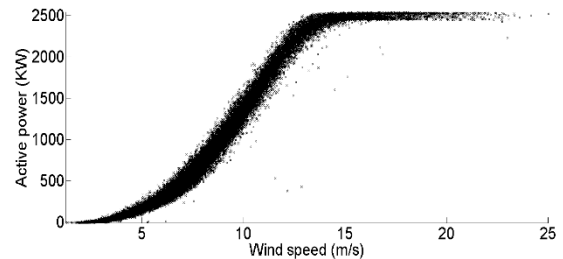
SCADA data used in this paper were obtained from an operational wind farm. The use of actual operational data of wind turbines is a good method in order to demonstrate the effectiveness of the proposed algorithms. The data cover 12 months' duration and consist of 128 parameters that contain various temperatures, pressures, vibrations, power outputs, wind speed and digital control signals. In order to reduce the amount of data gathered from the operating wind turbines, SCADA data are usually sampled at 10 minute interval.
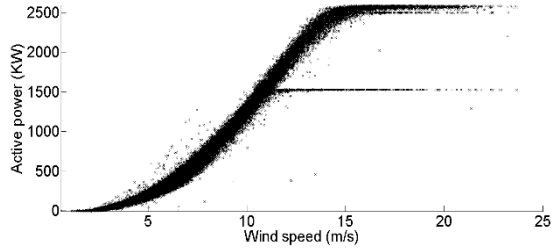
A. Model identification

Examples of the power curve of turbines are illustrated in Fig. 4. Fig. 4 (a) shows a normal power curve of the reference wind turbine; turbine power varies cubically with wind speed, and wind speed varies continuously on time-scales. When the wind speed is lower than the cut-in speed (4 m/s in this case), the turbine does not produce any power because the rotor torque is too low; while the wind speed is above the cut-out speed (15 m/s in this case), the turbine does not produce any power either because it is shut down to protect the turbine. If the wind speed is above the rated wind speed (15 m/s in this case) but below the cut-out speed, the turbine's output power is capped at the rated power. Fig. 4 (b) shows the power curve of a wind turbine operating for a period with reduced power output following a gearbox fault at some points. Other points shown in the curves might be because of the occurrence of maintenance periods or indicate the turbines were inactive when the wind speed was too high or too slow.

It has been found that gearbox oil sump temperature, gearbox oil exchange temperature and generator winding temperature change with both power output and cooling air temperature. Therefore in order to achieve an appropriate model identification, wind speed and power output are selected as the inputs while the temperature is considered as the output. This multiple-input and single-output (MISO) approach allows a more sensitive detection, which naturally mean that temperatures have a close relationship with both wind speed and power output. Furthermore, it is essential to choose the fault-free wind turbine as the reference turbine to train the model with the ELM method as proposed above.

a. Power curve of the fault-free turbine

b.   Power curve of the turbine with a gearbox fault
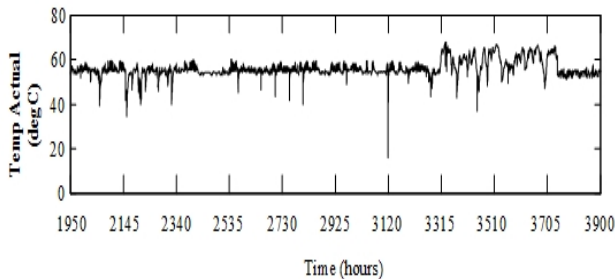
Fig. 4 Examples of power curve of the turbines

B. Detection of faults

The wind speed and the power output are selected as the input for the model identification to test whether the proposed ELM approach is able to detect faults of the wind turbine by comparing corresponding model predicting signals with actual temperatures of gearbox oil sump, gearbox oil exchange and generator winding from SCADA system.
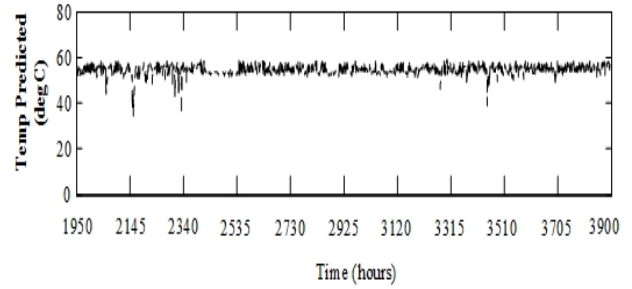
Let us first look at model prediction for the temperature of gearbox oil sump. Fig. 5 (a) illustrates the temperature of gearbox oil sump obtained from the SCADA system and Fig. 5 (b) shows the temperature of gearbox oil sump predicted by the ELM model. It can be seen from the residual signal as shown in Fig. 5 (c), the temperature of gearbox oil sump deviates from the model prediction around at hour 3315, which indicates the beginning of the gearbox fault.

Fig. 6 is also a good example to verify the effectiveness of the proposed ELM model. It shows the plots of actual SCADA data, the model predicting signal, and the residual signal of the temperature of gearbox oil exchange, respectively. Residual signal is shown in Fig. 6 (c). The temperature of gearbox oil exchange deviates from the model prediction also around at hour 3315, which happens at the same time as the temperature of gearbox oil sump. This means the gearbox fault of the wind turbine can be detected effectively.
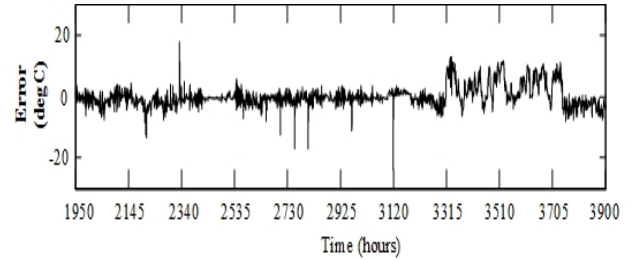
Apart from the detection of gearbox fault, the ELM model is also used to detect the generator winding faults. As can be seen from Fig. 7, the temperature of generator winding starts abnormal at around hour 3120.

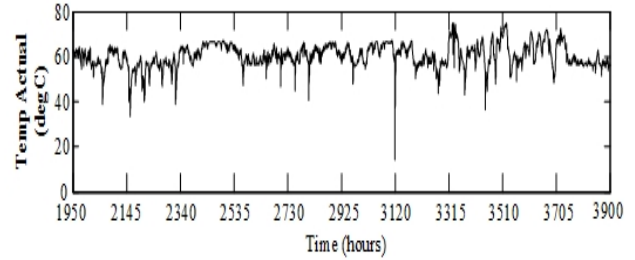a. Actual gearbox oil sump temperature from SCADA system

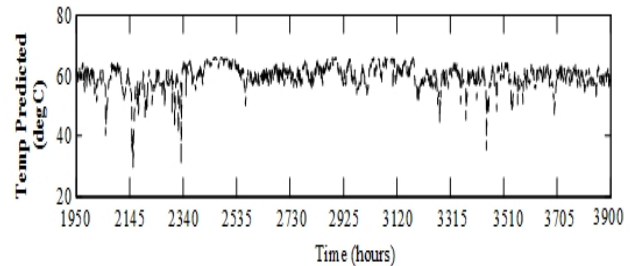b. Model output of the gearbox oil sump temperature

c. Residual signal of gearbox oil sump temperature between SCADA data and model prediction
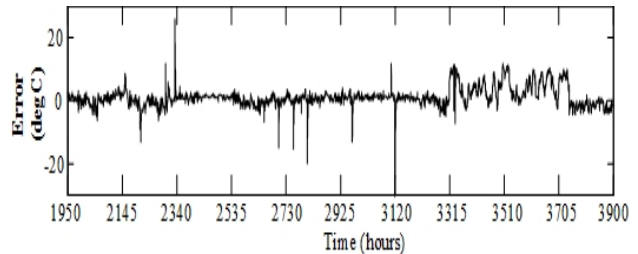
Fig. 5 ELM model response compared to SCADA data for gearbox oil sump temperature

a. Actual gearbox oil exchange temperature from SCADA system

b. Model output of the gearbox oil exchange temperature

c. Residual signal of gearbox oil exchange temperature between SCADA data and model prediction

Fig. 6 ELM model response compared to SCADA data for gearbox oil exchange temperature

a. Actual generator winding temperature from SCADA system

b. Model output of the generator winding temperature

c. Residual signal of generator winding temperature between SCADA data and model prediction

Fig. 7 ELM model response compared to SCADA data for generator winding temperature

The ELM algorithm was implemented on an ordinary PC platform equipped with a Xeon E3-1271 v3 3.6GHz CPU and 16GB RAM. The execution time was compared with the traditional BP neutral network. As can be seen from Table 1, the ELM algorithm significantly outperform the BP neural network in terms of the computation time required.

Table 1 Performance comparison between the ELM and the BP method

Type	Algorithm	Time(s)
Gearbox oil sump temperature	ELM	0.161s
Gearbox oil exchange temperature	ELM	0.158s
Generator winding temperature	ELM	0.165s
Gearbox oil sump temperature	BP	23.376s
Gearbox oil exchange temperature	BP	26.14s
Generator winding temperature	BP	25.384s

IV. CONCLUSION

This paper has presented ELM algorithm that has been used in condition monitoring of wind turbines. SCADA data obtained from an operational wind farm, including the temperatures of gearbox oil exchange, gearbox oil sump and generator winding t, are used to verify the effectiveness of the proposed method. Models developed from SCADA data have been used to identify faults in gearbox and generator winding in the turbines. The results have shown that differences between actual output signals and the model predicting signals are caused by a gearbox fault and a generator winding fault. Consequently, the proposed method can provide an early warning of the impending component failure.
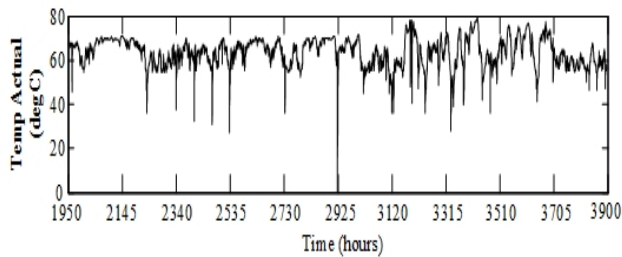
It is worth emphasising that the SCADA data used in the paper are mostly indicative of normal operations of the wind turbines and contain less fault information; therefore static models are only used in the paper. Dynamic models will be considered in our future work to investigate the effect of more values of the past inputs on the model output. Future work will also focus upon the development of the early warning system employing online sequential ELM that can predict faults of the wind turbines in real-rime. In addition, the research will consider using the experimental platform in our lab to further validate the models in real-time implementation.
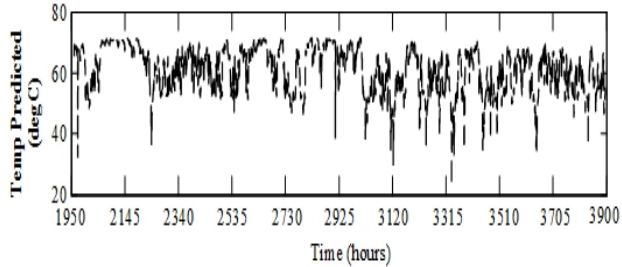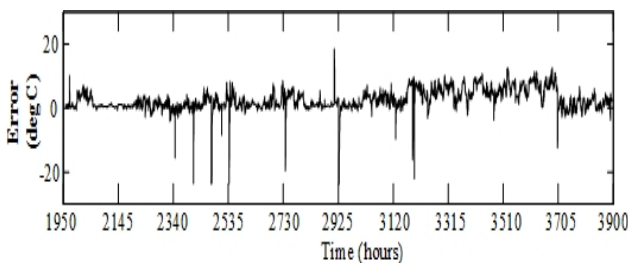
REFERENCES

[1] B. Snyder, M. J. Kaiser. A comparison of offshore wind power development in Europe and the US: Patterns and drivers of development. Applied Energy, vol. 86, no. 10, pp. 1845-1856, 2009.

[2] C. A. Walford. Wind Turbine Reliability: Understanding and Minimizing Wind Turbine Operation and Maintenance Costs, Technical Report SAND 2006-1100, Sandia National Laboratories, USA, 2006.

[3] G. J. W. van Bussel, C. Schontag. Operation and maintenance aspects of large offshore wind farms. In Proceedings of European Wind Energy Conference, Dublin, Ireland, 1997.

[4] Communication Technologies Inc. Supervisory control and data acquisition systems. USA: National Communications System; 2004.

[5] Dvorak D, Kuipers B. Model-based monitoring of dynamic systems. In: International Joint Conference on Artificial Intelligence, vol. 2; 1989. pp. 1238-43.

[6] Schlechtingen M, Ferreira Santos I. Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection. Mech Syst Signal Process 2011;25:1849-75

[7] G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine:Theory and applications," Neurocomputing, vol. 70, no. 1–3, pp. 489–501, Dec. 2006.

[8] A. H. Nizar, Z. Y. Dong, and Y.Wang, "Power utility nontechnical loss analysis with extreme learning machine method," IEEE Trans. Power Syst., vol. 23, no. 3, pp. 946–955, Aug. 2008.

[9] G. B. Huang, X. J. Ding, and H. M. Zhou, "Optimization method based extreme learning machine for classification," Neurocomputing, vol. 74, no. 1–3, pp. 155–163, Dec. 2010.

[10] G. Karniotakis, G. S. Stavrakakis, and E. F. Nogaret, "Wind power forecasting using advanced neural network models," IEEE Trans. Energy Convers., vol. 11, no. 4, pp. 762–767, Dec. 1996.

[11] T. G. Barbounis, J. B. Theocharis, M. C. Alexiadis, and P. S. Dokopoulos, "Long-term wind speed and power forecasting using local recurrent neural network models," IEEE Trans. Energy Convers., vol. 21, no. 1, pp. 273–284, Mar. 2006.

[12] K. Bhaskar and S. N. Singh, "AWNN-assisted wind power forecasting using feed-forward neural network," IEEE Trans. Sustain. Energy, vol. 3, no. 2, pp. 306–315, Apr. 2012.