

Article

Estimating Health Condition of the Wind Turbine Drivetrain System

Peng Qian ¹, Xiandong Ma ¹ and Dahai Zhang ^{2,*} 

¹ Engineering Department, Lancaster University, Lancaster LA1 4YW, UK; peng.qian@brunel.ac.uk (P.Q.); xiandong.ma@lancaster.ac.uk (X.M.)

² Ocean College, Zhejiang University, Hangzhou 310058, China

* Correspondence: zhangdahai@zju.edu.cn

Received: 4 September 2017; Accepted: 10 October 2017; Published: 12 October 2017

Abstract: Condition Monitoring (CM) has been considered as an effective method to enhance the reliability of wind turbines and implement cost-effective maintenance. Thus, adopting an efficient approach for condition monitoring of wind turbines is desirable. This paper presents a data-driven model-based CM approach for wind turbines based on the online sequential extreme learning machine (OS-ELM) algorithm. A physical kinetic energy correction model is employed to normalize the temperature change to the value at the rated power output to eliminate the effect of variable speed operation of the turbines. The residual signal, obtained by comparing the predicted values and practical measurements, is processed by the physical correction model and then assessed with a Bonferroni interval method for fault diagnosis. Models have been validated using supervisory control and data acquisition (SCADA) data acquired from an operational wind farm, which contains various types of temperature data of the gearbox. The results show that the proposed method can detect more efficiently both the long-term aging characteristics and the short-term faults of the gearbox.

Keywords: condition monitoring; online sequential extreme learning machine (OS-ELM); Bonferroni interval; health condition; drivetrain; wind turbine

1. Introduction

Over the past decades, wind energy has been widely regarded as an effective energy solution to reducing CO₂ emissions and producing sustainable energy because of its technology maturity and improved cost competitiveness. One of the key priorities identified by the wind industry is to reduce costs in the operation and maintenance of wind turbines, which currently accounts for 18% of the cost of offshore energy [1]. A cost-effective operation of the wind farm is therefore crucial due to the fierce competition in the global sustainable energy market. Monitoring of the operating conditions of the wind turbines has been considered as an effective method to enhance the reliability of wind turbines and implement cost-effective maintenance. Clearly, it is essential to develop effective CM techniques for wind turbines [1–3] to provide information regarding the past and current conditions of the turbines, and to enable the optimal scheduling of maintenance tasks [4].

From surveys concerning the reliability of the wind turbines, faults caused by the drivetrain system account for over 20% of total faults and contribute to approximately 30% of the downtime of doubly-fed induction generator (DFIG)-based wind turbines [5,6]. Thus, studies about fault diagnosis of the drivetrain system are necessary. Figure 1 shows a typical drivetrain system in a DFIG wind turbine that contains hub, main bearing, main shaft, gearbox, brake, generator shaft and generator. The main function of drivetrain system is to transmit kinetic energy from the turbine rotor to the electric generator by adjusting rotational speed and torque.

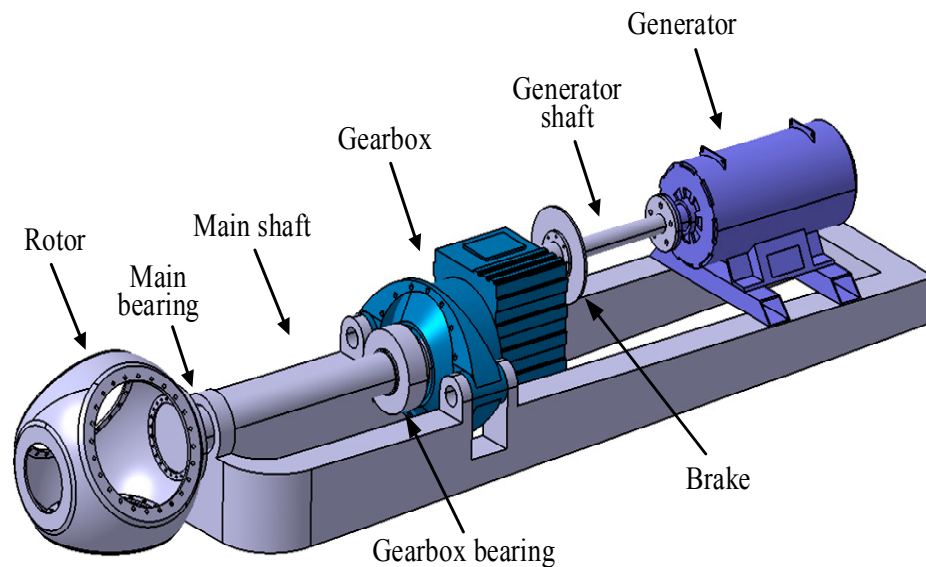


Figure 1. A typical drivetrain and its components in the doubly-fed wind turbine.

For the mechanical transmission system, monitoring and analysis of the vibration signals have been proven very effective, as it is easy to obtain the fault signature of a specific component in the frequency or time-frequency domains. However, it is difficult to obtain accurate vibration signals in the wind turbine under varying speed operation. Furthermore, condition monitoring based on vibration signals is a kind of component-specific technique, which lacks providing an inherent relationship between different subsystems [7]. For example, variable speed wind turbines have the function of maximum power point tracking that optimizes performance in relation to time changing wind speed. This means the rotor, gearbox and generator rotational speed may change significantly and frequently during operation. Due to the significant rotational speed change, the vibration frequency bandwidth varies, which may cause difficulties in locating the exact location of the fault.

Acoustic emission (AE) is another effective method that can be applied in condition monitoring of the wind turbine [8]. When materials bear external strain or stress, they may generate sound waves called AE. Even a tiny structural change will make AE signals to be excited, meaning that AE signal is very suitable to be applied to detect incipient structure defect and monitor its development. For wind turbines, AE signals are generally used for fault detection of the blade, gearbox, bearing, and generator. Compared with vibration signals, AE signals have high signal-to-noise ratio, which means that AE signals can be applied in high-noise environments. However, AE technique also has its own disadvantages. To monitor subsystems of the wind turbine, it is necessary to install a large number of AE sensors, and each sensor requires an independent data acquisition system for signal sensing, processing and transferring, which increases the cost and complexity of the condition monitoring system.

Monitoring techniques based on temperature signals have been developed for fault diagnosis of gearboxes, generators, and power converters. Furthermore, temperature signals can also provide key information on the health condition of mechanical transmission system in wind turbines [9–11]. However, in previous work, the relationship between the temperature rise and the operating power has not been considered yet. Although the same temperature change can be observed whilst under different operating power, its effect on the indicative damage of the drivetrain system might be different. For example, a temperature anomaly will cause different working efficiency for the drivetrain system when working at full power and half power conditions. A working efficiency decline as low as to 0.34% for each gear stage would lead to 10 KW power loss in a 1 MW wind turbine [12–14]. In [15], the relationship between the gearbox temperature and power generation is illustrated; however, it does

not consider correcting temperature changes under different power output condition for drivetrain condition monitoring.

A new data-driven model-based method is proposed in this paper for estimating the health condition of the drivetrain in wind turbines. The model to predict output values is built based on the OS-ELM algorithm [16]. The residual signal is then obtained by comparing the predicted values with those from measurements. Compared with other artificial intelligence (AI) methods, such as artificial neural networks (ANNs) [17], support vector machines (SVMs) [18], OS-ELM has a faster training speed and a better generalization performance [19]. The residual signal produced is then further assessed by the physical kinetic energy correction model of the drivetrain, which evaluates the degree of faults by investigating the relationship between the temperature rise and power output. Finally, the Bonferroni method, a cost-effective method used to counteract the problem of multiple comparisons, is used to adjust and assess the health condition of the drivetrain. Although a physical model was proposed and applied to investigate relationships between temperature, efficiency, rotational speed and power output in [15], one major contribution of this paper is that the temperature rise is normalized at the rated power output, thus providing a more sensitive diagnosis. This paper also assesses the health condition using multivariate data analysis by taking into account both the independence of variables and the relationship among variables, which is more appropriate when modeling a practical process than the univariate analysis. Essentially, a wind turbine is a complex multivariate system, resulting in strong coupling among variables.

The remainder of this paper is organized as follows. Working principle of the online sequential extreme learning machine is presented in Section 2. Section 3 describes the physical kinetic energy correction model for the wind turbine drivetrain while Section 4 calculates the health condition of gearbox based on the Bonferroni interval. A case study using SCADA (Supervisory Control and Data Acquisition) data is then performed and the results are shown and discussed in Section 5. Section 6 contains conclusions and suggestions for further research.

2. Online Sequential Extreme Learning Machine

Extreme learning machine (ELM) algorithm was first proposed by Huang [20] for single hidden layer feed forward neural networks (SLFNs). Compared with other traditional supervised batch learning algorithms in ANNs, ELM algorithm has the advantages of faster learning and better generalization capability [21–24]. However, the ELM algorithm assumes that all training data are available before the training begins. In real cases, this assumption cannot always be satisfied, as data are available for training on a chunk-by-chunk or one-by-one basis. Thus, this paper considers using a novel sequential extreme learning machine due to its advantages below.

- (1) OS-ELM learning algorithm can receive the training data sequentially, i.e., arriving chunk-by-chunk or one-by-one.
- (2) At any time, only newly arriving data are used as training data and transferred to the learning algorithm.

Thus, the application of OS-ELM algorithm is very suitable for condition monitoring of wind turbines. Nowadays, the operation of wind turbine follows the power curve designed by the wind turbine manufacturer. As an example, a normal power curve of turbines from SCADA data is illustrated in Figure 2a; 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. When the wind speed is above the cut-out speed (25 m/s in this case), the turbine does not produce any power either because it has to be shut down to protect it from overloading. 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.

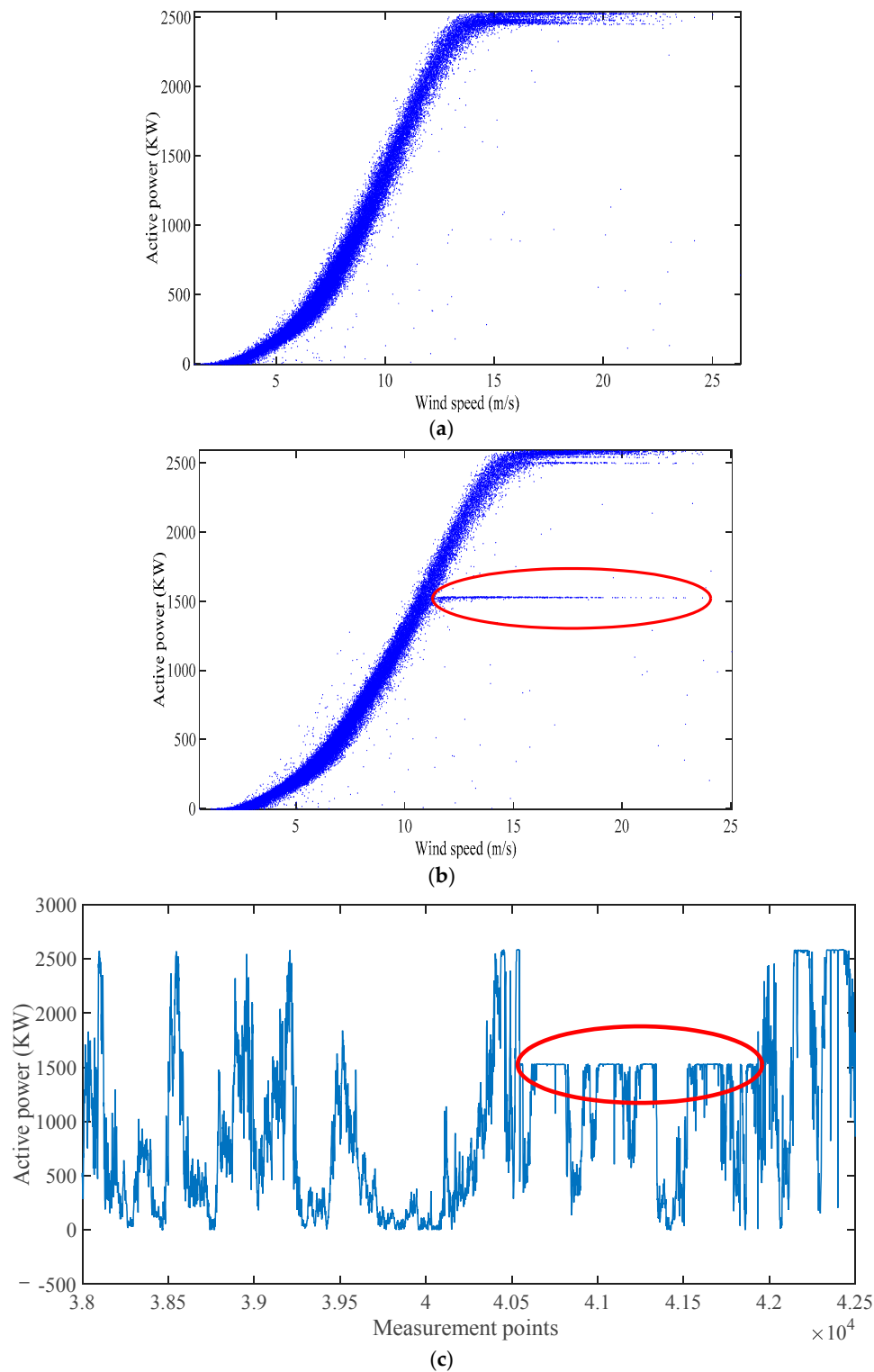


Figure 2. Examples of power curve of fault-free and faulty wind turbines: (a) power curve of the fault-free wind turbine; (b) power curve of the wind turbine with a gearbox fault; and (c) wind power time series data for the faulty turbine.

The normal power curve can represent the operation performance of a fault-free wind turbine. The change of operation performance, i.e., the change of the power curve, may indicate the onset of a turbine fault. The power curve of a wind turbine, as shown in Figure 2b, is an example of abnormal

operations. In this case, the wind turbine reduces to half of its rated power output, as shown in the red circle in Figure 2b and wind power time series data in Figure 2c, to prevent the development of more serious problems. Compared to an immediate shut-down once the fault is detected, the operation of the turbine by reducing power output would reduce the dynamic mechanical loads experienced by the turbine structure, whilst still maintaining its operation.

When the operation performance is changed, new training data should be refreshed into the prediction model to fit the new operation scenarios. Thus, the advantages of OS-ELM algorithm are able to update training data resulting from the new operation scenarios of the wind turbine. A full description of the OS-ELM algorithm is given as follows.

The schematic diagram of a single hidden-layer feed forward neural network is shown in Figure 3, which consists 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; x_1, x_2, \dots, x_n and y_1, y_2, \dots, y_m are input and output signals, respectively.

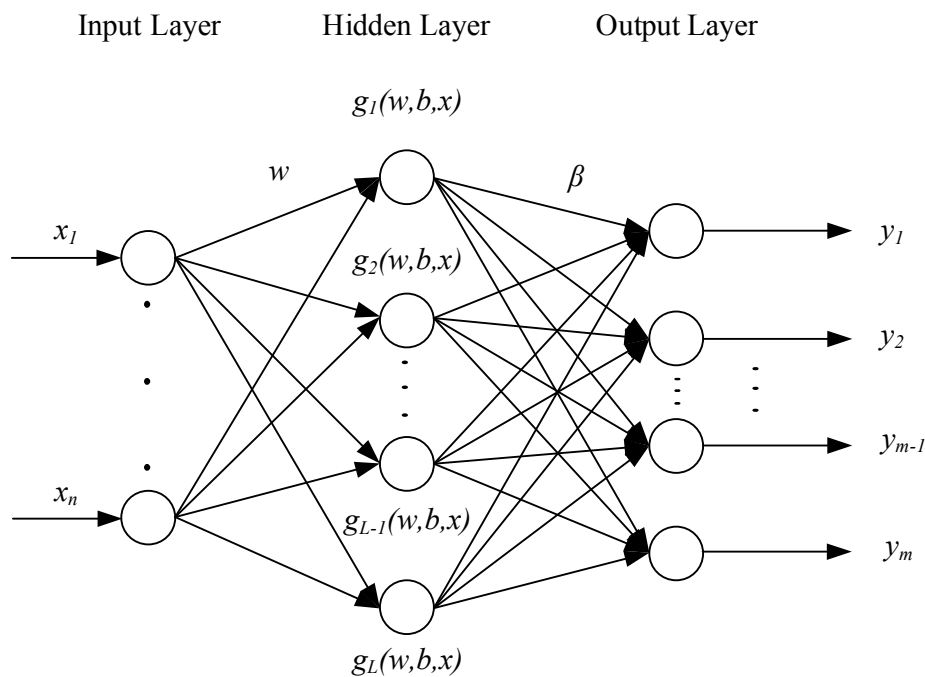


Figure 3. Schematic diagram of the single hidden-layer feed forward neural network (SLFN).

If there exists an ELM with L neurons in the hidden layer and an activation function $g(\cdot)$ can approximate the N samples with zero error, the output matrix M of ELM between inputs nodes and outputs nodes can be represented by

$$M = \begin{bmatrix} \sum_{i=1}^L \beta_{i1} g(w_i x_1 + b_i) & \sum_{i=1}^L \beta_{i2} g(w_i x_2 + b_i) & \cdots & \sum_{i=1}^L \beta_{in} g(w_i x_n + b_i) \\ \sum_{i=1}^L \beta_{i2} g(w_i x_1 + b_i) & \sum_{i=1}^L \beta_{i2} g(w_i x_2 + b_i) & \cdots & \sum_{i=1}^L \beta_{i2} g(w_i x_n + b_i) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^L \beta_{im} g(w_i x_1 + b_i) & \sum_{i=1}^L \beta_{im} g(w_i x_2 + b_i) & \cdots & \sum_{i=1}^L \beta_{im} g(w_i x_n + b_i) \end{bmatrix}_{m \times N} \quad \left(\begin{array}{l} i = 1, 2, \dots, L \\ j = 1, 2, \dots, N \end{array} \right) \quad (1)$$

where $w_i = \begin{bmatrix} w_{i1} & w_{i2} & \cdots & w_{in} \end{bmatrix}$ is the weight vector between the i th hidden node and input nodes; $\beta_i = \begin{bmatrix} \beta_{i1} & \beta_{i2} & \cdots & \beta_{im} \end{bmatrix}$ is output weight vector connecting the i th hidden

node and output nodes; $x_j = \begin{bmatrix} x_{1j} & x_{2j} & \cdots & x_{nj} \end{bmatrix}^T$ is the j th input samples; and $b_i = \begin{bmatrix} b_1 & b_2 & \cdots & b_L \end{bmatrix}^T$ represents the bias of the hidden layer matrix.

For simplicity, Equation (1) can be compactly described as,

$$H\beta = M^T \quad (2)$$

where M^T is the transpose matrix of M and H is the output matrix of the hidden layer of the ELM. The matrix H can be expressed as,

$$H = \begin{bmatrix} g(w_1 \cdot x_1 + b_1) & g(w_2 \cdot x_1 + b_2) & \cdots & g(w_L \cdot x_1 + b_L) \\ g(w_1 \cdot x_2 + b_1) & g(w_2 \cdot x_2 + b_2) & \cdots & g(w_L \cdot x_2 + b_L) \\ \vdots & \vdots & \ddots & \vdots \\ g(w_1 \cdot x_N + b_1) & g(w_2 \cdot x_N + b_2) & \cdots & g(w_L \cdot x_N + b_L) \end{bmatrix}_{N \times L} \quad (3)$$

When the input weight matrix w 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 calculated by minimizing the error function as follow,

$$\min_{\beta} \| H\beta - M^T \| \quad (4)$$

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

$$\hat{\beta} = H^+ M^T \quad (5)$$

The matrix H^+ is the generalized Moore-Penrose inverse of the matrix H , which can be found using the singular value decomposition method.

To make ELM online sequential, $\hat{\beta}$ can be transferred as follows:

$$H^+ = (H^T H)^{-1} H^T \quad (6)$$

$$\hat{\beta} = (H^T H)^{-1} H^T M^T \quad (7)$$

Suppose the training data has two sets; one is the chunk of initial training data N_0 and another is the chunk of new training data N_1 . Then, the Equation (5) can be updated to Equation (8) by minimizing the error function between two moments, where H_0 and M_0 are the output matrix of the hidden layer and the output matrix for the initial training data N_0 , while H_1 and M_1 are the output matrix of the hidden layer and the output matrix for first chunk of training data N_1 .

$$\left\| \begin{bmatrix} H_0 \\ H_1 \end{bmatrix} \beta - \begin{bmatrix} M_0^T \\ M_1^T \end{bmatrix} \right\| \quad (8)$$

The output weight matrix β that considers both initial block of training data N_0 and block of training data received in the next moment N_1 becomes

$$\beta^{(1)} = K_1^{-1} \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}^T \begin{bmatrix} M_0^T \\ M_1^T \end{bmatrix} \quad (9)$$

where

$$K_1 = K_0 + H_1^T H_1 \quad (10)$$

Therefore, the output weight matrix $\beta^{(1)}$ for the 1st chunk of training data N_1 is updated. Suppose $\beta^{(0)}$ is the output weight matrix for the chunk of initial training data N_0 , then

$$\beta^{(1)} = \beta^{(0)} + K_1^{-1} H_1^T (M_1^T - H_1 \beta^{(0)}) \quad (11)$$

As mentioned above, when a chunk of data arrives at step $K + 1$, the parameters are updated as follows:

$$K_{K+1} = K_K + H_{K+1}^T H_{K+1} \quad (12)$$

$$K_{K+1}^{-1} = K_K^{-1} - K_K^{-1} H_{K+1}^T (I + H_{K+1} K_K^{-1} H_{K+1}^T)^{-1} H_{K+1} K_K^{-1} \quad (13)$$

Using $P_{K+1} = K_{K+1}^{-1}$, the equation for $\beta^{(K+1)}$ can be updated

$$P_{K+1} = P_K - P_K H_{K+1}^T (I + H_{K+1} P_K H_{K+1}^T)^{-1} H_{K+1} P_K \quad (14)$$

The output weight matrix $\beta^{(K+1)}$ at step $K + 1$ therefore becomes

$$\beta^{(K+1)} = \beta^{(K)} + P_{K+1} H_{K+1}^T (M_{K+1}^T - H_{K+1} \beta^{(K)}) \quad (15)$$

Hence, this sequential ELM algorithm has the ability of achieving an online training in real time, if the sampling speed for updated training data is quick enough. However, it is worth noting that, in this paper, the main purpose of using OS-ELM is to achieve updated training data to ensure that the model is adapted to accommodate different operational behaviors of the wind turbines encountered during their operations. Real time online training capacity of the method is not considered in the paper. In this paper, one-year historical data are used as initial data to train initial weights. When new scene data are available, the new dataset is then transmitted to OS-ELM model to update the weights.

In our study, the updating duration is one month and the length of the data to be used is around between 4320 and 4464 points for each parameters depending on the calendar month; more information about data can be seen in the later Section 5.1. It is worth emphasizing that the data obtained from the fault-free turbine are selected for model training. Then, the data gained from a turbine with system aging only and a faulty turbine are used as the input data of the model to predict their monitoring variable output. Consequently, the established model can identify both the aging and faulty wind turbines.

As verified in our previous work [24], the ELM can learn much faster than the traditional back propagation (BP) neural network while still achieving similar model fit performance.

3. Physical Kinetic Energy Correction Model

As a key component of DFIG turbine drivetrain, the gearbox is used because turbine rotor cannot reach synchronous speed that satisfies the operating condition of DFIG generator. The use of gearbox is to transmit kinetic energy from the turbine rotor to the DFIG electric generator through the drivetrain system. The CM of temperature signal is a proven method to diagnose the faults and predict the residual life of the drivetrain system. Traditionally, a same threshold is applied to temperature monitoring regardless of the operating power, which means the same weight is assigned for the temperature change contributing to the damage of gearbox. However, the operating power could have a significant impact on the temperature changes; therefore, its effect on the temperature changes should be weighted differently.

Supposing σ is the drivetrain system efficiency, E is the input kinetic energy from the rotor to the drivetrain, P is the output kinetic energy from gearbox to generator, then $E = 1/\sigma \times P$.

Based on the first law of thermodynamics, we can have

$$Q = \left(\frac{1}{\sigma} - 1 \right) P \quad (16)$$

where Q represents the heat loss of gearbox, which leads to the temperature rise of drivetrain. If \varnothing is the compound heat transfer coefficient, the relationship between the heat loss of drivetrain Q and the gearbox temperature rise ΔT can be described by

$$Q = \varnothing \Delta T \quad (17)$$

Substituting Equation (17) into Equation (16) gives

$$\Delta T = \frac{1}{\varnothing} \left(\frac{1 - \sigma}{\sigma} \right) P \quad (18)$$

In the ideal conditions, the compound heat-transfer coefficient \varnothing is considered as constant. Figure 4 illustrates an example of the relationship between temperature rise and efficiency change of the drivetrain at different operating power outputs for a 2.5 MW wind turbine. As can be seen from the figure, the efficiency change varies at different power outputs. This implies that a fault occurring in drivetrain will lead to an increase in ΔT in response to a reduction of efficiency σ if the same power output is to be maintained. The higher the operating power output is, the smaller the efficiency varies under the same temperature change ΔT . This also means, although the faults may cause a same value of ΔT , their effects on the level of damage of drivetrain differs if the power output is different.

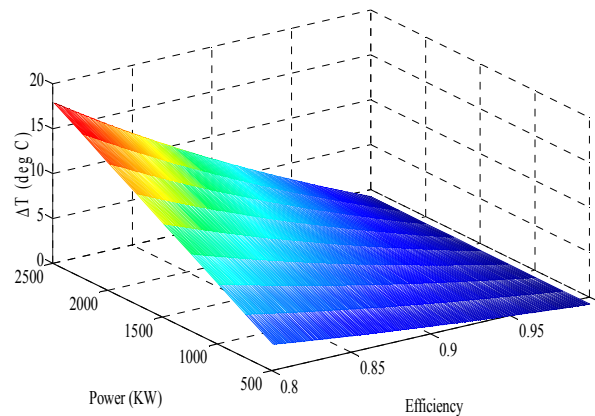


Figure 4. A 3D view of temperature rise, efficiency change and operating power of the drivetrain system.

Consequently, a temperature correction method should be considered. In our study, the temperature change ΔT is firstly obtained from the ELM model in response to the power output gained from SCADA data, and then the corresponding efficiency σ is calculated by using Equation (18). The temperature ΔT is further corrected to the value when the turbine is operating at the full power at the given efficiency. All the temperature changes at different power outputs are thus finally normalized to the value at the rated power output.

4. Estimating the Health Condition

The residual signals obtained from prediction model are now processed by the energy correction model. For the gearbox, its bearing temperature rise can be caused by either the gearbox aging or a potential fault or a failure. Figure 5 shows an example of temperature curve of gearbox bearing due to system aging in wind turbine. The temperature curve in red shows the trend of temperature

rise with active power in the wind turbine due to system aging only during first three months of one year. Temperature rises with increasing active power output. The temperature curve after six months of operation is also shown in the figure with blue color; apparently, temperature increases after the turbine operates for a period. An example of temperature rise due to a fault of the gearbox is shown in Figure 6. In this case, the temperature actually deviates from the curve randomly, indicating the onset of a fault. The fault is identified after checking the event logs recording user activities, exceptions and alarms in the SCADA system, which is related to the gearbox cooling system.

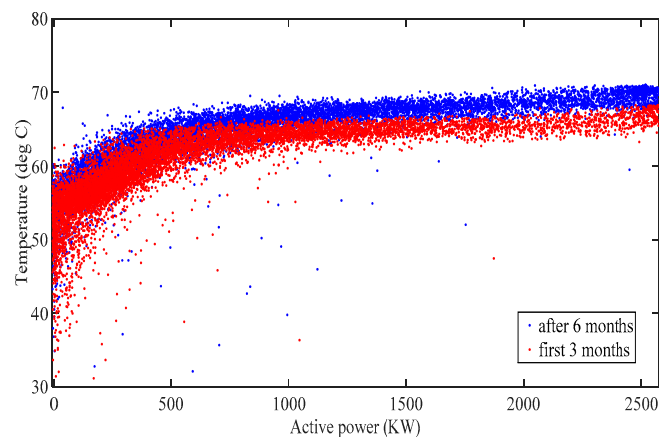


Figure 5. Gearbox bearing temperature rise trend in the first three months and after six months operation.

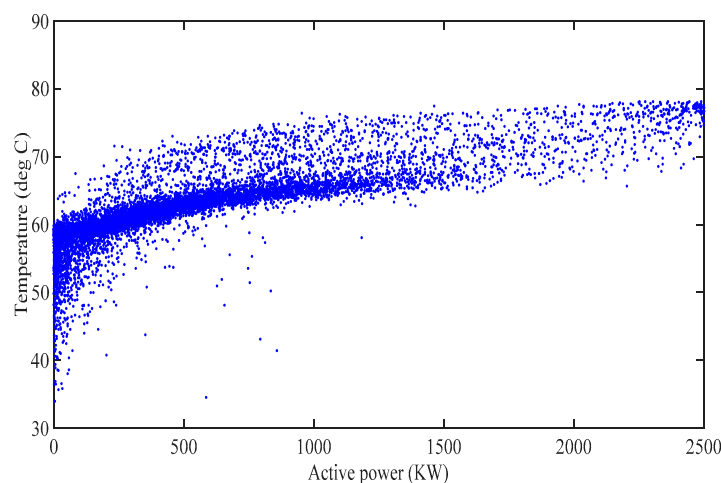


Figure 6. Gearbox bearing temperature rise trend resulting from a gearbox fault.

In the model-based CM systems, faults can be diagnosed by comparing the difference between the actually measured signal and the predicted value from the sequential extreme learning machine algorithm. Although a method relying on residual signals alone can detect faults effectively, it is not able to evaluate accurately the significance about the failure of components. Furthermore, the drivetrain in a wind turbine is generally composed of several components, which are, specifically, gears, bearings and the cooling system (usually oil cooling). Clearly, it would be desirable to use a more appropriate method in order to identify the health condition of the drivetrain by considering relationship between different components.

The estimation of system performance by using Hotelling's T-square, as described in Equation (19), has been proven to be effective in [25], where the method was used to achieve multivariate failure mode analysis of electronics. The method can provide the global information of deviation level in

wind turbines. Confidence intervals for the Hotelling's T-square method can be computed using Equation (20) and utilized to estimate deviation level for each variable.

$$T^2 = N(U - \mu)^T S^{-1} (U - \mu) \quad (19)$$

$$d_k(\alpha_k) = \bar{u}_k + \sqrt{\frac{p(N-1)}{(N-p)} F_{p,N-p}(\alpha_k)} \sqrt{\frac{s_{kk}}{N}} \quad (20)$$

where $U = (u_1, u_2, \dots, u_k)^T$ indicates a set of variables, for example, the temperatures of gearbox bearing u_1 , gearbox oil u_2 , and drivetrain main bearing u_3 in this study. The $\mu = (\bar{u}_1, \bar{u}_2, \dots, \bar{u}_k)^T$, where \bar{u}_k represents the mean value of the measurement parameter u_k . The distribution $F_{p,N-p}$ is a F distribution used in statistics; N is the number of the samples for each measurement parameter; the parameter p is the number of variables; S is the covariance matrix of U ; s_{kk} is diagonal value in the covariance matrix. In Equation (20), α_k represents the critical value for the F distribution; when α_k is determined, the value of $F_{p,N-p}$ can be found from the F distribution table [26] and the confidence interval is thus determined. It is worth noting that α_k indicates probability of occurrence of the residual signal values. If $\alpha_k < 0.01$, the monitoring data are considered to indicate a fault in the component [7], while, if α_k is larger than a particular value which can be application dependent, the component can be in a debilitating condition. In this case study, $\alpha_k = 0.25$ is selected as a threshold value for debilitating condition.

Despite of the global effect the above confidence interval demonstrates, the method lacks the ability to provide details concerning the effect of individual components on the overall operational conditions. Instead, Bonferroni intervals simply focuses on the means for each of the individual variables themselves, thereby enabling to build a more accurate confidence interval range [26]. Thus, Bonferroni intervals, as described in Equation (21), are applied in this study.

$$d_k(\alpha_k) = \bar{u}_k + t_{N-1} \left(\frac{\alpha_k}{2p} \right) \sqrt{\frac{s_{kk}}{N}} \quad (21)$$

The distribution t_{N-1} is a t distribution used in statistics and the value of t_{N-1} can be found from a t distribution table. In our study, the value of residual signal after processing with the energy correction model is considered as a fault if the value is higher than $d_k(0.01)$. Meanwhile, $d_k(0.25)$ is defined as the threshold value for debilitating condition.

5. Case Studies

5.1. SCADA Data

SCADA systems are available on most commercial wind farms, which utilize hardware and software elements and IT technologies to monitor, gather, and process data. The systems are used for a range of functions, including data acquisition, control, adjustment of parameters, and generating warning signals. To reduce the amount of data gathered from the operating wind turbines, SCADA data are usually stored at 10 min interval although sampled at the second level. SCADA data have been employed widely by researchers as the basis for CM systems and more detailed description of SCADA data and their applications can be found in Reference [27]. The SCADA data used for our study are acquired from an operational wind farm, sampled at 10 min interval, and consist of a number of parameters that contain various temperatures, pressures, vibrations, power outputs, wind speed and digital control signals. The SCADA data used in this paper cover nine months, and the data length for each parameter is approximately 38,880 points.

We select gearbox temperatures at different locations to monitor the condition of gearbox. The dataset contains temperature readings for gearbox bearing, i.e., the main speed shaft bearing connected to the rotor, the gearbox oil, i.e., the temperature of gearbox oil which is close to actual gear temperature,

and the main bearing temperature, as shown in Figure 1. Two wind turbines are selected to verify effectiveness of the proposed method, one being a turbine with system aging only and another one being a faulty turbine. To achieve an appropriate model identification, wind speed, ambient temperature and power output are selected as the inputs while the targeted temperature in the gearbox is considered as the output. This multiple-input and single-output (MISO) approach allows a more sensitive detection.

5.2. Model Predictions

This section now presents and compares the residual temperature signal produced by using OS-ELM model without and with energy correction method being applied. Figures 7 and 8 show the actual SCADA data, predicted values from the model and the residual signal of the gearbox bearing temperature for the system aging turbine and the faulty turbine respectively. As can be seen in Figure 7c, the temperature rises with time, which is caused by the system aging. Figure 9a,b illustrates residual signal of the gearbox bearing temperature due to the system aging, where the revised temperature is the residual signal processed by the physical kinetic energy correction model. Although both methods demonstrate temperature rises with time due to the system aging, the corrected gearbox bearing temperature exhibits a more obvious deviation trend than the one produced from the normal method because temperatures are converted to the values at the rated power output in the correction model. The maximum difference between corrected signal and normal signal for the turbine with system aging only can reach up to 4.7 °C. Figure 10a,b demonstrates residual signal of gearbox bearing temperature in the condition of a bearing fault. It can be seen that the gearbox bearing temperature deviates in April, indicating the onset of a fault, and reaches the highest at the beginning of September. It also appears that the corrected bearing temperature shows much clearer characteristics when the fault begins. The maximum difference between corrected signal and normal signal for the faulty turbine can reach up to 7.2 °C. With the correction model being applied, the residual signal is normalized at the rated power output, thus producing more fluctuations of the signal being monitored. This implies that the revised method has a better sensitivity than the normal method, thus facilitating a more accurate fault detection.

5.3. Fault Diagnosis

To assess further the condition of drivetrain components, the Bonferroni intervals and confidence intervals for Hotelling's T-square are compared, which are used to access each variable deviation in a multivariable system. Two sets of α values are selected for both Bonferroni intervals and confidence intervals for Hotelling's T-square. One is $\alpha = 0.01$ and the other one is $\alpha = 0.25$, as described above. The residual signal is considered as an anomaly if it is over $d_k(0.01)$ [7]. Figure 11 shows the residual signal between the actual temperatures and predicted ones of the gearbox bearing for the faulty turbine during nine months. It can be seen that the gearbox bearing temperature deviate from the prediction at April, indicating the onset of a fault. Although the residual signal value starts fluctuating between April and September, it is still within the tolerance zone. At the beginning of September, the fault leads to a dramatic increase in the residual value. The values of $d_k(0.01)$ and $d_k(0.25)$ for Bonferroni intervals using Equation (21) are 9.84 °C and 6.4 °C, respectively. Meanwhile, the values of $d_k(0.01)$ and $d_k(0.25)$ for Hotelling's T-square confidence intervals are 10.5 °C and 8 °C, respectively. Therefore, Bonferroni intervals are smaller than Hotelling's T-square confidence intervals.

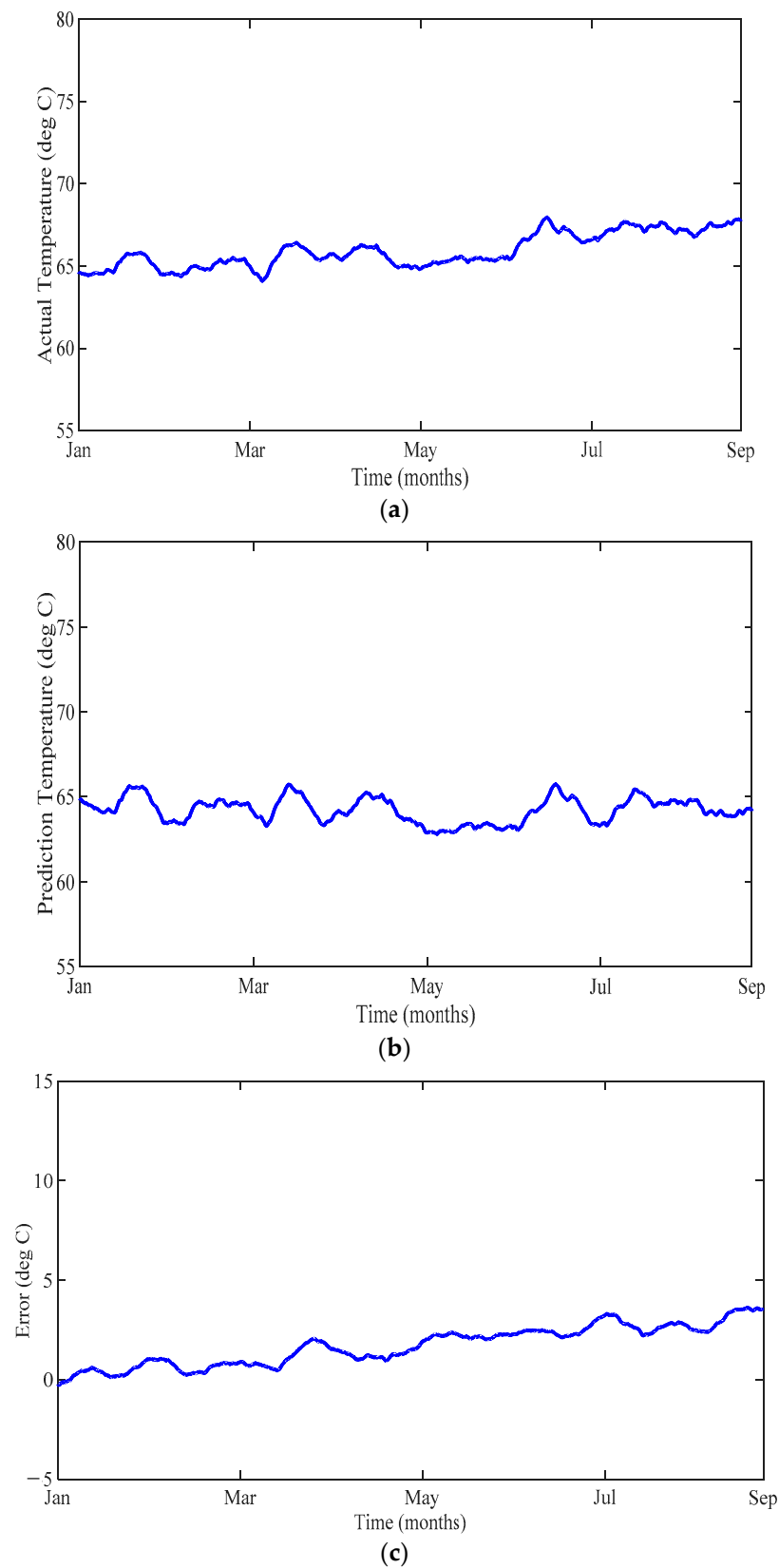


Figure 7. Model prediction compared to SCADA data for gearbox bearing temperature for the turbine with system aging only: (a) SCADA output; (b) model output; and (c) residual signal.

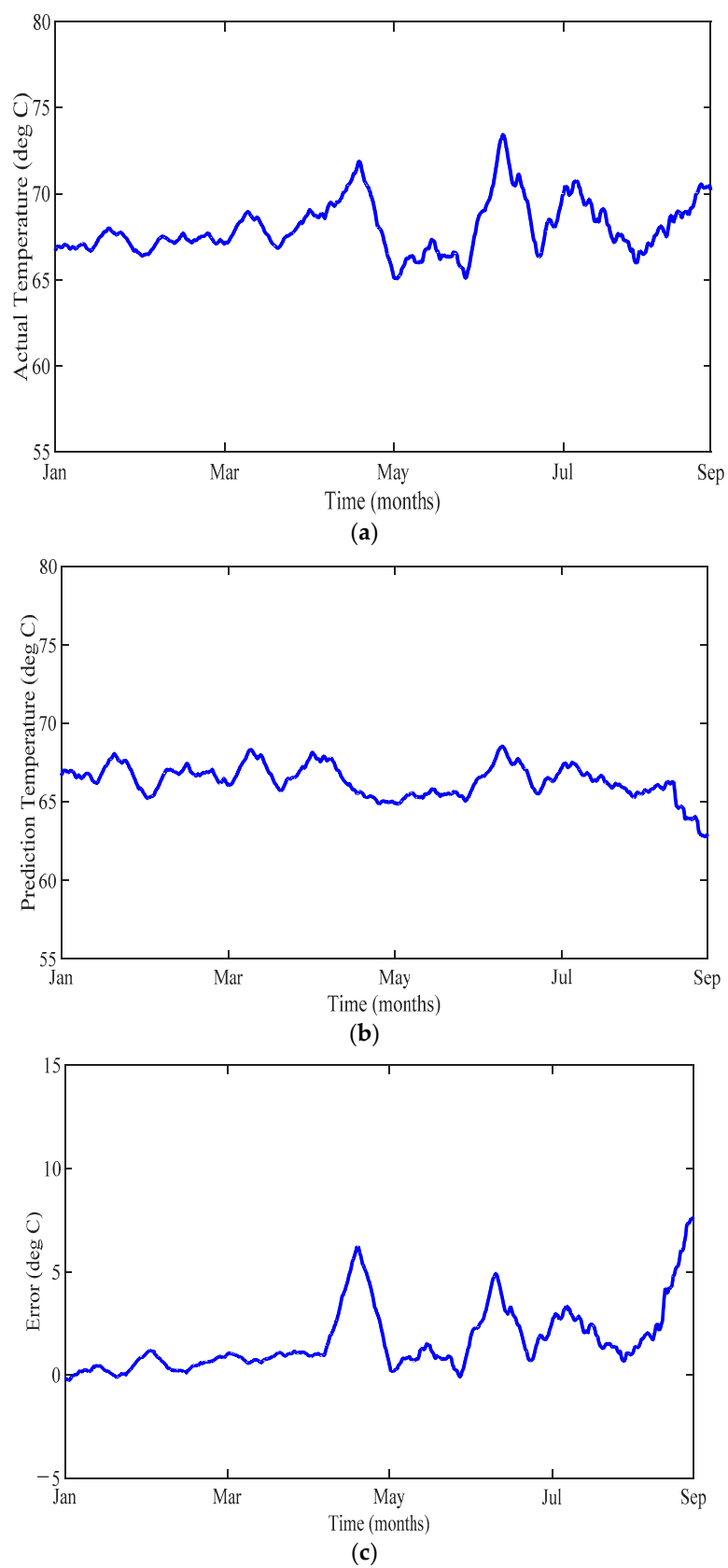


Figure 8. Model prediction compared to SCADA data for gearbox bearing temperature for the faulty turbine: (a) SCADA output; (b) model output; and (c) residual signal.

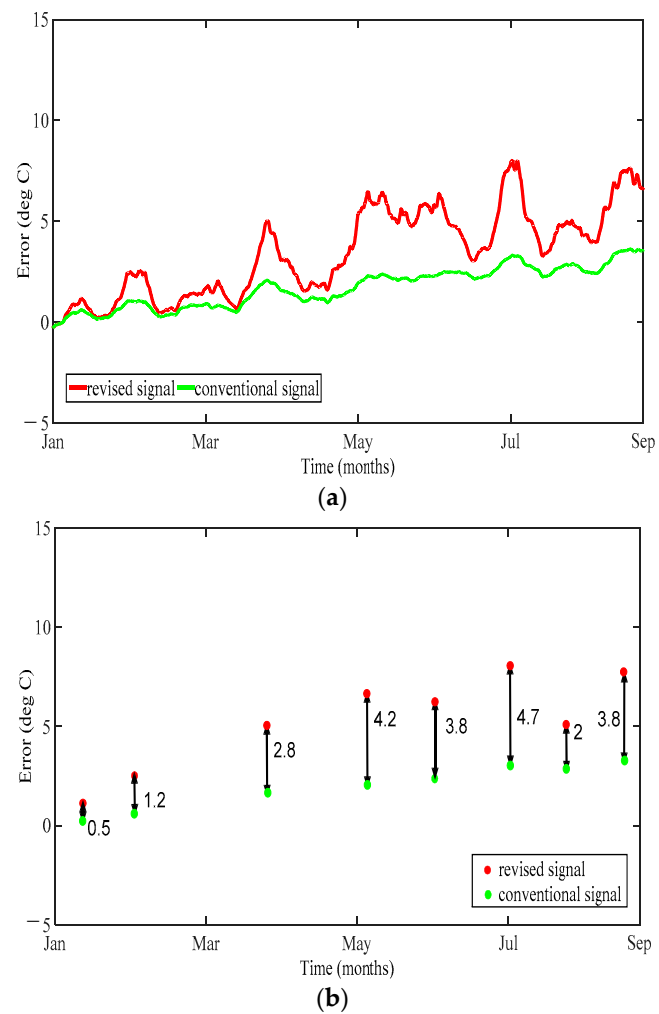


Figure 9. Residual temperatures from the ELM model with and without the energy correction model being applied the turbine due to the system aging only: (a) residual signal of the gearbox bearing temperature; and (b) temperature changes from the corrected signal and normal signal.

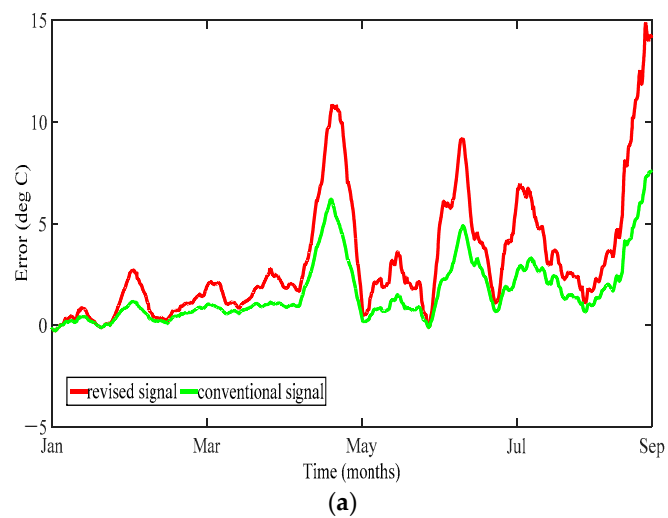


Figure 10. Cont.

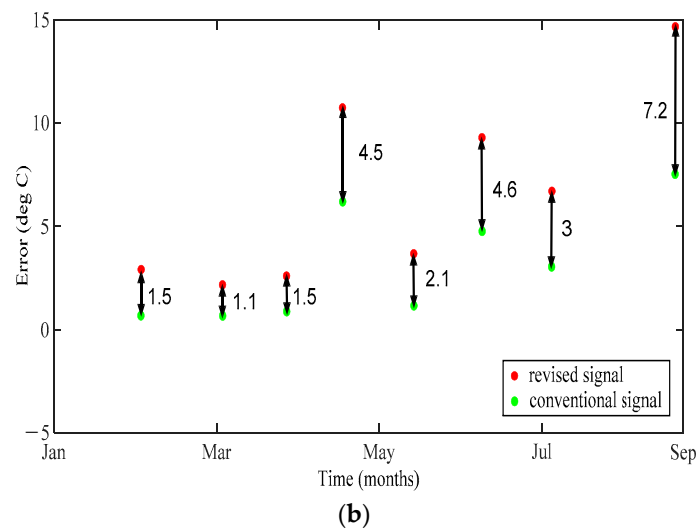


Figure 10. Residual temperatures from the ELM model with and without the energy correction model being applied for the faulty turbine: (a) residual signal of the gearbox bearing temperature; and (b) temperature changes from the corrected signal and normal signal.

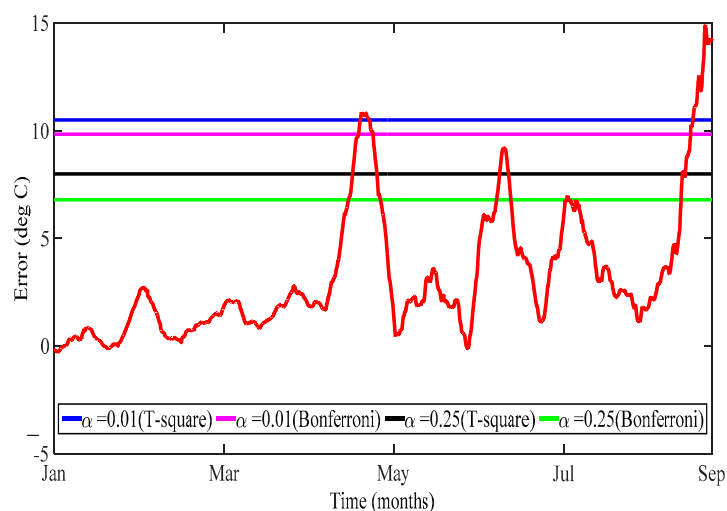


Figure 11. Comparison of Bonferroni intervals and confidence intervals for Hotelling's T-square for gearbox bearing in the faulty wind turbine.

The model predictions for gearbox oil and drivetrain main bearing using the OS-ELM model and physical kinetic energy correction model are described as follows. Figure 12a shows the residual signal between the actual temperatures and predicted ones of the gearbox oil for the faulty turbine during nine months. The values of $d_k(0.01)$ and $d_k(0.25)$ for Bonferroni intervals are 9.78 °C and 6.6 °C, respectively. It seems that the fault characteristic of the gearbox oil is the same as the gearbox bearing. Figure 12b illustrates the residual signal between the actual temperature and predicted temperature of the drivetrain main bearing. The values of $d_k(0.01)$ and $d_k(0.25)$ for Bonferroni intervals are 13.1 °C and 10.2 °C, respectively. The drivetrain main bearing deviates from the prediction between April and at the end of May with the values exceeding $d_k(0.01)$. It also begins to recover to normal temperature at the end of July. After checking the alarm log of the SCADA data, it was found that a fault in the gearbox cooling system caused the rising of temperatures of both the gearbox oil and the gearbox bearing. This result indicates that the root cause of the fault occur at the cooling system of the gearbox.

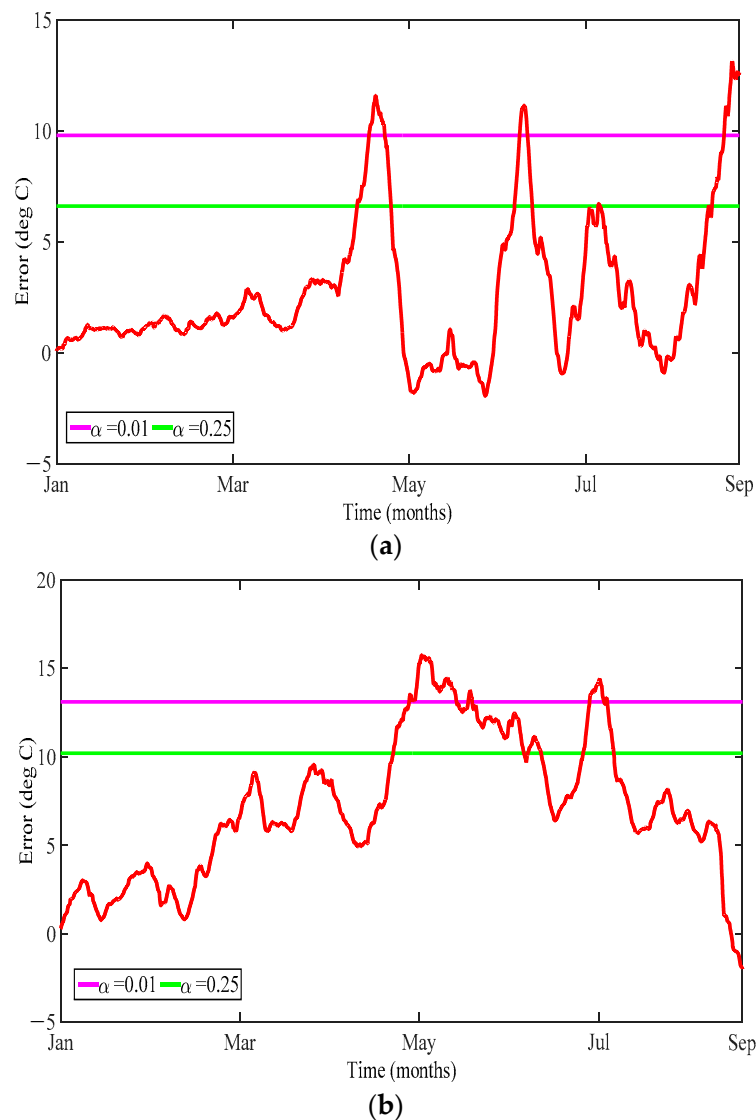
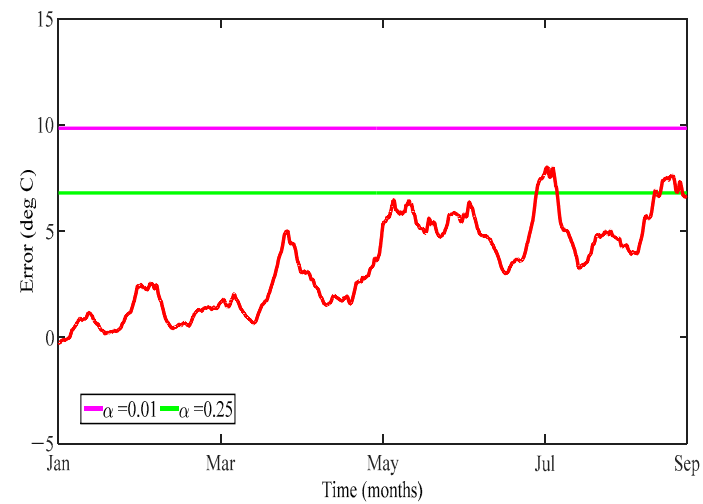
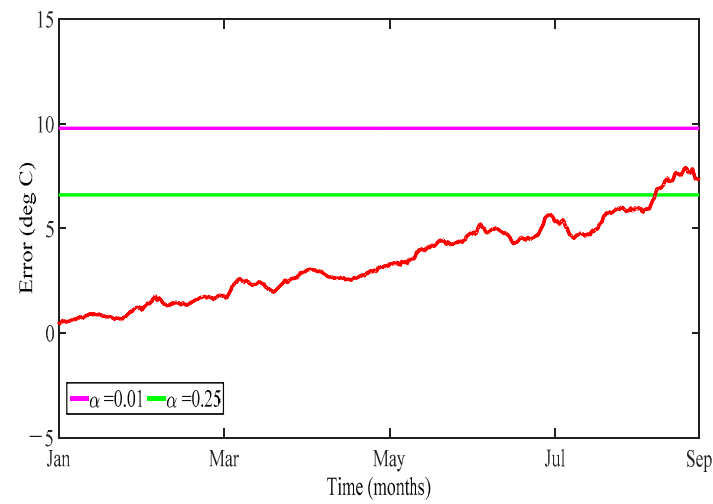


Figure 12. Residual signals of the temperature rise trend for gearbox in the faulty wind turbine: (a) residual signal of the gearbox oil temperature rise trend during nine months; and (b) residual signal of the drivetrain main bearing temperature rise trend during nine months.

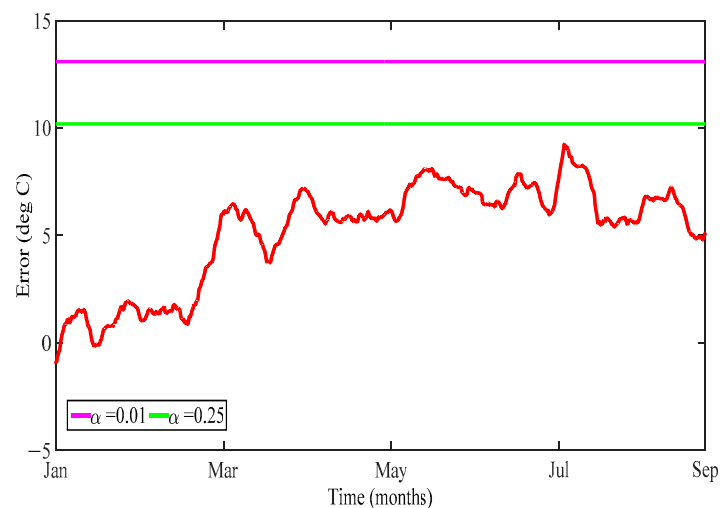
In comparison, we also assess the above temperatures for the wind turbine with aging problem only. Figure 13a shows the residual signal between the actual temperature and predicted temperature of the gearbox bearing for the aging turbine during nine months. Figure 13b illustrates the residual signal of gearbox oil temperature for the same wind turbine. It can be seen that, although the gearbox bearing and oil temperature residual signals have a rising trend, they are still below the $d_k(0.25)$, indicating the gearbox is in the healthy operational condition. The variation of power output has less effect on oil temperature than bearing temperature, because cooling oil has a higher specific heat capacity than metal. Figure 13c illustrates the residual signal between the actual temperature and predicted temperature of the drivetrain main bearing, showing it is also within a safe operational range.



(a)



(b)



(c)

Figure 13. Residual signals of the temperature rise trend in the system aging condition during nine months: (a) residual signal of the gearbox bearing temperature rise trend; (b) residual signal of the gearbox oil temperature rise trend; and (c) residual signal of the drivetrain main bearing temperature rise trend.

6. Conclusions

This paper has presented an online sequential extreme learning machine algorithm that has been applied in the model-based condition monitoring of the wind turbine drivetrain system. To demonstrate effectiveness of the proposed method, two types of representative condition (long-term system aging and short-term component fault) from two wind turbines are examined in this paper. The SCADA data, including the temperatures of gearbox oil, gearbox bearing and drivetrain main bearing obtained from an operational wind farm have been used. Fault identification is achieved by using data-driven models derived from these data. The residual signals are the difference between the actual output and the corresponding predicted value, which are caused by the system aging or the fault, thus providing an early warning of impending failure. The residual temperature signal is further revised by a physical kinetic energy correction model, which normalizes the temperature change at different power point to the value at the rated power output. The corrected values can provide a more sensitive trend indicative of signal changes. The results show that, using the Bonferroni method, a more accurate estimate of the health condition of drivetrain can be achieved, thus facilitating a more reliable fault diagnosis. Although the paper focuses on the gearbox fault detection due to the failure of the cooling system, the proposed method can be generic and used for other fault detection such as in generator bearings and windings. The temperature of these key drivetrain components in the wind turbine depend not only on the wind speed but also on the power demand from the grid. The normalization of the temperature change with regards to the rated power output can eliminate the effect of variable speed operation of the turbines, thus providing a more accurate fault detection. Future work will therefore consider online real time function of OS-ELM; and an online real time condition-monitoring device based on OS-ELM will be designed.

Acknowledgments: The authors would like to acknowledge the financial support from National Natural Science Foundation of China (Grant No. 51579222), the international collaboration and exchange program from the NSFC-RCUK/EPSRC (Grant No. 51761135011), Youth Funds of the State Key Laboratory of Fluid Power and Mechatronic Systems (Zhejiang University, No. KLoFP_QN_1604) and Fundamental Research Funds for the Central Universities (No. 2017XZZX001-02A).

Author Contributions: Peng Qian carried out most of the work presented here; Xiandong Ma revised the contents and reviewed the manuscript; and Dahai Zhang contributed to the writing and summarizing proposed ideas of Section 3.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhang, W.; Ma, X. Simultaneous Fault Detection and Sensor Selection for Condition Monitoring of Wind Turbines. *Energies* **2016**, *9*, 280. [[CrossRef](#)]
2. Yang, W.; Tavner, P.J.; Tian, W. Wind Turbine Condition Monitoring Based on an Improved Spline-Kernelled Chirplet Transform. *IEEE Trans. Ind. Electron.* **2015**, *62*, 6565–6574. [[CrossRef](#)]
3. Guo, P.; Infield, D. Wind Turbine Tower Vibration Modeling and Monitoring by the Nonlinear State Estimation Technique (NSET). *Energies* **2012**, *5*, 5279–5293. [[CrossRef](#)]
4. Yang, W.; Tavner, P.J.; Crabtree, C.J.; Wilkinson, M. Cost-Effective Condition Monitoring for Wind Turbines. *IEEE Trans. Ind. Electron.* **2010**, *57*, 263–271. [[CrossRef](#)]
5. Qiao, W.; Lu, D. A Survey on Wind Turbine Condition Monitoring and Fault Diagnosis; Part I: Components and Subsystems. *IEEE Trans. Ind. Electron.* **2015**, *62*, 6536–6545. [[CrossRef](#)]
6. Qiao, W.; Lu, D. A Survey on Wind Turbine Condition Monitoring and Fault Diagnosis; Part II: Signals and Signal Processing Methods. *IEEE Trans. Ind. Electron.* **2015**, *62*, 6546–6557. [[CrossRef](#)]
7. Bangalore, P.; Tjernberg, L.B. An Artificial Neural Network Approach for Early Fault Detection of Gearbox Bearings. *IEEE Trans. Smart Grid* **2015**, *6*, 980–987. [[CrossRef](#)]
8. Elasha, F.; Greaves, M.; Mba, D.; Fang, D. A comparative study of the effectiveness of vibration and acoustic emission in diagnosing a defective bearing in a planetary gearbox. *Appl. Acoust.* **2017**, *115*, 181–195. [[CrossRef](#)]
9. Gill, S.; Stephen, B.; Galloway, S. Wind Turbine Condition Assessment through Power Curve Copula Modeling. *IEEE Trans. Sustain. Energy* **2012**, *3*, 94–101. [[CrossRef](#)]

10. Zhang, Y.; Muljadi, E.; Kosterev, D.; Singh, M. Wind Power Plant Model Validation Using Synchrophasor Measurements at the Point of Interconnection. *IEEE Trans. Sustain. Energy* **2015**, *6*, 984–992. [[CrossRef](#)]
11. Wang, Y.; Infield, D. Supervisory control and data acquisition data-based non-linear state estimation technique for wind turbine gearbox condition monitoring. *IET Renew. Power Gener.* **2013**, *7*, 350–358. [[CrossRef](#)]
12. Fernandes, C.M.C.G.; Marques, P.M.T.; Martins, R.C.; Seabra, J.H.O. Gearbox power loss. Part I: Losses in rolling bearings. *Tribol. Int.* **2015**, *88*, 298–308. [[CrossRef](#)]
13. Fernandes, C.M.C.G.; Marques, P.M.T.; Martins, R.C.; Seabra, J.H.O. Gearbox power loss. Part II: Friction losses in gears. *Tribol. Int.* **2015**, *88*, 309–316. [[CrossRef](#)]
14. Fernandes, C.M.C.G.; Marques, P.M.T.; Martins, R.C.; Seabra, J.H.O. Gearbox power loss. Part III: Application to a parallel axis and a planetary gearbox. *Tribol. Int.* **2015**, *88*, 317–326. [[CrossRef](#)]
15. Feng, Y.; Qiu, Y.; Crabtree, C.J.; Long, H.; Tavner, P.J. Monitoring wind turbine gearboxes. *Wind Energy* **2013**, *16*, 728–740. [[CrossRef](#)]
16. Liang, N.Y.; Huang, G.B.; Saratchandran, P.; Sundararajan, N. A fast and accurate online sequential learning algorithm for feedforward networks. *IEEE Trans. Neural Netw.* **2006**, *17*, 1411–1423. [[CrossRef](#)] [[PubMed](#)]
17. Yang, S.; Li, W.; Wang, C. The Intelligent Fault diagnosis of Wind Turbine Gearbox Based on Artificial Neural Network. In Proceedings of the International Conference on Condition Monitoring and Diagnosis, Beijing, China, 21–24 April 2008.
18. Zeng, J.; Lu, D.; Zhao, Y.; Zhang, Z.; Qiao, W. Wind Turbine Fault Detection and Isolation Using Support Vector Machine and a Residual-Based Method. In Proceedings of the American Control Conference (ACC), Washington, DC, USA, 17–19 June 2013.
19. Tang, J.; Deng, C.; Huang, G.-B.; Zhao, B. Compressed-Domain Ship Detection on Spaceborne Optical Image Using Deep Neural Network and Extreme Learning Machine. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 1174–1185. [[CrossRef](#)]
20. Huang, G.; Zhu, Q.; Mao, K.Z.; Chee-Kheong, S.; Saratchandran, P.; Sundararajan, N. Can threshold networks be trained directly? *IEEE Trans. Circuits Syst. II Express Briefs* **2006**, *53*, 187–191. [[CrossRef](#)]
21. Golestaneh, F.; Pinson, P.; Gooi, H.B. Very Short-Term Nonparametric Probabilistic Forecasting of Renewable Energy Generation 2014—With Application to Solar Energy. *IEEE Trans. Power Syst.* **2016**, *31*, 3850–3863. [[CrossRef](#)]
22. Lu, X.; Liu, C.; Huang, M. Online Probabilistic Extreme Learning Machine for Distribution Modeling of Complex Batch Forging Processes. *IEEE Trans. Ind. Inform.* **2015**, *11*, 1277–1286. [[CrossRef](#)]
23. Sun, Z.-L.; Ng, K.M.; Soszynska-Budny, J.; Habibullah, M.S. Application of the LP-ELM Model on Transportation System Lifetime Optimization. *IEEE Trans. Intell. Transp. Syst.* **2011**, *12*, 1484–1494. [[CrossRef](#)]
24. Qian, P.; Ma, X.; Cross, P. Integrated data-driven model-based approach to condition monitoring of the wind turbine gearbox. *IET Renew. Power Gener.* **2017**, *11*, 1177–1185. [[CrossRef](#)]
25. Lall, P.; Gupta, P.; Goebel, K. Decorrelated Feature Space and Neural Nets Based Framework for Failure Modes Clustering in Electronics Subjected to Mechanical Shock. *IEEE Trans. Reliabil.* **2012**, *61*, 884–900. [[CrossRef](#)]
26. Krishnamoorthy, K.; Mathew, T. *Statistical Tolerance Regions: Theory, Applications, and Computation*; Wiley: Hoboken, NJ, USA, 2009.
27. Tavner, P.J. *Offshore Wind Turbines: Reliability, Availability & Maintenance*; IET: London, UK, 2012.

