Article

A Short-Term Outage Model of Wind Turbines with Doubly Fed Induction Generators Based on Supervisory Control and Data Acquisition Data

Peng Sun, Jian Li *, Junsheng Chen and Xiao Lei

The State Key Laboratory of Power Transmission Equipment and System Security and New Technology, College of Electrical Engineering, Chongqing University, Chongqing 400044, China; sunpeng@cqu.edu.cn (P.S.); cjsgiant@126.com (J.C.); cquleixiao@163.com (X.L.)
* Correspondence: lijian@cqu.edu.cn; Tel.: +86-139-0835-5616

Abstract: This paper presents a short-term wind turbine (WT) outage model based on the data collected from a wind farm supervisory control and data acquisition (SCADA) system. Neural networks (NNs) are used to establish prediction models of the WT condition parameters that are dependent on environmental conditions such as ambient temperature and wind speed. The prediction error distributions are discussed and used to calculate probabilities of the operation of protection relays (POPRs) that were caused by the threshold exceedance of the environmentally sensitive parameters. The POPRs for other condition parameters are based on the setting time of the operation of protection relays. The union probability method is used to integrate the probabilities of operation of each protection relay to predict the WT short term outage probability. The proposed method has been used for real 1.5 MW WTs with doubly fed induction generators (DFIGs). The results show that the proposed method is more effective in WT outage probability prediction than traditional methods.

1. Introduction

In consideration of the intermittency and randomness of wind power, the large-scale wind power integration has a great influence on the safe and stable operation of the electrical power system [1,2]. In view of this, the electrical power system in China has clearly claimed to bring the active output control of the wind farm into the power network dispatching management [3]. Currently, the electric power system dispatching is mainly based on the short-term wind speed and wind power prediction [4–6], which fails to consider the impact of wind turbine (WT) outages on the wind power variation. Meanwhile, there is still a lack of warning information of the WT outages to support the internal coordination control technologies of the wind farm [7,8]. In fact, due to the adverse natural environment, the outage rate of the WT is far higher than that of the traditional electric transmission and transformation equipment. Therefore, the WT short-term operational reliability evaluation is an important foundation of the optimal dispatching of electrical power systems and the economic operation of wind farms.

The short-term operational reliability evaluation of the power system equipment could provide a quantitative description of short-term outage risks of the equipment in the future (minute, hour and day) [9,10]. Because of the uncertainty of wind speed and wind direction, the operation points of the WT change quite significantly in the short term. Statistics show that the failure rate of the WT equipment has a strong correlation with the external environment [11,12]. Furthermore, the

WT outages are also influenced by factors such as equipment aging, maintenance strategies, and geographical environment. Some studies have been conducted to evaluate the reliability of WTs and the important WT subassemblies [12–14]. However, these long-term reliability evaluation methods cannot reflect the impacts of various operation conditions on WT reliability. It is difficult to accurately estimate the short-term WT outage probability based solely on the statistical data of WT faults.

Compared with the traditional electric transmission and transformation equipment, the WTs have abundant condition monitoring data, which could provide comprehensive condition information for the WT short-term operational reliability evaluation. The condition monitoring data of WTs are mainly obtained from the wind farm supervisory control and data acquisition (SCADA) system. The SCADA system can provide a large amount of monitoring parameters of the natural environment, equipment operation and electricity system. To ensure the safe operation of WTs, once any parameters exceed their threshold values of protection relays, the WT will be shut down. The short-term change in the SCADA monitoring parameters are closely related to the external environment and operation condition of WTs. By using advanced SCADA data mining methods, various condition parameter prediction models have been developed to detect the significant changes in WT behavior prior to fault occurrences [15–19]. These prediction models were established by employing data-driven approaches that involve neural networks (NNs) [20–22], a support vector machine (SVM) [23], adaptive neuro-fuzzy inference system (ANFIS) [24], and the nonlinear state estimate technique (NSET) [25]. In [26], different WT performance curves, such as the power curve, rotor curve, and blade pitch curve were modeled for monitoring the performance of WTs. A nonlinear data-based modeling approach was proposed in [27] to detect anomalies in WT generator winding and gearbox bearing. In [28], a NN based normal behavior model of generator bearing temperature was developed to analyze bearing faults in WTs. A comparative analysis of two NN-based models and a regression-based model was presented in [29] to detect anomalies in gearbox bearing temperature and generator stator temperature.

The above studies are mainly for anomaly identification of WT condition parameters. These WT anomaly identification models could provide a qualitative description of the WT health conditions. However, since these methods do not consider the WT outage risks due to the anomalies in WT condition parameters, it is difficult to quantitatively evaluate the short-term operational reliability of WTs. This challenge lead to the discovery of a new approach to predict the short-term WT outage probability based on comprehensive SCADA monitoring parameters.

Currently, the wind farm operators usually use a power curve to estimate the expected WT output power based on the predicted wind speeds. However, in order to reduce the outage probability and the number of outages of WTs, power generation schedule for the low-reliability WTs with higher outage probabilities should be adjusted by performing the active output power control of WTs. Hence, the short-term WT outage probability prediction is useful for the wind farm operators to monitor the WT condition and make proper active output power control strategies. Furthermore, since the WT outages could impact the output power and power variation of a wind farm, the wind farm operators should evaluate the potential wind power loss due to the WT outages according to the WT outage probability prediction results and formulate appropriate power generation schedules for WTs and wind farms.

This paper presents a short-term (i.e., 15-min-ahead) WT outage model based on condition parameters obtained from the wind farm SCADA system. Prediction models are established to predict condition parameters that are dependent on environmental conditions such as ambient temperature and wind speed. The calculation for probabilities of protection operation caused by the threshold exceedance of the component temperatures is based on the probability distributions of the prediction errors of the developed prediction models. The calculation for probabilities of protection operation for other monitoring parameters is based on the setting time of the protection operation. The union probability method is utilized to integrate the probabilities of each protection relay to predict the WT short term outage probability. The proposed method has been used for real 1.5 MW WTs with doubly fed induction generators (DFIGs).

The remainder of the paper is organized as follows: The SCADA data that can be used for the WT outage model development are discussed and grouped in Section 2. Section 3 proposes the methodology for developing the short-term outage model. The details in developing condition parameter prediction models and the calculation for POPR are given in Sections 4 and 5. The procedure for predicting the short-term WT outage probability is given in Section 6. Three cases are investigated to validate the proposed short-term WT outage model in Section 7. The conclusions are given in Section 8.

2. Parameter Description and Classification

Table 1 shows the typical condition parameters measured and delivered by the SCADA system of a wind farm. The positions of corresponding sensors are shown in Figure 1. These condition parameters can be grouped into two types:

Type 1 parameters include various component temperatures which are strongly influenced by environmental conditions. For example, the relationship between the gearbox input shaft temperature of a variable speed constant-frequency (VSCF) WT and the wind speed is shown in Figure 2. For the VSCF WT, when the wind speed is below its rated speed, a faster rotational speed, due to a higher wind speed, will evidently raise the temperature of the mechanical components. When the wind speed is over the rated wind speed, the WT will be kept at its rated output power by variable pitch control and the component temperatures will be less affected by the wind speed. Figure 3 shows the relationship between the gearbox input shaft temperature and the ambient temperature. The heat dissipation of mechanical components of a WT is also affected by the ambient temperature. At a lower ambient temperature, the heat dissipation is faster and the component temperatures can vary in a larger range. Conversely, at a higher temperature, the component temperature tolerable variation range is smaller with slower heat dissipation. It is noted that the gearbox will produce a large amount of heat in the operation of WTs [30]. In order to ensure the secure and stable operation of WT, an effective gearbox cooling system has to be implemented. Therefore, the gearbox cooling water temperature, which has a large impact on the gearbox component temperatures, is also considered as one of Type 1 condition parameters. Moreover, wind speed has a major impact on the wind power output and WT component temperatures, which is also classified into Type 1 parameters in this paper.

Type 2 parameters include yaw position, yaw angle error, and hydraulic oil pressure. Yaw position represents the rotation angle of the nacelle. Yaw angle error is the angle between the wind and the nacelle position. Type 2 parameters do not have an obvious relationship with environmental conditions.

Table 1. Wind turbine (WT) condition parameters studied in this paper. Temp.: Temperature.

Number	WT Condition Parameters	Normal Range		Type
		Upper Limit	Lower Limit	
1	Temp. of gearbox input shaft	100 °C	-	1
2	Temp. of gearbox output shaft	100 °C	-	1
3	Temp. of gearbox oil	80 °C	10 °C	1
4	Temp. of gearbox cooling water	50 °C	-	1
5	Temp. of main bearing *a* (on the rotor side)	70 °C	-	1
6	Temp. of main bearing *b* (on the gearbox side)	70 °C	-	1
7	Temp. of generator winding (U∣V∣W)	165 °C	-	1
8	Temp. of generator bearing *a* (front)	95 °C	-	1
9	Temp. of generator bearing *b* (back)	95 °C	-	1
10	Temp. of generator cooling air	50 °C	35 °C	1
11	Temp. of control cabinet	55 °C	-	1
12	Wind speed	25 m/s	3 m/s	1
13	Yaw position	750°	−750°	2
14	Yaw angle error	15°	−15°	2
15	Hydraulic oil pressure for yaw	150 bar	135 bar	2
16	Hydraulic oil pressure for rotor brake	120 bar	-	2

The SCADA data used in this paper are obtained from an onshore wind farm in northern China. The wind farm has 34 1.5 MW WTs with doubly fed induction generators (DFIGs), labeled WT 1 to WT 34. All the WTs are the same type in the wind farm. The SCADA data have been collected since 15 February 2011. The sampling rate of the SCADA system is one sample per second, i.e., 1 S/s.
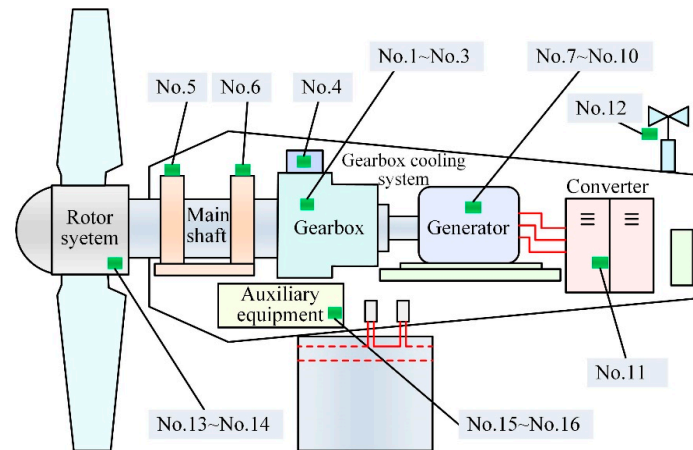

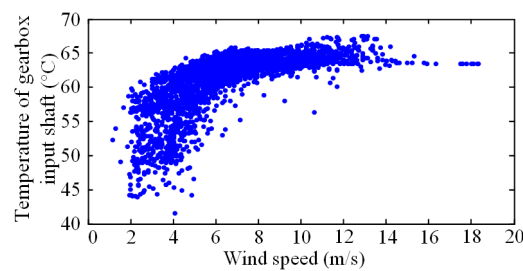**Figure 1.** The main components and sensor positions of the considered wind turbine (WT).

Figure 2. Relationship between the gearbox input shaft temperature of a variable speed constant-frequency (VSCF) WT and the wind speed.
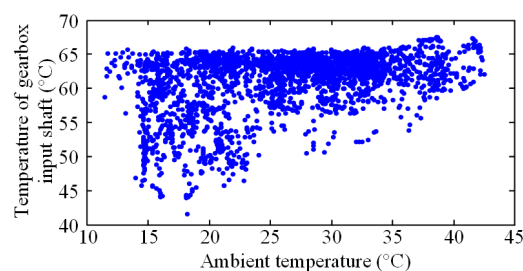

Figure 3. Relationship between the gearbox input shaft temperature of a VSCF WT and the ambient temperature.

3. Framework of the Wind Turbine Outage Model

Most of the wind farm SCADA monitoring parameters reflect the current condition of WTs and power grid. In order to assure the safe operation of WTs, several outage protections are designed for the WTs. Once any monitoring parameters exceed their threshold value, the corresponding protection relays will be activated to stop WTs. Hence, based on the logical relationship between the WT outage and the operation of each protection relay, the short-term operational reliability of WTs can be evaluated by applying reliability theory of the series networks [31]. Therefore, the WT outage probability can be calculated by using the union probability method [31,32] as follows:

$$F = 1 - \prod_{i=1}^{N} (1 - F_i) \tag{1}$$

where F_i represents the probability of the operation of the protection relay (POPR) for condition parameter i and N is the number of condition parameters.

The POPR for Type 1 parameters: Figure 4 shows the framework of calculating the POPR for the environmentally sensitive parameters (i.e., Type 1 parameters). First, the short-term wind speed prediction model is established to predict the wind speed in the next 15 min. Since the WT outages caused by the lower wind speeds have little influence on the WT output power fluctuation, the WT outage probability due to the wind speed lower than the cut-in wind speed is not considered. Then, the 15-min-ahead prediction models are developed to predict the Type 1 parameters at each predicted value of wind speed. Finally, the threshold exceedance probabilities for Type 1 parameters are calculated based on the prediction error distribution of the wind speed prediction model and Type 1 parameter prediction models. Since the protection relays of Type 1 parameters do not have the time delay for the protection action, the threshold exceedance probabilities of the Type 1 parameters are approximate to the POPR.

The POPR for Type 2 parameters: It is difficult to predict the threshold exceedance probabilities of Type 2 parameters by using the environmental parameters such as wind speed and ambient temperature. In this paper, the POPRs for Type 2 parameters are estimated based on operation principles of the protection relays.
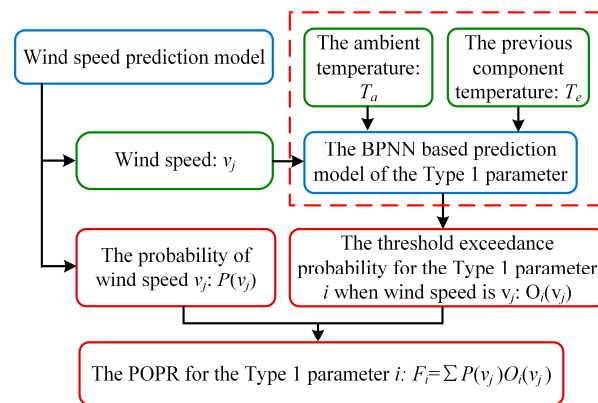

Figure 4. The calculation method of the probabilities of the operation of the protection relays (POPR) for the environmentally sensitive parameters.

4. The Probabilities of the Operation of the Protection Relays for Type 1 Parameters

4.1. Wind Speed Prediction Model

Various studies have been done on the short-term wind speed prediction, however, the current prediction models still have non-negligible prediction errors. To accurately calculate the POPR for the Type 1 parameters, the error distribution of the short-term wind speed prediction model should be analyzed. Most study results indicate that the wind speed prediction errors follow the normal distribution [33,34]. In this paper, the autoregressive moving average (ARMA) model is used for 15-min-ahead wind speed prediction and the normal distribution parameters of prediction errors are obtained. The 10-min averaged wind speed data are used to develop the prediction models. Figure 5 shows the normal distribution based probability density function (PDF) of the wind speed prediction errors. Based on this, the prediction errors are divided into nine regions (i.e., $-\infty$ to -1.75 m/s, -1.75 m/s to -1.25 m/s, ... , 1.75 m/s to ∞), as seen in Figure 5. To simplify the calculation of POPR for the Type 1 parameters, the wind speed prediction errors are discretized into nine discrete values (i.e., 0 m/s, ± 0.5 m/s, ± 1 m/s, ± 1.5 m/s, ± 2 m/s) corresponding to the nine regions. The probability for each discrete prediction error can be calculated based on (2) and the results are shown in Table 2.

$$P(\xi = k) = \begin{cases} F_{Nw}(k + 0.25) - F_{Nw}(k - 0.25) & k = 0, \pm 0.5, \pm 1, \pm 1.5 \\ 1 - F_{Nw}(k - 0.25) & k = 2 \\ F_{Nw}(k + 0.25) & k = -2 \end{cases} \qquad (2)$$

where ξ is the discrete prediction error of wind speed and F_{Nw} is the cumulative distribution function (CDF) of the wind speed prediction error.
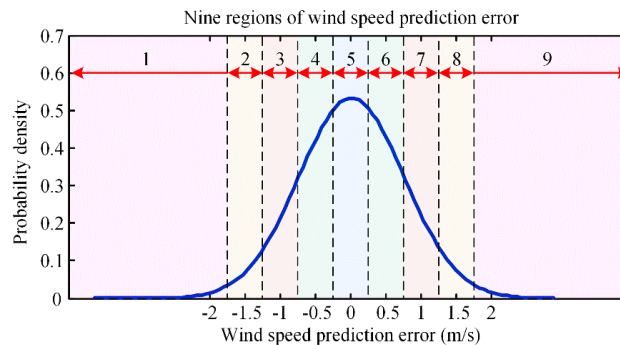

Figure 5. Nine regions of wind speed prediction errors.

Table 2. The probability for each discrete wind speed prediction error.

Discrete Prediction Error (m/s)	Probability
0	0.234
±0.5	0.197
±1	0.118
±1.5	0.05
±2	0.018

4.2. Prediction Models of Component Temperature Parameters

The 15-min-ahead prediction models for component temperature parameters are established in this section. To develop the prediction models, the following three factors have been carefully considered: (1) input parameters; (2) training algorithm; and (3) sample data.

Input parameters: For the VSCF WTs, the wind speed is one of the key factors that influence the component temperatures [35]. Faster rotational speed due to higher wind speed will evidently raise the mechanical component temperatures. Consequently, the output power and the electrical component temperatures increase [35]. Moreover, the ambient temperature can impact on the heat dissipation efficiency of the WT components and thus influence the component temperatures. Therefore, based on the operating characteristics of the VSCF WTs, the wind speed $v(t)$, the previous ambient temperature $T_a(t-1)$ and the previous component temperature $T_e(t-1)$ are selected as input parameters of the prediction models of component temperature parameters $T_e(t)$.

Training algorithm: Back-propagation NN (BPNN), a widely used NN for prediction, is applied in this study for the prediction of the component temperature parameters. In consideration of the size of the training set and the training time, the NN is chosen to have only one hidden layer and the number of neurons in the hidden layer ranges from 2 to 10. The transfer function used in the hidden layer is tan-sigmoid while the output layer transfer function is log-sigmoid based. The actual number of nodes (i.e., neurons) in the hidden layer is determined by trial-and-error to find the right number at which the NN has the best generalization performance. The relationship between the BPNN prediction performances of four component temperature parameters (i.e., gearbox input shaft temperature, generator bearing *a* temperature, generator winding U temperature and main bearing *a* temperature) of a WT and the number of hidden layer nodes (i.e., neurons) are taken as an example, as shown in Figure 6. Based on this, the optimal hidden layer node numbers of the four prediction models are chosen as 10, 8, 7, and 6 respectively.
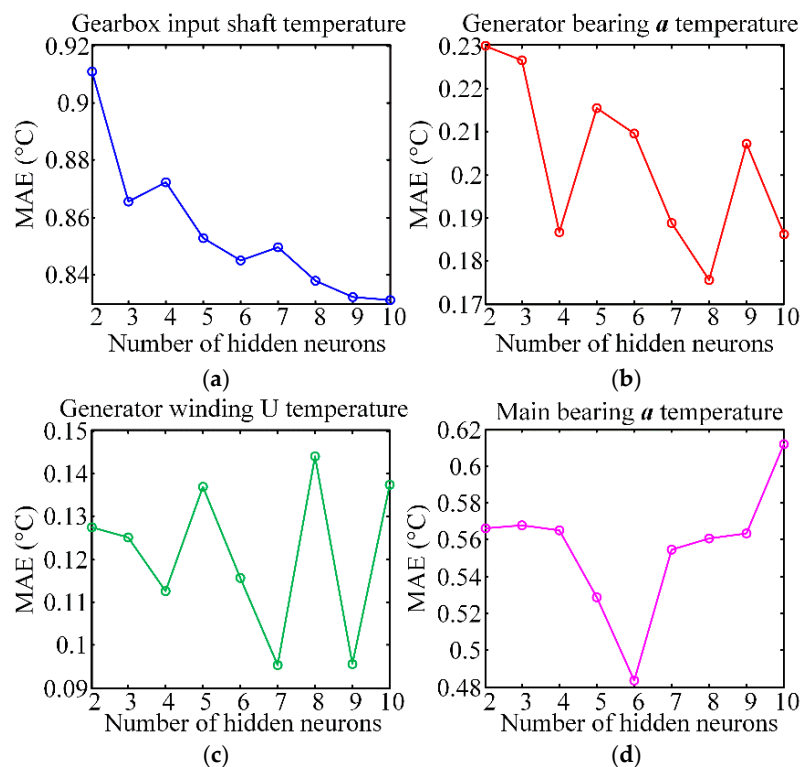
Figure 6. The back-propagation neural network (BPNN) performances over the numbers of hidden nodes. (**a**) Gearbox input shaft temperature; (**b**) Generator bearing *a* temperature; (**c**) Generator winding U temperature; and (**d**) Main bearing *a* temperature.

Data sampling and data pre-processing: There are no generalized rules for selecting training samples of the WT condition parameter prediction models. Two points have been carefully considered to select the sample data. Firstly, the training samples must be collected when the WT is under normal condition. Secondly, ambient temperature and wind speed vary with the seasons. Figure 7 shows the boxplot of ambient temperature and wind speed of four seasons in a year. The data are collected from the SCADA dataset of a WT. The median, 25th and 75th percentiles are shown with boxes, while the maximum, minimum and outliers are shown with whiskers and crosses, respectively. It is clear that there are significant differences among the data distributions of the four seasons.
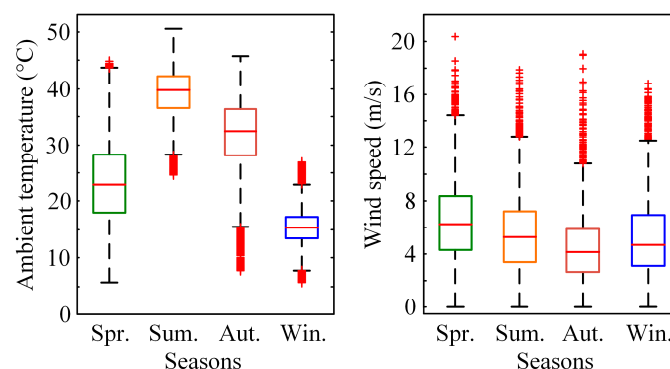

**Figure 7.** Boxplot of ambient temperature and wind speed of four seasons.

Table 3 shows the training and test data for the prediction models of the Type 1 parameters. The samples are collected from the SCADA data in 2011 and 2012. The 10-min averaged SCADA data of each season are used to develop the prediction models of component temperature parameters.

Table 3. Data sampling for the prediction models of Type 1 parameters.

Seasons	Sample Data		Sample Type
	Start Time	Stop Time	
Spring	15 February 2011	23 April 2011	training
	24 April 2011	30 April 2011	test
Summer	1 May 2011	24 July 2011	training
	25 July 2011	31 July 2011	test
Autumn	1 August 2011	24 October 2011	training
	25 October 2011	31 October 2011	test
Winter	1 November 2011	24 January 2012	training
	25 January 2012	31 January 2012	test

When the wind speed varies within its low value region, it may cause frequent WT startup or shutdown. Thus, the data collected during these periods of time have different stochastic characteristics and should be filtered out. To solve this problem, a lower limit of output power is set at 100 kW for the 1.5 MW WTs for sample data selection. In other words, only the data when the output power is greater than 100 kW are used to train the NNs.

Furthermore, during certain instances in the operation of a WT, the output power can be curtailed by adjusting the blade pitch angle. Figure 8 shows the measured power curve of a WT during two time periods (i.e., 1–30 May 2011 and 1–5 June 2011) for example. The power curve in May is considered to be in normal condition. However, the WT output power between 1 and 5 June was curtailed through increasing the blade pitch angle, which can be seen in Figure 9. This reduction in output power for a particular wind speed is due to change in control by the pitch regulation module and is not related to a degraded condition of the WT [36]. Moreover, the prediction models could not provide an accurate prediction since few SCADA data for this situation can be collected. Consequently, the data instances in which the WT pitch control mechanism was engaged (based on the pitch blade angle reading) are filtered out.
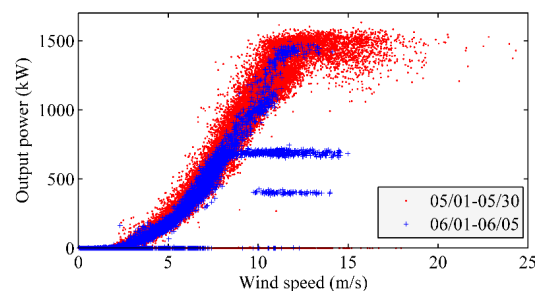
Figure 8. Relationship between WT output power and wind speed during two time periods.
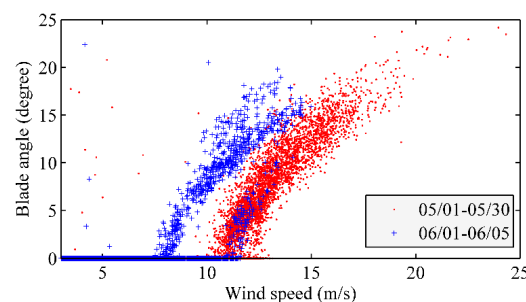
Figure 9. Relationship between pitch blade angle and wind speed during two time periods.

Furthermore, scaling and missing data processing are performed before training the BPNNs.

(a) Data scaling: Data scaling is an important step when training the BPNN based prediction models. The data scaling method can be expressed as follows:

$$y_s = \frac{y - y_{\min}}{y_{\max} - y_{\min}} \tag{3}$$

where y is a variable and y_s is the corresponding normalized variable.

(b) Missing data processing: Unknown or missing values are particularly harmful when training the BPNNs. If they are not classified as not a number (i.e., NaN) but 999 for example, the BPNN may try to fit these values. This will reduce the prediction accuracy of the BPNN models. Hence, it is necessary to clarify how the missing values are labeled in the SCADA system. Since there is a large amount of SCADA data available, no approximation was performed in this paper. Instead missing input and target values are processed by neglecting the data set [21].

Prediction performance analysis: To analyze the performance of the prediction models, metrics such as mean absolute error (*MAE*), mean squared error (*MSE*), and mean absolute percentage error (*MAPE*) have been used, as shown in (4)–(6).

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |\hat{y}_i - y_i| \tag{4}$$

$$MSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2} \tag{5}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\% \tag{6}$$

where n is the number of test samples, $\hat{y}_i$ is the predicted value for a time period i, and y_i is the measured value at the same time.

Table 4 shows the prediction performance of the component temperature parameters of a WT. Although the prediction accuracy of each component temperature is different, the maximum of *MAE* is lower than 1 °C. The trained models are tested by using the test data in Table 3. The error residuals and prediction error distributions for the gearbox input shaft temperature, the main bearing *a* temperature, the generator bearing *a* temperature and the generator winding U temperature are given as an example, as shown in Figures 10 and 11. It can be seen that the prediction errors coming from the successfully trained prediction models are normally distributed with a mean around zero.

Table 4. Prediction performance of the component temperature parameters (i.e., Type 1 parameters). *MSE*: mean squared error; *MAE*: mean absolute error; *MAPE*: mean absolute percentage error.

| Condition Parameters	Index		
	MSE (°C)	*MAE* (°C)	*MAPE* (%)
Temp. of gearbox input shaft	0.62	0.46	0.77
Temp. of gearbox output shaft	0.85	0.62	1.03
Temp. of gearbox oil	0.66	0.39	0.83
Temp. of gearbox cooling water	0.77	0.51	0.95
Temp. of main bearing *a*	0.19	0.13	0.34
Temp. of main bearing *b*	0.13	0.09	0.22
Temp. of generator winding U	0.83	0.57	1.07
Temp. of generator winding V	0.77	0.48	0.86
Temp. of generator winding W	0.80	0.43	0.74
Temp. of generator bearing *a*	0.36	0.19	0.50
Temp. of generator bearing *b*	0.73	0.42	0.73
Temp. of generator cooling air	0.72	0.35	1.10
Temp. of control cabinet	0.81	0.55	1.22
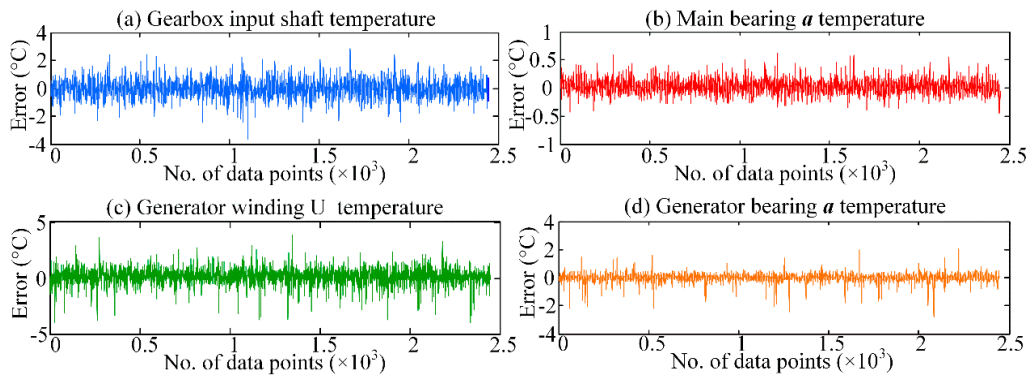
Figure 10. Error residuals for the prediction models of four component temperature parameters. (**a**) Gearbox input shaft temperature; (**b**) Main bearing *a* temperature; (**c**) Generator winding U temperature; and (**d**) Generator bearing *a* temperature.
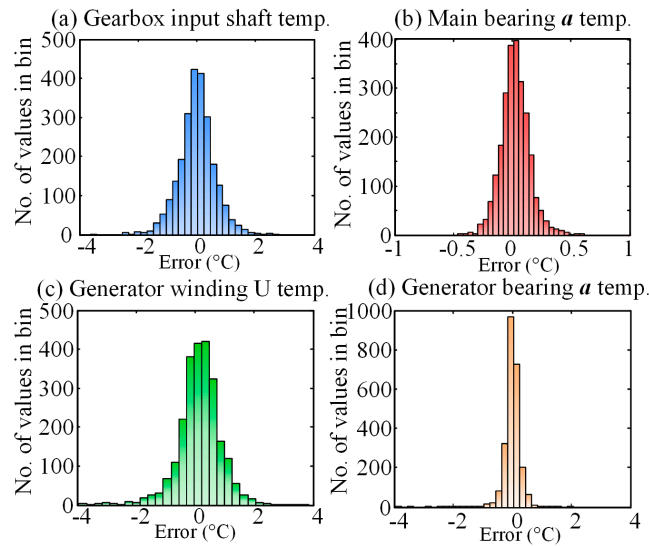

Figure 11. Prediction error distributions of four component temperature parameters. (**a**) Gearbox input shaft temperature; (**b**) Main bearing *a* temperature; (**c**) Generator winding U temperature; and (**d**) Generator bearing *a* temperature.

4.3. The Probabilities of the Operation of the Protection Relays for Type 1 Parameters

The POPR for component temperature parameters: According to the normal distribution characteristics of the prediction error for component temperature parameters, the threshold exceedance probabilities of the Type 1 parameters at the predicted wind speed can be calculated as:

$$O_i(v_j) = P(T_e > T_{lim} - T_{vj}) = 1 - F_{Nc}(T_{lim} - T_{vj}) \tag{7}$$

where T_e is the prediction error of the target component temperature; T_{vj} represents the predicted value of the component temperature parameter *i* when the wind speed is v_j; T_{lim} represents the upper limit of the Type 1 parameter *i*; and $F_{Nc}(\cdot)$ represents the CDF of the prediction error.

The anomaly of WT condition parameters can be detected in advance through the normal distribution characteristics of prediction error [22]. The existing statistical anomaly identification methods are mainly based on the assumption that normal instances occur in the high probability region of a stochastic model, while abnormal conditions happen in the low probability regions [37]. In this paper, when the predicted residuals are out of the 99% confidence interval of the normal distribution, it counts as one invalid prediction.

Three months of SCADA data from 1 April 2012 to 31 July 2012 are collected when the WTs are under healthy condition, which are used as the test data of the prediction models. The prediction error series of five WT temperature parameters (i.e., gearbox input shaft temperature, gearbox cooling water temperature, generator bearing *a* temperature, generator winding U temperature, and main bearing *a* temperature) are obtained. The prediction error series of gearbox input shaft temperature are taken as an example, as shown in Figure 12. Based on the 99% confidence interval of the prediction error, a total of 209 invalid predictions are detected from the error series of the five temperature parameters. Then, a fixed-size time window, which contains data in one hour, is used to collect the prediction error data when the invalid prediction occurs (seen in Figure 12). A total of 182 groups of the prediction error series are obtained by using the 1-hour fixed-size time window. It is found that 84.62% (154/182) of the error groups contain one invalid prediction, 14.84% (27/182) of them contain two invalid predictions, and only 0.55% (1/182) of them contain three invalid predictions. Therefore, in order to avoid misdiagnosis of the anomalies in component temperature parameters, the anomalies are detected based on the following criterion: if there are at least three invalid predictions in one hour, the corresponding parameter is considered as abnormal.
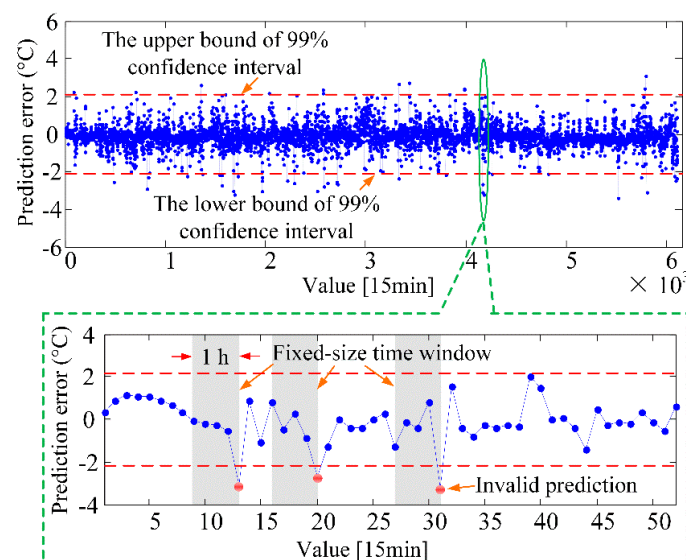

Figure 12. Collection of the prediction error series for gearbox input shaft temperature by using a fixed-size time window.

A case is used for further explanation. A WT suffered a sudden breakdown on 20 July 2012 because of the severe oxidation of its carbon brush. Figure 13a,b shows the prediction results and error residuals of the generator bearing *a* temperature from 8 July 2012 to 10 July 2012. It can be seen that the original prediction model could not realize the efficient prediction. Thus, it is necessary to adjust the prediction results according to the abnormal level of the component temperature parameters. One simple and efficient method is to set the mean value of original normal distribution of prediction error as the previous prediction error. Based on this, the prediction results of the above case are adjusted after occurrence of the parameter anomaly (more than three invalid predictions in one hour), as shown in Figure 13c. It can be seen that errors of the adjusted predictions (*MAE* from 19:30, 10 July to 21:30, 10 July is 1.06 °C) are lower than original errors (*MAE* from 19:30, 10 July to 21:30, 10 July is 2.34 °C). The condition parameter prediction results for the other eight WT outages are used to further verify the proposed method. Table 5 compares the *MAE* errors for the original predictions and the adjusted predictions during the period from the occurrences of the condition parameter anomaly to the WT outage. It can be seen that the *MAE* errors of the adjusted predictions are lower than the original *MAE* errors, which demonstrates the effectiveness of the prediction adjustment method.
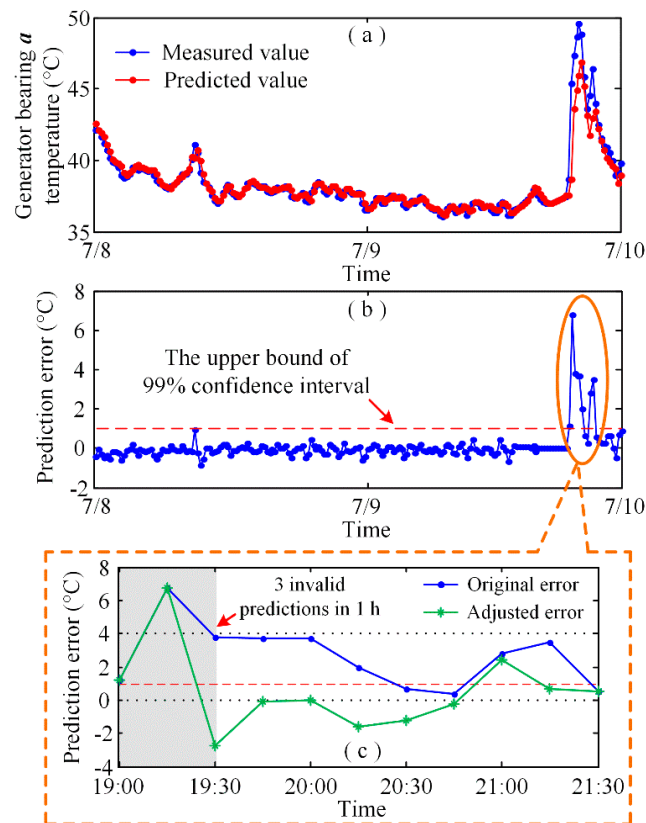
Figure 13. Prediction results of generator bearing *a* temperature: (**a**) Measurement and prediction results; (**b**) Prediction error series; and (**c**) Error series for the adjusted predictions.

Table 5. WT outage cases and corresponding *MAE* for the original predictions and the adjusted predictions.

Number	WT Outage Cases	Abnormal Condition Parameters	*MAE* (°C)	
			Original Prediction	**Adjusted Prediction**
1	Gearbox output shaft overheating of WT 9	Gearbox output shaft temperature	1.15	0.86
2	Generator bearing overheating of WT 14	Generator bearing *b* temperature	1.04	0.55
3	Generator carbon brush oxidation of WT 18	Generator bearing *a* temperature	0.71	0.69
4	Gearbox input shaft overheating of WT 18	Gearbox input shaft temperature	1.84	1.35
5	Gearbox oil over-temperature of WT 20	Gearbox oil temperature	0.91	0.58
6	Generator carbon brush oxidation of WT 22	Generator bearing *a* temperature	0.52	0.47
7	Gearbox output shaft overheating WT 24	Gearbox output shaft temperature	1.96	1.46
8	Generator bearing overheating of WT 31	Generator bearing *b* temperature	0.88	0.83

Based on the prediction adjustment, when the WT is under abnormal condition, the POPR for the component temperature parameters at the predicted wind speed can be calculated as:

$$O_i(v_j) = 1 - F_{Nc}(T_{lim} - T_{vj} - \varepsilon_{t-1}) \tag{8}$$

where ε_{t-1} represents the previous prediction error.

Finally, the POPR for the condition parameter *i* is calculated as follows:

$$F_i = \sum_{j=1}^{N} P(v_j)O_i(v_j) \tag{9}$$

where *N* represents the number of the discrete wind speed prediction errors and $P(v_j)$ represents the probability of the wind speed v_j.

The POPR for wind speed parameter: The WT outages due to exceedance of the cut-out wind speed (i.e., 25 m/s for the 10 min average wind speed) of WTs also have a major impact on the wind power output variability, which should be included in the WT short-term outage model.

Based on the CDF of the wind speed prediction errors, the POPR for wind speed can be calculated as:

$$F_i = P(v_e > v_{lim} - v) = 1 - F_{Nw}(v_{cut-out} - v) \tag{10}$$

where v_e is the prediction error of wind speed; *v* is the predicted value of wind speed; $v_{cut-out}$ is the cut-out wind speed (i.e., 25 m/s); and $F_{Nw}(\cdot)$ is the CDF of the wind speed prediction error.

5. The Probabilities of the Operation of the Protection Relays for Type 2 Parameters

The POPRs for Type 2 parameters are estimated based on operation principles of the protection relays. In order to avoid the frequent WT startup or stop, the WT outage protections usually have a specified setting time. It means that the protection relay operates only when threshold exceedance duration of the parameter exceeds the setting time. The setting time is decided by the WT producers and wind farm operators, which is usually less than 1 min. The POPR will increase with the increase of threshold exceedance duration. However, it is difficult to obtain the threshold exceedance probability density function by statistics technology due to the lack of outage data. In this paper, the linear function is used to describe the relationship between the POPR and the duration of parameter threshold exceedance. The instantaneous data of Type 2 parameters are used for POPR prediction. The POPR for the Type 2 parameters can be calculated as follows:

$$F_i = \begin{cases} 1 & t < t_{\lim} \\ t/t_{\lim} & 0 < t < t_{\lim} \\ 0 & t \leq 0 \end{cases} \tag{11}$$

where *t* is the threshold exceedance duration of the parameter and $t_{\lim}$ is the setting time. Since the setting time for the protection relays of Type 2 parameters is short, the time interval of the POPR prediction is also short, usually within 1 min.

6. Procedure for Predicting the Short-Term Wind Turbine Outage Probability

Following the discussion in the previous sections, a procedure for dealing with the short-term WT outage probability prediction is established, as shown in Figure 14. The procedure can be summarized into the following two steps:

Step 1 (Prediction model development): The wind speed prediction model is developed and the probabilities of the predicted wind speeds (i.e., $P(v_j)$) are obtained based on Table 2. The BPNN based prediction models of component temperature parameters (i.e., Type 1 parameters) are established based on what was discussed in Section 4.2 and the threshold exceedance probabilities of the Type 1 parameters (i.e., $O_i(v_j)$) are calculated according to (7) and (8).

Step 2 (POPR and WT outage probability calculation): Based on the probabilities of the predicted wind speeds and the threshold exceedance probabilities at each predicted wind speed, the POPRs for

Type 1 parameters are calculated based on (9). The POPRs for Type 2 parameters are obtained by using (11). Finally, the short-term WT outage probability can be calculated according to (1).
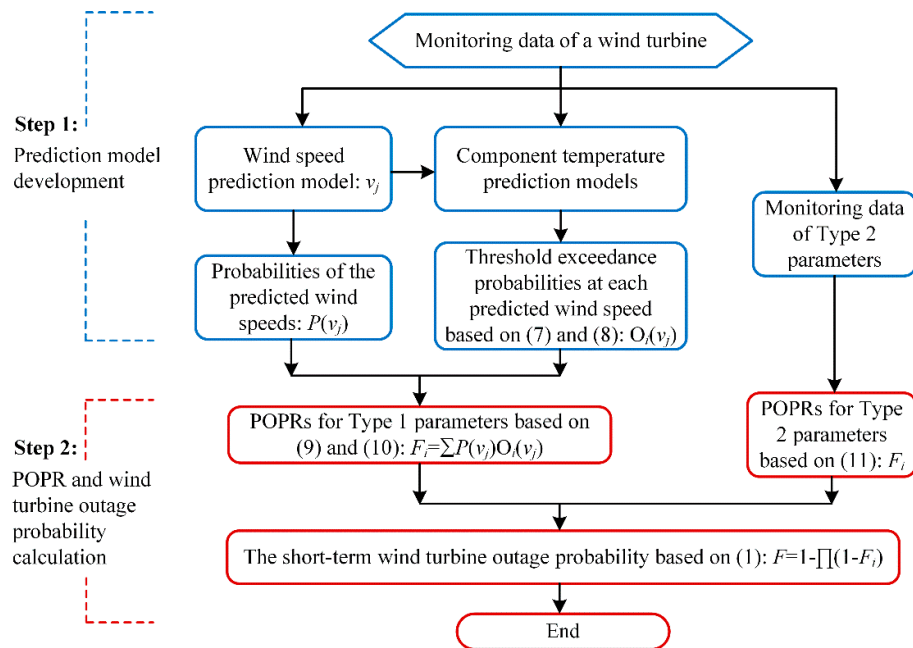

Figure 14. Flowchart of the short-term WT outage probability prediction.

7. Case Study

7.1. Case 1

In this case, a WT suffered a breakdown because the temperature of generator bearing b, which belongs to the Type 1 parameters, exceeded its threshold value (i.e., 95 °C). The WT stopped at 15:30 and restarted at 17:00. Figure 15 shows the time series of wind speed, ambient temperature, and generator bearing b temperature from 05:00 to 17:00. It can be seen that wind speed fluctuated around the rated value 11.5 m/s and ambient temperature kept increasing within two hours before the WT outage. Measured values of the generator bearing b temperature are always higher than the predicted values. According to the proposed anomaly detection method, the temperature of generator bearing b has been found abnormal. Therefore, under the high-temperature and high-wind speed natural environment, it was the continuous increase of the bearing temperature that led to the WT outage.

The data at 15:30 (i.e., the moment of outage) are used to elaborate the calculation process of the WT outage probability. The prediction result of wind speed at this moment is 11.2 m/s by using the ARMA model. According to the probability of the discrete prediction error (Table 2), the predicted wind speeds and the corresponding probabilities can be obtained, as shown in Table 6.

For each predicted wind speed, the generator bearing b temperature is predicted by using the wind speed, ambient temperature, and the previous generator bearing b temperature as input parameters of the prediction models. In normal condition, the mean of the prediction error for generator bearing temperature b is 0. Because the WT is in abnormal condition, the mean of the prediction error should be superimposed by 2.5 °C on the basis of the previous prediction error (i.e., the prediction error at 15:15. The POPR for generator bearing b temperature at each predicted wind speed is calculated according to (8), as shown in Table 7. Based on the probabilities of predicted wind speeds (i.e., $P(v_j)$ in Table 6) and the parameter threshold exceedance probabilities at each predicted wind speed (i.e., $O(v_j)$ in Table 7), the POPR for generator bearing b temperature can be obtained by using (9), which is calculated to be 0.66. Because the POPR for other condition parameters are very low at this time (i.e., 15:30), based on (1), the WT outage probability is approximated to be 0.66.
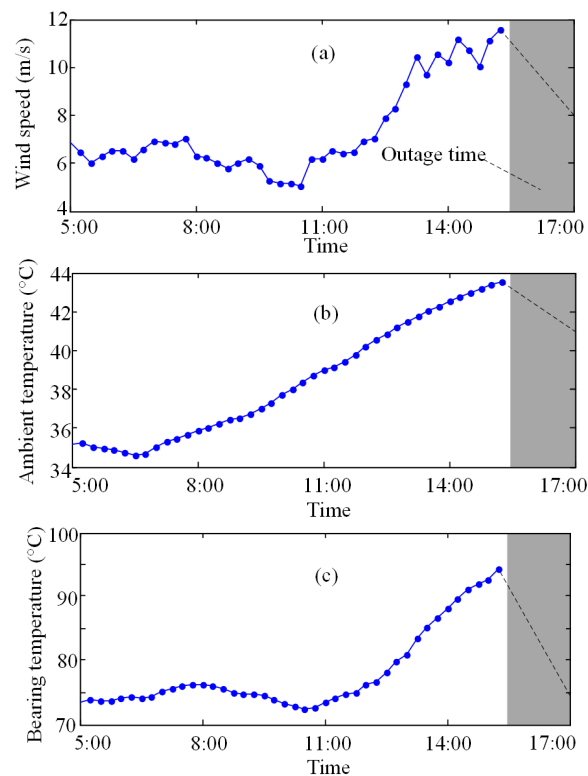
Figure 15. Time series of (**a**) wind speed; (**b**) Ambient temperature; and (**c**) Generator bearing b temperature.

Table 6. Prediction results of wind speed.

Predicted Wind Speeds (m/s): $v_j(j = 1, 2, \ldots, 9)$	Probabilities of the Predicted Wind Speeds: $P(v_j)$
13.2	0.018
12.7	0.05
12.2	0.118
11.7	0.197
11.2	0.2347
10.7	0.197
10.2	0.118
9.7	0.05
9.2	0.018

Table 7. The threshold exceedance probabilities for generator bearing b temperature under each predicted wind speed.

Predicted Wind Speeds (m/s): $v_j(j = 1, 2, \ldots, 9)$	Prediction Result under Each Predicted Wind Speed (°C)	Threshold Exceedance Probabilities under Each Predicted Wind Speed: $O(v_j)$
13.2	93.13	0.66
12.7	93.15	0.67
12.2	93.11	0.66
11.7	93.20	0.68
11.2	93.34	0.71
10.7	93.26	0.69
10.2	93.01	0.63
9.7	92.59	0.52
9.2	92.06	0.38

For the traditional statistic-based reliability assessment method, the component random outage probability during a period of Δt can be calculated according to $P = 1 - e^{-\lambda \Delta t}$ [31], where λ represents

the outage rate of a component. Usually, the component outage rate λ can be considered as the statistical average outage rate which is used in traditional power system reliability evaluation [32]. Since the wind speed has a significant impact on the WT short-term operational reliability [11], the WT outage rate varies depending on the wind speed. Figure 16 shows a relationship between the WT outage rate and wind speed in the studied wind farm. Based on the multiple predicted wind speeds and the corresponding WT outage rate as shown in Figure 16, the traditional statistic-based method is modified and the WT outage probability during Δ*t* (i.e., 15min) can be calculated as follows:

$$P = \sum_{j=1}^{N} P(v_j)(1 - e^{-\lambda(v_j)\Delta t}) \tag{12}$$

where P is the WT outage probability obtained by the statistics-based method; N represents the number of the discrete wind speed prediction errors; v_j is the predicted wind speed and $\lambda(v_j)$ is the corresponding WT outage rate; $P(v_j)$ represents the probability of predicted wind speed v_j.
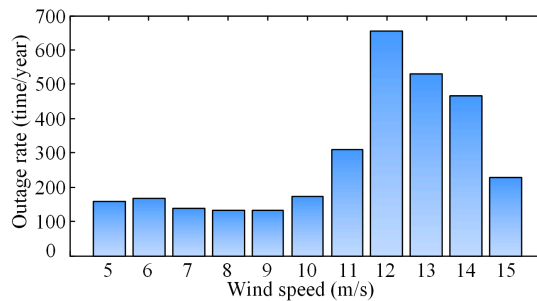

Figure 16. Relationship between WT outage rate and wind speed.

The time series of WT outage probabilities obtained by the proposed method and the statistics-based method from 05:00 to 17:00 are compared in Figure 17. It can be seen that the WT outage probabilities obtained by the proposed method are much higher than those calculated by the traditional statistic-based method during the period of high wind speed (i.e., from 14:00 to 15:30). Since the traditional method fails to consider the WT outage risks caused by the threshold exceedance of WT condition parameters, the proposed method is more accurate than the statistic-based method.
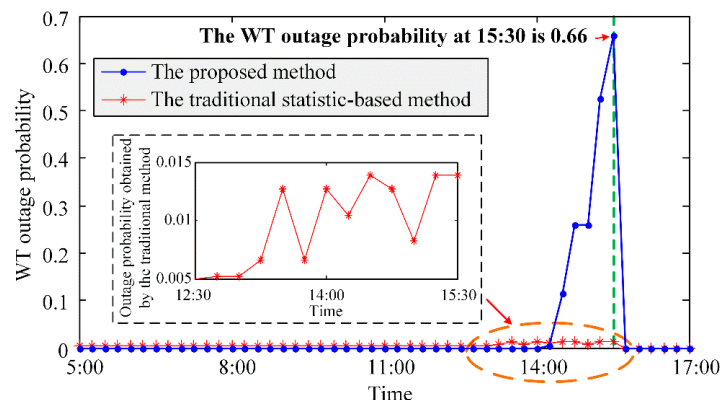

Figure 17. A comparison of WT outage probabilities obtained by the proposed method and the traditional statistic-based method.

7.2. Case 2

In this case, a WT stopped on 17 July 2012 because of the yaw system fault. Figure 18 shows the time series of yaw angle error in one hour on 17 July 2012. The upper and lower limit of the yaw

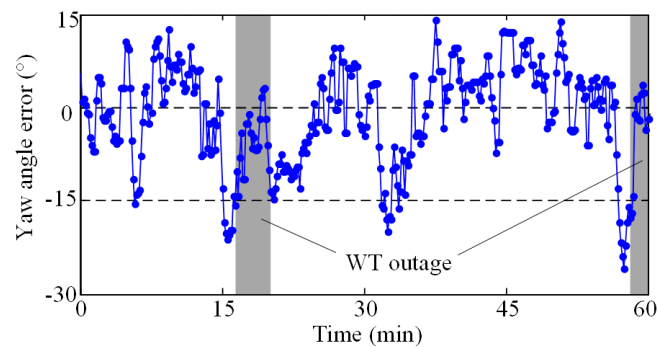angle error is −15° and 15°. The parameter exceeded the lower limit several times, which caused two WT outages.

Figure 18. Time series of yaw angle error.

The setting time of yaw angle error protection relay is 1 min and the POPR is calculated every 10 s, as show in Figure 19. It is noticed that, when the WT is stopped (as shown in the grey part of Figure 18), the POPR is set to be 0. At 15 and 56 min, the threshold exceedance duration of the yaw angle error keeps raising and finally exceeds the setting time of protection. The yaw angle error exceeded the threshold value twice at 32 min without the protection relay operation because the threshold exceedance duration is shorter than the setting time (i.e., 1 min). Since other condition parameters did not exceed their threshold values, the WT outage probability is equal to the POPR for yaw angle error. It is proved that the proposed method is able to effectively estimate the WT outage probabilities.
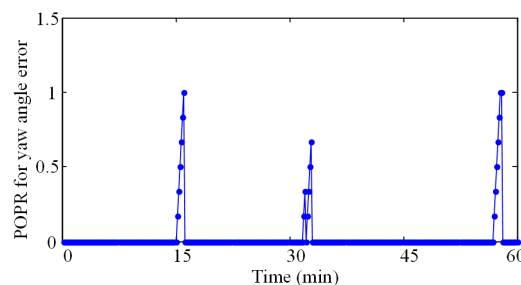

Figure 19. POPR for yaw angle error.

7.3. Case 3

Figure 20 shows the output power and power variation of the studied wind farm from 18 July to 25 July 2012. It can be seen that the wind farm output power showed a fierce fluctuation at constantly high wind speed on 21 July 2012, with the maximum of 9 MW. However, according to the Chinese national standard "Technical Regulations of Wind Power Integration on Electrical Power System" [3], the maximum limit of wind farm active power variation within 1 min shall be no more than 3.3 MW. It is clear that the output power fluctuation of the wind farm during this period cannot satisfy the requirements of the electrical power system. If only the fluctuation of the wind speed is considered, the output power variation shall be in normal condition. However, after the analysis of the operation condition of each WT during this period, it is found that many WTs repeatedly started and stopped during this period due to the threshold exceedance of condition parameters. The low reliability of the WTs at high wind speed is one of the major reasons for the fierce fluctuation of output power of the whole wind farm.

A WT is taken as an example for the outage probability analysis. This WT stopped nine times from 11:00, 21 July to 06:00, 22 July due to the threshold exceedance of the gearbox oil temperature.

Figure 21 shows the time series of the gearbox oil temperature, where the time points of the WT outages are marked. The WT outage probabilities during this period are obtained by using the proposed outage model. The calculation process is similar to Case 1 and the results are shown in Figure 22. Compared with the prediction results of outage probabilities only in consideration of the statistical information, the proposed outage model can accurately predict the short-term outage probability of the WT. If the wind farm operators could predict the short-term outage probability of each WT and formulate the proper power generation schedules, the reliability of the wind farm can be improved. The proposed short-term WT outage model is helpful for the safe and economical operation of the wind farms and the electrical power systems.
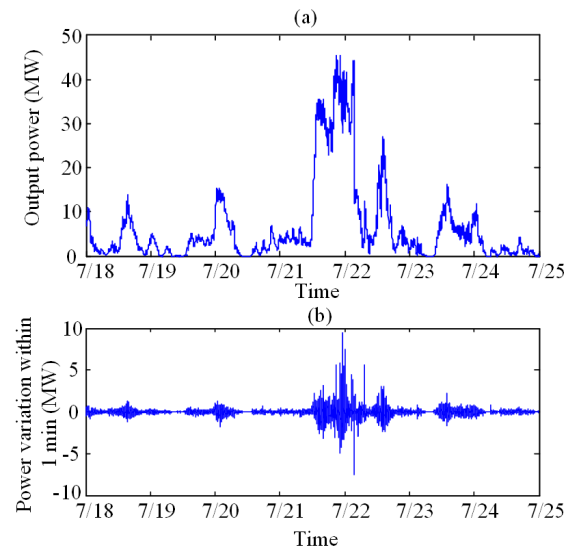

Figure 20. Wind farm output power (**a**) and output power variation within 1 min (**b**).
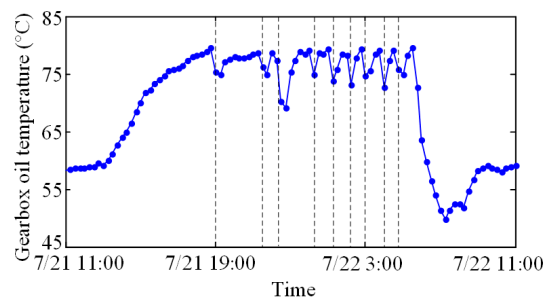

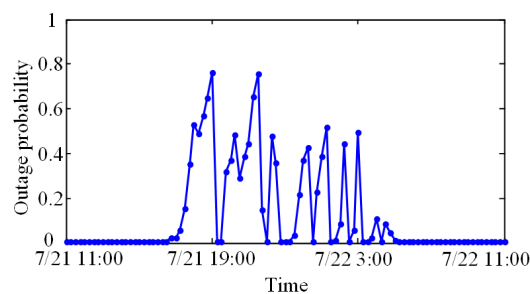**Figure 21.** Time series of gearbox oil temperature.

Figure 22. WT outage probabilities of Case 3.

8. Conclusions

This paper presents a novel approach for short-term WT outage probability prediction. The conclusions of the study can be summarized as follows:

(1) The BPNN based prediction models are established for the WT component temperature parameters. The test results show that although the prediction accuracy of each component temperature is different, the maximum of *MAE* is lower than 1 °C.

(2) The component temperature prediction models cannot realize the efficient prediction when the WT is under abnormal condition. The proposed prediction adjustment method could effectively decrease the prediction errors when the WT is abnormal.

(3) A novel approach is proposed to calculate the short-term WT outage probability by considering the threshold exceedance of the SCADA condition parameters. Three case studies for an onshore wind farm in northern China have been carried out and analyzed, which demonstrates the effectiveness of the short-term WT outage model.

(4) Since the traditional statistic-based reliability evaluation method fails to consider the WT outage risks caused by the threshold exceedance of WT condition parameters, the proposed method is more accurate than the statistic-based method.

(5) The WT outages could impact on the power output variation of a wind farm. The prediction results of short-term WT probabilities are useful for the wind farm operators to estimate the potential wind power loss due to the WT outages and formulate proper power generation schedules.

Since the proposed short-term WT outage model has not been used in wind farms, there is still a lack of actual WT outage cases to support a statistical analysis to show how much the WT outage probability can be reduced by using the proposed method. In future research we will focus on enhancing the applicability of the proposed short-term WT outage model in large wind farms. Furthermore, active power control strategies of wind farms based on the short-term WT outage probability prediction will be investigated. The condition based maintenance strategies with respect to different WT faults, also deserve further study.

Author Contributions: The research presented in this paper was a collaborative effort among all authors. Peng Sun and Jian Li proposed the methodology of the model and contributed to the manuscript at all stages. Junsheng Chen and Xiao Lei performed the design of the model, discussed the results and revised the manuscript critically.

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

References

1. Gil, M.D.P.; Gomis-bellmunt, O.; Sumper, A. Technical and economic assessment of offshore wind power plants based on variable frequency operation of clusters with a single power converter. *Appl. Energy* **2014**, *125*, 218–229.

2. Rona, B.; Guler, O. Power system integration of wind farms and analysis of grid code requirements. *Renew. Sustain. Energy Rev.* **2015**, *49*, 100–107. [CrossRef]

3. *Technical Rule for Connecting Wind Farm to Power System*; GB/T 19963-2011; China Standards Press: Beijing, China, 2012.

4. Chen, W.; Wu, W.C.; Zhang, B.N. An on-line rolling generation dispatch method and model for accommodating large-scale wind power. *Autom. Electr. Power Syst.* **2012**, *35*, 136–140. (In Chinese)

5. Ummels, B.C.; Brand, A.J.; Gibescu, M.; Kling, W.L. Impacts of wind power on thermal generation unit commitment and dispatch. *IEEE Trans. Energy Convers.* **2007**, *22*, 44–51. [CrossRef]

6.   Sideratos, G.; Hatziagyriou, N.D. An advanced statistical method for wind power forecasting. *IEEE Trans. Power Syst.* **2007**, *22*, 258–265. [CrossRef]

7.   Ebrahimi, F.M.; Khayatiyan, A.; Farjah, E. A novel optimizing power control strategy for centralized wind farm control system. *Renew. Energy* **2016**, *86*, 339–408. [CrossRef]

8.   Park, J.; Law, K.H. Cooperative wind turbine control for maximizing wind farm power using sequential convex programming. *Energy Convers. Manag.* **2015**, *101*, 295–316. [CrossRef]

9.   He, J.; Cheng, L.; Sun, Y.Z. Power system short-term operational reliability evaluation considering weather forecast. *Power Syst. Prot. Control* **2010**, *38*, 31–38. (In Chinese)

10. He, J.; Cheng, L.; Sun, Y.Z. Condition dependent short-term reliability models of transmission equipment. *Proc. CSEE* **2009**, *29*, 39–46. (In Chinese)

11. Cheng, L.; Lin, J.; Sun, Y.Z. A model for assessing the power variation of a wind farm considering the outages of wind turbines. *IEEE Trans. Sustain. Energy* **2012**, *3*, 432–444. [CrossRef]

12. Tavner, P.J.; Greenwood, D.M.; Whittle, M.W.G.; Gindele, R.; Faulstich, S. Study of weather and location effects on wind turbine failure rate. *Wind Energy* **2013**, *16*, 175–187. [CrossRef]

13. Tavner, P.J.; Xiang, J.; Spinato, F. Reliability analysis for wind turbines. *Wind Energy* **2007**, *10*, 1–18. [CrossRef]

14. Spinato, F.; Tavner, P.J.; van Bussel, G.J.W.; Koutoulakos, E. Reliability of wind turbine subassemblies. *IET Renew. Power Gener.* **2009**, *3*, 387–401. [CrossRef]

15. Schlechtingen, M.; Santos, I.F.; Achiche, S. Wind turbine condition monitoring based on SCADA data using normal behavior models. Part 1: System description. *Appl. Soft Comput.* **2013**, *13*, 259–270. [CrossRef]

16. Schlechtingen, M.; Santos, I.F. Wind turbine condition monitoring based on SCADA data using normal behavior models. Part 2: Application examples. *Appl. Soft Comput.* **2014**, *14*, 447–460. [CrossRef]

17. Sanz-Bobi, M.A.; Pico, J.D.; Garcia, M.C. SIMAP: Intelligent system for predictive maintenance application to the health condition monitoring of a wind turbine gearbox. *Comput. Ind.* **2006**, *57*, 552–568.

18. Zaher, A.; McArthur, S.D.J. A multi-agent fault detection system for wind turbine defect recognition and diagnosis. In Proceedings of the IEEE Lausanne Power Tech, Lausanne, Switzerland, 1–5 July 2007; pp. 22–27.

19. Yang, W.X.; Court, R.; Jiang, J.S. Wind turbine condition monitoring by the approach of SCADA data analysis. *Renew. Energy* **2013**, *53*, 365–376. [CrossRef]

20. Ata, R. Artificial neural networks applications in wind energy systems: A review. *Renew. Sustain. Energy Rev.* **2015**, *49*, 534–562. [CrossRef]

21. Yan, Y.L.; Li, J.; Gao, D.W.Z. Condition parameter modeling for anomaly detection in wind turbines. *Energies* **2014**, *7*, 3104–3120. [CrossRef]

22. Zaher, A.; McArthur, S.D.J.; Infield, D.G. Online wind turbine fault detection through automated SCADA data analysis. *Wind Energy* **2009**, *12*, 574–593. [CrossRef]

23. Liu, Y.Q.; Shi, J.; Yang, Y.P.; Lee, W.J. Short-term wind-power prediction based on wavelet transform-support vector machine and statistic-characteristics analysis. *IEEE Trans. Ind. Appl.* **2012**, *48*, 1–7. [CrossRef]

24. Chen, B.; Peter, C.M.; Peter, J.T. Automated on-line fault prognosis for wind turbine pitch systems using supervisory control and data acquisition. *IET Renew. Power Gener.* **2015**, *9*, 503–513. [CrossRef]

25. Guo, P.; Infield, D.; Yang, X.Y. Wind turbine generator condition-monitoring using temperature trend analysis. *IEEE Trans. Sustain. Energy* **2012**, *3*, 124–133. [CrossRef]

26. Kusiak, A.; Verma, A. Monitoring wind farms with performance curves. *IEEE Trans. Sustain. Energy* **2013**, *4*, 192–199. [CrossRef]

27. Cross, P.; Ma, X.D. Nonlinear system identification for model-based condition monitoring of wind turbines. *Renew. Energy* **2014**, *71*, 166–175. [CrossRef]

28. Kusiak, A.; Verma, A. Analyzing bearing faults in wind turbines: A data-mining approach. *Renew. Energy* **2012**, *48*, 110–116. [CrossRef]

29. Schlechtingen, M.; Santos, I.F. Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection. *Mech. Syst. Signal Process.* **2011**, *25*, 1849–1875. [CrossRef]

30. Yuan, W.W.; Jiang, Y.L. Cooling systems in wind turbine. *World Sci-Tech R&D* **2007**, *29*, 80–85.

31. Li, W.Y. *Risk Assessment of Power Systems: Models, Methods, and Applications*; IEEE Press Series on Power Engineering; John Wiley & Sons: New York, NY, USA, 2005.

32. Billinton, R.; Allan, R.N. *Reliability Evaluation of Engineering Systems: Concepts and Techniques*, 2nd ed.; Plenum Press: New York, NY, USA, 1992.

33. Ogimi, K.; Kamiyama, S.; Yona, M.; Yona, A.; Senju, T. Optimal operation planning of wind farm installed bess using wind power forecast data of wind turbine generators considering forecast error. *Int. J. Emerg. Electr. Power Syst.* **2013**, *14*, 207–218. [CrossRef]

34. Wang, S.Y.; Yu, J.L. Joint conditions probability forecast method for wind speed and wind power. *Proc. CSEE* **2011**, *31*, 7–15. (In Chinese)

35. Spera, D.A. *Wind Turbine Technology: Fundamental Concepts of Wind Turbine Engineering*; American Society of Mechanical Engineers (ASME): New York, NY, USA, 1994.

36. Lapira, E.; Brisset, D.; Ardakani, H.D.; Siegel, D.; Lee, J. Wind turbine performance assessment using multi-regime modeling approach. *Renew. Energy* **2012**, *45*, 86–95. [CrossRef]

37. Chandola, V.; Banerjee, A.; Kumar, V. Anomaly detection: A survey. *ACM Comput. Surv.* **2009**, *41*, 75–79. [CrossRef]