

# A Generalized Model for Wind Turbine Faulty Condition Detection Using Combination Prediction Approach and Information Entropy

J. S. Chen<sup>1</sup>, W. G. Chen<sup>1</sup>, J. Li<sup>1\*</sup>, and P. Sun<sup>2</sup>

<sup>1</sup>State Key Laboratory of Power Transmission Equipment & System Security and New Technology, College of Electrical Engineering, Chongqing University, Chongqing 400030, China

<sup>2</sup>State Grid Henan Electrical Power Research Institute, Zhengzhou 450000, China

Received 11 October 2015; revised 28 September 2017; accepted 13 February 2018; published online 20 September 2018

**ABSTRACT.** A generalized model for detecting the incipient wind turbine (WT) faulty condition based on the data collected from wind farm supervisory control and data acquisition (SCADA) system is proposed in this paper. The linear combination prediction approach and the information entropy are integrated to develop the generalized model, in which the linear combination prediction approach improves the accuracy and generalization performance of the model, and the information entropy of prediction residual quantifies the abnormal level of the condition parameter. SCADA datasets were selected to establish the prediction models of WT condition parameters that are dependent on environmental conditions such as ambient temperature and wind speed. The combination prediction models of WT condition parameters were developed based on different data mining algorithms such as Back propagation neural network (BPNN) algorithm, radial basis function neural network (RBFNN) algorithm and least square support vector machine (LSSVM) algorithm. The information entropy was utilized to extract useful information from residuals of the prediction models for WT faulty condition detection. Finally, the proposed method has been used for real 1.5 MW WTs with doubly fed induction generators (DFIG). Through investigation of cases of actual WT faults, the effectiveness of the proposed WT imminent fault identification approach was verified.

**Keywords:** Combination prediction model, faulty condition detection, information entropy, SCADA, Wind turbines

## 1. Introduction

The rapid expansion of wind farms in recent years have drawn great attention to operation and maintenance (O&M) issues (Johan et al., 2007; Gil et al., 2014; Yang et al., 2014). The O&M costs of wind turbines (WTs) account for approximately 25 to 30% of the overall wind energy generation cost (Milborrow, 2006). Various condition monitoring and fault diagnosis approaches of WTs have been proposed to improve the availability of WTs and reduce unscheduled downtime as well as O&M costs (e.g. Tavner et al., 2010; Djurovic et al., 2012; GarcíaMárquez et al., 2012; Yang et al., 2013; Souza et al., 2013; Lee et al., 2013). However, most of these approaches have not been widely used in wind farms because of the high additional investments of installing extra equipments or sensors for condition monitoring systems of WT and generalization performance of the fault detection approaches. Most commercially available WT condition monitoring systems are vibration analysis-based systems to achieve the condition monitoring of WT blades and gearboxes (Ribrant et al., 2007). However, they

are not good at detecting WT electrical, electronic and hydraulic problems, and the WT electrical, electronic and hydraulic subassemblies also suffer failures (Kusiak et al., 2011; Yang et al., 2013). They would lead to long downtime as well. As one of the most important parts of the WT condition monitoring system, the supervisory control and data acquisition (SCADA) system of a wind farm collects data extensively from key WT subassemblies. The SCADA system records comprehensive WTs condition parameters which contain rich information concerning the health of the WTs (Kusiak et al., 2012). Furthermore, SCADA systems have been installed in the majority of MW-scale WTs. No more hardware investment is needed when developing a SCADA-based WT fault detection. Therefore, detection of incipient WT faults based on SCADA data is a cost-effective way to monitor the health of WT.

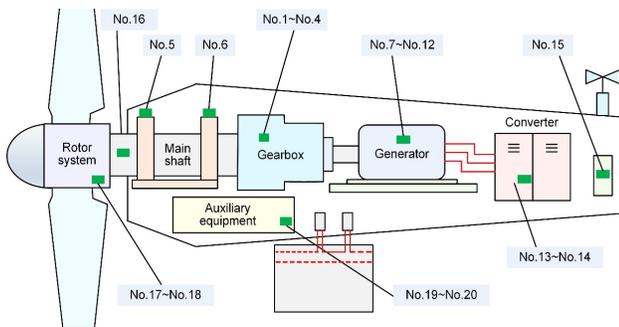
Values of some WT condition parameters obtained from SCADA system, such as rotor speed, output power, and component temperature, fluctuate quite obviously with the environmental conditions and vary over wide ranges under varying operational conditions. Hence, it is hard to detect incipient faults from the raw SCADA data. An effective data-driven method is necessary to address this problem. Various WT condition parameter prediction models and anomaly identification methods (e.g. Zaher et al. 2009; Schlechtingen and Santos, 2011;

\*Corresponding author. Tel.: +86 23 65106880; fax: +86 23 65102442.  
E-mail address: lijian@cqu.edu.cn (J. Li)

Kusiak and Verma, 2012; Schlechtingen et al., 2013) have been developed to extract the useful information hidden within the raw data and mitigate the impacts of environmental conditions upon the real WT condition parameters. For instance, an anomaly identification method was presented based on the prediction residuals for WT condition parameters. It was claimed that the prediction residuals coming from successfully trained models were normally distributed with a mean around zero (Schlechtingen, 2013). Various data-driven approaches, such as neural network (NN) (Zaher et al., 2009; Ata, 2015), support vector machine (SVM) (Liu et al., 2012; Yang et al., 2006) and nonlinear state estimate technique (Cross and Ma, 2014) have been used for developing WT condition parameter prediction models. Performance of the prediction models established by six data mining algorithms was studied, and input parameter selection method was investigated to improve the prediction accuracy (Kusiak and Li, 2010). Different algorithms were used to obtain the relevant input parameters for the predicting target parameters (Kusiak and Verma, 2012).

In the above work, the condition parameter was predicted with one single prediction model, and the model was developed based on SCADA data collected from one single WT. Since the sensitivity of WT faulty condition detection is affected by the precision of the condition parameter prediction model, a single-model-based method may cause problems on the accuracy and generalization performance of the anomaly identification. Instead of using a single-model-based prediction method, the combination prediction approach was utilized in this paper. The combination prediction approach can take advantages of different data mining algorithms to improve the accuracy and generalization performance of condition parameters prediction models.

The existing statistical anomaly identification methods are 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 (Chandola et al., 2009). Thus, the prediction residual (i.e. the difference between the measured value and model output) could provide an indication of incipient WT faults. Certain thresholds of prediction residuals are usually set to identify the anomalies in



**Figure 1.** The main components and sensor positions of the considered WT.

WTs (Kusiak and Verma, 2012; Schlechtingen et al., 2013). However, a certain threshold-based method may cause misdiagnosis (Sun et al. 2016). To accurately quantify the abnormal level of condition parameter is another key point for detecting incipient WT fault. Recently, the information entropy has proved its suitability for condition monitoring, fault detection and risk assessment (Cabal et al., 2010; Zhang et al., 2010; Li et al., 2012; Ai et al., 2013; Yang et al., 2015; Nourani et al., 2015). For instance, faulted rotor bars were detected through the analysis of vibration signals based on information entropy (Cabal et al., 2010). The utility of multiscale entropy in diagnosing faulty bearings in a rotating machine was investigated (Zhang et al., 2010). Therefore, the information entropy was employed to quantify the abnormal level of the condition parameters in this study.

The motivation of the present paper lies in the detection of WT faulty condition based on wind farm SCADA data. In this study, the linear combination prediction approach and the information entropy of prediction residual were integrated to develop a generalized model for incipient WT fault detection. The SCADA datasets were selected through the analysis of the cumulative probability distribution of wind speed, and the relationship between output power and wind speed. The combination prediction models of WT condition parameters were developed based on linear combination prediction approach. The information entropy of prediction residual was utilized to extract the useful information from the residual of the prediction model for WT fault detection.

The remainder of the paper is organized as follows. The SCADA data that can be used for WT faulty condition detection is discussed and grouped in Section 2. Section 3 presents the selection of wind farm SCADA data. The combination condition parameter prediction models are developed in Section 4. The anomaly identification method based on the information entropy is illustrated in Section 5. Two cases are investigated to validate the proposed method in Section 6. The conclusions are drawn in Section 7.

## 2. Parameter Description and Classification

Table 1 shows 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:

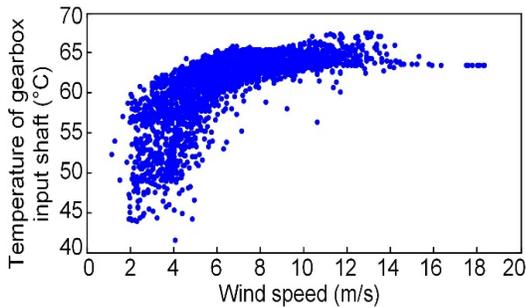
Type 1: Type 1 parameters include various component temperatures, WT output power and rotor speed, which are strongly influenced by environmental conditions. For example, the relationship between the gearbox input shaft temperature of a variable speed constant-frequency (VSCF) WT and the wind speed is 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 relation-

**Table 1.** WT Condition Parameters Studied in this Paper

No.	WT condition parameters	Unit	Type
1	Temp. of gearbox input shaft	°C	1
2	Temp. of gearbox output shaft	°C	1
3	Temp. of gearbox oil	°C	1
4	Temp. of gearbox cooling water	°C	1
5	Temp. of main bearing a (on the rotor side)	°C	1
6	Temp. of main bearing b (on the gearbox side)	°C	1
7	Temp. of generator winding ph.1 (U)	°C	1
8	Temp. of generator winding ph.2 (V)	°C	1
9	Temp. of generator winding ph.3 (W)	°C	1
10	Temp. of generator bearing A	°C	1
11	Temp. of generator bearing B	°C	1
12	Temp. of generator cooling air	°C	1
13	Temp. of control cabinet	°C	1
14	Temp. of converter controller	°C	1
15	Output power	kW	1
16	Rotor speed	rpm	1
17	Yaw angle error	°	2
18	Pitch angle error	°	2
19	Hydraulic oil pressure for yaw	bar	2
20	Hydraulic oil pressure for rotor brake	bar	2

ship 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. For the parameters of output power and rotor speed, the wind speed is the most influential factor and wind farm operators usually use a power curve to estimate the power generated by a WT at different wind speeds.

Type 2: Type 2 parameters include yaw angle error, pitch angle error and hydraulic oil pressure. Yaw angle error is the angle between the wind and the nacelle position. Pitch angle error represents the pitch angle that deviates from the set point. Type 2 condition parameters do not have an obvious relationship with environmental conditions. Since the anomaly identification for condition parameters of Type 2 can be easily done by setting certain threshold values, only the WT anomaly identi-



**Figure 2.** Relationship between the gearbox input shaft temperature of a VSCF WT and the wind speed.

fication based on condition parameters of Type 1 will be discussed.

### 3. SCADA Data Collection

The SCADA data used in this paper are obtained from an onshore wind farm in Northern China. The wind farm has 34 WTs with 1.5 MW, labeled WT 1 to WT 34. All the WTs are the same type with doubly fed induction generators (DFIG) in the wind farm. The SCADA data have been collected since February 15, 2011. The SCADA data of WTs 3, 10, 17, 23, and 31 are randomly selected for analysis of condition parameters. The SCADA data with a time interval of 10 min are used for the work.

The SCADA data are selected with a criteria of wind speed. The Bin method is used to analyze the SCADA data. Figure 4 shows the wind speed probability distributions of five WTs in the studied wind farm. The cumulative probability of the wind speed of WT 17, as shown in Figure 4a, indicates 99.22% of wind speed values is lower than 13.5 m/s. Figure 4b shows that the wind speed probability distributions of WTs 3, 10, 23, and 31 that are identical with that of WT 17. The wind speed range between 3 and 13.5 m/s is thus determined for SCADA data selection, which covers most output power of each WT in the wind farm.

### 4. Combination Prediction Model Development

The sensitivity of anomaly identification is affected by the accuracy of the condition parameter prediction models. In this paper, the linear combination prediction approach is used to improve the accuracy and generalization performance of the prediction models.

#### 4.1. Principle of the condition parameter combination prediction model

It is assumed that there are  $m$  kinds of individual prediction methods to predict an index sequence  $\{x_t, t = 1, 2, \dots, N\}$  of the same predicting object.  $x_{it}$  represents the predicted value of the  $i$ th individual prediction method at time  $t$  ( $i = 1, 2, \dots, m; t = 1, 2, \dots, N$ ). The  $e_{it} = (x_t - x_{it})$  is the prediction residual.

$J_1$  is defined as the sum-of-squares of the prediction residual. The linear combination prediction model is developed based on the following optimization problem:

$$\begin{aligned} \min J_1 &= \sum_{t=1}^N \sum_{i=1}^m \sum_{j=1}^m l_i l_j e_{it} e_{jt} \\ \text{s.t. } \sum_{i=1}^m l_i &= 1 \end{aligned} \tag{1}$$

where  $e_{jt}$  is the prediction residual of the  $j$ th prediction method at time  $t$ , and  $l_1, l_2, \dots, l_m$  represent the weight of each

individual prediction method. The summation of the weights is equal to 1. The matrix form of Equation (1) can be expressed as:

$$\begin{aligned} \min J_1 &= L^T E L \\ \text{s.t.} \begin{cases} R^T L = 1 \\ L \geq 0 \end{cases} \end{aligned} \quad (2)$$

where  $L = [l_1 \ l_2 \ \dots \ l_m]^T$  is the column vector of weight of each individual prediction method;  $R$  is  $m$ -dimensional column vector with all the elements equal to 1; and  $E$  is  $m \times m$  information square matrix of combination prediction residual. When  $i \neq j$ ,  $E_{ij}$  represents the covariance of the prediction residual of the  $i$ th and the  $j$ th individual prediction method. When  $i = j$ ,  $E_{ii}$  represents the sum-of-squares of prediction residual for the  $i$ th individual prediction method.

Equation (2) is a quadratic convex programming problem, which has a unique optimal solution on the feasible region or the boundary of the feasible region. However, it is difficult to obtain the direct-form of the solution for Equation (2). In this paper, the iterative algorithm for optimal combination prediction of nonnegative weights is utilized to find the optimal weights. Kuhn-Tucker (K-T) condition is the necessary and sufficient condition of the existence of optimal solution according to the quadratic programming theory (Wismar and Chattergy, 1978). Therefore, the solving of Equation (2) can be transformed into the solving of K-T problem. Equation (2) is equivalent to:

$$\begin{cases} 2EL - \lambda_1 R + \lambda_2 R - U = 0 \\ R^T L = 1 \\ u_i l_i = 0 (i = 1, 2, \dots, m) \\ \lambda_1, \lambda_2, u_i \geq 0 \end{cases} \quad (3)$$

where  $U = (u_1, u_2, \dots, u_m)^T$ ;  $u_i$  and  $l_i$  are the complementary slack variables;  $\lambda_1$  and  $\lambda_2$  are the Lagrange multipliers of equation  $R^T L = 1$ .

Equation (3) is a linear programming problem. It can be

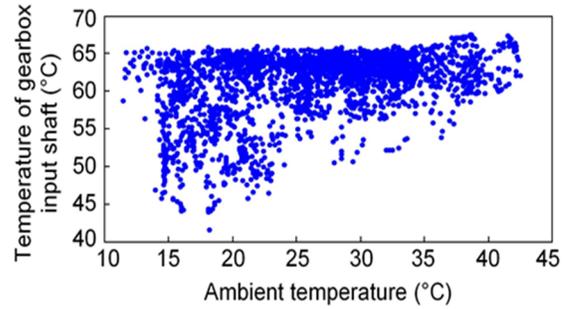
**Table 2.** Input Parameters of the Prediction Models.

Target condition parameters	Input parameters (SCADA data)
Output power (t)	Wind speed (t) Yaw angle error (t) Pitch angle (t)
Rotor speed (t)	Wind speed (t) Yaw angle error (t) Pitch angle (t)
Component temperatures (t)	Wind speed (t) Output power (t) Ambient Temperature (t) Component temperatures (t-1)

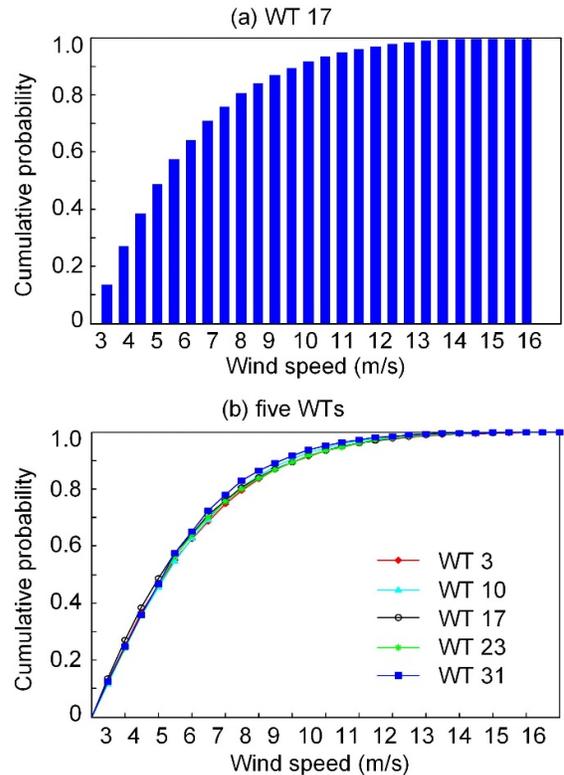
solved through adding artificial variables  $Z = (z_1, z_2, \dots, z_m)^T$  and  $z$ . Hence, the auxiliary linear programming model is obtained as follows:

$$\begin{aligned} \min f &= z + z_1 + \dots + z_m \\ \text{s.t.} \begin{cases} 2EL - \lambda_1 R + \lambda_2 R - U + Z = 0 \\ R^T L + z = 1 \\ \lambda_1, \lambda_2, u_i, z, z_i \geq 0 \end{cases} \end{aligned} \quad (4)$$

where  $u_i \cdot l_i = 0$  ( $i = 1, 2, \dots, m$ ). In order to apply this non-linear constraint condition, the following modifications are required when selecting a base variable: if  $l_i$  is the base variable, the corresponding  $u_i$  cannot be selected as a base variable.



**Figure 3.** Relationship between the gearbox input shaft .



**Figure 4.** Wind speed distribution of wind turbine: (a) WT 17; (b) Five WTs.

iable. Similarly, if  $u_i$  is the base variable, the corresponding  $l_i$  cannot be selected as a base variable. Therefore, the optimal weights of the optimal combined prediction of nonnegative weights are obtained by Equation (4).

**4.2. Development of the condition parameter combination prediction model**

The BPNN model, RBFNN model and LSSVM model are selected as the individual prediction model. The linear combination prediction models of WT condition parameters are developed based on the principle of minimizing the sum-of-squares of the prediction residual. The development of the prediction model for generator bearing B temperature of WT 17 is taken as an example.

The selection of input parameters of prediction model is a premise in simplifying models and ensuring prediction accuracy (Yan et al., 2014; Sun et al., 2016). The input parameters of different types of condition parameter prediction models are shown in Table 2. A fixed pitch is usually applied at the VSCF WT and the generator torque control is used to maximize the captured power from the wind (Assareh and Biglari, 2015). The wind speed, yaw error, and the pitch angle are selected as the input parameters to improve the accuracy of the models of output power and rotor speed. The component temperatures depend on the wind speed, the ambient temperature, the WT’s output power and the previous component temperatures. Based on this, four condition parameters such as the wind speed; the output power, the ambient temperature, the previous temperature of generator bearing B are selected as the input parameters of the prediction model for

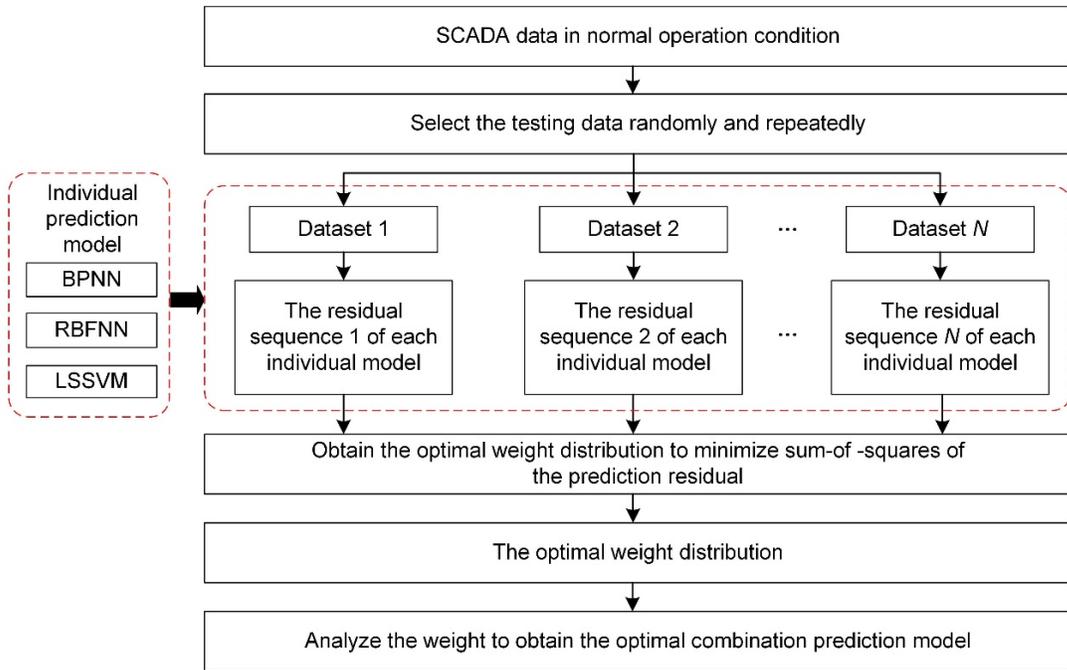
generator bearing B temperature. The modeling progress is shown as follows:

1) A total of 30,000 data obtained under normal operating conditions are randomly selected from the 1-year SCADA data of WT 17 to form the sample dataset. The 10-fold cross-validation method is used for parameter optimization of individual prediction models to ensure the better generalization ability of the prediction models.

2) Figure 5 shows the modeling procedures of the combination prediction model. A total of 3000 data is randomly selected from the sample data (30,000 data) to form a dataset. The selection process is repeated for N times to obtain N data groups. The BPNN model, the RBFNN model and the LSSVM model are tested for N times by using the N sets datasets. For each time, the prediction residuals of the three individual models are recorded respectively. The iterative algorithm for optimal combination forecasting of non-negative weight is applied to obtain the weight distribution for each individual model to develop the combination prediction model. Figure 6 shows the weight distribution when N = 15.

As shown in Figure 6, the weight of BPNN model is greater than that of RBFNN model and LSSVM model, which indicates that the output of BPNN model contains more useful information. The average of the weights that are obtained by each test is determined as the optimal weight of the corresponding individual prediction model. Based on this, the combination prediction model can be obtained as follows:

$$\hat{F} = l_1 \cdot \hat{f}_1 + l_2 \cdot \hat{f}_2 + l_3 \cdot \hat{f}_3 \tag{5}$$



**Figure 5.** Sketch of the combination prediction model.

where  $\hat{F}$  is predicted value of the combination prediction model;  $l_1$ ,  $l_2$  and  $l_3$  are the optimal weights of the BPNN model, the RBFNN model and the LSSVM model;  $\hat{f}_1$ ,  $\hat{f}_2$  and  $\hat{f}_3$  are predicted value of each individual model, respectively. The optimal weights for each individual prediction model of generator bearing B temperature are  $l_1 = 0.645$ ,  $l_2 = 0.2018$ , and  $l_3 = 0.1532$ .

The mean absolute error (MAE), mean absolute percentage error (MAPE), root-mean-square error (RMSE) and mean squared percentage error (MSPE) are used to analyze the performance of each prediction model, as shown in Equations (6) ~ (9).

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{r}_i - r_i| \quad (6)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{r}_i - r_i}{r_i} \right| \times 100\% \quad (7)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{r}_i - r_i)^2} \quad (8)$$

$$MSPE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left( \frac{\hat{r}_i - r_i}{r_i} \right)^2} \quad (9)$$

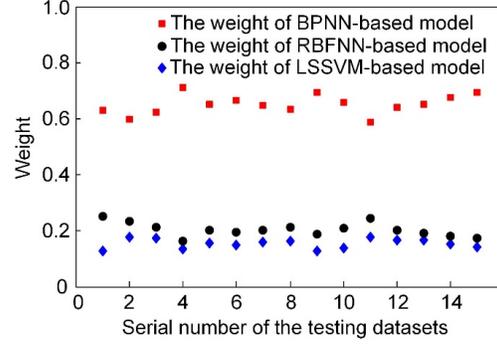
where  $n$  is number of test samples;  $r_i$  is the measured value;  $\hat{r}_i$  is the predicted value.

The prediction accuracy of the combination prediction model and three individual models (the BPNN model, the RBFNN model and the LSSVM model) is compared. The temperature of generator bearing B of WT 17 in the studied wind farm is taken as an example. 15 test datasets that are selected randomly from the 30000 sample data.

Table 3 shows the mean value of each index of prediction models for generator bearing B temperature of WT 17. The combination prediction model has the highest accuracy. The accuracy of BPNN model is obvious higher than that of RBFNN model and LSSVM model. Therefore, the combination prediction model of WT condition parameters based on the principle of minimizing the sum-of-squares of the prediction residual is effective to improve the prediction accuracy.

**Table 3.** Prediction Performance of Different Prediction Models for Generator Bearing B Temperature

Type	Index			
	MAE (°C)	MAPE (%)	RMSE (°C)	MSPE (%)
BPNN model	1.6015	3.5618	2.3869	6.2083
RBFNN model	1.7232	4.2419	2.5381	6.8371
LSSVM model	1.7549	4.3627	2.6016	7.0946
Combination prediction model	1.5861	3.3217	2.1593	5.6192



**Figure 6.** Weight distribution for the individual models.

### 5. Anomaly Identification

Anomaly identification aims to detect data patterns that do not conform to the principle of expectation. Based on the developed combination prediction model, the residuals between the predicted value and measured value of the condition parameters can be calculated. In normal condition, the prediction residuals change in a small range centered at 0. In the abnormal condition, the residuals show a high fluctuation with large amplitude. Effective quantization for the abnormal level of the prediction residuals is the basis for WT anomaly identification. The information entropy could effectively quantify the order degree of the systems or the complexity of the signals (Zhou, 2012). In this paper, the information entropy is utilized to quantify the intensity of the change of prediction residuals.

The information entropy of prediction residual with the interval 1 day can be calculated as follows:

$$H_d = - \sum_{i=1}^N \left( \frac{n_i}{T_d} \right) \ln \left( \frac{n_i}{T_d} \right) \quad (10)$$

where  $H_d$  is the information entropy with the interval 1 day;  $N$  is the number of the statistical intervals;  $n_i$  is the number of prediction residuals in the  $i$ th statistical interval. For example, if  $[-2, -1]$ ,  $[-1, 0]$ ,  $[0, 1]$  and  $[1, 2]$  °C are the statistical intervals for the temperature parameter,  $n_i$  represents the number of prediction residuals that fall into the  $i$ th interval.  $T_d$  is the number of prediction residuals in one day. Entropy is calculated only when  $n_i$  is not equal to 0.

Figure 7 shows the flowchart of anomaly identification of the WT condition parameters, which is based on the condition parameter prediction model and the information entropy. The procedure can be summarized into the following steps:

Step 1: The appropriate wind farm SCADA data are selected to form the training and testing datasets.

Step 2: The individual prediction models of the target condition parameter are developed. The optimal weight distribution is calculated and the combination prediction model is developed.

Step 3: Target parameter is predicted with the combination prediction model. The predicted values of the target parameter are compared with the actual monitoring value so as to obtain the prediction residuals.

Step 4: For reflecting the change tendency of the RMSE, continuous equivalent interval RMSE are calculated. The statistical interval is 1 day when calculating RMSE in the paper. Change situation of RMSE is obtained with calculating RMSE of each day.

Step 5: When the RMSE is smaller than the threshold, the condition parameter is considered as normal.

Step 6: When the RMSE is greater than the threshold, the information entropy of prediction residual data is calculated.

Step 7: The thresholds of information entropy for each condition parameter are determined through analyzing the distribution of information entropy of each parameter when the WT is under normal conditions. The condition parameter is considered as normal if the entropy is smaller than the threshold. In this case, the prediction residuals change slightly. Although RMSE is greater than the threshold, the corresponding condition parameter cannot be considered as abnormal.

Step 8: The condition parameter is considered abnormal if the entropy is greater than the threshold.

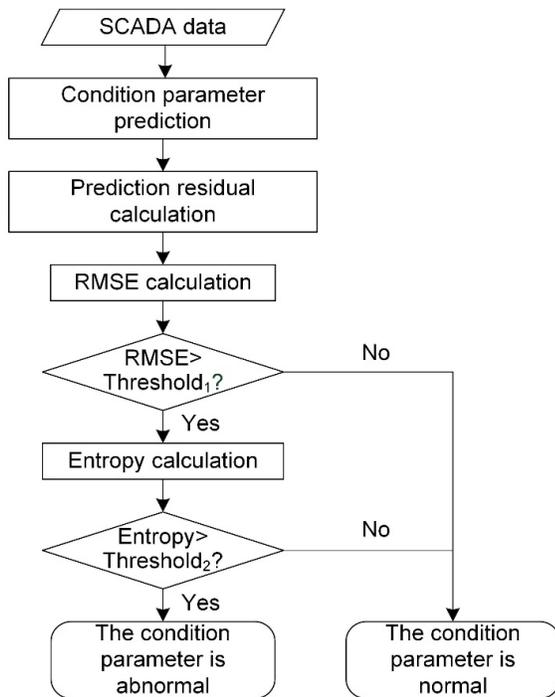


Figure 7. Flowchart of anomaly identification of condition parameters.

## 6. Verification and Analysis

Two cases are selected to validate the proposed WT fault detection method. In Case 1, the proposed method is verified when WT operates at normal condition. The proposed method is validated in case 2 when WT occurs fault.

### 6.1 Verification under normal condition

In this section, SCADA data of WT 17 during a period of 20 days (3000 data points) are collected during normal condition, which are used to verify the information entropy based anomaly identification method. Figure 8 shows the run chart of the generator bearing B temperature. It can be seen that the component temperature shows obvious dynamic characteristics under varying wind speed. The highest temperature is 85 °C and the lowest temperature is 27 °C. Both the highest and the lowest temperature do not exceed the threshold value ( $\pm 95$  °C).

The generator bearing B temperature is predicted with the developed WT condition parameter prediction model. Figure 9a shows the prediction residuals of the combination prediction model of the generator bearing B temperature. It can be seen that most of the residuals are distributed in the range between -5 to 5 °C. The values of few residuals exceed this range mainly due to the collected raw SCADA data used to train the prediction models are disturbed. Figure 9b shows the histogram of the prediction residuals data. It can be seen that the prediction residuals coming from the trained prediction models are normally distributed with a mean around zero. The distribution of residual data is consistent with the Gaussian distribution.

Figure 9c shows the run chart of the RMSEs on a daily basis. The RMSEs change slightly and are mainly distributed between 2 and 3 °C during the normal condition of WTs. The information entropy of the residual is calculated according to Equation (5). The run chart of the information entropy is shown in Figure 9d. The entropy changes in a small range centering at 2 when the WTs operate at normal condition.

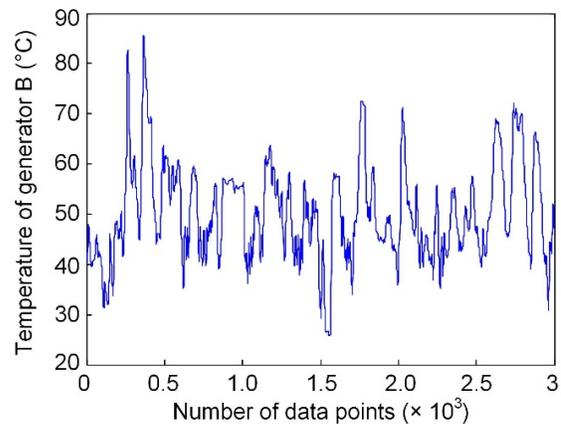
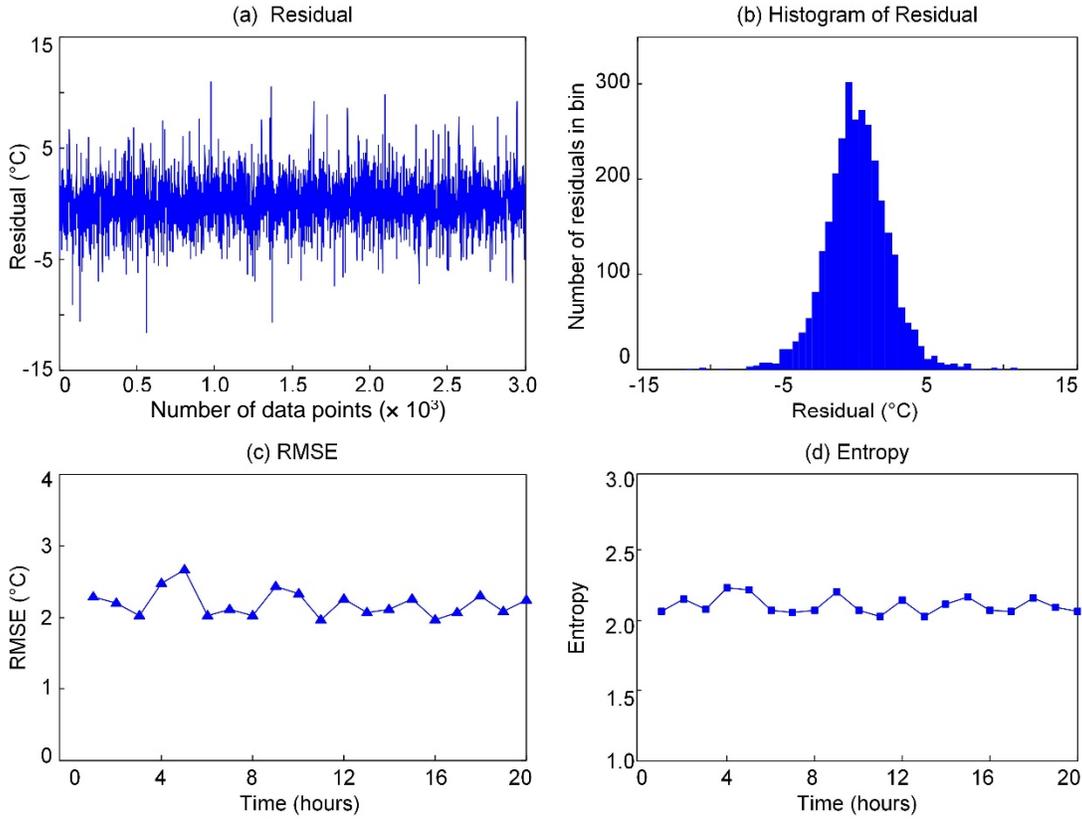


Figure 8. Temperature of generator bearing B of WT 17.



**Figure 9.** Residual, RMSE and entropy of parameter model for temperature of generator bearing B in WT17 under normal condition: (a) Residual; (b) Histogram of Residual; (c) RMSE; (d) Entropy.

The proposed model for incipient WT fault detection performs well without misdiagnosis when WT operates during normal condition. The combination prediction model for generator bearing B temperature has a good performance on prediction accuracy. Besides, the information entropy distribution of prediction residual is consistent with the situation at WT normal condition.

### 6.2 Verification under Abnormal Condition

The SCADA data of WTs 17 and 31 are used to verify the proposed incipient WT fault detection method in this section.

According to maintenance records, WT 17 suffered a sudden breakdown on May 30, 2012 due to the overheating of generator bearing B. The SCADA data from March 1, 2012 (i.e., 90 days before the fault occurrence) to May 30, 2012 are analyzed. WT 31 experienced an overheating fault in bearing B on July 30, 2012. The SCADA data from May 16, 2012 (i.e., 73 days before the fault occurrence) to July 30, 2012 are analyzed. These two WTs have been in operation for three months before the fault.

The process of anomaly identification of the generator bearing B is given in this case study. The upper limit of the generator bearing B temperature of is 95 °C according to the setting values of WT parameters in the studied wind farm.

Figure 10 shows the analysis results of the WT 17's generator bearing B temperature. Figure 10a shows the run chart of the generator bearing B temperature. Based on the traditional prediction residual or RMSE threshold-based diagnosis method, the generator bearing B temperature is normal before the fault occurrence. Therefore, misdiagnosis occurs when using the threshold-based method. Figure 10b presents the prediction residuals of the combination prediction model. From the beginning to around the 10,000<sup>th</sup> data point, the prediction residuals are mainly distributed between -5 and 5 °C, and the WT 17's generator bearing B temperature is at normal condition. At about the 10,000<sup>th</sup> data point, the residual rapidly increases to 20 °C and decreases to -18 °C. The residuals become smaller during a short time after the 10,810<sup>th</sup> data point, and then show a great change of magnitude at about the 12,000<sup>th</sup> data point.

It can be seen from Figure 10b that data points with the absolute value of the residual greater than 5 °C are less. However, the data points with the absolute value of the residuals greater than 10 °C exist. Thus, it is difficult to identify these points are abnormal data or disturbance. Figure 10c shows the run chart of RMSEs with the interval 1 day. From the first day to the 70<sup>th</sup> day, the amplitudes of RMSE during normal condition are mainly smaller than 4 °C. It nearly reaches to 6.6 °C at the 75<sup>th</sup> day, then drops to around 4,

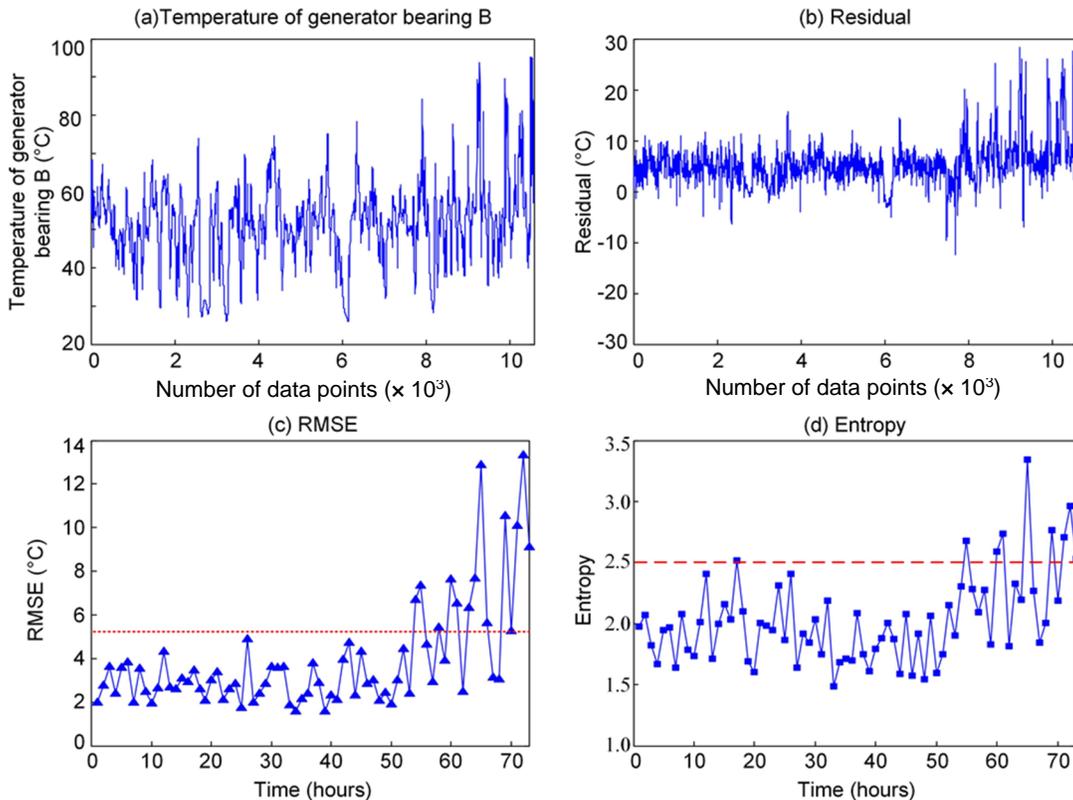
and increases rapidly two days before the fault occurrence. It can be observed that the RMSEs around the 75<sup>th</sup> day and before fault occurrence are significantly greater than the situations in the normal condition.

According to the analysis in Section 4.1, the distribution of the generator bearing B temperature prediction residual submit to the normal distribution during the normal condition. Therefore, the information entropy of prediction residual changes slightly. Figure 10d presents the run chart of the daily information entropy. It can be observed that the information entropy is smaller than 2.5 before the 75<sup>th</sup> day, and then decrease to 2. The information entropy increases more than 2.5 suddenly and rapidly before the fault occurs. The bearing fault of WT 17 leads to the anomalies in generator bearing B temperature. There is a great different between the predicted value and the measured value with the generator bearing B temperature. Compared with the residual distribution during normal condition, the residual distribution during the abnormal condition changes disorderedly. To detect the WT fault, the information entropy is employed to quantify the abnormal level of the residual. It can be observed that the information entropy is effective to detect the WT faulty condition by quantifying the abnormal levels of the residuals from Figure 10d.

Figure 11 shows the analysis results of the generator bearing B temperature of WT 31, which presents a similar

change tendency with that in Figure 10. The prediction residuals deviate without fluctuating around 0 °C mainly because the training data of the prediction model are collected from other WT. Due to the deviation of prediction residual in WT 31, the RMSE amplitudes in WT 31 are greater than the RMSE amplitudes in WT 17. Although the anomalies in generator bearing B temperature of WT 31 also can be seen from the change tendency of the prediction residual and RMSE in Figs. 11b and 11c, the abnormal level of condition parameter is unquantifiable. The incipient WT faulty cannot be detected with traditional residual or RMSE threshold-based method because the thresholds of prediction residual and RMSE are difficult to determine. The run chart of daily information entropy is presented in Figure 11d. It can be seen that the entropy is still smaller than 2.5 before the 55<sup>th</sup> day, and then increases rapidly more than 2.5 before the fault occurrence. The values of information entropy are both lower than 2.5 through analyzing the distribution of information entropy of generator bearing B temperature when the WT is under normal conditions. The condition parameter generator bearing B temperature can be considered abnormal. The generator bearing fault of WT 31 can be detected effectively through the combination prediction model and the analysis of information entropy.

However, the threshold values of information entropy for each condition parameter are different because of the different



**Figure 10.** Analysis results for SCADA data of WT 17: (a) Temperature of generator bearing B; (b) Residual; (c) RMSE; (d) Entropy.

distributions of each condition parameter. Therefore, it is viable to develop the uniform quantization standard for detecting anomalies in a certain condition parameter of different WTs.

### 7. Conclusions

In this work, a generalized model for WT faulty condition detection was developed using the combination prediction approach and the information entropy. SCADA datasets were selected for developing the prediction models of WT condition parameters that are dependent on environmental conditions such as ambient temperature and wind speed. The combination prediction models were developed for these WT condition parameters. The modeling method was illustrated by using the temperature of generator bearing B as an example. The BPNN model, the LSSVM model, and the RBFNN model were used as the individual prediction models. The weights of the individual prediction models were optimized to improve the prediction accuracy based on the principle of minimizing the sum-of-squares of the prediction residual. The information entropy was utilized to quantify the abnormal level of condition parameters. Two case studies for an onshore wind farm in Northern China have been carried out and analyzed. The results revealed that the proposed method is more effective in detecting the parameter anomalies prior to fault occurrence than traditional methods such as prediction residual or RMSE threshold-based method and a single-model-based method.

The linear combination prediction approach was demonstrated to be effective at developing the prediction models for WT condition parameters with better accuracy and generalization performance in this paper. The information entropy of prediction residual was also proved to be useful in establishing the uniform quantization standard for detecting faulty conditions of different WTs of a wind farm.

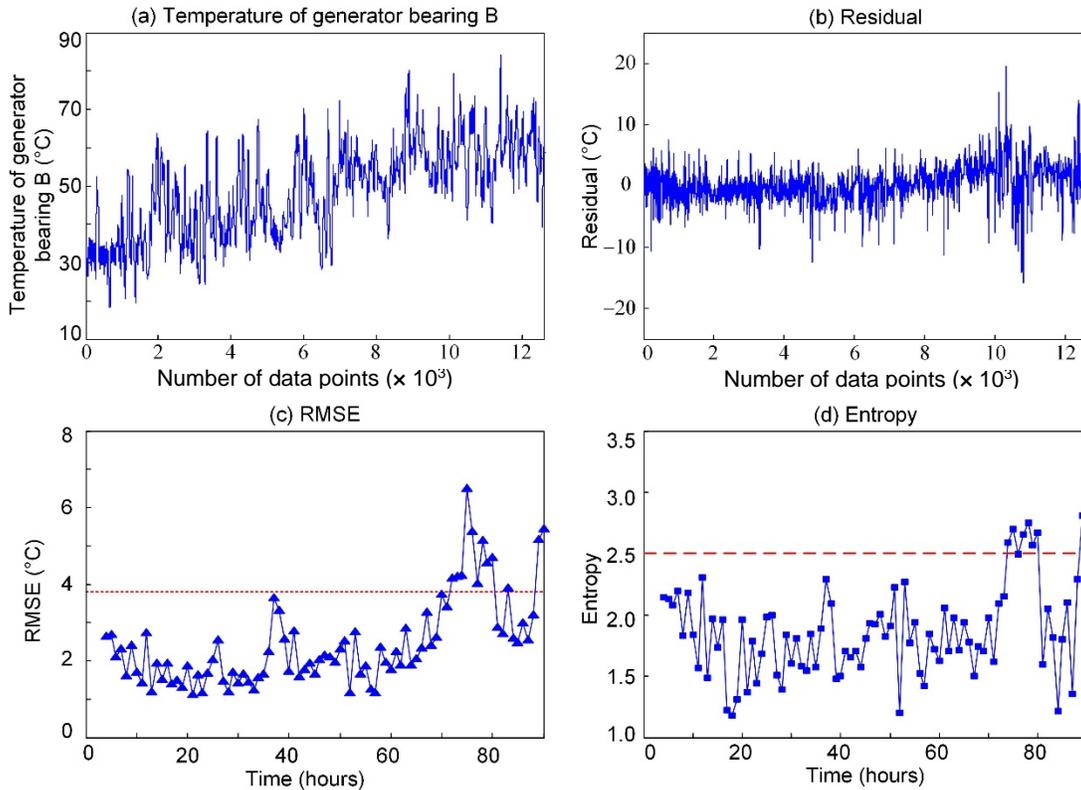
Due to the limited data that we collected at present, the impact of different training data (e.g., current SCADA data, historical SCADA data and SCADA data obtained on other WTs with different geographical positions) on the effectiveness of proposed WT fault detection method will be investigated in our future work.

We hope our work is useful for researchers who are interested in WT condition monitoring, fault detection and helpful for wind farm operators to monitor the health condition of WTs.

**Acknowledgments.** This research was supported by the Chinese Ministry of Science and Technology 973 Research Program (No. 2012CB215205) and the National “111” Project of the Ministry of Education of China (No. B08036).

### References

Ai Y.T., Fu Q. and Tian J. (2013). Diagnosis method for crack-



**Figure 11.** Analysis results for SCADA data of WT 31: (a) Temperature of generator bearing B; (b) Residual; (c) RMSE; (d) Entropy.

- rubbing coupled fault in rotor system based on integration of information entropy distance. *Journal of Aerospace Power*, 28 (10), 2161-2166. <https://doi.org/10.13224/j.cnki.jasp.2013.10.004>
- Assareh E. and Biglari M. (2015). A novel approach to capture the maximum power from variable speed wind turbines using PI controller, RBF neural network and GSA evolutionary algorithm. *Renewable Sustainable Energy Rev.*, 51(5-6), 1023-1037. <https://doi.org/10.1016/j.rser.2015.07.034>
- Ata R. (2015). Artificial neural networks applications in wind energy systems: a review. *Renewable Sustainable Energy Rev.*, 49, 534-562. <https://doi.org/10.1016/j.rser.2015.04.166>
- Cabal-Yepey E., Romero-Troncoso R. J., Garcia-Perez A. and Rodriguez-Donate C. (2010). Novel hardware processing unit for dynamic on-line entropy estimation of discrete time information. *Digital Signal Process.*, 20(2), 337-346. <https://doi.org/10.1016/j.dsp.2009.06.012>
- Chandola V., Banerjee A. and Kumar V. (2009). Anomaly detection: a survey. *ACM Comput. Surv.*, 41(3), 75-79. <https://doi.org/10.1145/1541880.1541882>
- Cross P. and Ma X.D. (2014). Nonlinear system identification for model-based condition monitoring of wind turbines. *Renewable Energy*, 71(11), 166-175. <https://doi.org/10.1016/j.renene.2014.05.035>
- Djurovic S., Crabtree C.J., Tavner P.J. and Smith A.C. (2012). Condition monitoring of wind turbine induction generators with rotor electrical asymmetry. *IET Renewable Power Generation*, 6(4), 207-216. <https://doi.org/10.1049/iet-rpg.2011.0168>
- GarcíaMárquez F.P., Tobias A.M., Pinar Pérez J.M. and Papaelias M. (2012). Condition monitoring of wind turbines: Techniques and methods. *Renewable Energy*, 46(5), 169-178. <https://doi.org/10.1016/j.renene.2012.03.003>
- Gil M.D.P., Gomis-bellmunt O. and Sumper A. (2014). Technical and economic assessment of offshore wind power plants based on variable frequency operation of clusters with a single power converter. *Appl. Energy*, 125(21), 218-229. <https://doi.org/10.1016/j.apenergy.2014.03.031>
- Johan R. and Lina M.B. (2007). Survey of Failures in Wind Power Systems With Focus on Swedish Wind Power Plants During 1997-2005. *IEEE Trans. Energy Conversion.*, 22(1), 167-173. <https://doi.org/10.1109/TEC.2006.889614>
- Kusiak A. and Li W.Y. (2010). Virtual models for prediction of wind turbine parameters. *IEEE Trans. Energy Conversion*, 25(1), 245-252. <https://doi.org/10.1109/TEC.2009.2033042>
- Kusiak A. and Verma A. (2011). The prediction and diagnosis of wind turbine faults. *Renewable Energy*, 36(1), 16-23. <https://doi.org/10.1016/j.renene.2010.05.014>
- Kusiak A. and Verma A. (2012). A data-mining approach to monitoring wind turbines. *IEEE Trans. Sustainable Energy*, 3(1), 150-157. <https://doi.org/10.1109/TSTE.2011.2163177>
- Lee J.R. and Kim H.C. (2013). Feasibility of in situ blade deflection monitoring of a wind turbine using a laser displacement sensor within the tower. *Smart Mater. Struct.*, 22(2), 100-104. <https://doi.org/10.1088/0964-1726/22/2/027002>
- Li R.Q. and Su H.Y. (2012). Risk assessment of spinning reserve scheme based on maximum entropy principle. *Power Syst. Technol.*, 36(3), 241-245. <https://doi.org/10.13335/j.1000-3673.pst.2012.03.006>
- Liu Y.Q., Shi J., Yang Y.P. and Lee W. (2012). Short-term wind-power prediction based on wavelet transform-support vector machine and statistic-characteristics analysis. *IEEE Trans. Ind. Appl.*, 48(48), 1-7. <https://doi.org/10.1109/ICPS.2011.5890873>
- Milborrow D. (2006). Operation and maintenance costs compared and revealed. *Windstats Newsl.*, 19(3): 1-3.
- Nourani V., Khanghah T.R. and Baghanam A.H. (2015). Application of Entropy Concept for Input Selection of Wavelet-ANN Based Rainfall-Runoff Modeling. *J. Environ. Inf.*, 26(1), 52-70. <https://doi.org/10.3808/jei.201500309>
- Ribrant J, Bertling L. (2007). Survey of failures in wind power systems with focus on Swedish wind power plants during 1997-2005. *IEEE Trans. Energy Conversion*, 22(1), 167-173. <https://doi.org/10.1109/TEC.2006.889614>
- Schlechtingen M., Santos I.F. and Achiche S. (2013). Wind turbine condition monitoring based on SCADA data using normal behavior models. Part I: system description. *Appl. Soft Comput.*, 13 (1), 259-270. <https://doi.org/10.1016/j.asoc.2012.08.033>
- Schlechtingen M. and Santos I.F. (2011). Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection. *Mech. Syst. Signal Process*, 25(5), 1849-1875. <https://doi.org/10.1016/j.ymssp.2010.12.007>
- Soua S., Lieshout P.V., Perera A., Gan T.H., and Bridge B. (2013). Determination of the combined vibrational and acoustic emission signature of a wind turbine gearbox and generator shaft in service as a pre-requisite for effective condition monitoring. *Renewable Energy*, 51(2), 175-181. <https://doi.org/10.1016/j.renene.2012.07.004>
- Sun P., Li J., Wang C.S., and Lei X. (2016). A generalized model for wind turbine anomaly identification based on SCADA data. *Appl. Energy*, 168, 550-567. <https://doi.org/10.1016/j.apenergy.2016.01.133>
- Tavner PJ, Xiang J and Spinato F.(2007). Reliability analysis for wind turbines. *Wind Energy*, 10(1), 1-18. <https://doi.org/10.1002/we.204>
- Wisner D.A. and Chattergy R. (1978). *Introduction to nonlinear optimization: a problem solving approach*, North-Holland Pub. Co.
- Yan Y.L., Jian L. and David W.Z.G. (2014). Condition parameter modeling for anomaly detection in wind turbines. *Energies*, 7(5), 3104-3120. <https://doi.org/10.3390/en7053104>
- Yang W.X., Richard C. and Jiang J.S. (2013). Wind turbine condition monitoring by the approach of SCADA data analysis. *Renewable Energy*, 53(9), 365-376. <https://doi.org/10.1016/j.renene.2012.11.030>
- Yang W.X., Tavner P.J. and Court R. (2013). An online technique for condition monitoring the induction generators used in wind and marine turbines. *Mech. Syst. Signal Process.*, 38(1), 103-112. <https://doi.org/10.1016/j.ymssp.2012.03.002>
- Yang W.X., Tavner P.J., Crabtree C.J., Feng Y. and Qiu Y. (2014). Wind turbine condition monitoring: technical and commercial challenges. *Wind Energy*, 17(5), 673-693. <https://doi.org/10.1002/we.1508>
- Yang Y.H., Guergachi A. and Khan G. (2006). Support Vector Machines for Environmental Informatics: Application to Modelling the Nitrogen Removal Processes in Wastewater Treatment Systems. *J. Environ. Inf.*, 7(1), 14-23. <https://doi.org/10.3808/jei.200600063>
- Yang Y., Zhang C.T. and Zhang R.X. (2015) Improving Environmental Prediction by Assimilating Auxiliary Information. *J. Environ. Inf.*, 26(2), 91-105. <https://doi.org/10.3808/jei.201500316>
- Zaher A., McArthur S.D.J., and Infield D.G. (2009). Online wind turbine fault detection through automated SCADA data analysis. *Wind Energy*, 12(6), 574-593. <https://doi.org/10.1002/we.319>
- Zhang L., Xiong G., Liu H., Zou H., and Guo W. (2010). Bearing fault diagnosis using multi-scale entropy and adaptive neuro-fuzzy inference. *Expert Syst. Appl.*, 37(8), 6077-6085. <https://doi.org/10.1016/j.eswa.2010.02.118>
- Zhou Y.Q. (2012). *Information Science Theory*. Beijing University of Aeronautics and Astronautics Press, Beijing.