



Article Diagnosis of Blade Icing Using Multiple Intelligent Algorithms

Xiyun Yang ^{1,2}, Tianze Ye ^{1,*}, Qile Wang ³ and Zhun Tao ¹

- ¹ School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China; yangxiyun916@ncepu.edu.cn (X.Y.); 18618162564@163.com (Z.T.)
- ² Key Laboratory of Condition Monitoring and Control for Power Plant Equipment of Ministry of Education, North China Electric Power University, Beijing 102206, China
- ³ Zhong-Neng Power-Tech Development Co., Ltd., Beijing 100089, China; 12070097@chnenergy.com.cn
- * Correspondence: yetianze613@163.com or yetianze@ncepu.edu.cn

Received: 11 May 2020; Accepted: 5 June 2020; Published: 9 June 2020



Abstract: The icing problem of wind turbine blades in northern China has a serious impact on the normal and safe operation of the unit. In order to effectively predict the icing conditions of wind turbine blades, a deep fully connected neural network optimized by machine learning (ML) algorithms based on big data from the wind farm is proposed to diagnose the icing conditions of wind turbine blades. This study first uses the random forest model to reduce the features of the supervisory control and data acquisition (SCADA) data that affect blade icing, and then uses the K-nearest neighbor (KNN) algorithm to enhance the active power feature. The features after the random forest reduction and the active power mean square error (MSE) feature enhanced by the KNN algorithm are combined and used as the input of the fully connected neural network (FCNN) to perform and an empirical analysis for the diagnosis of blade icing. The simulation results show that the proposed model has better diagnostic accuracy than the ordinary back propagation (BP) neural network and other methods.

Keywords: random forest algorithm; k-nearest neighbor; fully connected neural network; blade icing recognition

1. Introduction

Recently, with the continuous development of the renewable energy industry, the cumulative installed wind power capacity in this word has greatly increased, and wind power has become a major contributor to power generation [1]. In order to make better use of wind energy, wind turbines (WTs) are widely built in high altitude areas with cold climates and high humidity. However, in such operating environment, the WT is prone to the phenomenon of blade icing, which may cause many problems [2]. On the one hand, after the ice accumulates on the blade, the airfoil changes, which reduces the ability to capture wind energy, and leads to increased consumption of energy to drive the blade to rotate, that ultimately reduces the power generation efficiency. On the other hand, icing changes the modal parameters of the corresponding area on the blade, which may cause the blade to break, leading to more serious operating accidents. Therefore, when ice accumulation on the blade is detected, the deicing equipment should be started immediately. Accordingly, timely detection of icing is of great significance to enhance the power generation efficiency and service life of WTs in wind farms.

Due to the convenient access and huge amount of data provided, the research of WT fault detection based on supervisory control and data acquisition (SCADA) data has been extensively studied. Reference [3] proposed a WT fault detection model based on SCADA data, which used a variety of data mining algorithms. This model can predict failures within 5–60 min before they

occur. In a subsequent study, Reference [4] used principal component analysis (PCA) to decrease the data dimensionality, and then the random forest (RF) is used to identify early faults. Reference [5] proposed an alarm processing and diagnosis scheme for WT SCADA systems based on Artificial Neutral Network (ANN) for identifying faults in the pitch system. Their simulation results showed that the method can quickly identify faults.

It can be seen from the above references that the WT status monitoring and fault detection based on SCADA data has a good effect. However, many current research focuses on the monitoring and diagnosis of generators, gearboxes, and pitch systems. Few studies have used SCADA data to detect icing on WT blades. Recently, Reference [6] proposed a blade icing detection method based on random forest algorithm, which contains 29 kinds of characteristic parameters in a SCADA system. Reference [7] proposed a hybrid fault detection system that integrates multiple intelligent algorithm. This failure detection strategy can accurately detect the early failure of the blade, and can improve maintenance costs and system availability. Reference [8] uses SCADA data to construct the features of wind power, wind speed, and generator speed. And uses them as the input of support vector machine (SVM), and combined with particle swarm optimization (PSO) algorithm to establish icing detection model. Reference [9] proposed a SVM fault diagnosis method based on big data analysis, and used multiple wind turbine data for verification, effectively proving the effectiveness and generalization ability of the model.

Since the SCADA data cannot intuitively obtain the health status of wind turbines, feature enhancement is particularly critical when using SCADA data to assess the health status of various components of wind turbines. According to reference [10–12], common research is to extract fault features based on expert experience. The lack of expertise and slow manual selection will affect the performance of the model. In recent years, deep learning algorithms have been widely used in many fields [13–15], and its performance is also constantly improving. Since deep learning is good at extracting features, recently fully connected neural network (FCNN) algorithm have been widely used in the field of machine health monitoring and fault diagnosis. Reference [16] used the FCNN method to extract frames from a video and ran a classifier to perform supervised learning and classification of objects in order to obtain different classes of probabilities, thereby classifying the subject matter and detecting any objects in the video. Compared with earlier similar methods, its accuracy has been improved. Reference [17] studied the generalization ability of FCNNs trained in the context of time series detection, and studied how to control the generalization ability of the network by adjusting control variables. With these hyperparameters, the complexity of the output function can be effectively controlled without imposing explicit constraints. Reference [18] proposed a blade icing detection model using a deep autoencoders network, and compared with the traditional machine learning (ML) models, demonstrated the high accuracy and generalization ability of the proposed method.

Although deep learning has been successfully applied in unsupervised feature extraction, it can be further improved. The methods proposed in the above studies for selecting feature subsets from the original feature sets all achieve to some extent effective dimension reduction of high-dimensional data. However, the current method to select input features for WT blade icing prediction models still requires effective comprehensive research. Accordingly, in this study an RF model is used to reduce the features of the SCADA data that affect blade icing, and then a K-nearest neighbor (KNN) algorithm is used to enhance the active power feature. The features after the RF reduction and the active power mean square error (MSE) feature enhanced by the KNN algorithm are combined and used as the input of the FCNN. Then, an empirical analysis is performed for the diagnosis of blade icing. The simulation results show that the proposed model has better diagnostic accuracy than the ordinary back propagation (BP) neural network and other methods, and FCNN–MSE method has excellent chronergy and applicability.

The rest of the paper is arranged as follows: Section 2 introduces the theoretical method of the RF algorithm. Section 3 introduces the application of the KNN algorithm in this study. Section 4 describes details of the structure of the FCNN model and optimization methods used in this study. Section 5

introduces the blade icing detection model and steps proposed in this study. Section 6 provides a detailed experimental analysis. Section 7 gives the conclusions drawn from this study.

2. Random Forest (RF)

2.1. RF Classification (RFC) Theory

RFC is a classification model that contains multiple decision trees, and each decision tree votes to select the best result. The basic procedure of RFC is as follows: First, the bootstrap sampling is used to extract k samples from the original training set, where the sample size of each sample is the same size as the original training set. Second, establish the k decision trees for k samples, and obtain the k categories results. Lastly, vote for the final classification result based on the k classification results. In order to deal with the multi-dimensional feature signals in WT blade icing recognition and enhance the detection ability of the model, the RF method is used in this study to reduce the feature dimension in the SCADA data. The classification process is shown in Figure 1.



Figure 1. Flow chart of random forest.

2.2. The Feature Selection of RF

Let the sample set be $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$, the input sample set $X = (x^{(1)}, x^{(2)}, \dots, x^{(n)}) \subseteq \mathbb{R}^n$ be the input space, and the class label set $Y = \{c_1, c_2, \dots, c_L\}$ be the output space, where the i-th sample is $x_i = (x_i^{(1)}, x_i^{(2)}, \dots, x_i^{(n)})$, and $x_i^{(j)}$ represents the value of the sample x_i on the j-th feature.

The RF feature selection algorithm measures the importance of each feature, uses it as a basis to rank the features, and then selects features based on the minimum out of bag (OOB) error rate criterion. The basic procedure of the RF feature selection algorithm for single feature importance measurement is as follows: After adding noise to a related feature, the accuracy of the prediction decreases, and the change in the accuracy rate measures the importance of this feature [19]. The RF feature selection removes the redundant features of the original data in the WT SCADA system, reduces noise interference, makes the selected feature indicators more representative, and effectively improves the accuracy of classification. RF construction:

For the sample set $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$, K times bootstrap sampling are performed to generate K self-help sample sets B_k and OOB sample sets $(OOB_k), k = 1, 2, \dots, k$, and a meta classifier

 $C_k(x)$ is established for the self-help sample set B_k . Then, the classification result of any sample x_i on the combination classifier C * (x) is:

$$\hat{y} = C * (x_i) = \arg \max_{y \in Y} \sum_i \delta(C_k(x_i) = y)$$
(1)

where $\delta(\cdot)$ is an illustrative function. When the parameter is true, $\delta(\cdot) = 1$ or $\delta(\cdot) = 0$. C * (x) is called RF.

3. Wind Power Feature Enhancement Using the KNN Method

3.1. The KNN Theoretical Background and Case Analysis

In the regression problem, the y_{out} output after KNN analysis for a given input x_{in} is obtained after a priori calculation based on the known k inputs $x_{k,in}$ and k outputs $y_{k,out}$. Reference [20] used the KNN method in power curve representation research, and emphasized the superiority of this method compared to other data mining algorithms. The power generation and wind speed of a WT installed in a cold climate in one year is shown in Figure 2. During early February 2018, as shown in the enlarged area in Figure 2, there was a situation where the power was 0 for a long time, and the WT stopped for a long time, which was caused by the icing of the blades. Two sets of wind speed-power sets are presented in Figure 3. The blue dots displayed follow the normal WT power curve, while the red dots represent a considerable deviation. When icing forms on the WT blades, its power output will be different from the normal power curve value, and the icing will cause the unit to stop. Therefore, the degree of deviation between the output power of the icing unit and the normal power output is an important feature of icing WT units. Under the same icing condition, the output power fluctuates under different wind speed conditions. Thus, this feature variable is clearly better than the icing information contained in the value of the output power of the WT unit. The main idea of the KNN introduced in this paper is to reconstruct the power curve based on the blue dots, and then compare the red dots to calculate the error between them, in order to accurately extract the information about the degree of deviation of the output power of the icing WT unit. Considering the impact of the site environment, the reconstructed new power curve is limited to specific WT units.



Figure 2. Wind speed and power curve of a wind turbine in 2018.



Figure 3. Wind turbine (WT) speed-power scatter plot (red is the freezing point and blue is the normal point).

The power points based on data from Figure 2, all of which were collected when the WT is in operating mode, are shown in Figure 3. The scattered dots in red indicate that ice has accumulated on the WT blades, but are still in operation. Therefore, the main goal of the KNN-based power curve analysis is to distinguish those outliers in order to start the deicing device or stop the wind turbine in time.

3.2. Calculation of the Best K-Nearest Neighbor

The KNN method enhances the weak features in the original data to become strong features after calculation, thereby playing an important role in the prediction of blade icing, by enhancing the stability of the prediction results and the accuracy of icing recognition. The KNN method largely depends on the best choice of the amount of best neighbors K_{opt} . The K_{opt} is calculated based on two data sets, namely the training set S_{train} and validation set S_{train} , and there are no failure points in the training set. The first step is to sort the training and validation sets in ascending order according to the wind speed, and calculate the error between the validation and training sets. Then, the MSE between the power of each verification point and the training point from the nearest neighbor number from 1 to K_{max} is calculated. The sum of the errors is averaged by the amount of verification points N_{valid} . The amount of best neighbors K_{opt} corresponds to the minimum error E_{min} , as shown in Equation (2).

$$E_{\min} = \min\{\frac{1}{N_{valid}} \sum_{i=1}^{N_{valid}} \sum_{k=1}^{K_{\max}} (t_{train,k} - t_{valid,i})^2\}$$
(2)

4. Deep Fully Connected Neural Network Prediction Model for Blade Icing

The traditional BP neural network algorithm has the following drawbacks: (1) the requirements for feature selection are high. The introduction of irrelevant variables will increase noise data and reduce model accuracy. (2) When the number of hidden layers is increased, gradient disappearance and gradient explosion problems occur, thereby resulting in a partial optimal solution. (3) Overfitting problems occur. Various scientists have proposed the concepts of deep learning and deep neural network (DNN) on this basis.

In the blade icing prediction model based on deep FCNN, the structure and related optimization methods adopted by deep learning can overcome the above problems, which can find hidden features of deep-level WT information, and fewer iterations, and have a more powerful nonlinear fitting and self-learning ability.

Deep Neural Network Structure

The structure of the deep FCNN in this paper is basically similar to that of the BP neural network. Each neuron weights and sums the input components and selects the corresponding activation function f. Too few hidden neurons and hidden layers result in a model with poor non-linear learning ability, which cannot deeply explore the hidden features of the WT information. Too many neurons and hidden layers result in a model that is highly redundant. Too many parameters are difficult to train, and at the same time may cause overfitting problems. The amount of neurons and hidden layers is mainly determined by experience and cut-and-trial method.

The internal neural network structure of the deep FCNN is shown in Figure 4, where the relationship between layers is a fully connected relationship. The input layer neurons are set to the determined number of features, and the pre-processed WT information dataset is used as input. After the cut-and-trial method, the amount of neurons in the hidden layers 1–3 are all set to 50, and the output layer neurons are set to 2. The following formula is used to determine whether the blade is frozen.

$$a_{j}^{l} = f(\sum_{k=1}^{m} w_{jk}^{l} a_{k}^{l-1} + b_{j}^{l})$$
(3)

where *m* is the number of neurons in layer l - 1; a_k^{l-1} represents the output of the k-th neuron in layer l - 1 as input the output of the j-th neuron in layer *l* is represented by a_j^l ; w_{jk}^l denotes the weight of the k-th neuron in layer l - 1 to the j-th neuron in layer *l*; b_j^l is the bias of j-th neuron in layer *l*; and *f* is the activation function.



Figure 4. The structure of fully connected neural network (FCNN) method.

The final activation function of the output layer for the icing classification of wind power blades is the softmax function. The purpose is to convert the output value of the output layer into a probability value in the interval (0,1), which is expressed as:

$$o^{i} = \operatorname{softmax}(y_{i}) = \frac{e^{y_{j}}}{\sum\limits_{k=1}^{K} e^{y_{k}}}$$

$$(4)$$

where o^i is the softmax layer output of the i-th sample point; y_j is the j-th neuron output of the output layer; and *K* is the number of output neurons of the output layer.

5. Research Method Proposed in this Paper—The FCNN-MSE Model

The main steps of WT blade icing prediction based on deep FCNN are as follows:

- (1) Wind power unit data preprocessing: process the missing values, abnormal values, and duplicates in WT data, and data normalization except wind power;
- (2) Use the feature selection function of the RF method to reduce the feature of the SCADA operating data;
- (3) Use the KNN method to calculate the best K value of the experimental data and the active power MSE. Then, use it as the enhanced power feature to replace the active power feature in step 2 and normalize it;
- (4) Establish a blade icing prediction model based on deep FCNNs, determine the neural network structure, and select the activation functions, weights initialization, iterations, batch size under batch gradients and batch size;
- (5) Input the training set containing enhanced power feature to the model for training, and obtain the blade icing prediction model based on deep FCNN; and
- (6) Input the test set to the model for prediction, and compare with the actual icing state of the blades to obtain the model detection accuracy.

The simulation process is shown in Figure 5.



Figure 5. Overview of the proposed icing detection model.

6. Experimental Simulation

6.1. Data Acquisition

The SCADA data of a 2.5 MW wind turbine located in Yunnan Province is used in this paper, China. The data has been collected since November 2010 according to the industrial standard SCADA system. Each piece of data is a moment in time (Each moment contains multiple continuous numerical monitoring variables, collected every 7 s) marked with normal or icing conditions, during which icing events of WT blades occur multiple times. The various features of the data are shown in Table 1.

This paper selected the SCADA data for the whole year of 2018 for processing. After screening, 6000 pieces of data are taken as experimental data. Among them, 4800 pieces of data are data pieces collected during normal operation of the WT blades, and 1200 pieces are operation data collected when the WT blades had icing accumulation.

Number	ID	Description	Number	ID	Description
1	WIND_SPEED	Wind speed	11	GENAPHSA	A-phase current
2	CONVERTER_MOTOR_SPEED	Generator speed	12	GENAPHSB	B-phase current
3	ROTOR_SPEED	Impeller speed	13	GENAPHSC	C-phase current
4	WIND_DIRECTION	Wind angle	14	GENVPHSA	A-phase voltage
5	TURYAWDIR	Yaw angle	15	GENVPHSB	B-phase voltage
6	GBXOILTMP	Gearbox oil temperature	16	GENVPHSC	C-phase voltage
7	GBXSHFTMP	Gearbox bearing temperature	17	GENHZ	Motor frequency
8	EXLTMP	Ambient temperature	18	TURPWRREACT	Reactive power
9	TURINTTMP	Cabin temperature	19	REAL POWER	Active power
10	GENGNTMP	Generator temperature			

Table 1. Nineteen feature parameters.

6.2. Evaluation Criteria

After constructing the experimental model and determining its structure, the model was evaluated with performance indicators. According to the predicted category and the actual category, the prediction results are divided into four types, namely the true positive (TP) type, which correctly identifies the blade in the icing state; the false positive (FP) type, which incorrectly assesses blades that are not frozen as frozen blades; the false negative (FN) type, which assesses that the blades in the frozen state are in the non-iced state; and the true negative (TN) type correctly recognizes that the blades are in the normal state. The confusion matrix results are shown in Table 2.

Table 2. Confusion matrix classification results.

Forecast State		0	1
A . (] . (0	True Positive (TP)	False Negative (FN)
Actual state	1	False Positive (FP)	True Negative (TN)

In order to obtain a stable generalization algorithm model, the hold-out method is used to divide the overall dataset into a training set (80%) and a testing set (20%) and then input them into the model. Four evaluation indicators can be obtained:

$$P = \frac{TP}{TP + FP} \tag{5}$$

$$R = \frac{TP}{TP + FN} \tag{6}$$

$$A = \frac{TP + TN}{TP + TN + FP + FN}$$
(7)

$$F_1 = 2\frac{P \cdot R}{P + R} \tag{8}$$

Among them, the precise rate (P) indicates the ratio of the actual blade icing state time to the predicted icing state time; the recall rate (R) represents the ratio of the predicted icing state time to the actual icing state time; the accuracy rate (A) indicates the ratio of the amount of correct classifications to the amount of test sets. F1 is a kind of harmonious score, which is an evaluation parameter for the classification effect of binary classification problems, taking into account the precise rate and accuracy rate.

6.3. Case Study

6.3.1. Feature Reduction of the WT SCADA Data Using RF

Excessive data sometimes does not lead to good experimental results, and instead will interfere with the results. Eliminating redundant features and reducing the noise interference factors, not only

simplifies the model structure, reduces the calculation complexity and makes the selected feature indicators more representative, but also effectively improves the accuracy of the experiment. The RF algorithm is used to reduce the features of the original dataset, and uses the 19 kinds of feature parameters collected by the SCADA system as inputs to obtain the weights of the different features in the prediction, as shown in Figure 6.



Figure 6. Weight of each feature in the diagnosis of blade icing.

The algorithm implemented in this paper uses a sequence backward search strategy when searching for a subset of features that achieves the maximum classification accuracy. The results of the feature selection process, shown in Figure 7 reveal that as the unimportant features (the features that are ranked last in the ranking of the importance of RF variables) are sequentially deleted, the classification accuracy rate as a whole gradually increases, mainly due to the improvement of the performance of the classifier after the elimination of unrelated and redundant features. When the classification accuracy reaches the highest value of 0.9561, it starts to show a downward trend, as the useful features are eliminated, which reduces the performance of the classifier. This result shows that the algorithm developed in this study can effectively identify and eliminate redundant and irrelevant features, thereby improving the classification performance of the classifier.



Figure 7. Relationship between features combination and blades icing recognition accuracy.

6.3.2. Feature Enhancement Based on the KNN Algorithm

The training dataset corresponds to data from the beginning of June 2018 and is limited to 50 points, which is $K_{max} = 50$. In addition, the validation dataset, selected in late June 2018, consists of approximately 200 wind speed–power points. As the state is underfitted before reaching the optimal value, the corresponding MSE value is larger when there are fewer neighbors. Usually the best neighbor number of the wind power is 20, but due to the phenomenon of overfitting, the error trend gradually increases. Experimental verification shows that the optimal K value of this dataset is equal to 13, as shown in Figure 8. However, if different training or validation data sets are used, the results may be slightly different than expected.



Figure 8. Mean square error (MSE) at different K values.

6.3.3. Feature Enhancement Verification of Active Power

After the RF algorithm performs feature reduction, nine of the 19 features of the original data remain. As shown in Table 3, the input data for this set of features is named RF Simplify (RFS). After applying the KNN algorithm to enhance the active power, the original active power feature is converted into the active power MSE feature, and the active power feature in the nine feature parameters is replaced with the active power MSE feature to obtain a set of new feature input data. This set of input data is named the RFS-KNN Refine (RFS–KNNR).

Table 3.	Final	features	combina	ation after	random	forest	(RF)	features	reduction.
----------	-------	----------	---------	-------------	--------	--------	------	----------	------------

Optimal Feature Combination						
Wind speed	Gearbox bearing temperature	Cabin temperature				
Reactive power	Ambient temperature	Impeller speed				
Active power	Generator temperature	Wind angle				

In order to verify the effect of the enhancement of the active power MSE feature, the above two RFS and RFS–KNNR input data are separately calculated by the RF algorithm, and the weight coefficients of each feature parameter in the icing state recognition are obtained, as shown in Table 4.

The data in Table 4 reveal that, through the RF algorithm, the weight of the reactive power feature in the RFS method ranks first with a value of 0.350813 Also, the three features of the impeller speed, gearbox bearing temperature, and generator temperature also play important roles. The weight ratios are 0.187269, 0.152058, and 0.149669, respectively, and the active power feature is ranked seventh, at 0.0286996. After the KNN method is used to enhance the active power feature, its effect is fully demonstrated in the results of the RFS–KNNR method. Among them, the reactive power feature still has the largest weight, at 0.313183. The second is the active power MSE feature, at 0.182323.

The generator temperature, impeller speed, and gearbox bearing temperature are the next three features, with weights of 0.161385, 0.149864, and 0.112182, respectively. Among them, the weight ratio of the active power MSE is 0.182323, which is higher than the original power weight ratio of 0.0286996. The above results show that solving the MSE of the active power feature and extracting useful information from the data improves the active power feature weights by 15.36%, and increases the sensitivity to icing recognition. It provides higher quality input data for the subsequent icing recognition process of deep FCNNs.

Number	Feature Name	Weights Comparison of the Two Methods				
	i cuture i vuine	RFS	Ranking	RFS-KNNR	Ranking	
А	Wind speed	0.0381619	6	0.0169116	7	
В	Impeller speed	0.187269	2	0.149864	4	
С	Wind angle	0.0579268	5	0.0446717	6	
D	Gearbox bearing temperature	0.152058	3	0.112182	5	
Е	Ambient temperature	0.0223962	8	0.0117826	8	
F	Cabin temperature	0.0130077	9	0.0076973	9	
G	Generator temperature	0.149669	4	0.161385	3	
Н	Reactive power	0.350813	1	0.313183	1	
Ι	Active power (used in RFS)/Active power-MSE (used in RFS-KNNR)	0.0286996	7	0.182323	2	

Table 4. Comparison of icing weights for each feature of RF Simplify (RFS) and RFS-KNN Refine (RFS–KNNR).

6.3.4. Selection of Model Parameters in Deep FCNNs

When using a deep learning optimization algorithm, that is, a deep FCNN, the results of using the rectified linear unit (ReLU) function as the activation function and the Tanh function as the activation function are shown in Figure 9. When the Tanh function is used as the activation function, the accuracy curve rises quickly, but the accuracy does not increase significantly with the number of iterations, and may fall into a local optimum. When the ReLU function is used as the activation function, the accuracy curve is at a high level, and the curve is relatively smooth when it reaches stability, the accuracy of the testing set is significantly higher than that of the Tanh function curve.



Figure 9. Accuracy curves using the rectified linear unit (ReLU) activation function or Tanh activation function.

Different batch sizes also affect the convergence accuracy of the model, that is, the average error of the training set, and the average error of the testing set. In the simulation, 10 experiments are set

for different batch sizes, and the number of iterations is set to 100, and eventually the F1 scores are averaged. The results, shown in Figure 10, reveal that when the batch size is 40, the F1 value of the final testing set is relatively good, at 0.9658.



Figure 10. Curve of batch size and F1 scores.

6.3.5. Analysis of Icing Recognition Results Based on FCNN with Active Power MSE

In order to verify the diagnostic performance of the proposed method, this paper uses BP, FCNN, SVM, KNN, and RF to perform simulation and comparison experiments. When the RFS data set is used as the input of the classification model, the performance indicators of the above five models are obtained as shown in Table 5. When the RFS–KNNR data set is used as the input of the classification model, the performance indicators of the corresponding models are obtained as shown in Table 6. It is shown that in order to distinguish the models, the name of the model using the RFS–KNNR data set is suffixed with "-MSE". For the above five models, the testing set (5000 groups) and the training set (1000 groups) are divided, and the simulation score results are shown in Tables 5 and 6. The two FCNN models using deep learning algorithms are more accurate in the recognition of blades icing conditions, both reaching 95% in accuracy and F1 scores, and the FCNN–MSE method in F1 scores is 97.69%.

-	Accuracy	Precision	Recall	F1
BP	97.30%	83.33%	99.24%	90.59%
FCNN	99.10%	96.21%	96.96%	96.58%
SVM	96.71%	90.38%	89.77%	90.07%
KNN	94.12%	87.33%	90.64%	88.95%
RF	93.53%	85.07%	90.11%	87.51%

Table 5. Classification results of five algorithms.

Table 6. Classification results of five algorithms combined with RFS-KNNR data set.

-	Accuracy	Precision	Recall	F1
BP-MSE	97.50%	84.87%	98.47%	91.17%
FCNN-MSE	99.40%	98.45%	96.96%	97.89%
SVM-MSE	97.10%	91.86%	92.06%	91.96%
KNN-MSE	95.49%	89.08%	91.48%	90.26%
RF-MSE	94.65%	87.27%	91.81%	89.48%

When the input is the RFS–KNNR data set, the performance indicators of each classification algorithm are shown in Table 6. The results show that after using the RFS–KNNR data set, the classification performance of each classifier has been greatly improved.

In the diagnosis of blade icing by the RF model, the recall rate R, the accuracy rate A, and the precise rate P are 90.11%, 93.53%, and 85.07% respectively, and the F1 score is 87.51. Compared with the RF model, all indicators of the RF-MSE model have improved, with the F1 score increasing to 89.48%. When the KNN model algorithm is used for icing diagnosis, its characteristic is that the algorithm is simple. At the same time, the evaluation score of the model is quite different from the deep FCNN and the SVM algorithm model, and the F1 score is about 87%. The indicators of the KNN–MSE algorithm are slightly improved, and the F1 score is about 90%. When using the SVM model algorithm for blade icing diagnosis, the F1 score is 90.07%, the model evaluation is good, but the calculation time is too long, which is suitable for non-linear problems with a small number of samples. The accuracy rate A of the SVM–MSE model reached 97.10%, and the F1 score rose to 91.96%.

The BP neural network algorithm has a simple structure and has strong non-linear mapping capabilities. According to the blade icing diagnosis results obtained with the BP model, shown in Tables 5 and 6, the recall rate R and the accuracy rate A are 99.24% and 97.30%, respectively, but the precise rate P is only 83.33%. Overall, the F1 only scores 90.59%. In the diagnosis results obtained with the BP-MSE method after processing the active power feature, the recall rate R and the accuracy rate A are still high, the precise rate P is also increased to reach 0.8487, and the value of F1 is increased to 91.17%. The deep FCNN and BP neural network algorithms are used for icing prediction. Their characteristics are that the model algorithms are relatively complicated, and more parameters need to be adjusted. The BP neural network algorithm does not require high feature selection, and the calculation overhead is smaller than that of other algorithms. However, in the evaluation score, the deep FCNN is better than the BP neural network. The results in Tables 5 and 6 show that although the recall rate R of the FCNN method is slightly lower than that of the BP neural network method, their accuracy P and accuracy A are significantly higher than those of the BP neural network method, at 96.21% and 99.10%, respectively. In the diagnosis results of the FCNN-MSE method, after processing the active power feature, except for the similar recall rate R, the precise rate P, accuracy rate A, and F1 are higher than those of the ordinary FCNN method, which are 98.45%, 99.40%, and 97.69%, respectively. The results show that the method of FCNN optimized by deep learning methods is better than the BP neural network method. Additionally, the accuracy of its testing set is significantly better than the accuracy of the testing set of the BP neural network algorithm. The evaluation score of the FCNN–MSE model is also better than that of the BP-MSE, and was significantly improved. In addition, the FCNN-MSE model proposed in this paper exhibited better overall diagnosis accuracy.

6.3.6. Chronergy of Blade Icing Detection by FCNN-MSE Method

Based on the above experimental results and the SCADA data of a single wind turbine from 31 January to 1 February 2018, we verified and analyzed the chronergy and the applicability of the FCNN–MSE method. As shown in Figure 11, A result value of 1 means icing, and a result value of 0 means normal. The blade has two icing phenomena within two days. For the first time icing, the original system detected icing at 2:10 on 31 January, and the FCNN–MSE method detected icing at 1:30 on 31 January. Considering the detection accuracy rate of the FCNN–MSE method, at a confidence level of 99.40%, we believe that the detection result of the FCNN–MSE method is 40 min earlier than the original system. Similarly, for the second time icing, the original system detected icing at 0:00 on 1 February, while the FCNN–MSE method detected icing at 23:30 on 31 January. Therefore, FCNN–MSE has excellent chronergy and applicability.



Figure 11. Comparison of timeliness of FCNN–MSE method and the original system for blade icing detection (The blue line represents the detection result of the FCNN–MSE method, while the green line represents the detection result of the original system).

7. Conclusions

This paper proposes a novel blade icing detection scheme based on SCADA data integrating FCNN and ML algorithms. Using deep FCNN algorithm to extract effective fault features from the SCADA data. ML algorithms, such as RF and KNN algorithms are used to enhance the accuracy and generalization of the models. And using the actual operation data from the wind farm to verify the performance of the proposed method. The comparison with other traditional ML methods showed that the proposed method in this paper has higher detection accuracy, generalization ability, excellent chronergy, and applicability.

Some conclusions drawn from this paper can be summarized as follows:

- 1. The WT blade icing detection model proposed in this paper comprehensively uses the RF algorithm, KNN algorithm, and deep FCNN model to separately perform feature reduction and feature enhancement on the SCADA data, and processes information at different feature levels to ultimately achieve higher detection accuracy and better performance.
- 2. FCNN is a useful deep learning method for adaptively extracting features from SCADA data. It can use more abstract functions to automatically extract features after the original data is preprocessed. This can greatly reduce reliance on expert experience.
- 3. The KNN method is used to analyze the wind speed–power dataset, in order to enhance the original power feature. The results show that in order to extract useful information, the squared sum of the obtained power errors needs to be averaged, which can enhance the stabilization effect and prediction accuracy.
- 4. Through the use of the RF algorithm, the data of the redundant features are removed without affecting the classification accuracy. As a result, the key features are extracted, showing a good processing ability in feature selection.

Author Contributions: Conceptualization, X.Y., T.Y. and Q.W.; Data curation, T.Y., Q.W. and Z.T.; Formal analysis, X.Y., T.Y. and Q.W.; Funding acquisition, X.Y.; Methodology, X.Y. and T.Y.; Project administration, X.Y. and Q.W.; Resources, X.Y. and Q.W.; Software, T.Y.; Supervision, X.Y. and Q.W.; Validation, X.Y., T.Y. and Z.T.; Visualization, T.Y.; Writing–original draft, T.Y.; Writing–review & editing, X.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (grant Number 51677067) and the Fundamental Research Funds for the Central Universities (grant Number 2018MS27).

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

References

- 1. *Global Wind Statistics* 2017; Global Wind Energy Council: Brussels, Belgium, 2018; Available online: https://gwec.net/newsroom/press-releases/ (accessed on 8 June 2020).
- 2. Wei, K.; Yang, Y.; Zuo, H.; Zhong, D. A review on ice detection technology and ice elimination technology for wind turbine. *Wind Energy* **2020**, *23*, 433–457. [CrossRef]
- 3. Zhen, C.; Yue, L.; Wu, Y. Research on cloud platform for wind turbine fault diagnosis. *Adv. Mater. Res.* 2013, 765, 2255–2258. [CrossRef]
- Kusiak, A.; Verma, A. Prediction of status patterns of wind turbines: A data-mining approach. J. Sol. Energy Eng. 2011, 133, 011008. [CrossRef]
- Chen, B.; Qiu, Y.; Feng, Y.; Tavner, P.; Song, W. Wind Turbine SCADA Alarm Pattern Recognition. In Proceedings of the IET Conference on Renewable Power Generation (RPG 2011), Edinburgh, UK, 6–8 Sepember 2011; pp. 1–6. [CrossRef]
- 6. Ge, Y.; Yue, D.; Chen, L. Prediction of wind turbine blades icing based on MBK-SMOTE and random forest in imbalanced data set. In Proceedings of the 2017 IEEE Conference on Energy Internet and Energy System Integration (El2), Beijing, China, 26–28 November 2017.
- Rezamand, M.; Kordestani, M.; Carriveau, R.; Ting, D.S.-K.; Saif, M. A New Hybrid Fault Detection Method for Wind Turbine Blades Using Recursive PCA and Wavelet-Based PDF. *IEEE Sens. J.* 2020, 20, 2023–2033. [CrossRef]
- 8. Guangfei, Z.; Wen, T.; Da, Z. Ice detection for wind turbine blades based on PSO-SVM method. *J. Phys. Conf. Ser.* **2018**, *1087*, 022036.21. [CrossRef]
- 9. Chuyang, L.I.; Mengzhao, Z.; Jian, J.; Wei, Z.; Yubo, Z. Fault diagnosis of wind turbine blade ice based on large data analysis. *Autom. Instrum.* 2020, *3*, 12–16.
- 10. Aijun, H.; Wanli, M.; Guiji, T. Rolling Bearing Fault Feature Extraction Method Based on Ensemble Empirical Mode Decomposition and Kurtosis Criterion. *Proc. CSEE* **2012**, *32*, 106–111.
- 11. Godse, R.; Bhat, S. Mathematical Morphology-Based Feature-Extraction Technique for Detection and Classification of Faults on Power Transmission Line. *IEEE Access* **2020**, *8*, 38459–38471. [CrossRef]
- 12. Chen, Y.Q.; Fink, O.; Sansavini, G. Combined Fault Location and Classification for Power Transmission Lines Fault Diagnosis with Integrated Feature Extraction. *IEEE Trans. Ind. Electron.* **2018**, *65*, 561–569. [CrossRef]
- 13. Gong, M.; Liu, J.; Li, H.; Cai, Q.; Su, L. A Multiobjective Sparse Feature Learning Model for Deep Neural Networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *26*, 3263–3277. [CrossRef] [PubMed]
- 14. Liao, J.; Liu, T.; Liu, M.; Wang, J.; Wang, Y.; Sun, H. Multi-Context Integrated Deep Neural Network Model for Next Location Prediction. *IEEE Access* **2018**, *6*, 21980–21990. [CrossRef]
- 15. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from ovefiting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
- Mazumdar, M.; Sarasvathi, V.; Kumar, A. Object recognition in videos by sequential frame extraction using convolutional neural networks and fully connected neural networks. In Proceedings of the International Conference on Energy Communication, Data Analytics and Soft Computing (ICECDS-2017), Chennai, India, 1–2 August 2017.
- 17. Borovykh, A.; Cornelis, W.; Sander, O.; Bohté, M. Generalization in fully-connected neural networks for time series forecasting. *J. Comput. Sci.* 2019, *36*, 101020, ISSN 1877-7503. [CrossRef]
- 18. Liu, Y.; Cheng, H.; Kong, X.; Wang, Q.; Cui, H. Intelligent wind turbine blade icing detection using supervisory control and data acquisition data and ensemble deep learning. *Energy Sci. Eng.* **2019**, *7*, 2633–2645. [CrossRef]
- 19. Zhang, L.; Liu, K.; Wang, Y.; Omariba, Z. Ice detection model of wind turbine blades based on random forest classifier. *Energies* **2018**, *11*, 2548. [CrossRef]
- 20. Park, J.; Lee, J.; Oh, K.; Lee, J. Development of a Novel Power Curve Monitoring Method for Wind Turbines and Its Field Tests. *IEEE Trans. Energy Convers.* **2014**, *29*, 119–128. [CrossRef]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).