

Article

Research on Misalignment Fault Isolation of Wind Turbines Based on the Mixed-Domain Features

Yancai Xiao *, Yujia Wang, Huan Mu and Na Kang

School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing 100044, China; shiyanshi10071@163.com (Y.W.); shiyanshi10073@163.com (H.M.); shiyanshi10072@163.com (N.K.)

* Correspondence: ycxiao@bjtu.edu.cn; Tel.: +86-10-5168-4273

Academic Editor: Javier Del Ser Lorente

Received: 2 May 2017; Accepted: 8 June 2017; Published: 10 June 2017

Abstract: The misalignment of the drive system of the DFIG (Doubly Fed Induction Generator) wind turbine is one of the important factors that cause damage to the gears, bearings of the high-speed gearbox and the generator bearings. How to use the limited information to accurately determine the type of failure has become a difficult study for the scholars. In this paper, the time-domain indexes and frequency-domain indexes are extracted by using the vibration signals of various misaligned simulation conditions of the wind turbine drive system, and the time-frequency domain features—energy entropy are also extracted by the IEMD (Improved Empirical Mode Decomposition). A mixed-domain feature set is constructed by them. Then, SVM (Support Vector Machine) is used as the classifier, the mixed-domain features are used as the inputs of SVM, and PSO (Particle Swarm Optimization) is used to optimize the parameters of SVM. The fault types of misalignment are classified successfully. Compared with other methods, the accuracy of the given fault isolation model is improved.

Keywords: misalignment; fault isolation; mixed-domain features; SVM; PSO

1. Introduction

With the rapid development of wind power, a lot of wind turbines have faults in the operation because most of them are installed in remote areas, and their loads are unstable. The operation and maintenance cost of up to 25–30% seriously restrict the development of wind power industry [1,2]. Therefore, how to reduce the risk of wind turbine operation and decrease the cost of wind power generation has become a research topic for many scholars. The fault isolation of wind turbine is the bottleneck to solve the above problems; undoubtedly, it is the key to ensuring long-term stable operation of the wind turbine, and it is the necessary means to realize the condition-based maintenance of the wind turbines [3].

Variable speed constant frequency doubly fed wind turbine is the main type in wind farms [4], and its simplified model is shown in Figure 1. The fault rate of gearbox and generator is relatively high [5]. There are many reasons for the faults of gearbox and generator, and the misalignment between them is one of the most important [6]. This is mostly because:

1. The precise alignment in the wind turbine is very difficult;
2. The wind speed fluctuation can cause the start-stop of wind turbine frequently, and, as time goes on, it will cause the shift of generator or the deformation of some parts, resulting in the misalignment of generator and gearbox.

When the misalignment occurs, it is easy to cause damage to gearbox gears, bearings and generator bearings and so on. Therefore, the study and fault isolation of the transmission system in the wind turbine is the key to ensuring its long stable operation.

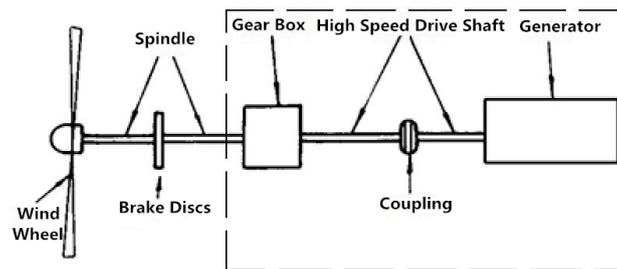


Figure 1. Simplified model of a doubly fed wind turbine.

The misalignment belongs to a latent fault. When the fault is accumulated to a certain extent, the equipment is damaged seriously, leading to the outage of the wind turbine, affecting the power generation seriously, resulting in huge economic losses. For many years, how to isolate the latent fault accurately has been a difficult problem. This is mainly due to the fault feature of latent faults not being obvious, and researchers lack the awareness of the developing rules for latent fault. In order to identify and isolate the existence of the misalignment fault accurately and to overcome the difficulty of obtaining a large number of misalignment fault samples in reality, in this paper, Solidworks and Adams models are established according to the 1.5 MW wind turbine. The simulation of the transmission system is carried out in Adams under the normal operation, the parallel misalignment, the angle misalignment and the comprehensive misalignment. The vibration signal (angular acceleration) of a high speed shaft is extracted as the information (the established model and its effectiveness are shown in [7]).

The methods of extracting fault features from vibration signals are mainly based on time-domain analysis, frequency-domain analysis and time-frequency domain analysis:

1. The time-domain analysis is one of the most simple and direct analysis methods, and it is effective in isolating the fault in a certain extent. Except for the poor anti-interference of serious faults, the numerical values in time-domain are close to that in normal state, so it is easy to make misjudgments only using time-domain analysis.
2. The frequency-domain analysis of signal is the most commonly used method of mechanical equipment fault analysis. It can be used to get more intuitive fault information than by the time-domain analysis. However, the theoretical basis of frequency-domain analysis is the method of Fourier analysis, so it has sidedness, and cannot extract the features of vibration signal comprehensively.
3. The vibration signal is expressed in the time-domain and frequency-domain at the same time for the time-frequency domain analysis, and it has very prominent advantages. The main methods of feature extraction based on time-frequency analysis are as follows: short time Fourier Transform, Wigner-Ville Distribution, Wavelet Transform, Blind Source Separation, Empirical Mode Decomposition (EMD) and so on.

Due to the complexity of the fault mechanism of the wind turbine transmission system, the state information is different with different feature indexes, the sensitivity and regularity of the running state, as well as the clustering and separability in the model space not being the same. Single features or single domain features are difficult to fully reflect the different conditions of wind turbines with different degrees and types of faults. Therefore, it is necessary to construct a mixed feature library with time-domain, frequency-domain and time-frequency domain, which becomes one of the development trends of fault feature extraction.

Many scholars have applied one or two of the above domains to extract the fault isolation feature based on vibration signal. For example, Ref. [8] extracted root mean square (RMS), peak value and kurtosis coefficient of the vibration signal, and combined with the time-domain waveform to judge whether there was fault in equipment. In Ref. [9], the kurtosis and peak value of the vibration signal

were selected as the time-domain indexes, and then the wavelet packet algorithm was used to extract the frequency band energy and the 2-norm as the time-frequency domain indexes, the PCA (Principal Component Analysis) was used to confirm the principal component, and the crack fault of the wind turbine gearbox was diagnosed effectively. The EMD method was used to judge the working states and fault types of the gear by calculating EMD energy entropy of different vibration signals as the input features of Support Vector Machine in [10], but only the time-frequency domain feature was considered. Ref. [11] used a testing system for a wind turbine vibration test to collect vibration signal, and the time-domain and frequency-domain analysis were carried out to get the fault features of the rotor misalignment through the analysis of wind turbine generator, but the time-frequency analysis was not considered. In Ref. [12], the wavelet analysis was used to denoise the vibration signals of the wind turbine gearbox under normal, wear fault and broken tooth fault. Five feature parameters were extracted into the LVQ (Learning Vector Quantification) neural network, and the results showed that the method can identify the fault quickly and accurately. In Ref. [13], the time-domain of the peak, RMS and kurtosis of the vibration signal were extracted as the time-domain features of the fault gearbox. Combining observation and analysis of the normal signal and fault signal frequency-domain feature, the possible location of the failure was determined. In Ref. [14], the vibration signals of three typical states of normal conditions, tooth wear and tooth breakage of the gearbox in the wind turbine are analyzed, and the margin index, kurtosis index, peak index, pulse index, power spectrum entropy in the frequency-domain were extracted, the time-domain and frequency-domain features were the inputs in the fault isolation. In Refs. [15] and [16], the mixed-domain feature set is constructed to completely characterize the property of each fault by combining Empirical Mode Decomposition (EMD) with the Autoregression (AR) model coefficients.

A large amount of literature shows that there are few reports on the method of combining the time-domain, frequency-domain and time-frequency domain features together to construct a mixed-domain feature library when researching the fault of the wind turbine transmission system. In order to extract the feature parameters which reflect the vibration signal as much as possible, and to make the fault isolation more reliable and accurate, in this paper, firstly, the time-domain, frequency-domain and time-frequency domain of the wind turbine vibration signal are combined to construct the mixed-domain feature library to obtain more comprehensive and accurate fault isolation information. The information entropy of the intrinsic mode function (IMF) component decomposed by EMD is used as the time-frequency feature. Then, the support vector machine that is suitable for small samples is used as the isolation tool, and the PSO algorithm is used to optimize the relevant parameters of the support vector machine to obtain better classification performance with high diagnostic accuracy. The results show that the proposed method can identify the types of misalignment effectively compared with other methods.

2. Establishment of Mixed-Domain Feature Library

2.1. Feature Extraction of Vibration Signal in the Time-Domain

The vibration signal of the transmission system (assuming that the signal is a discrete time series $x_0, x_1, x_2, \dots, x_{N-1}$ with finite length) contains the running status information, so a representative time-domain index can be selected as the fault feature parameter to help determine whether the drive system has a fault and what type of misalignment has occurred. The time-domain features of the signal include the dimensional and the dimensionless time-domain index [17].

2.1.1. Dimensional Index

In this paper, the dimensional time-domain indexes selected include the variance, the square root amplitude, the RMS, the standard deviation (SD) and the kurtosis. Their formulas are as follows:

Variance:

$$\sigma^2 = \frac{1}{N} \sum_{i=0}^{N-1} (x_i - \bar{x})^2 = x_a - \bar{x}^2, \quad (1)$$

where \bar{x} is the mean of the signal.

Square root amplitude:

$$x_r = \left(\frac{1}{N} \sum_{i=0}^{N-1} |x_i| \right)^{1/2}, \quad (2)$$

RMS:

$$x_{rms} = \left(\frac{1}{N} \sum_{i=0}^{N-1} x_i^2 \right)^{1/2}, \quad (3)$$

SD:

$$x_{std} = \left(\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \right)^{1/2}, \quad (4)$$

Kurtosis:

$$\beta = \frac{1}{N} \sum_{i=1}^N x_i^4. \quad (5)$$

2.1.2. Dimensionless Index

The dimensionless index is not so sensitive to the change of amplitude and frequency of the signal, that is, it is not directly related to the working parameters of the wind turbine. The kurtosis index, the waveform index, the peak index, the pulse index and the margin index are commonly used as dimensionless indexes [18].

The impact component will be produced when there is fault in the transmission system of wind turbine, and the kurtosis index can reflect the degree of steepness, and the expression is:

$$x_q = \frac{1}{N} \sum_{i=0}^{N-1} x_i^4 / x_a^2. \quad (6)$$

The waveform index reflects the shape of the time-domain waveform, which is sensitive to the waveform change caused by the fault, and the expression is:

$$K = \frac{x_{rms}}{x'}. \quad (7)$$

The peak index reflects the distribution of large data components in the vibration signal, which is the deviation degree of the vibration amplitude distribution relative to the normal Gaussian distribution, and the expression is:

$$C = \frac{x_p}{x_{rms}}. \quad (8)$$

The pulse index represents the magnitude of the impact component in the signal, which is sensitive at the beginning of the fault, and the expression is:

$$I = \frac{x_p}{x'}. \quad (9)$$

The margin index is sensitive to the impact signal, and the expression is:

$$L = \frac{x_p}{x_r}. \quad (10)$$

2.2. Feature Extraction of Vibration Signal in the Frequency-Domain

For a complex vibration signal, the information extracted in the time-domain is limited. Therefore, the time-domain signal is often transformed into a frequency-domain signal by a mathematical method to reveal the frequency component of the signal, so as to extract more information from the signal. The mathematical basis for the transition from time-domain to frequency-domain is Fourier Transform. The process of Fourier Transform is called spectrum analysis. The commonly used spectral analysis includes amplitude spectrum analysis, phase spectrum analysis, logarithmic spectrum analysis and power spectrum analysis. Power spectrum is a commonly used method for frequency-domain feature extraction. It can be used to describe the change of signal frequency-domain features by describing the change of power centroid and the dispersion degree of power distribution [19].

For a discrete time series $(x_0, x_1, x_2, \dots, x_{N-1})$ with finite length, assuming the sampling frequency is f_s , three commonly used frequency-domain parameters are as follows [20]:

Centroid frequency (FC):

$$FC = \frac{1}{2\pi f_s} \cdot \frac{\int_0^\pi \omega S(\omega) d\omega}{\int_0^\pi S(\omega) d\omega}, \tag{11}$$

Mean square Frequency (MSF):

$$MSF = \frac{1}{4\pi^2 f_s^2} \cdot \frac{\int_0^\pi \omega^2 S(\omega) d\omega}{\int_0^\pi S(\omega) d\omega}, \tag{12}$$

Variance frequency (VF):

$$VF = \frac{1}{4\pi^2 f_s^2} \cdot \frac{\int_0^\pi (\omega - 2\pi f_s FC) S(\omega) d\omega}{\int_0^\pi S(\omega) d\omega} = MSF - FC^2, \tag{13}$$

where $S(\omega)$ is power spectrum of the discrete signal, $S(\omega) = X(\omega) \cdot \overline{X(\omega)}$, $X(\omega) = \sum_{i=0}^{N-1} x(i) e^{-j\pi\omega}$; and ω is the angular frequency.

2.3. Feature Extraction of Vibration Signal in the Time-Frequency Domain

In this paper, EMD is improved by the method of mirror extension, and the energy entropy of the vibration signal is extracted by the improved EMD (IEMD).

Assuming [21]:

$$E_i = \int |c_i(t)|^2 dt = \sum_{k=1}^n |x_{ik}^2|, \tag{14}$$

$$E = \sum_{i=1}^n E_i, \tag{15}$$

$$p_i = \frac{E_i}{E}, \tag{16}$$

where $i = 1, 2, 3, \dots, n$; x_{ik} is the amplitude of each discrete point.

The expression of IEMD energy entropy is

$$P_i = -\sum_{i=1}^n p_i \lg p_i. \tag{17}$$

According to Ref. [7], the energy entropy of the first eight IMF components of IEMD decomposition is the time-frequency domain indexes.

Collecting the vibration signal of the high-speed shaft at normal and three misalignment conditions with the speeds of $81.3^\circ/s$, $104.4^\circ/s$, $110.7^\circ/s$ and $117.0^\circ/s$, respectively, one gets 114 samples at each speed with each case, so the total number of samples is $4 \times 4 \times 114$. The features of

time-domain, frequency-domain and time-frequency domain of the samples are extracted to construct the mixed-domain fault feature library, which are shown in Table 1.

Table 1. Mixed-domain feature library.

Feature Library	Feature	Index
Mixed-domain feature library	Time-domain	root mean square, square root amplitude, variance, standard deviation, kurtosis, waveform index, peak index, pulse index, margin index, kurtosis index
	Frequency-domain	centroid frequency, mean square frequency, variance frequency
	Time-frequency domain	energy entropy of the first eight IMF components of IEMD decomposition

Tables 2–4 are partial data of the extracted time-domain, frequency-domain and time-frequency domain indexes. Type 0, 1, 2, 3 represent normal state, parallel misalignment, angle misalignment and comprehensive misalignment, respectively.

Table 2. The partial data of the time-domain indexes.

Type	RMS ($\times 10$)	Square Root Amplitude	Variance ($\times 10^6$)	SD ($\times 10$)	Kurtosis	Waveform Index	Peak Index	Pulse Index	Margin Index	Kurtosis Index
0	10,107	71,069	10,212	10,105	2.49	1.21	3.42	4.15	4.86	2.50
	10,112	70,651	10,223	10,111	2.53	1.22	3.02	3.68	4.32	2.54
	10,117	69,920	10,229	10,114	2.55	1.22	3.28	4.01	4.74	2.56
	10,135	70,627	10,261	10,130	2.52	1.22	3.44	4.19	4.93	2.54
	10,136	69,844	10,267	10,132	2.59	1.23	3.03	3.72	4.40	2.60
1	68,743	203,770	47,184	68,690	14.19	1.89	6.26	11.81	21.11	14.40
	70,048	213,120	49,046	70,033	15.37	1.85	5.44	10.07	17.89	15.51
	70,194	208,350	49,209	70,149	14.69	1.90	5.21	9.90	17.57	14.94
	70,720	217,070	49,931	70,662	16.04	1.91	5.11	9.75	16.66	16.30
	71,651	227,770	51,201	71,555	12.08	1.84	5.57	10.24	17.53	12.23
2	176,820	689,720	30,935	175,880	7.41	1.64	4.42	7.24	11.34	7.58
	177,180	702,010	31,389	177,170	6.72	1.63	4.87	7.92	12.28	6.72
	177,930	680,850	31,660	177,930	7.52	1.67	4.78	8.00	12.50	7.52
	178,990	679,150	31,589	177,730	7.71	1.69	3.84	6.48	10.13	8.28
	179,010	706,130	32,048	179,020	6.77	1.64	5.04	8.26	12.77	6.77
3	601,360	1,681,900	36,139	601,160	14.83	1.95	5.82	11.35	20.81	15.09
	625,910	1,801,800	39,127	625,520	13.28	1.91	6.70	12.76	23.26	13.58
	656,970	2,154,200	43,123	656,680	11.11	1.79	7.30	13.09	22.27	11.22
	664,860	2,069,300	44,194	664,780	11.38	1.82	6.80	12.37	21.85	11.51
	671,750	2,101,200	45,063	671,290	12.30	1.83	6.82	12.50	21.81	12.50

Table 3. The partial data of the frequency-domain indexes.

Type	FC	MSF	VF
0	−458	7,083,800	6,873,700
	−455	7,047,800	6,841,100
	−457	7,077,400	6,868,500
	−453	6,958,700	6,753,100
	−448	6,988,700	6,787,900
1	−750	12,950,000	12,387,000
	−647	11,048,000	10,629,000
	−697	12,436,000	11,950,000
	−728	13,462,000	12,933,000
	−715	12,794,000	12,283,000

Table 3. Cont.

Type	FC	MSF	VF
2	−440	8,587,500	8,393,500
	−443	8,652,700	8,456,200
	−475	9,280,300	9,053,800
	−447	8,707,000	8,507,200
	−451	8,785,300	8,581,700
3	−376	7,338,600	7,196,600
	−355	7,033,700	6,907,400
	−388	8,190,100	8,039,200
	−369	7,400,400	7,264,100
	−346	6,958,600	6,838,700

Table 4. The partial data of the energy entropy of the first eight IMF components of IEMD.

Type	P_1	P_2	P_3	P_4	P_5	P_7	P_7	P_8
0	0.3109	0.3429	0.1709	0.0747	0.0476	0.072	0.097	0.0521
	0.3166	0.352	0.1553	0.0784	0.0477	0.0688	0.0575	0.0176
	0.3186	0.3498	0.1657	0.0814	0.0478	0.0452	0.069	0.0172
	0.3256	0.341	0.2031	0.0895	0.0564	0.037	0.0933	0.0148
	0.3302	0.3386	0.1669	0.0748	0.0574	0.0522	0.0788	0.0129
1	0.1484	0.3676	0.3623	0.193	0.1201	0.0863	0.0477	0.1012
	0.1489	0.3669	0.3678	0.2267	0.0938	0.0481	0.028	0.1249
	0.1513	0.3678	0.3662	0.1839	0.1077	0.0665	0.0517	0.2216
	0.1558	0.3533	0.3675	0.2133	0.1479	0.0683	0.0637	0.1977
	0.1574	0.3601	0.3639	0.157	0.0807	0.062	0.2031	0.1429
2	0.2335	0.3067	0.2325	0.1467	0.0843	0.0763	0.1386	0.3648
	0.2361	0.3024	0.3679	0.289	0.1934	0.1157	0.1594	0.0478
	0.2423	0.3665	0.3417	0.2574	0.1862	0.0897	0.0395	0.0279
	0.2455	0.3476	0.3679	0.3096	0.1083	0.0873	0.0337	0.0536
	0.2487	0.3298	0.3134	0.2565	0.1254	0.0804	0.1309	0.3447
3	0.1456	0.2285	0.2912	0.3322	0.2521	0.1672	0.2001	0.1162
	0.1061	0.2784	0.3393	0.2876	0.2762	0.0813	0.2059	0.1677
	0.1074	0.2581	0.3357	0.3582	0.1606	0.1101	0.1396	0.1396
	0.1082	0.2853	0.3271	0.3022	0.1685	0.1765	0.0751	0.1556
	0.111	0.3268	0.3654	0.2737	0.2681	0.0886	0.0773	0.1014

3. Fault Isolation of Transmission System Based on PSO-SVM

After establishing the mixed-domain fault feature library of the vibration signal, the samples can be divided into two groups, one is the training set and the other is the testing set. In this paper, the training samples are 4×304 , and the testing samples are 4×152 . SVM is selected to be the classifier that will be trained by the training samples.

3.1. The Principles of SVM

SVM was proposed by Vapnik in 1995. The principle of SVM can be briefly described as follows [7].

The giving training samples $(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)$ (where $x_i \in R^n, y_i \in \{1, -1\}$, and l is the number of training samples) can be separated by a hyper plane, and the hyper plane can be expressed as:

$$\omega x + b = 0, \tag{18}$$

where ω and b , respectively, represent the normal vector and the constant term. When ω and b are the best, it means the optimal hyper plane has been found, which makes the distance between the positive and negative samples the largest.

For linearly inseparable problems, the samples can be mapped to the hyper plane by the kernel function $K(x_i, y_j)$ to realize linear separability. Its objective function is as follows:

$$W(a) = \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l a_i a_j y_i y_j K(x_i, x_j),$$

$$\text{s.t. } \sum_{i=1}^l a_i y_i = 0, 0 \leq a_i \leq C,$$
(19)

where a_i is the Lagrange operator; C is the punishment factor, and its basic function is to control the penalty of the wrong samples; $K(x_i, x_j)$ is the kernel function, and its basic function is to transform the vectors of low-dimensional to the inner product in the high-dimensional.

The kernel functions used in SVM are mainly linear kernel function, polynomial kernel function, radial basis kernel function and sigmoid kernel function. Many studies and experiments show that the radial basis function (RBF) is a better choice. Therefore, in this paper, the RBF is used as the kernel function. The function is as follows:

$$K(x_i, x_j) = \exp(-g \|x_i - x_j\|^2), g > 0,$$
(20)

where g is the kernel parameter.

In SVM, parameter C determines the training error and the generalization ability of the classifier; and parameter g affects the distribution form of the samples in the feature space. The selection of them is very important for the performance of the classifier. In this paper, PSO is used to select the parameters of C and g to improve the classification performance.

3.2. Particle Swarm Optimization

PSO is a kind of swarm intelligence optimization algorithm in addition to the ant colony algorithm and the fish swarm algorithm, and it was first proposed by Kennedy and Eberhart in 1995 [7]. The principle can be briefly described as follows.

Supposing in a D -dimensional search space, there is a population $X = (X_1, X_2, \dots, X_n)$, consisting of n particles. The i -th particle is represented as a D -dimensional vector of $X_i = [x_{i1}, x_{i2}, \dots, x_{iD}]$, which represents the position of the i -th particle in the D -dimensional search space, and it also represents a potential solution to the problem. According to the objective function, the fitness value of each particle's position X_i can be calculated. If the velocity of the i -th particle is $V_i = [V_{i1}, V_{i2}, \dots, V_{iD}]$, the individual extremum is $P_i = [P_{i1}, P_{i2}, \dots, P_{iD}]$, the population extremum is $P_g = [P_{g1}, P_{g2}, \dots, P_{gD}]$, and then the velocity and position of the particle are updated according to the following equations:

$$V_{id}^{k+1} = V_{id}^k + c_1 r_1 (P_{id}^k - X_{id}^k) + c_2 r_2 (P_{gd}^k - X_{id}^k),$$
(21)

$$X_{id}^{k+1} = X_{id}^k + V_{id}^{k+1},$$
(22)

where $d = 1, 2, \dots, D$; $i = 1, 2, \dots, n$; k is the number of iterations; and c_2 are learning factors, which are non negative constants, making the particles learn from their own or other better particles to achieve the purpose of being close to the better position of itself or the whole group; r_1 and r_2 are random numbers distributed in $[0, 1]$ to maintain the diversity of group. In order to prevent the blind search of particles, it is generally recommended to limit the position and velocity to a certain range of $[-X_{\max}, X_{\max}]$ and $[-V_{\max}, V_{\max}]$.

The realization of the fault isolation based on PSO-SVM (in the model, PSO is used to select the parameters of SVM classifier) is shown in Figure 2.

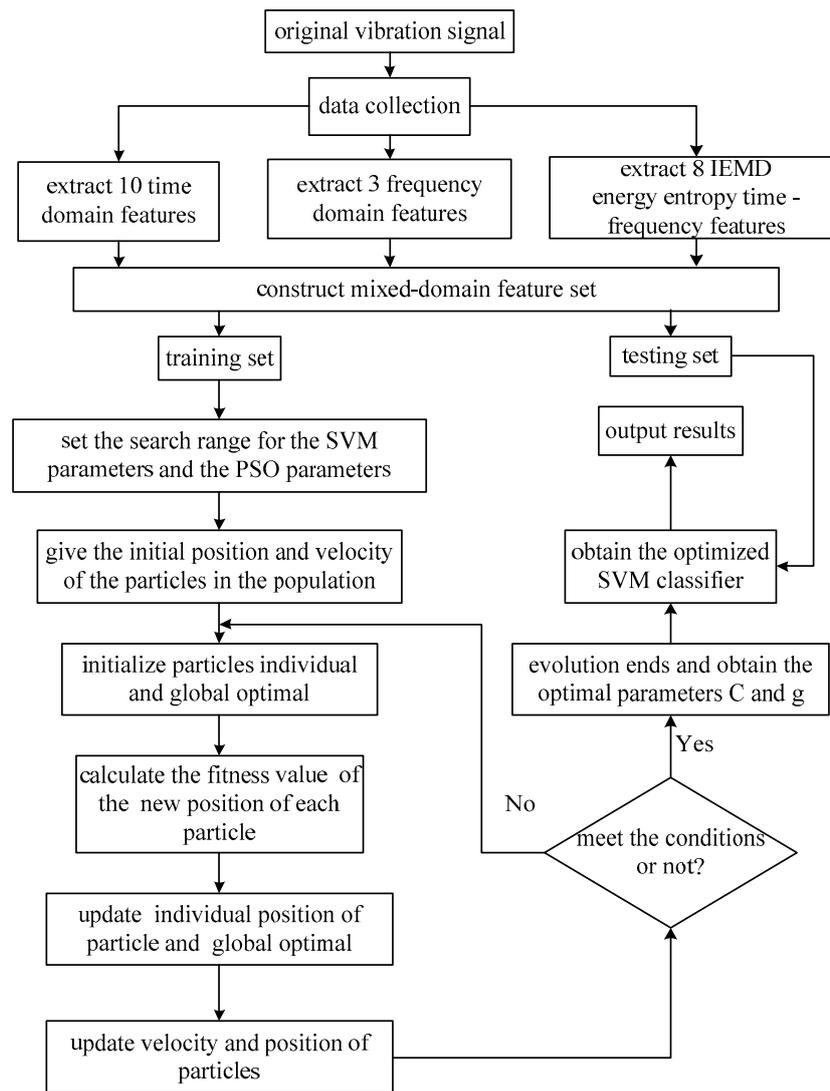


Figure 2. The PSO-SVM fault isolation model.

4. PSO-SVM Fault Isolation Results Based on Mixed-Domain Features

The results of parameters optimizing of SVM using PSO and the mixed-domain features of the wind turbine are shown in Figure 3, that is, the optimal fitness of PSO algorithm is 83.3882%, and the obtained optimal parameters are $C = 3.1787$, $g = 26.2761$. The classification results of the testing set using the optimized SVM classifier are shown in Figure 4, where the category label "0" indicates the normal condition; "1" indicates the parallel misalignment; "2" indicates the angle misalignment; and "3" indicates the comprehensive misalignment.

The classification accuracy rate of training set obtained by using an optimized SVM classifier achieves 97.9441%, and the classification accuracy rate of testing set is 92.1053%, so the fault classification accuracy is very high. In order to better illustrate the superiority of the PSO-SVM algorithm proposed in this paper, the same fault features are adopted by SVM, GridSearch-SVM (the parameters of SVM are optimized by GridSearch) and GA-SVM (the parameters of SVM are optimized by Genetic Algorithm), and the testing results are shown in Figure 5a–c and Table 5.

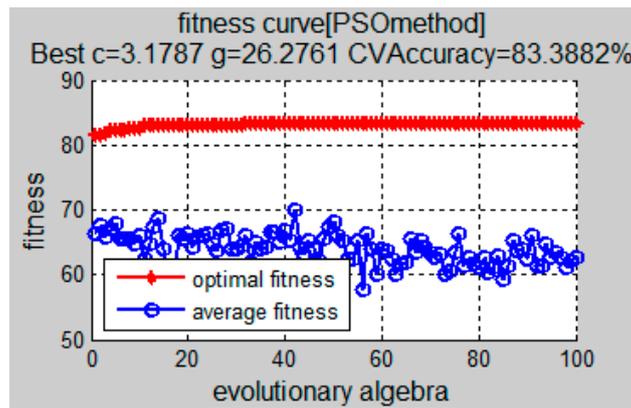


Figure 3. Particle fitness curves.

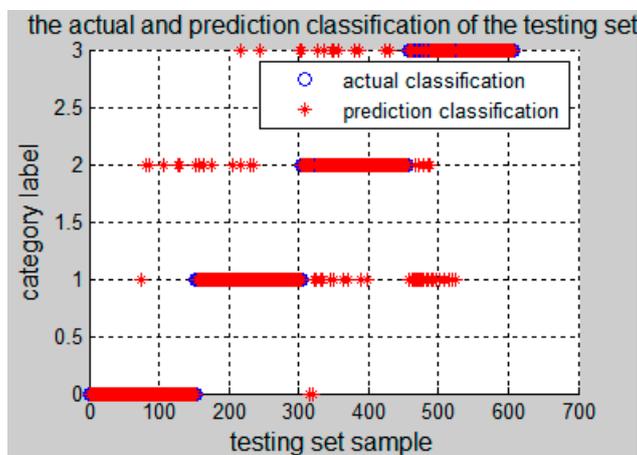


Figure 4. PSO-SVM testing results.

Table 5. Comparison of PSO-SVM with other commonly used classifiers.

Item	C	g	Accuracy of Training Set	Accuracy of Testing Set
SVM	1	0.01	63.2401% (769/1216)	63.6513% (387/608)
GridSearch-SVM	16	1.4142	93.4211% (1136/1216)	90.9539% (553/608)
GA-SVM	26.2168	6.7162	99.5066% (1210/1216)	91.1184% (554/608)
PSO-SVM	3.1787	26.2761	97.9441% (1191/1216)	92.1053% (560/608)

From the simulation results, it can be seen that the fault isolation by SVM only has a poor recognition effect. When using GridSearch-SVM, the training accuracy and testing accuracy are better, but not higher than PSO-SVM. The training accuracy of GA-SVM is very high, but the promotion ability is less than PSO-SVM. Therefore, it is better to use the PSO-SVM algorithm to isolate the fault.

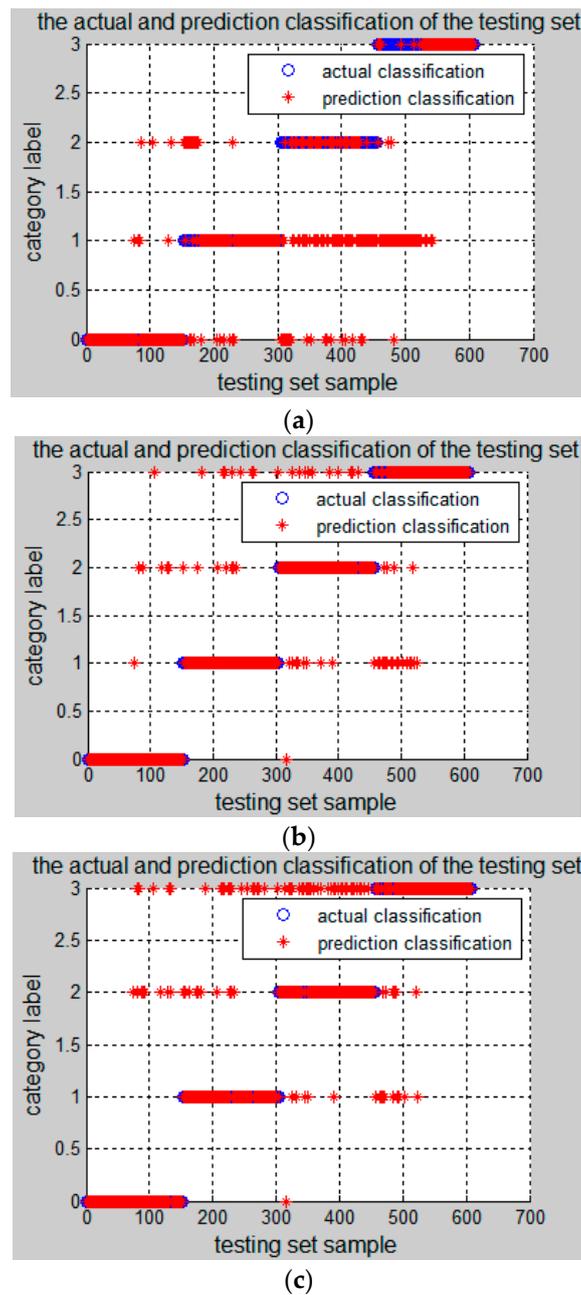


Figure 5. Fault isolation results of other algorithms. (a) SVM testing results; (b) GridSearch-SVM testing results; (c) GA-SVM testing results.

5. Comparison of Fault Isolation Results with Different Fault Features

In this paper, the time-domain, frequency-domain and time-frequency domain features are used to isolate the misalignment fault. In order to compare the superiority of the extracted fault features, the paper also adopts time-domain indexes, frequency-domain indexes, IEMD energy entropy, and their combinations of any two kinds as the inputs of SVM, and the recognition results are shown in Figure 6 and Table 6.

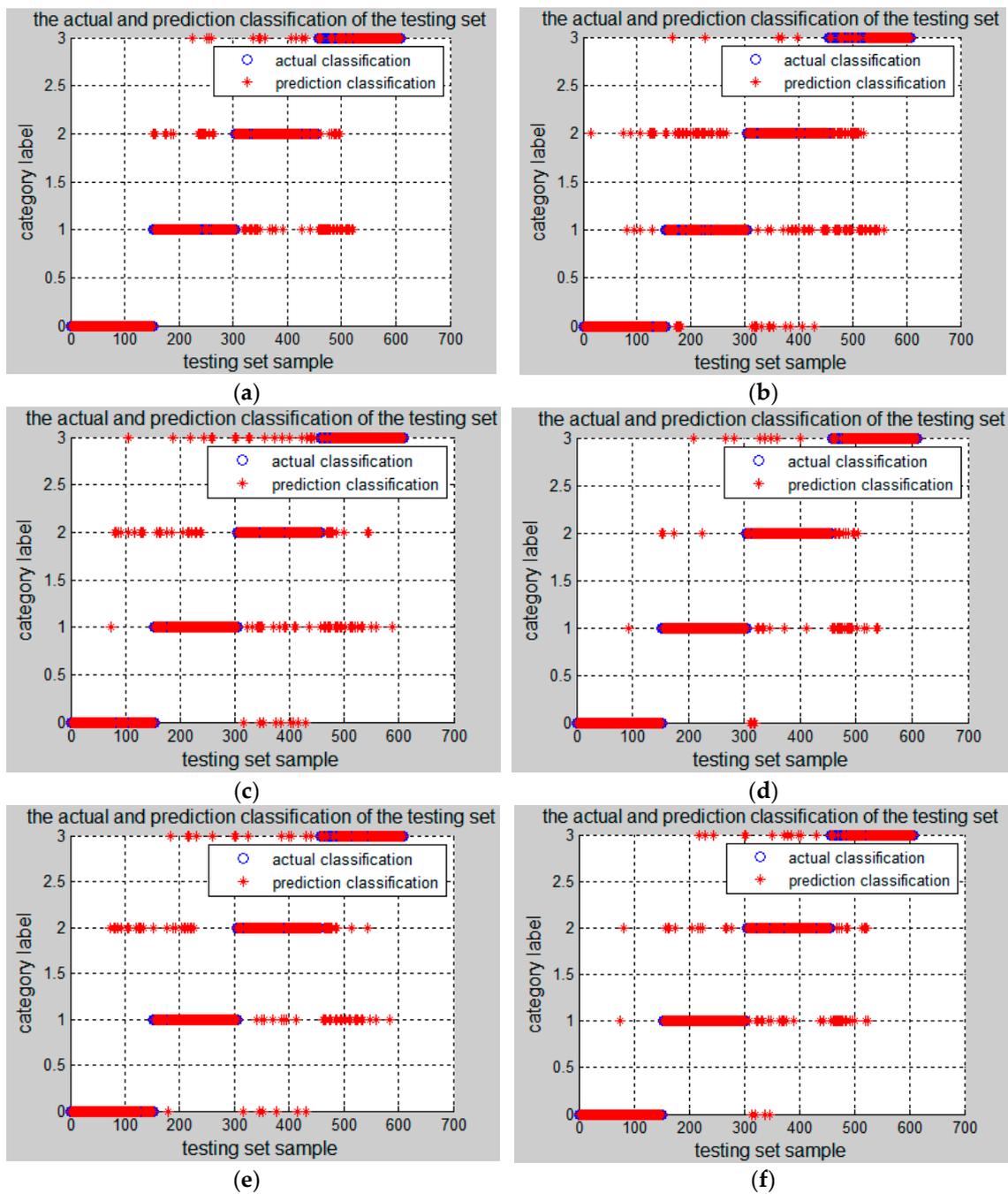


Figure 6. Testing results of PSO-SVM using different input features. (a) using time-domain features only; (b) using frequency-domain features only; (c) using IEMD energy entropy features only; (d) using time-domain + frequency-domain features; (e) using frequency-domain + IEMD energy entropy features; and (f) using time-domain + IEMD energy entropy features.

From Figure 6 and Table 6, it can be seen that when the time-domain, frequency-domain and time-frequency domain features are used separately as inputs of the PSO-SVM isolation model, the accuracy of the training set and the testing set are low, and the recognition results are not good. When two kinds of features are combined as the inputs, the training set accuracy is above 91%, the testing set accuracy is more than 84%, and the recognition rate is increased. When the three mixed-domain features are combined, the correct rate of the training set reaches 97.9441%, and the accuracy of the testing set reaches 92.1053%. Therefore, it can be concluded that the mixed-domain

features combining three kinds of domain extractions proposed in this paper can effectively reflect the characteristics of the misalignment fault, and can be used to classify the fault types. It is also interesting to note that, whether used individually or combined with energy entropy, the time-domain features seem to provide better identification results than the frequency-domain data. This is because, when the wind turbine transmission system fails, it will produce impact components, and time-domain features can well reflect them, so better results can be obtained when using time-domain features.

Table 6. PSO-SVM classification results of different fault features.

Fault Features	C	g	Accuracy of Training Set	Accuracy of Testing Set
Time-domain	22.5396	74.8823	88.0757% (1071/1216)	84.2105% (512/608)
Frequency-domain	7.71812	346.316	77.7138% (945/1216)	74.5066% (453/608)
IEMD energy entropy	7.88573	6.12414	87.4178% (1063/1216)	82.4013% (501/608)
Time-domain + frequency-domain	73.0786	20.0281	93.5033% (1137/1216)	89.9671% (547/608)
Time-domain + IEMD energy entropy	58.3106	1.42904	93.2566% (1134/1216)	87.8289% (534/608)
Frequency-domain + IEMD energy entropy	2.1093	10.0231	91.1184% (1108/1216)	84.8684% (516/608)
Time-domain + frequency-domain + IEMD energy entropy	3.1787	26.2761	97.9441% (1191/1216)	92.1053% (560/608)

6. Conclusions

This paper is based on the time-domain, frequency-domain and time-frequency domain analysis of vibration signals of the wind turbine in four different operating conditions with four kinds of rotating speeds. The mixed-domain features are extracted in the paper, which are the inputs of SVM classifiers. The parameters of SVM are optimized using the PSO algorithm. The testing results show that the proposed model for misalignment fault isolation is effective compared with other commonly used algorithms.

When the wind turbine operates in the condition of misalignment, there will be some changes in the temperature and electrical signals in addition to the reflection of the mechanical aspects, so ANSYS, MATLAB, and other software should be included in the simulation further to get temperature and electrical parameters, so as to study the fault characteristics more comprehensively.

At the same time, it is necessary to point out that the method presented in this paper is a general method, and it can be used elsewhere.

Acknowledgments: This work was supported by the National Natural Science Foundation of China (51577008). The authors are grateful to the anonymous reviewers for their constructive comments on the previous version of the paper.

Author Contributions: Yancai Xiao and Yujia Wang contributed to paper writing and the whole revision process. Huan Mu and Na Kang performed simulation and experiments. All the authors have read and approved the final manuscript.

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

References

- Zhang, T.; Dwight, R.; El-Akruti, K. Condition based maintenance and operation of wind turbines. In Proceedings of the 8th World Congress on Engineering Asset Management & 3rd International Conference on Utility Management & Safety, Hong Kong, China, 30 October–1 November 2013.
- Zappalá, D.; Tavner, P.J.; Crabtree, C.J.; Sheng, S. Side-band algorithm for automatic wind turbine gearbox fault detection and diagnosis. *IET Renew. Power Gener.* **2014**, *8*, 380–389. [[CrossRef](#)]
- Kusiak, A.; Li, W. The prediction and diagnosis of wind turbine faults. *Renew. Energy* **2011**, *36*, 16–23. [[CrossRef](#)]
- Bak-Jensen, B.; Kawady, T.A.; Abde1-Rahman, M.H. Coordination between fault-ride-through capability and over-current protection of DFIG generators for wind farms. *J. Energy Power Eng.* **2010**, *4*, 20–29.

5. Amirat, Y.; Benbouzid, M.E.H.; Al-Ahmar, E.; Bensaker, B.; Turri, S. A brief status on condition monitoring and fault diagnosis in wind energy conversion systems. *Renew. Sustain. Energy Rev.* **2009**, *13*, 2629–2636. [[CrossRef](#)]
6. Liao, M.; Liang, Y.; Wang, S.; Wang, Y. Misalignment in Drive Train of Wind Turbines. *Mech. Sci. Technol.* **2011**, *30*, 173–180.
7. Xiao, Y.; Kang, N.; Hong, Y.; Zhang, G. Misalignment Fault Diagnosis of DFWT Based on IEMD Energy Entropy and PSO-SVM. *Entropy* **2017**, *19*, 6. [[CrossRef](#)]
8. Chen, S. Application of Time Domain Analysis in Mechanical Equipment Fault Diagnosis. *J. Mech. Transm.* **2007**, *31*, 79–83.
9. Zhen, L. Study on Fault Diagnosis Methods of Generator and Gearbox for Wind Turbines. Master's Thesis, North China Electric Power University, Beijing, China, 2014.
10. Zhang, C.; Chen, J.; Guo, X. A gear fault diagnosis method based on EMD energy entropy and SVM. *J. Vib. Shock* **2010**, *29*, 216–220.
11. Long, J.; Wu, J. Fault Diagnosis of Rotor Misalignment for Windmill Generators. *Noise Vib. Control* **2013**, *33*, 222–225.
12. Ding, S.; Chang, X.; Wu, Q.; Wei, H.; Yang, Y. Study on wind turbine gearbox fault diagnosis based on LVQ neural network. *Mod. Electron. Tech.* **2014**, *37*, 150–152.
13. Ren, S. The Study of Fault Diagnosis Methods for Wind Turbine Gearbox Based on the Analysis Vibration Characteristics. Master's Thesis, Hebei University of Technology, Tianjin, China, 2015.
14. Wang, B. Research on Fault Diagnosis System for Gearbox of Wind Turbine. Master's Thesis, North China Electric Power University, Beijing, China, 2012.
15. Tang, B.; Song, T.; Li, F.; Deng, L. Fault diagnosis for a wind turbine transmission system based on manifold learning and Shannon wavelet support vector machine. *Renew. Energy* **2014**, *62*, 1–9. [[CrossRef](#)]
16. Li, F.; Tang, B.; Yang, R. Rotating machine fault diagnosis using dimension reduction with linear local tangent space alignment. *Measurement* **2013**, *46*, 2525–2539. [[CrossRef](#)]
17. Xu, Z. Research on Transmission Chain State Monitoring and Fault Diagnosis of Wind Turbine Based on Vibration Method. Master's Thesis, Zhejiang University, Hangzhou, China, 2012.
18. Ding, K.; Li, W. *Practical Technology for Fault Diagnosis of Gear and Gearbox*; China Machine Press: Beijing, China, 2006.
19. Zhang, L.; Zhu, L.; Zhong, B. Study on the Features of Some Specific Parameters Utilized for Machine Condition Monitoring. *J. Vib. Shock* **2001**, *20*, 20–21.
20. Yi, L. *On-Site Practical Technology of Simple Vibration Diagnosis*; China Machine Press: Beijing, China, 2003.
21. Zhang, C.; Chen, J.; Guo, X. Gear fault diagnosis method based on EEMD energy entropy and support vector machine. *J. Cent. South Univ.* **2012**, *3*, 932–939.

