

Article

The Application of Dual-Tree Complex Wavelet Transform (DTCWT) Energy Entropy in Misalignment Fault Diagnosis of Doubly-Fed Wind Turbine (DFWT)

Yancai Xiao ^{1,*}, Yi Hong ¹, Xiuhai Chen ² and Weijia Chen ³

¹ School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing 100044, China; shiyanshi10071@163.com

² State Grid Beijing Haidian Electric Power Supply Company, Beijing 100195, China; chenxiuhai@sina.com

³ Department of Electrical Engineering, North China Electric Power University, Baoding 071003, China; shiyanshi10073@163.com

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

Received: 20 September 2017; Accepted: 1 November 2017; Published: 2 November 2017

Abstract: Misalignment is one of the common faults for the doubly-fed wind turbine (DFWT), and the normal operation of the unit will be greatly affected under this state. Because it is difficult to obtain a large number of misaligned fault samples of wind turbines in practice, ADAMS and MATLAB are used to simulate the various misalignment conditions of the wind turbine transmission system to obtain the corresponding stator current in this paper. Then, the dual-tree complex wavelet transform is used to decompose and reconstruct the characteristic signal, and the dual-tree complex wavelet energy entropy is obtained from the reconstructed coefficients to form the feature vector of the fault diagnosis. Support vector machine is used as classifier and particle swarm optimization is used to optimize the relevant parameters of support vector machine (SVM) to improve its classification performance. The results show that the method proposed in this paper can effectively and accurately classify the misalignment of the transmission system of the wind turbine and improve the reliability of the fault diagnosis.

Keywords: misalignment; dual-tree complex wavelet; energy entropy; fault diagnosis; PSO; SVM

1. Introduction

In recent years, doubly-fed wind turbines (DFWT) have become the main units for large-capacity wind farms. It has attracted wide attention to reduce the failure frequency and maintenance costs of DFWT by applying state monitoring technology [1,2]. Misalignment is a common fault of DFWT. Because the transmission chain is long and being installed in tens of meters, or even hundreds of meters high altitude cabin, the exact alignment of the crew is very difficult. Moreover, the unit starts and stops frequently because of the fluctuations of wind speed. As time goes on, there will be a shift or deformation in some components, resulting in misalignment between the generator and the gearbox [3]. When misalignment happens, heavy dynamic load will be produced between the gearbox high-speed shaft and the generator bearings. These will lead to the increasing axial and radial vibration. Also, it will cause bearing oil leakage, high temperature and the fastening bolts loosening, affecting the power generation seriously, resulting in huge economic losses. The accumulation of eccentric error even causes damage to bearings of the high-speed end and generator. Because misalignment is an important cause of early failure of large DFWT and may cause damage to two core parts including the gearbox and generator, it is necessary to study the misalignment fault diagnosis methods to ensure long and stable running of the DFWT.

At present, the vibration-based monitoring is the main mean to detect misalignment failure. For example, in literature [4], the vibration test system is used to monitor the abnormal working of the wind turbine, the misalignment fault signal of the rotor is obtained, which provides the basis for the rotor misalignment fault diagnosis of the wind turbine. However, the fault alarm rate and equipment costs of this method are high and the sensor installation is inconvenient. By analyzing the generator stator current to diagnose the fault, it can overcome these shortcomings. Now, some scholars have conducted research in this area. For example, MATLAB/Simulink was used to establish an electromechanical coupling model of the asynchronous motor to simulate misalignment, imbalance and other failures by Hou et al. [5], and the characteristics of the stator current were analyzed, which provided a new theoretical basis for improving the rotor system dynamics and fault diagnosis. However, the traditional simulation model of the wind turbine transmission system established in the MATLAB environment cannot reflect the real operating conditions. Of course, many fault parts models cannot be established in MATLAB [6]. In view of the three-dimensional simulation model of the transmission system established by ADAMS, it can better reflect the actual physical model and MATLAB has some advantages in establishing the control simulation model of wind turbine. In this paper, the 3D model of 1.5 MW wind turbine is established using Solidworks, and the model is then imported into ADAMS, where the normal working condition, parallel misalignment, angle misalignment and comprehensive misalignment are simulated. The wind turbine models and its control system are established by MATLAB. The mechanical input of the wind turbine models is the mechanical output of the high-speed shaft of the transmission system. The advantages of MATLAB and ADAMS are combined together in this way. The corresponding stator current can be obtained for analysis. Because the fault signals contained in the stator current have the characteristics of nonlinear and non-stationary, the complex transmission path and the strong on-site noise have caused great difficulty in extracting the fault characteristic information [7]. Therefore, how to remove the strong background noise while preserving the weak fault information is the key to the fault feature extraction and diagnosis of the transmission system. At present, the main methods of processing stator currents are traditional Fourier transform, wavelet transform [8] and dual-tree complex wavelet transform (DTCWT) [9]; the comparison of these is shown in Table 1.

Table 1. Comparison of stator currents processing methods.

Item	Methods	Characteristic
Traditional Signal Processing	Fourier Transform	Not applicable to time-varying signals and non-stationary signals
Modern Signal Processing	Wavelet Transform	Using variable size windowing technology, with strong local analysis ability of frequency band; prone to frequency aliasing and energy leaking
	Dual-Tree Complex Wavelet Transform	With the advantages of wavelet transform, little frequency aliasing and energy leaking

Among them, the DTCWT is a new type of wavelet transform, which has the advantages of little frequency aliasing, approximate shift invariance and so on, compared with wavelet transform [10]. In this paper, the DTCWT is used to extract the energy entropy of the stator current, and then the support vector machine (SVM) is used as the diagnostic tool. As a more and more widely used diagnostic tool, SVM have prominent advantages compared with other diagnostic tools [11–14]. The characteristics are shown in Table 2. For SVM, the choice of parameters is important. The commonly used methods of parameters selection are grid search, genetic algorithm (GA), particle swarm optimization (PSO) and so on. Grid search is to arrange the possible values of the various parameters in a row and list all possible combinations of results to generate a “grid”. The combinations are then used for SVM training and the performance is evaluated using cross validation. After the fitting function tries all the combinations of parameters, it returns a suitable classifier with the best combination of parameters. GA, also known as evolutionary algorithm, is a heuristic search algorithm proposed based on the

biological evolution process. Compared with grid search and GA, PSO has the merits of high search speed, high efficiency, easy to find the optimal solution, less set parameters, etc. It has been widely utilized due to its simple structure [15]. In this paper, PSO algorithm is used to optimize the relevant parameters of SVM to get better classification performance and improve the accuracy of diagnosis. The results show that the method can effectively identify the misalignment fault types.

Table 2. Comparison of diagnostic methods.

Methods	Characteristic
Neural network	The adjustment methods of weight coefficients of neural network are limited; easy to get into local optimum; over-reliance on learning samples and requiring a lot of samples
Fuzzy	Determining the membership and corresponding laws based on experience and experiment, which has certain limitations
Expert system	With expert knowledge and experience; but the knowledge acquisition and development are difficult
SVM	Good generalization ability; suitable for a few samples; with global optimality

2. The Related Theories

2.1. Dual-Tree Complex Wavelet Transform

DTCWT was first proposed by Kingsbury in 1998. It is composed of two parallel discrete wavelet transforms using different low-pass and high-pass filters, which are called real-tree and imaginary-tree [16,17]. DTCWT is not only with the excellent properties of real wavelet, but also in the form of dual-tree filter, which ensures the complete reconstruction of the signals. Meanwhile, it also solves the problems of shift variance and low directional selectivity in two and higher dimensions found with the commonly used discrete wavelet transform [18]. Therefore, DTCWT has the advantages of approximate shift invariance, good directional selectivity (for two-dimensional image processing) and explicit phase information. The frequency response of the filter has approximate analyticity, that is, there is almost no negative frequency component in the frequency response of the dual-tree complex wavelet filter. Compared with real wavelet transform, it has low computational complexity and can be used for on-line monitoring and mechanical fault diagnosis. Especially, it is suitable for the feature extraction of weak faults in mechanical equipment [19].

The wavelet function of the dual-tree complex wavelet is a complex, as shown in Equation (1):

$$\psi(t) = \psi_h(t) + i\psi_g(t) \quad (1)$$

where $\psi_h(t)$ and $\psi_g(t)$ represent two real wavelets, i is the complex unit, $\psi_h(t)$ and $\psi_g(t)$ are required to be a Hilbert transform pair [20].

Different from complex wavelet transform, the real and imaginary portions of the dual-tree complex wavelet transform are performed independently, i.e., using two parallel and independent real wavelet transform, 1/2 sampling interval is required between those two transform. While avoiding a large number of calculations, the translation invariance of the complex wavelet transform can be easily obtained [21]. Decomposition and reconstruction process of the dual-tree complex wavelet transform is shown in Figure 1. The two wavelet transform is respectively called the real tree and imaginary tree. h_0 and h_1 are low-pass filter and high-pass filter corresponding to the real tree, respectively, and g_0 and g_1 are corresponding to the imaginary tree, respectively.

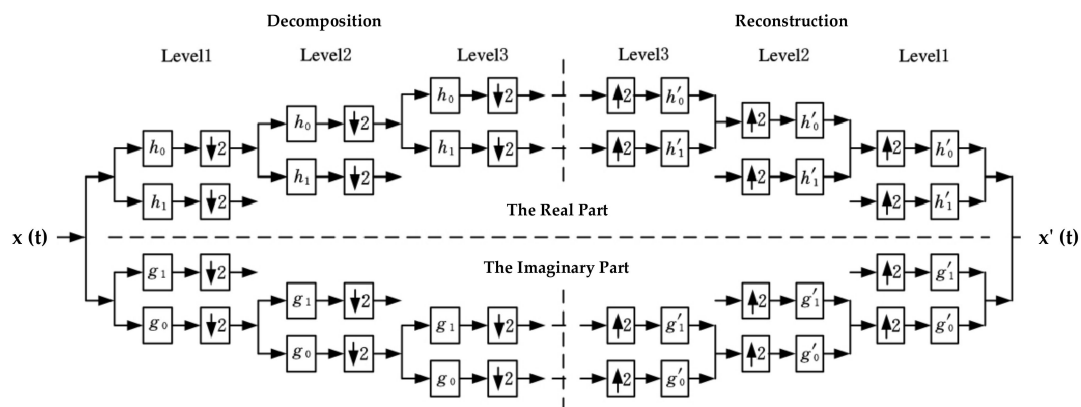


Figure 1. Decomposition and reconstruction process of the dual-tree complex wavelet transform (DTCWT).

2.2. The Principles of SVM

The principle of SVM can be briefly described as: giving l sets of training samples (x_n, y_n) , where $n = 1, 2, \dots, l$; the input $x_n \in R^n$; the output $y_n \in \{1, -1\}$. For linearly inseparable problems, the samples can be mapped to the hyper plane by the kernel function $K(x_i, x_j)$ to realize linear separability.

The SVM optimization problem is defined as follows [22,23]:

$$W(a) = \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l a_i a_j y_i y_j K(x_i, x_j), \text{ s.t. } \sum_{i=1}^l a_i y_i = 0, 0 \leq a_i \leq C \quad (2)$$

where a_i is the Lagrange operator; C is the penalty factor; $K(x_i, x_j)$ is the kernel function that satisfies Mercer's condition [24]. The commonly used kernel functions are polynomials, RBF, Sigmoid and so on. The parameter of the RBF kernel function is easy to choose. The complexity of the space will not become more serious when the kernel parameter is changed within the effective range. Moreover, it is easy to implement and the identification effect is also good. Thus, the RBF kernel function is selected in this paper.

$$K(x_i, x_j) = \exp[-|x_i - x_j|^2 / (2\sigma^2)] \quad (3)$$

where σ is the kernel width, which reflects the radius enclosed by the boundary.

SVM was originally designed for binary classification problems. When dealing with multiple kinds of problems, it is necessary to construct a suitable classifier. At present, there are two main ways to construct multiple kinds SVM classifier: one-against-all and one-against-one. One-against-all is to take a kind of samples as a category, the other remaining samples into another category, so k SVM classifiers are constructed. The unknown sample will be classified as the kind with the largest classification function values. This classification is not very practical. One-against-one is to recombine any two classes among multiple classes, and then design a binary classifier for each combination, thus for M categories need to construct $M(M-1)/2$ binary classifiers. After that, the voting method is used to screen those classifiers, and the majority of them is the class of the sample to be classified [25,26]. One-against-one algorithm is adopted in this paper.

For SVM system, the selection of kernel function parameter σ and penalty factor C have a significant influence on the establishment of the model. The traditional parameter selection method is based on experience and trial practice; accuracy and speed of calculation cannot be guaranteed. In order to make the model a better result, the particle swarm optimization algorithm is adopted to optimize the SVM parameters in this paper.

2.3. Particle Swarm Optimization

Particle swarm optimization is a swarm intelligence optimization algorithm, proposed by Eberhart and Kennedy in 1995 [27]. The feasible solution space is firstly set and a certain scale of particle swarm are initialized in PSO algorithm, then a set of possible solutions can be represented by the coordinate vector of each particle. Position, velocity, and fitness value are shared among particles, and the particle iterates and updates its position and velocity information by tracking individual and group extreme value to find the optimal solution in the solution space [28].

Assuming the number of particles is M , dimension of the solution space is D , the initial position vector of the i th particle is $P_i^0 = (p_{i1}^0, p_{i2}^0, \dots, p_{iD}^0)$, its velocity vector is $V_i^0 = (v_{i1}^0, v_{i2}^0, \dots, v_{iD}^0)$, its historical optimal position vector is $P_{besti}^0 = (p_{best1}^0, p_{best2}^0, \dots, p_{bestD}^0)$, and the historical optimal position vector of the particle swarm is $G_{besti}^0 = (g_{best1}^0, g_{best2}^0, \dots, g_{bestD}^0)$.

In the process of optimization, the position and velocity information of the particles are updated based on the current individual extremum P_{best} and the group extremum G_{best} . The particle velocity vector V_i^k is determined by three parts: the current velocity vector of the particle, the own cognitive part and the group cognitive part. The three parts determine the global searching ability of the particle.

$$V_i^k = V_i^{k-1} + c_1 r_1 (P_{best} - P_i^{k-1}) + c_2 r_2 (G_{best} - G_i^{k-1}) \quad (4)$$

$$P_i^k = P_i^{k-1} + V_i^k \quad (5)$$

where V_i^k is the velocity vector of i th particle at k iteration; P_i^k is the position vector of i th particle at k iteration. c_1, c_2 are learning factors, they are generally taken as constant values according to experience [29], usually, let $c_1 = c_2 = 2$ [30,31]. r_1, r_2 are random values between 0 and 1 [32].

To prevent the particle from blind searching, the position and speed are generally limited to a certain range $[-P_{max}, P_{max}]$ and $[-V_{max}, V_{max}]$.

Particle swarm optimization algorithm sets the condition to end the search. That is, when the number of iterations achieves the setting or the searching results achieve the performance indexes, the group extreme value G_{best} is the global optimal solution.

3. The Establishment of DFWT Model

3.1. Transmission System Modeling by ADAMS

The 3D model was established by Solidworks for a 1.5 MW doubly-fed wind turbine in this paper, and then the model is imported into ADAMS, where the normal working, the parallel misalignment, the angle misalignment and the comprehensive misalignment of the system are simulated, which can be referred in literature [11], not repeated here.

3.2. Doubly-Fed Generator and Its Control System Modeling by MATLAB

The doubly-fed generator is similar to the winding asynchronous induction generator in structure. Its stator side is directly connected to the three-phase power network, and the rotor side is connected to the power grid through the rectifying inverter. Compared with the general asynchronous generator, the speed of doubly-fed wind turbine generator is allowed to fluctuate in a certain range. Since the magnitude and frequency of the current in the rotor side can be adjusted by the rectifying inverter, the constant output power frequency of the stator side can be maintained when the rotational speed changes. The structure of doubly-fed wind turbine system is shown in Figure 2.

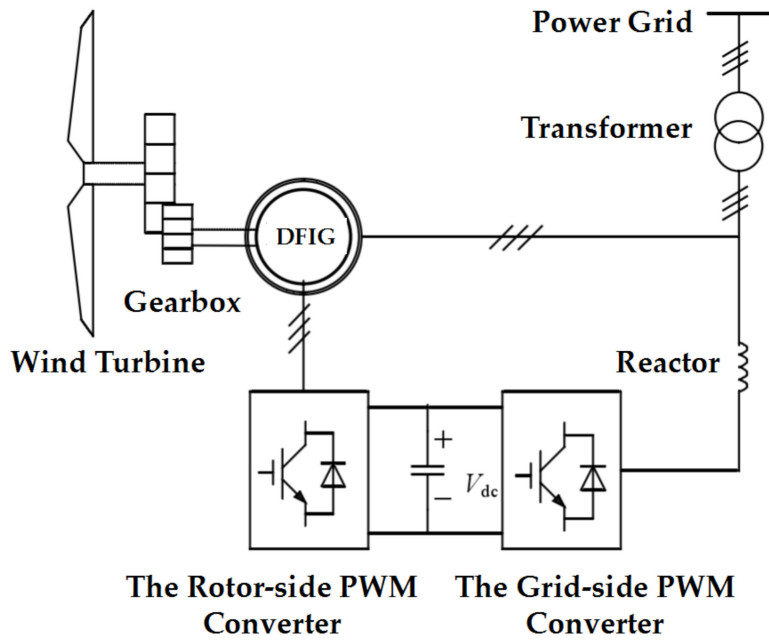


Figure 2. The structure of doubly-fed wind turbine (DFWT) system.

3.2.1. Mathematical Model of Doubly-Fed Induction Generator (DFIG)

The doubly-fed induction generator is actually a wound induction generator after removing the converter. Taking the direction of the current that flows into the circuit as the positive current direction and the direction of the electromagnetic torque that is in accordance with the generator rotation as the positive torque direction, the dynamic model of DFIG can be established in a synchronous rotating coordinate system.

Stator and rotor voltage equations:

$$\begin{cases} u_{ds} = R_s i_{ds} + \frac{d\psi_{ds}}{dt} - \omega_1 \psi_{qs} \\ u_{qs} = R_s i_{qs} + \frac{d\psi_{qs}}{dt} + \omega_1 \psi_{ds} \end{cases} \quad (6)$$

$$\begin{cases} u_{dr} = R_r i_{dr} + \frac{d\psi_{dr}}{dt} - \omega_s \psi_{qr} \\ u_{qr} = R_r i_{qr} + \frac{d\psi_{qr}}{dt} + \omega_s \psi_{dr} \end{cases} \quad (7)$$

Stator and rotor flux linkage equations:

$$\begin{cases} \psi_{ds} = L_s i_{ds} + L_m i_{dr} \\ \psi_{qs} = L_s i_{qs} + L_m i_{qr} \end{cases} \quad (8)$$

$$\begin{cases} \psi_{dr} = L_r i_{dr} + L_m i_{ds} \\ \psi_{qr} = L_r i_{qr} + L_m i_{qs} \end{cases} \quad (9)$$

Stator output power equations:

$$\begin{cases} P_s = \frac{3}{2} (u_{ds} i_{ds} + u_{qs} i_{qs}) \\ Q_s = \frac{3}{2} (u_{qs} i_{ds} - u_{ds} i_{qs}) \end{cases} \quad (10)$$

Electromagnetic torque equation:

$$T_e = 1.5 n_p L_m (i_{qs} i_{dr} - i_{ds} i_{qr}) \quad (11)$$

where T_e is the electromagnetic torque of DFIG; u, i, R, ψ, L, P, Q represent voltage, current, resistance, flux linkage, inductance, active power, reactive power, respectively; subscript s and r represent stator and rotor, respectively; subscript d and q represent the d and q axis components of the d - q coordinate system, respectively; L_m is the mutual inductance; ω_1 is the synchronous angular velocity; ω_r is the angular velocity of the rotor; s is the slip; $\omega_s = \omega_1 - \omega_r = s\omega_1$ is the rotation angular velocity of the d - q coordinate system relative to the rotor, that is, the slip angular velocity; n_p is the number of pole pairs.

3.2.2. Double-PWM Control Strategy

The AC-DC-AC converter is mainly composed of grid-side converter, rotor-side converter and energy storage capacitor. The grid voltage oriented vector control technique is adopted in the grid-side converter, and the basic control target is to ensure the stability of DC voltage and the decoupling control of active and reactive power. The main control target of the rotor-side converter is to control the rotor speed and reactive power (active power and power factor) of DFIG independently by the d -axis and q -axis components of the rotor current to realize the variable speed constant frequency operation of the generator. In order to achieve the control target, the stator voltage orientation vector control is adopted in the rotor-side converter to realize the decoupling control of active and reactive power of the stator. This can be seen in reference [33], not repeated here.

The model of DFIG and its control system established in MATLAB is shown in Figure 3.

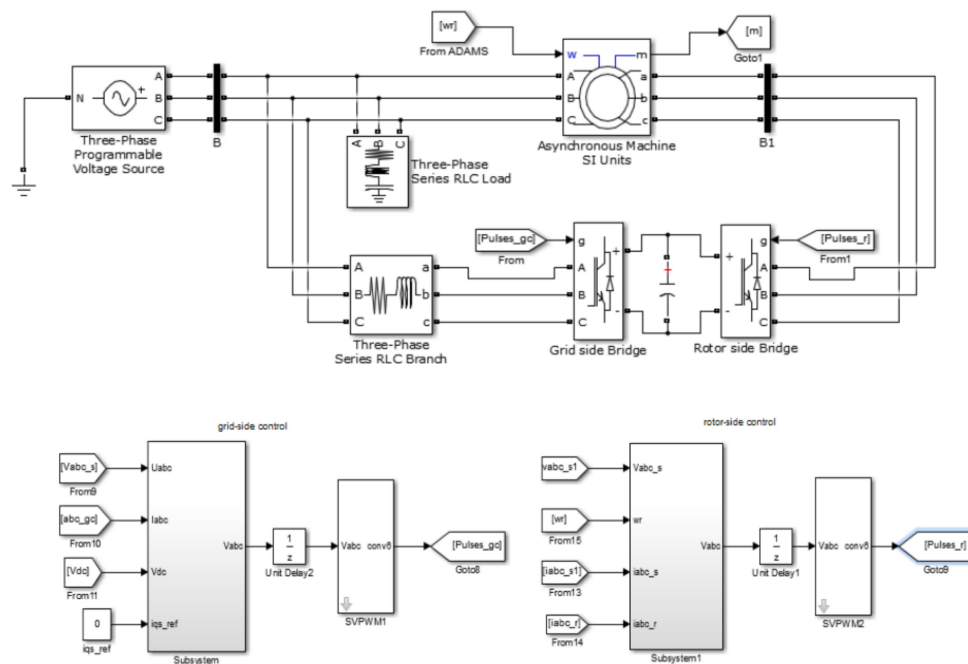


Figure 3. The model of doubly-fed induction generator (DFIG) and its control system.

3.2.3. Model Validation

Step speeds are input into the system; the curves of grid-side current, active and reactive power and DC voltage are shown in Figures 4–7, respectively.

From these figures, it can be seen that the grid-side current and active power can change with the input speeds of the system. The reactive power can be maintained around zero when the wind turbine is running at the maximum power point tracking stage. The decoupling control of active and reactive power on the grid-side is realized. In addition, the DC voltage can also be maintained near the given value of 1200 V. The control purpose is achieved, which shows that the model is established correctly.

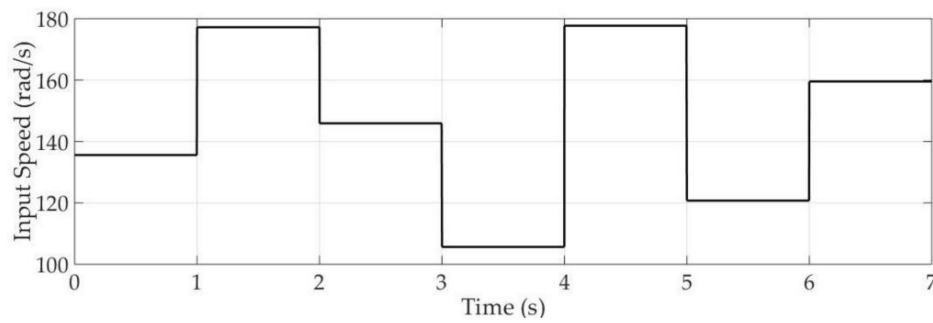


Figure 4. The varying of input speed.

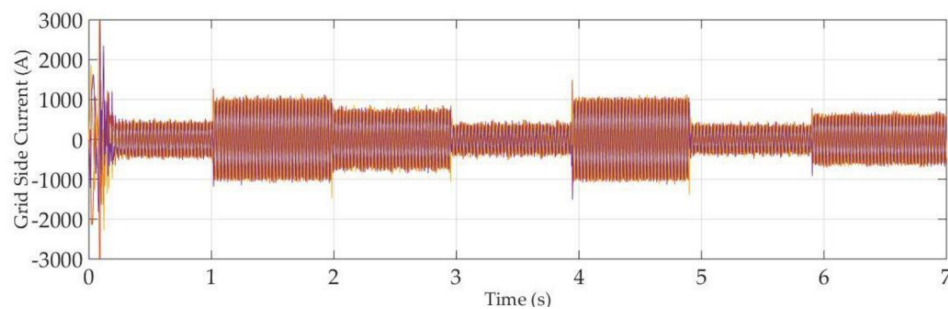


Figure 5. The varying of grid-side current.

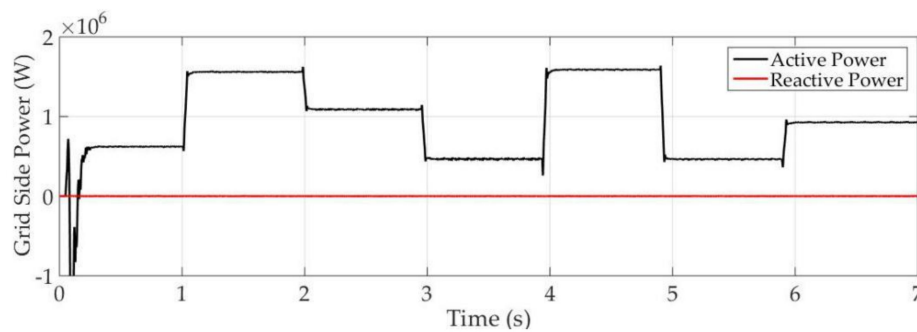


Figure 6. The varying of grid-side active and reactive power.

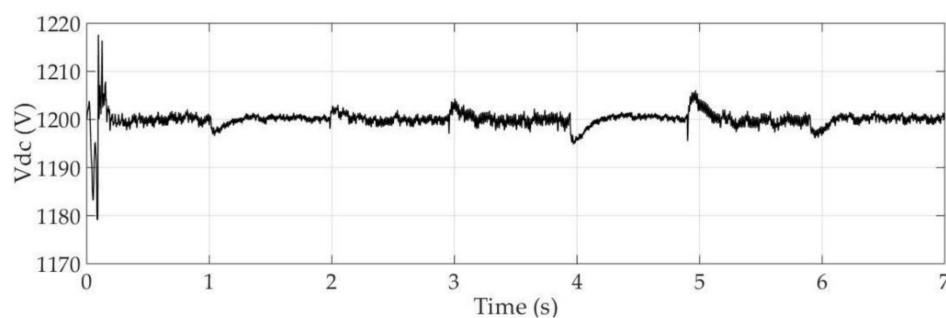


Figure 7. The varying of DC voltage.

3.3. The Joint Simulation of ADAMS and MATLAB

The established ADAMS model and MATLAB model are very large and complex. To realize the joint simulation between them, taking into account the simulation efficiency, the method of importing data is used. By inputting speed to the ADAMS model to simulate the normal working, parallel

misalignment, angle misalignment and comprehensive misalignment of the system, taking the speed of high-speed shaft of the gearbox as the input of the generator model in MATLAB, the corresponding stator currents can be obtained.

4. Fault Diagnosis Based on DTCWT Energy Entropy and SVM

4.1. Dual-Tree Complex Wavelet Energy Entropy Feature Extraction

When misalignment occurs in DFWT, the fault information will be mapped into the stator current, and the energy of each frequency component in the current will be changed accordingly and so will the energy distribution. Therefore, after performing the DTCWT on the stator current, the energy of each sub-band signal, which is called DTCWT energy entropy, can be calculated. The energy entropy can reflect the characteristics of the stator current. It is defined as follows:

Assuming

$$E_i = \int |c_i(t)|^2 dt = \sum_{k=1}^n |x_{ik}^2|, E = \sum_{i=1}^n E_i, p_i = \frac{E_i}{E} \quad (12)$$

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

Then, the energy entropy is the following:

$$P_i = -\sum_{i=1}^n p_i \lg p_i \quad (13)$$

Five frequency bands can be obtained through four-level dual-tree complex wavelet decomposition of the stator current, then they are reconstructed and the reconstructed signals contain the information of the fault from low to high frequency, respectively. The DTCWT energy entropy under four conditions (normal condition, parallel misalignment, angle misalignment and comprehensive misalignment) of the simulated transmission system is extracted, and the partial data are shown in Table 3.

Table 3. The partial data of DTCWT energy entropy of stator current in different working conditions.

Fault Type	$P_1 (\times 10^{-5})$	$P_2 (\times 10^{-5})$	$P_3 (\times 10^{-5})$	$P_4 (\times 10^{-5})$	$P_5 (\times 10^{-5})$
Normal Condition	6.9166	12.0458	0.006	0.0441	1.6267
	6.9187	12.0495	0.0061	0.0451	1.6273
	6.9196	12.0511	0.0059	0.0436	1.6274
	6.9199	12.0515	0.0058	0.0428	1.6274
	6.9254	12.0611	0.006	0.0445	1.629
Parallel Misalignment	5.9174	10.3131	0.0046	0.0371	1.3723
	5.9314	10.3375	0.0046	0.037	1.3758
	5.9437	10.3589	0.0047	0.0371	1.3789
	5.9529	10.3748	0.0047	0.0371	1.3813
	5.9712	10.4066	0.0047	0.0375	1.3859
Angle Misalignment	7.1076	11.5805	2.8627	0.2311	1.8312
	7.134	12.0723	1.3482	0.1344	1.7517
	7.1454	11.8373	2.225	0.2062	1.8049
	7.1698	11.4662	3.59	0.2938	1.8913
	7.2189	11.5697	3.5182	0.2481	1.8964
Comprehensive Misalignment	10.7387	16.5385	7.3899	1.4555	3.1362
	10.8266	16.4356	8.1625	1.7814	3.2307
	10.832	17.0982	6.1198	0.757	3.0325
	10.8764	17.1712	6.1879	1.5956	3.1145
	10.9098	14.4996	14.4082	2.7266	3.7177

4.2. Fault Diagnosis Method Based on PSO-SVM

The process of PSO optimization algorithm is as follows:

1. Initialize the size of the particle swarm, the acceleration factor, the maximum iterative number, the initial velocity and position of each particle;
2. Initialize the fitness values of each particle according to the fitness function for all the particles in the whole population;
3. Calculate the new velocity of the particle according to Formula (4) and limit the new velocity of each particle to the interval $[-V_{\max}, V_{\max}]$;
4. Calculate the new position of the particle according to Formula (5) and limit the new position of each particle to the interval $[-P_{\max}, P_{\max}]$;
5. For each particle, update the best position P_{best} ;
6. Update the group extreme value G_{best} ;
7. If the termination condition is satisfied, the solution can be gotten. Otherwise, return to step (3) to continue the search until the termination condition is satisfied.

After extracting the DTCWT energy entropy of the current, the samples are divided into two groups; one is the training set, the other is the testing set. The sample number of the training set is 100 (the number of each kind is 25), the testing sample number is also 100 (the number of each kind is also 25). PSO algorithm is used to optimize the penalty factor C and kernel function parameter σ of SVM. The specific implementation process is shown in Figure 8.

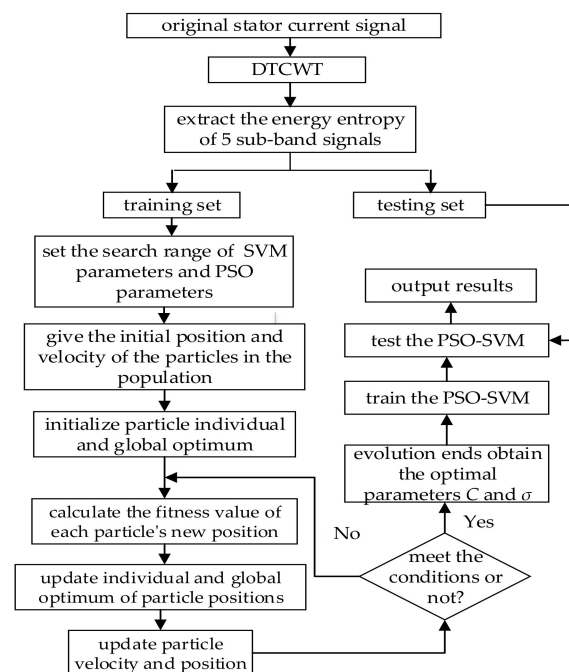


Figure 8. The particle swarm optimization (PSO)-support vector machine (SVM) fault diagnosis mode.

The value of C_1 and C_2 is important to the PSO algorithm. According to literature [30,31], some simulation results comparisons of other values of C_1 and C_2 are shown in Table 4 and the fitness curves corresponding to each group's values are shown in Figure 9.

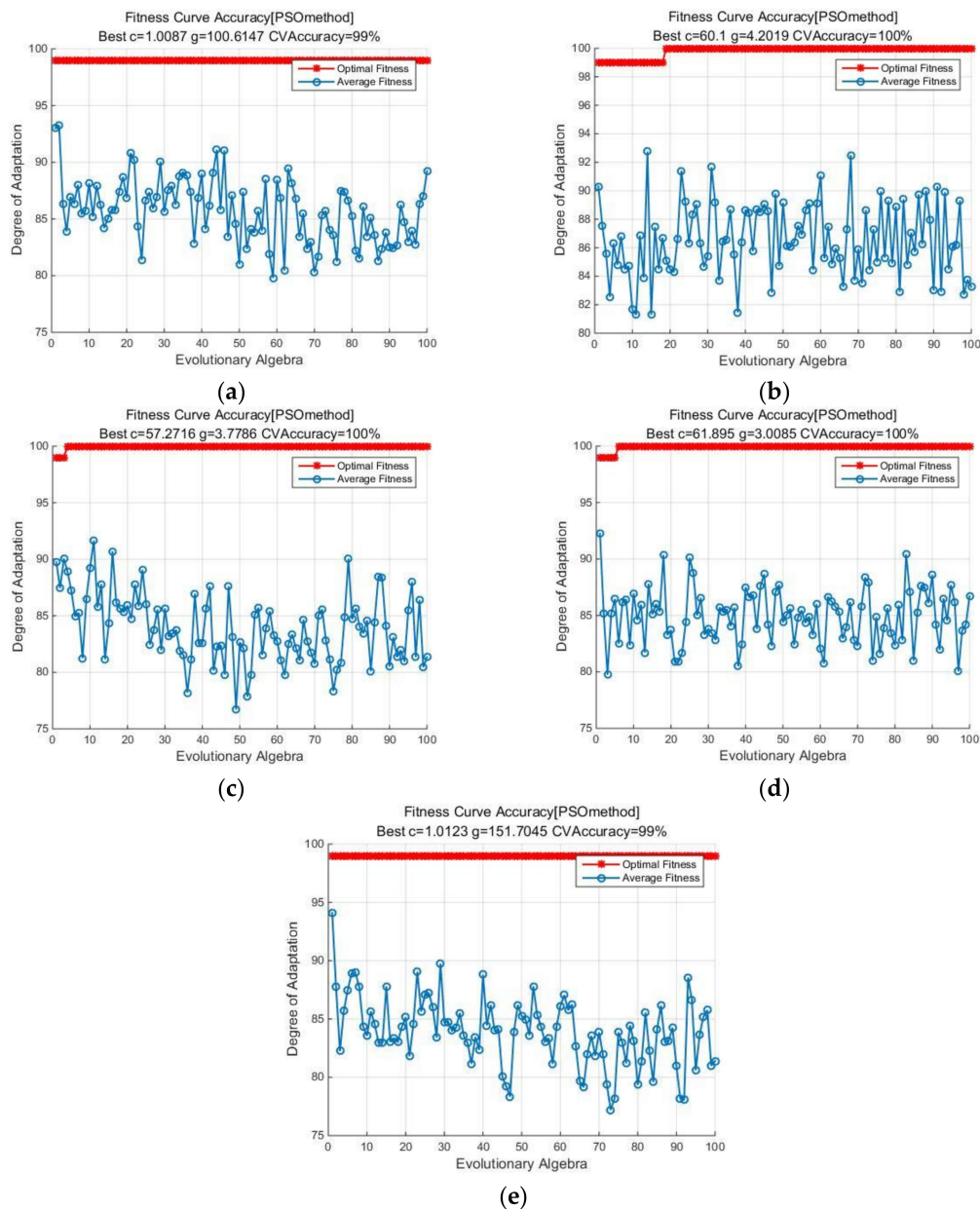


Figure 9. Fitness curves. (a) $C_1 = C_2 = 3$; (b) $C_1 = C_2 = 2.2$; (c) $C_1 = C_2 = 2$; (d) $C_1 = C_2 = 1.8$; (e) $C_1 = C_2 = 1$.

Table 4. Simulation results comparison of the values of C_1 and C_2 .

C_1	C_2	Evolution Algebra of Convergence (Reach 100%)	Optimal Fitness	Accuracy of Training Set	Accuracy of Testing Set
3	3	-	99%	99% (99/100)	94% (94/100)
2.2	2.2	18	100%	100% (100/100)	96% (96/100)
2	2	4	100%	100% (100/100)	96% (96/100)
1.8	1.8	5	100%	100% (100/100)	96% (96/100)
1	1	-	99%	100% (100/100)	96% (96/100)

From the above results, it can be seen that when $C_1 = C_2 = 2$, the fitness function converges fastest, and the optimal fitness, accuracy of training set and accuracy of testing set are the highest. So in this paper, let $c_1 = c_2 = 2$.

The testing set classification results are shown in Figure 10, where the category label “0” indicates normal working condition, “1” means parallel misalignment, “2” means angle misalignment, “3” means comprehensive misalignment.

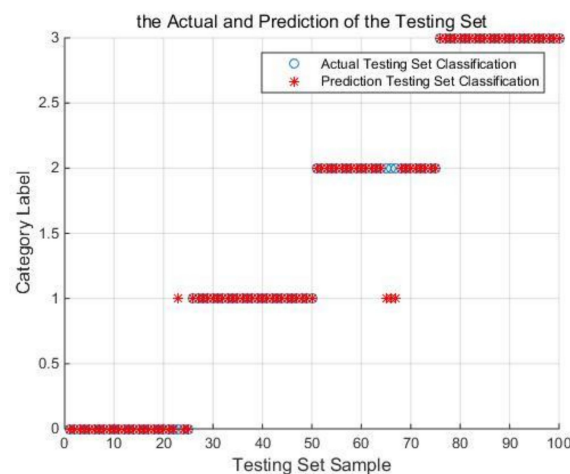


Figure 10. PSO-SVM testing results.

It can be seen from Figures 9c and 10 that the optimal fitness is 100%, the optimal parameters are $C = 57.2716$, $\sigma = 3.7786$ and the testing accuracy is 96%, which shows that the accuracy of fault classification is high.

In order to illustrate the advantages of the proposed PSO-SVM algorithm in this paper, SVM optimized by grid search and SVM optimized by genetic algorithm are also used to identify the faults of the same data, the specific recognition results of each diagnostic method are shown in Figure 11 and Table 5. It can be seen from the results that the training accuracy of GridSearch-SVM and GA-SVM is also high, but their promotion ability is less than PSO-SVM. Thus, it can be said that PSO-SVM algorithm is better in diagnostic performance.

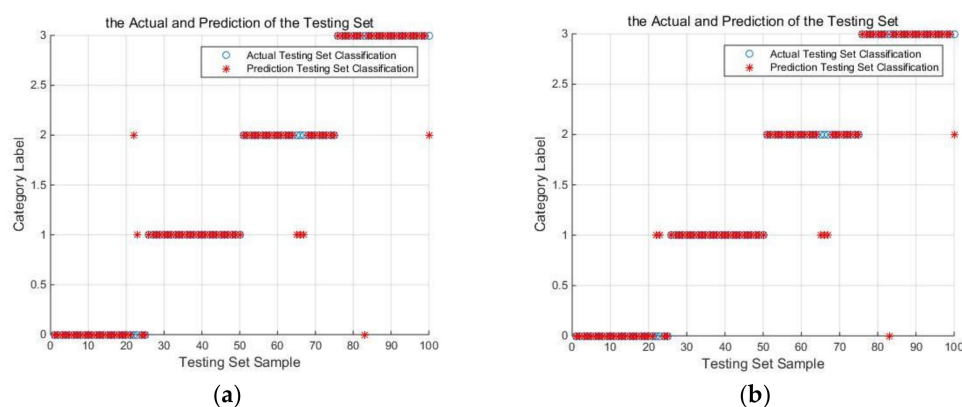


Figure 11. Diagnostic results of other algorithms. (a) GridSearch-SVM testing results; (b) Genetic algorithm (GA)-SVM testing results.

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

Item	C	σ	Accuracy of Training Set	Accuracy of Testing Set
GridSearch-SVM	16	4	100% (100/100)	93% (93/100)
GA-SVM	25.558	243.242	100% (100/100)	93% (93/100)
PSO-SVM	57.2716	3.7786	100% (100/100)	96% (96/100)

5. Conclusions

For actual situations where the monitoring equipment cost is high and the sensor installation is inconvenient when studying the misalignment of the transmission system of wind turbines based on traditional vibration signals, the joint simulation method of ADAMS and MATLAB is adopted to simulate the working conditions of the wind turbine transmission system in this paper. The corresponding stator current is firstly obtained; then the DTCWT is used to extract DTCWT energy entropy as the fault features; the SVM, which is good at a few samples learning, is used as classifier, and the parameters of SVM are optimized by PSO algorithm. The results show that the method is effective to diagnose the misalignment fault of the wind turbine transmission system, which is better than other commonly used algorithms. In the future, practical signals will be considered to verify the accuracy of the algorithm further.

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

Author Contributions: Yancai Xiao and Yi Hong contributed to paper writing and the whole revision process. Xiuhai Chen and Weijia Chen helped organize the articles. All the authors have read and approved the final manuscript.

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

References

1. Xiao, Y.; Wang, Y.; Mu, H.; Kang, N. Research on Misalignment Fault Isolation of Wind Turbines Based on the Mixed-Domain Features. *Algorithms* **2017**, *10*, 67. [[CrossRef](#)]
2. Hameed, Z.; Hong, Y.S.; Cho, Y.M.; Ahn, S.H.; Song, C.K. Condition monitoring and fault detection of wind turbine and related algorithms: A review. *Renew. Sustain. Energy Rev.* **2009**, *13*, 1–39. [[CrossRef](#)]
3. Liao, M.; Liang, Y.; Wang, S.; Wang, Y. Failure analysis of misalignment of wind turbine failure. *Mech. Sci. Technol.* **2011**, *30*, 173–190.
4. Long, J.; Wu, J. Fault Diagnosis of Rotor Misalignment for Windmill Generators. *Noise Vib. Control* **2013**, *3*, 222–225.
5. Hou, J.; Ding, H.; Yang, Z. Fault Diagnosis of Rotor Bearing System Based on Stator Current. *Coal Technol.* **2015**, *34*, 268–271.
6. Zhang, J. Research on Gear Fault Diagnosis of Wind Turbine Based on Stator Current Method. Master's Thesis, Harbin University of Science and Technology, Harbin, China, March 2016.
7. Yang, M.; Li, G.; Dong, C.; Chai, N.; Xu, D. Gear fault diagnosis method based on MCSA. *J. Beijing Jiaotong Univ.* **2015**, *39*, 87–91.
8. Zhang, K.; Wan, N. Application of stator current method in fault diagnosis of asynchronous motor. In Proceedings of the 12th National Conference on Equipment Monitoring and Diagnosis, Hainan, China, November 2005.
9. Zhang, T. Fault Detection in Locomotive Traction Motors by Stator Current Dual-Tree Complex Wavelet Analysis. Master's Thesis, Beijing Jiaotong University, Beijing, China, March 2016.
10. Xu, Y.; Meng, Z.; Zhao, G. Study on compound fault diagnosis of rolling bearing based on dual-tree complex wavelet transform. *Chin. J. Sci. Instrum.* **2014**, *35*, 447–452.
11. 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](#)]
12. Zhu, X. Research on Rotating Machine Fault Diagnosis and State Prediction Method Based on Support Vector Machine. Ph.D. Thesis, North China Electric Power University, Beijing, China, March 2013.
13. Xu, X. Research on Rotating Machinery Vibration Fault Diagnosis Based on Support Vector Machine. Master's Thesis, Northeast Petroleum University, Daqing, China, May 2016.
14. Li, N. The Fault Diagnosis Method of Support Vector Machine Based on fuzzy C-Mean and Particle Swarm Optimization. Master's Thesis, University of Electronic Science and Technology of China, Chengdu, China, May 2011.

15. Xiao, Y.; Zhang, T.; Ding, Z.; Li, C. The Study of Fuzzy Proportional Integral Controllers Based on Improved Particle Swarm Optimization for Permanent Magnet Direct Drive Wind Turbine Converters. *Energies* **2016**, *9*, 343. [CrossRef]
16. Kingsbury, N.G. The dual-tree complex wavelet transform: A new technique for shift invariance and directional filters. *IEEE Digit. Signal Process. Workshop* **1998**, *98*, 2–5.
17. Selesnick, I.W.; Baraniuk, R.G.; Kingsbury, N.G. The dual-tree complex wavelet transform. *IEEE Digit. Signal Process. Mag.* **2005**, *22*, 123–152. [CrossRef]
18. Vermaak, H.; Nsengiyumva, P.; Luwes, N. Using the Dual-Tree Complex Wavelet Transform for Improved Fabric Defect Detection. *J. Sens.* **2016**, *2016*, 9794723. [CrossRef]
19. Meng, Z. Applications of Electromechanical Equipment Fault Diagnosis Based on Dual-Tree Complex Wavelet. Master's Thesis, Beijing University of Technology, Beijing, China, March 2014.
20. Chakraborty, S.; Chatterjee, A. A dual-tree complex wavelet transform-based approach for recognition of power system transients. *Expert Syst.* **2015**, *32*, 132–140. [CrossRef]
21. Rajesh, D.S.; Shekar, B.H. Undecimated dual tree complex wavelet transform based face recognition. In Proceedings of the 2016 International Conference on Advances in Computing, Communications and Informatics, Jaipur, India, 21–24 September 2016.
22. Suykens, J.A.K. Support Vector Machines: A Nonlinear Modelling and Control Perspective. *Eur. J. Control* **2001**, *7*, 311–327. [CrossRef]
23. Suykens, J.A.K.; Vandewalle, J. Recurrent least squares support vector machines. *IEEE Trans. Circuits Syst. I Fundam. Theory Appl.* **2008**, *47*, 1109–1114. [CrossRef]
24. Zhang, X.G. Introduction to Statistical Learning Theory and Support Vector Machines. Available online: http://en.cnki.com.cn/Article_en/CJFDTOTAL-MOTO200001005.htm (accessed on 2 November 2017).
25. Wei, W. The Research of Gear Fault Diagnosis Based on Image Recognition. Master's Thesis, Zhengzhou University, Zhengzhou, China, May 2016.
26. Xu, K.; Zheng, X. Diagnosis on Wind Turbine Faults Based on Multi-classification Support Vector Machine. *Zhejiang Electr. Power* **2015**, *4*, 54.
27. Zhang, S.; Zhang, X.; Lei, T.; Di, Y. Research on Fault Diagnosis Based on PSO and SVM. *Comput. Meas. Control* **2008**, *16*, 1573–1575.
28. Zhang, W.; Ma, D.; Wei, J.J.; Liang, H.F. A parameter selection strategy for particle swarm optimization based on particle positions. *Expert Syst. Appl.* **2014**, *41*, 3576–3584. [CrossRef]
29. Ma, R.; Yang, L.; Zhang, Z. Analysis the Characteristic of C1, C2 Based on the PSO of Iterative Shift and Trajectory of Particle. *Math. Comput.* **2013**, *4*, 109–115.
30. Yuan, K. Improved Particle Swarm Optimization and Parameter Inversion in Internal Variable. *Coal Geol. Explor.* **2017**, *45*, 112–117.
31. Wang, W.; Fang, X.; Zhong, Q. Weighted Statistical Model of Dam Monitoring Based on Improved Particle Swarm Optimization. *J. Yangtze River Sci. Res. Inst.* **2017**, *34*, 41–46.
32. Pousinho, H.M.I.; Catalao, J.P.S.; Mendes, V.M.F. Wind power short-term prediction by a hybrid PSO-ANFIS approach. In Proceedings of the 15th IEEE Mediterranean Electrotechnical Conference, Valletta, Malta, 26–28 April 2010.
33. Ren, Y.; An, Q. *Flexible Grid Connected Operation and Control of Doubly Fed Wind Power Generation Unit*, 1st ed.; China Machine Press: Beijing, China, 2011; pp. 12–144.

