Fault Diagnosis of a Wind Turbine Gearbox Based on Improved Variational Mode Algorithm and Information Entropy

Fan Zhang, Wenlei Sun *, Hongwei Wang and Tiantian Xu

School of Mechanical Engineering, Xinjiang University, Urumqi 830047, China; zhangfan@stu.xju.edu.cn (F.Z.); wanghongwei@stu.xju.edu.cn (H.W.); xutiantian@stu.xju.edu.cn (T.X.)

* Correspondence: sunwenxj@163.com

Abstract: The working environment of wind turbine gearboxes is complex, complicating the effective monitoring of their running state. In this paper, a new gearbox fault diagnosis method based on improved variational mode decomposition (IVMD), combined with time-shift multi-scale sample entropy (TSMSE) and a sparrow search algorithm-based support vector machine (SSA-SVM), is proposed. Firstly, a novel algorithm, IVMD, is presented for solving the problem where VMD parameters (K and α) need to be selected in advance, which mainly contains two steps: the maximum kurtosis index is employed to preliminarily determine a series of local optimal decomposition parameters (K and α), then from the local parameters, the global optimum parameters are selected based on the minimum energy loss coefficient (ELC). After decomposition by IVMD, the raw signal is divided into K intrinsic mode functions (IMFs), the optimal IMF(s) with abundant fault information is (are) chosen based on the minimum envelopment entropy criterion. Secondly, the time-shift technique is introduced to information entropy, the time-shift multi-scale sample entropy algorithm is applied for the analysis of the complexity of the chosen optimal IMF and extract fault feature vectors. Finally, the sparrow search algorithm, which takes the classification error rate of SVM as the fitness function, is used to adaptively optimize the SVM parameters. Next, the extracted TSMSEs are input into the SSA-SVM model as the feature vector to identify the gear signal types under different conditions. The simulation and experimental results confirm that the proposed method is feasible and superior in gearbox fault diagnosis when compared with other methods.

Keywords: wind turbine gearbox; variational mode decomposition; time-shifting multi-scale sample entropy; sparrow search algorithm; support vector machine; fault diagnosis

1. Introduction

In recent decades, the installed scale and grid-connected capacity of wind turbines have increased significantly. As of the end of 2019, the installed capacity of wind turbines in the world exceeded 690 million kW, and more than 100,000 wind turbines had been built [1]. Behind such a huge installed capacity, the daily maintenance of wind turbines is essential. The efficient and accurate fault diagnosis method not only can real-time monitor the operation of the wind turbines, and also can detect the potential fault of wind turbines. Wind turbines' fault diagnosis methods mainly include the method based on the model and the method based on data [2]. Casau et al. [3] used fault detection and isolation (FDI) and fault-tolerant control (FTC) methods based on the model falsification technique using set-valued observers (SVOs) to ensure the safe operation of wind turbines. Odgaard et al. [4] proposed a standard wind turbine benchmark model, which was used to detect the running state of the wind turbines. Badihi et al. [5] used fuzzy mathematical language to model wind turbines and proposed an integrated fault detection and diagnosis (FDD) and FTC scheme. Apart from providing fault tolerance of wind turbine sensors, the scheme also greatly improves the overall performance of wind turbines under fault-free and fault-free conditions. The above methods have achieved certain results, but with the continuous development of rotating equipment, the mechanical composition is increasingly complex,

therefore, it is hard to build a model which can comprehensively reflect the mechanical equipment. With the wide application of big data, data mining technology has been maturing. As a result, the data-driven fault diagnosis method of rotating equipment has been favored by many researchers. The time-frequency characteristics of vibration signals in different states have clear differences, additionally, the vibration signals are easy to be collected. So, the fault diagnosis method of rotating mechanical equipment based on vibration signal analysis is the most widely used method at present.

The gearbox is a critical component in the drive system of a wind turbine; due to the influence of its working environment, local faults such as cracks or spalling frequently occur. According to research, gearbox faults account for one-quarter of the total wind turbine faults, and the operation and maintenance costs account for about 25%–35% of the total cost of wind turbines. The downtime caused by gearbox faults is longer than that caused by other faults [6,7]. Therefore, it is essential to effectively monitor the running condition of the gearbox.

The working environment of a wind turbine gearbox is complex, which is often exposed to heavy loads, high temperatures, and speed changes. Additionally, due to the influence of load fluctuations and sliding friction between teeth, the collected gearbox vibration signal is non-stationary and nonlinear. Consequently, researchers have proposed various methods for the analysis of non-stationary and nonlinear signals. Ridha et al. [8] used empirical mode decomposition (EMD) combined with a shock detector (SD) to detect impact damage to bevel gearboxes in the non-static state. Amrinder et al. [9] presented ensemble empirical mode decomposition (EEMD) to initially decompose vibration signals of rolling bearings and combined it with fuzzy entropy to predict and analyze the severity of faults in the inner and outer rings of rolling bearings. Yao et al. [10] proposed a complementary fully integrated empirical mode decomposition based on adaptive noise (CCEEMDAN) and applied it to the fault diagnosis of rolling bearings. These researches have achieved certain results, however, EMD is a signal analysis method based on a recursive algorithm. Due to the limitation of its principle, signals decomposed by EMD and some of its derivative algorithms may induce several problems such as mode mixing and end effects. The variational mode decomposition (VMD) algorithm proposed by Dragomiretskiy et al. [11] can effectively overcome the above problems. Additionally, the algorithm outperforms EMD in terms of noise robustness; therefore, it has been widely used in signal decomposition. Zhang et al. [12] used VMD combined with sample entropy to extract the features of weak faults of planetary gearboxes. Pan et al. [13] proposed the combination of VMD and wavelet packet for gear fault diagnosis. VMD has merits compared with other algorithms in signal decomposition, whereas its signal decomposition performance strongly depends on the selection of mode decomposition parameter K and the mode frequency bandwidth control parameter α of the VMD algorithm. When α is small, and the bandwidth of the Wiener filter is large, which can better detect the impact components in the signal; however, simultaneously, the decomposed modes are more likely to contain noise components [14]. When K is small, the number of decomposed mode components is small, which may cause partially missing information. When the value of K is large, it may induce problems such as over-decomposition and frequency mixing. In previous research work [15,16], most scholars used the default VMD algorithm parameters (K, α) to decompose the signal according to their own professional knowledge or previous experience. Due to the different signal characteristics of different objects, decomposing the signal using default parameters cannot achieve adaptive decomposition according to the characteristics of the signal, which significantly reduces the efficiency and precision of VMD algorithm signal decomposition. With the rapid development of meta-heuristics, many scholars have selected the VMD parameters automatically using the population intelligence algorithm to achieve the best VMD signal decomposition effect. Anil et al. [17] applied a genetic algorithm (GA) to achieve the adaptive optimization of VMD decomposition parameters. Yan et al. [18] presented the cuckoo search algorithm (CSA) to determine the decomposition parameters of VMD, yet the search effect of the population intelligent

search algorithm heavily relies on the preset parameters of the optimization algorithm and the fitness function selection. Accordingly, in this paper, an improved variational mode decomposition (IVMD) algorithm is proposed, which combines the kurtosis index with the energy loss coefficient (ELC) to determine the VMD parameters.

With the rapid development of nonlinear dynamics, nonlinear dynamic analysis methods have been increasingly applied to the fault diagnosis of mechanical equipment to extract fault features, for instance, entropy [19]. Approximate entropy (AE) was the entropy method first introduced in the field of mechanical fault diagnosis [20], followed by permutation entropy (PE), sample entropy (SE), fuzzy entropy (FE), discrete entropy (DE), and their derived algorithms [21–26]. Among these entropies, SE has been widely used in feature extraction due to its simple calculation and low time cost. The original sample entropy can only analyze a time series on a single scale, which leads to missing information to some extent. Multi-scale sample entropy (MSE) uses coarse graining based on the mean value to reconstruct the signal on the basis of SE, which enables the analysis of signals on different scales. MSE can be thought of as down-sampling processing; according to sampling theory, the sampling frequency is reduced and the bandwidth of the signal narrows, which leads to MSE ignoring the high-frequency information of the original signal. In addition, with the increase in the scale factor, the length of coarse-grained time series shortens; therefore, it is difficult to comprehensively obtain the information of the original signal. Considering the above issues, we propose the entropy method termed TSMSE, which combines time-shift technology and sample entropy to extract the features of gearbox signals to enable more accurate analysis and recognition of the running state of a gearbox.

At present, a mainstream fault diagnosis method involves inputting the extracted feature vector into the classifier to recognize the fault type. Xu et al. [27] presented a new method of intelligent fault diagnosis called deep convolutional nearest neighbor matching network (DC-NNMN) based on few-shot learning and used that method for the fault diagnosis of a gear with a small labeled sample. Yu et al. [28] proposed wavelet transform combined with an extreme learning machine (ELM) for automatic fault diagnosis of wind turbines. Liang et al. [29] presented the combination of convolutional neural network (CNN) and multi-label classification (ML) for the efficient compound fault diagnosis of a gearbox. Sheng et al. [30] obtained the features vectors of a bearing vibration signal by VMD-phase space reconstruction (PSR)-singular value decomposition (SVD) and used the improved k-nearest neighbor algorithm for the fault diagnosis of rolling bearings. Among numerous classifiers, support vector machine (SVM) has been widely used in fault diagnosis as a result of its strong generalization ability on nonlinear problems and its excellent reliability given small sample data. The SVM fault-recognition effect is heavily dependent on penalty factor c and kernel parameter σ; therefore, some SVM optimization methods were proposed to resolve this issue. Yan et al. [31], Li et al. [32], and Luo et al. [33] used particle swarm optimization (PSO), the cuckoo search algorithm (CSA), and gray wolf algorithm (GWA), respectively, to optimize the crucial SVM parameters; however, the above algorithms are prone to problems such as slow convergence, local optimal solutions, etc. The sparrow search algorithm (SSA) is a population intelligence algorithm proposed in 2020. Studies showed that the SSA is superior to the existing algorithms in terms of search accuracy, convergence speed, and stability [34]. Therefore, we chose to use the sparrow search algorithm to adaptively optimize the key support vector machine parameters (SSA-SVM).

To summarize, in this paper, we propose a neoteric fault diagnosis method for wind turbine gearboxes. The method contains three stages in general. First, the kurtosis index and energy loss coefficient (ELC) are used to select VMD parameters K and α, and the signal is decomposed by VMD with the determined parameters. The best IMF following the minimum envelope entropy criterion is chosen. Secondly, time-shift multi-scale sample entropy (TSMSE) is used to extract features of the optimal IMF. Ultimately, the sparrow search algorithm is applied to optimize the support vector machine parameters (SSA-SVM),

then the extracted feature entropy vector is input into the optimized SVM model to analyze and identify the gearbox fault types.

The rest of this paper is organized as follows: Section 2 introduces the basic principle of the VMD algorithm and the process of the IVMD algorithm. Section 3 briefly reviews the MSE and TSMSE algorithm implementation process. Section 4 describes the sparrow search algorithm and the process of the sparrow search algorithm-based support vector machine. Section 5 outlines the fault diagnosis model of a wind turbine gearbox proposed in this paper. Section 6 describes the experiment used to verify the feasibility and effectiveness of the model proposed in this paper. Section 7 provides the conclusion of this study.

2. VMD and IVMD Algorithms

2.1. VMD Algorithm

Vibration signals of wind turbine gearboxes are characterized by strong randomness, poor stability, and aliasing of physical information. Therefore, signal decomposing should be preliminarily carried out to extract the signal features. As stated in the Introduction, EMD and other signal decomposition methods experience problems such as the end effect and mode aliasing. The VMD algorithm is a new, adaptive, non-recursive, signal-decomposition method that analyzes and decomposes the signal in the time-frequency domain using the classical Wiener filter, Hilbert transform, and frequency mixing principle. The VMD algorithm decomposes a real signal into several intrinsic mode functions (IMFs) with certain center frequencies and limited bandwidth through iteration; these modes are independent of each other and have sparseness. Therefore, VMD can effectively preprocess nonlinear signals. By estimating the bandwidth of each mode, the variational constraint equation model, as expressed in Formula (1), can be solved, which minimizes the sum of the modes' bandwidths under the constraint that the sum of all modes is the original signal.

$$
\min_{\{u_k\},\{\omega_k\}} \left\{ \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\}
$$
$$
s.t. \sum_{k=1}^{K} u_k(t) = f(t) \tag{1}
$$

Here, $\{u_k\}$ denotes decomposed IMF components, $\{\omega_k\}$ represents the center frequency of the IMF components, ∂t is the partial derivative of time t, $\delta(t)$ represents the unit pulse function, j is the imaginary unit, and $f(t)$ indicates the raw signal in time-domain.

The Lagrangian multiplier $\lambda(t)$ and penalty parameter α are introduced to transform Formula (1) into an unconstrained variational model, expressed as:

$$
L(\{u_k\},\{\omega_k\},\lambda(t)) = \alpha \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 +
$$
$$
\left\| f(t) - \sum_{k=1}^{K} u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_{k=1}^{K} u_k(t) \right\rangle \tag{2}
$$

The alternative direction method of the multiplication operator is applied to update u_k, ω_k, and λ_k, iteratively until the variational model obtains the optimal solution. The outputs are the mode components obtained by VMD.

2.2. IVMD Algorithm

To solve the existing problem in the VMD algorithm, we propose a method that applies the kurtosis index combined with the energy loss coefficient (ELC) to adaptively select the decomposition parameters (K, α) of the VMD algorithm. The flow chart is depicted in Figure 1.
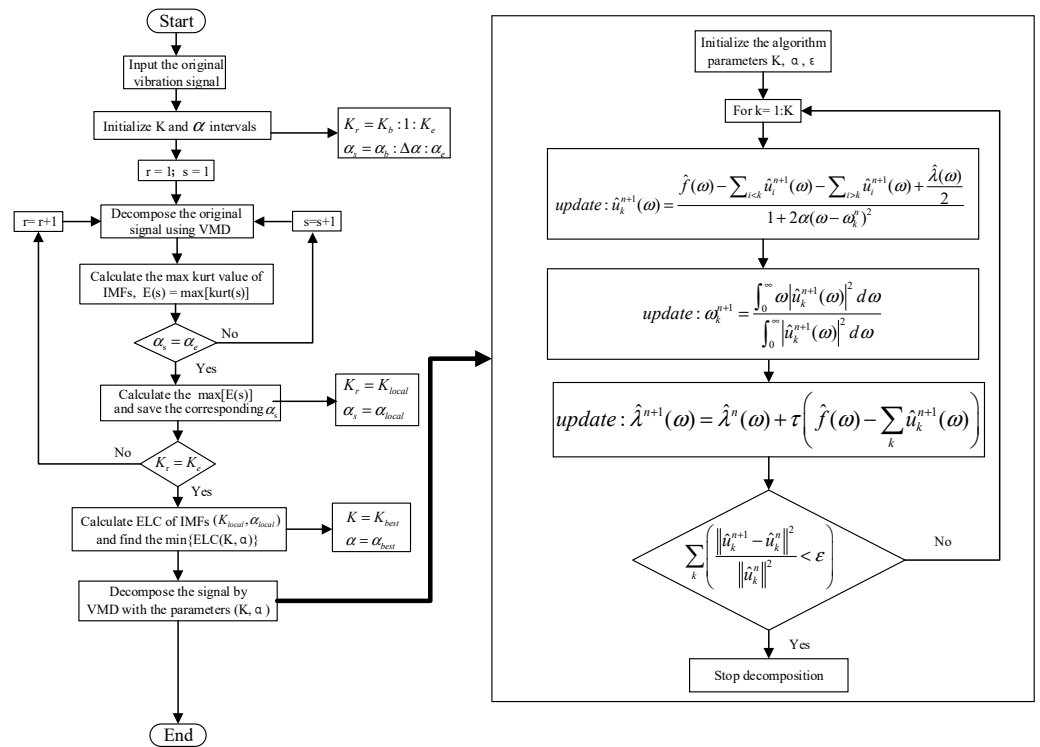
Figure 1. Flowchart of the IVMD algorithm.

The kurtosis index reflects the impact characteristics of the signal and is sensitive to weak faults in the early stage. Consequently, a series of local optimal parameter pairs are preliminarily selected according to the maximum kurtosis criterion. The calculation formula of kurtosis is as follows:

$$e(i) = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{u_i - \overline{u}}{\sigma}\right)^4 \tag{3}$$

where u_i denotes the mode component IMF_i, N is the mode length, $\overline{u}$ represents the mean of u_i, and σ is the standard deviation of u_i.

For the search range and step size of the VMD decomposition parameter, $K_r \in [K_b, K_e]$, the search step size is 1, $\alpha_s \in [\alpha_b, \alpha_e]$, and the search step size is set to $\Delta\alpha$. Suppose the mode decomposition number is k, the bandwidth control parameter is α_s, and the signal is decomposed by VMD with this parameter pair. Then, the kurtosis value of each IMF is calculated using Formula (3) and the maximum kurtosis value is considered the kurtosis value under (k, α_s). Then, the $(k, \alpha_{s+\Delta\alpha})$ parameter is used to decompose the signal using VMD, and the kurtosis value of each IMF is calculated again. The maximum kurtosis value is taken as the kurtosis value under $(k, \alpha_{s+\Delta\alpha})$. The cycle continues until $\alpha_s = \alpha_e$. The α_{lobal}^k corresponding to the maximum kurtosis value in the series of obtained kurtosis values is taken as the optimal α under this mode number k. Then, the mode number is set to $k + 1$, and the next cycle is continued until $k = K_e$. Then a series of local optimal parameters can be obtained. The mathematical expression is expressed as follows:

$$\begin{aligned}
\overrightarrow{e_s^k} &= \max(e_s^1, e_s^2, \ldots, e_s^k)\\
e_{lobal}^k &= (\overrightarrow{e_b^k}, \overrightarrow{e_{b+\Delta\alpha}^k}, \ldots, \overrightarrow{e_s^k}, \ldots, \overrightarrow{e_e^k})\\
\alpha_{lobal}^k &= \text{argmax}(e_{lobal}^k)
\end{aligned} \tag{4}$$

After obtaining a series of local optimal parameter pairs, the signal is decomposed by the combination of local parameters. Then, the ELC of the mode component is calcu-

lated, and the parameter with the least energy loss is taken as the global optimal VMD decomposition parameter. The ELC calculation formula is expressed as:

$$ELC = \frac{\left\| f(t) - \sum\limits_{k=1}^{K} u_k \right\|_2^2}{\|f(t)\|_2^2} \tag{5}$$

where *f(t)* denotes the signal.

3. TSMSE Algorithm

When wind turbine gearbox faults occur, the complexity and chaos characteristics of the collected vibration signal change. Therefore, the linear signal processing method cannot be used to accurately analyze the signal, so the nonlinear dynamic analysis method is introduced to extract and analyze the nonlinear characteristics of the signal. As stated in the Introduction, sample entropy and multi-scale sample entropy are widely used nonlinear dynamic parameter that can reflect the irregularity and complexity of signals, but they experience some issues, such as sample entropy can only reflect the information of a signal on a single scale, multi-scale sample entropy cannot measure the complexity of a signal at high scale. To solve the problems faced by sample entropy and multi-scale sample entropy for extracting fault features, time-shift multi-scale sample entropy (TSMSE) is proposed.

3.1. Multi-Scale Sample Entropy Algorithm

The multi-scale sample entropy (MSE) based on the sample entropy uses the coarse-graining technology to obtain a series of the coarse-graining time series, and then the sample entropy of each coarse-graining time series is calculated, the implementation process is described as follows:

For a given raw time series {$x(i)$, i = 1,2, … , N}, the coarse-graining process follows the Formula (6).

$$u_j^k = \frac{1}{k} \sum\limits_{i=(j-1)k+1}^{jk} x_i \tag{6}$$

where *k* represents the scale factor; *j* = 1, 2, 3, … , N/k; N is the length of the series.

$$MSE(x,k,m,r) = SE(u^k,m,r)$$
$$r = (0.2 \sim 0.25)^* std(u^k) \tag{7}$$

where *r* is the similarity tolerances; *m* is pattern dimension; std represents the standard deviation. SE means the sample entropy value of the time series. The detailed sample entropy calculation steps are described in [22].

3.2. Time-Shift Multi-Scale Sample Entropy Algorithm

In the time-shift multi-scale sample entropy algorithm, raw time series {$x(i)$, i = 1,2, … , N} are reconstructed according to Formula (8). The sample entropy is calculated for each time shift series after reconstruction.

$$u^\beta{}_k = (x_\beta, x_{\beta+k}, x_{\beta+2k}, \ldots, x_{\beta+ck}) \tag{8}$$

where *k* denotes the scale factor; β = 1, 2, 3, … , k, $c = \frac{N-\beta}{k}$, which indicates the largest integer not exceeding $\left(\frac{N-\beta}{k}\right)$, where N is the length of the series.

The TSMSE is expressed as:

$$TSMSE(x,m,r,k) = \frac{1}{k} \sum\limits_{\beta=1}^{k} SE(u^\beta{}_k,m,r)$$
$$r = (0.2 \sim 0.25)^* std(u^\beta{}_k) \tag{9}$$

We use an example to clearly present the process of signal reconstruction, where N is 10 and the scale is set to 3. Given a time series $\{x(i), i = 1, 2, \ldots, 10\}$, the calculation process of its TSMSE value is shown in Figure 2.
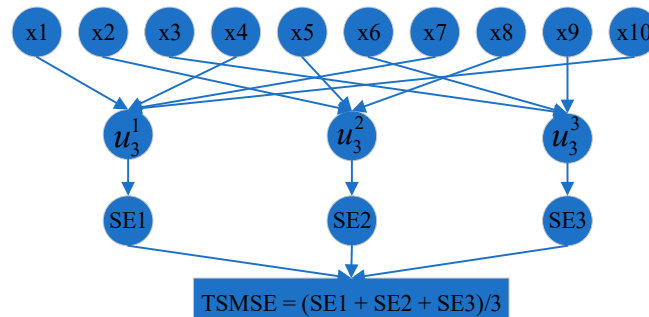
Figure 2. Calculation process of TSMSE.

It is worth noting that when the scale factor is 1, the time series reconstructed by the MSE and TSMSE algorithms are the same as the original time series, therefore, the entropy value calculated by the MSE and TSMSE algorithms also are the same as that value calculated by the SE algorithm.

4. SSA-SVM Algorithm

The SVM algorithm maps sample space data into high-dimensional space through the kernel function and then classifies the data in high-dimensional space. As such, the nonlinear separable problem can be transformed into a linear separable problem. Therefore, the SVM algorithm performs well with nonlinear data classification problems and it is widely used in the fault diagnosis of rotating machinery equipment [35–37], but the problem of parameter selection must be addressed. According to the above analyses, for solving the difficulties experienced with SVM classification model application, we selected the sparrow search algorithm (SSA) to search for the optimal SVM parameters (c, σ) and set the SVM classification error rate as the fitness function; the flow chart and mathematical model are shown in Figure 3 and Formula (10), respectively.

$$fit(c, \sigma) = \frac{1 - accuracy}{100} \tag{10}$$

Here, *accuracy* denotes the SVM classification accuracy.

The SSA is inspired by the behavior of sparrow groups when foraging and when facing predation, which divides the sparrow population into producers and scroungers, while each sparrow individual position represents a solution of an SVM parameter. The producers are actively looking for food sources, while scroungers obtain food from producers. Naturally, when a predator is found by a sparrow, it will send an alarm, then the population will fly away. Through the above process, the optimal solution to a problem can be determined. The positions of the producer, scrounger, and the sparrow who identifies the danger are updated according to Formulas (11)–(13).
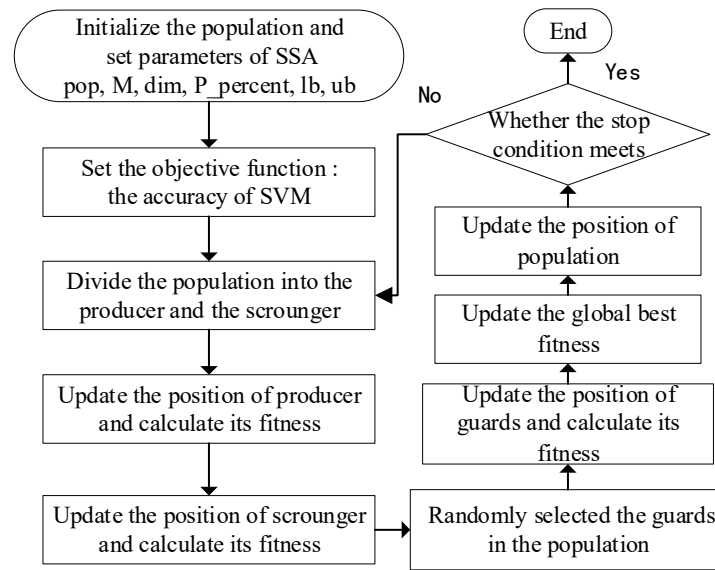
Figure 3. Flowchart of SSA-SVM algorithm.

$$x_{i,j}^{t+1} = \begin{cases} x_{i,j}^t * \exp\left(\frac{-i}{\alpha * iter_{\max}}\right) & \text{if } R_2 < ST \\ x_{i,j}^t + Q^*L & \text{if } R_2 \geq ST \end{cases} \quad (11)$$

Here, t denotes the current iteration; $j = 1, 2, \ldots, d$. $x_{i,j}^t$ represents the value of the jth dimension of the ith sparrow at iteration t; $iter_{\max}$ is the maximum number of iterations; α is a random number in $(0, 1]$; Q is a random number that obeys normal distribution; L represents a unit matrix of $1 \times d$; R_2 indicates the value of alarm, which usually is taken from $[0,1]$; and ST ($ST \in [0.5, 1.0]$) represents the safety threshold.

$$x_{i,j}^{t+1} = \begin{cases} Q^* \exp\left(\frac{x_{worst}^t - x_{i,j}^t}{i^2}\right) & \text{if } i > \frac{n}{2} \\ x_P^{t+1} + \left|x_{i,j}^t - x_P^{t+1}\right| * A^{+*}L & \text{other} \end{cases} \quad (12)$$

Here, x_p denotes the best position of the producer, x_{worst} represents the current global worst position, A is a $1 \times d$ matrix, and the elements in the matrix are randomly assigned from $[1, -1]$, and N indicates the population size. $A^+ = A^T(AA^T)^{-1}$.

$$x_{i,j}^{t+1} = \begin{cases} x_{best}^t + \beta^* \left|x_{i,j}^t - x_{best}^t\right| & \text{if } f_i > f_g \\ x_{i,j}^t + K^* \left(\frac{\left|x_{i,j}^t - x_{worst}^t\right|}{(f_i - f_w) + \varepsilon}\right) & \text{if } f_i = f_g \end{cases} \quad (13)$$

Here, x_{best} denotes the global best position at present; β is a random number, which follows a mean of 0 and a variance of 1 in the normal distribution; K represents a random number within $[-1, 1]$; f_i, f_g, and f_w indicate the fitness value of the current individual, the fitness value of the global optimal solution and the fitness value of the worst solution, respectively. In addition, an arbitrary small constant ε is introduced to avoid zero division error [34].

5. The Proposed Wind Turbine Gearbox Fault Diagnosis Model

Based on the above theories, we constructed a novel method for fault diagnosis of wind turbine gearboxes that combines IVMD-TSMSE with SSA-SVM. The flowchart of the proposed method is described in Figure 4 and the specific steps are as follows:

Step 1: The integrated wind turbine power transmission fault diagnosis platform collects the vibration signals of gears under different working conditions.

Step 2: The kurtosis index and energy loss coefficient are used to determine the optimal parameter pairs (K_{best}, α_{best}) of VMD, then the original vibration signals are decomposed into several modes.

Step 3: In line with the minimum envelopment entropy criterion, the optimal mode is chosen for subsequent analysis.

Step 4: The feature vectors that contain rich fault information are extracted from the optimal mode using the TSMSE algorithm.

Step 5: The best parameters' penalty factor c and kernel parameter σ of SVM are determined by the SSA.

Step 6: The extracted feature vectors are randomly divided into training samples and testing samples. The training samples are used to train the optimized SVM model with the SSA, whereas the testing samples are used to test the trained SVM model for proving the effectiveness and superiority of the method proposed in this paper on the classification of different conditions of wind turbine gearboxes.
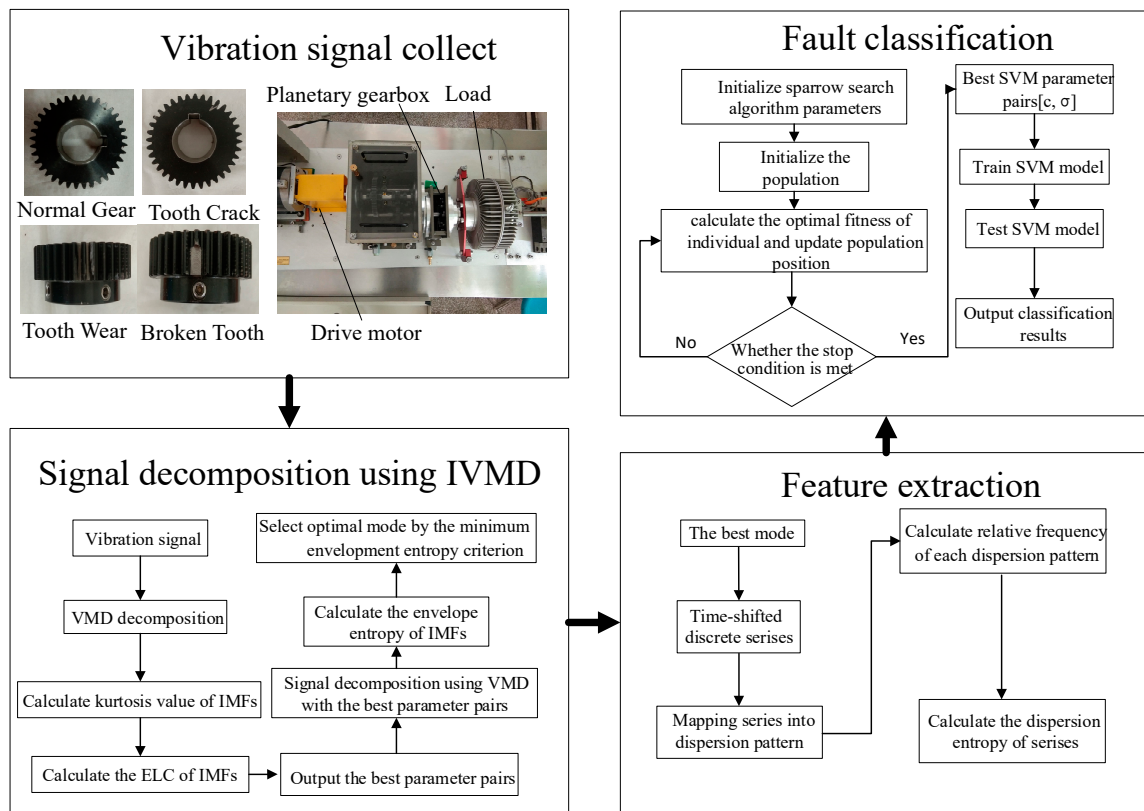

Figure 4. Flowchart of fault diagnosis.

6. Experimental Validation of the Proposed Model

In this study, the integrated wind turbine power transmission fault diagnosis platform was adopted for data acquisition and the experimental platform is described in Figure 4. We used the three-axis acceleration sensor produced by the American PCB company to collect the gear signals, set the motor speed to 1000 rpm, the load to 0.8 hp, the sampling frequency to 20,480 Hz, and the sampling time to 4 s.

The vibration signals of the input shaft of the secondary parallel shaft gearbox under four conditions (i.e., normal gear (NG), tooth crack (TC), tooth wear (TW), and broken tooth (BT)) were collected separately. Forty sets of vibration data were collected for each condition and each set included 2048 sampling points. Additionally, the ratio of training samples to testing samples was 1:1. The time-domain waveform diagram of the original vibration signals under the four conditions is shown in Figure 5. To better display the

characteristics of the different signals, the waveform diagram only shows one set of data (2048 data points) for each work condition.
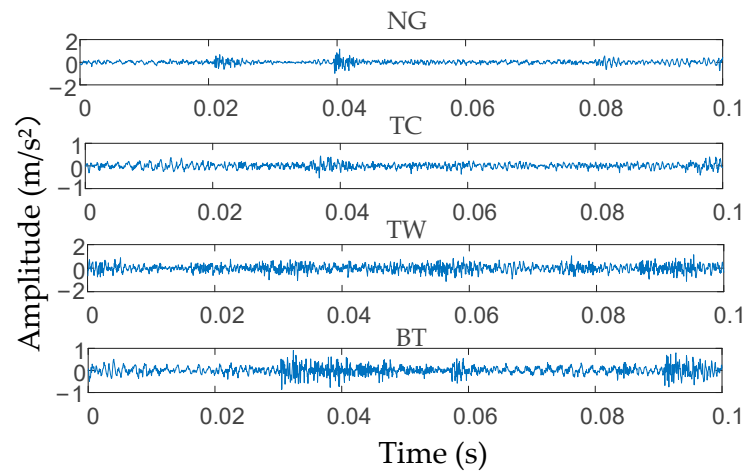

Figure 5. Time-domain waveform of the original signal.

6.1. Experimental Signal Decomposition Based on the IVMD Algorithm

The simulation signal obtained from [38] was used to validate the signal decomposition effectiveness of the proposed IVMD algorithm, and the simulation signal expression and its time-domain waveform are shown in Formula (14) and Figure 6.

$$
\begin{aligned}
s_1(t) &= \sin(20\pi t + \tfrac{100}{3}\pi), \\
s_2(t) &= \begin{cases}
0.6\sin(480\pi t), & t \in [0.2, 0.25], \\
0.5\sin(480\pi t), & t \in [0.4, 0.45], \\
0.4\sin(480\pi t), & t \in [0.6, 0.65], \\
0.3\sin(480\pi t), & t \in [0.8, 0.85],
\end{cases} \\
s_3(t) &= 0.65\sin(160\pi t), \\
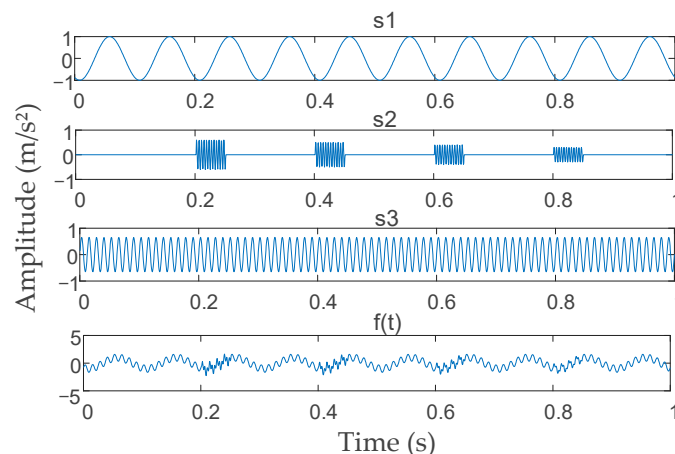f(t) &= s_1 + s_2 + s_3
\end{aligned}
\tag{14}
$$

Figure 6. Time-domain waveform of the simulation signal.

The search range and step of K and α were preliminarily set to $K \in [2, 7]$ and $\alpha \in [1000, 10{,}000]$, respectively, with a search step $\Delta\alpha = 500$. Selecting the best VMD parameters according to the IVMD algorithm process described above. The parameter optimization process of the IVMD algorithm for the simulation signal is shown in Figure 7.
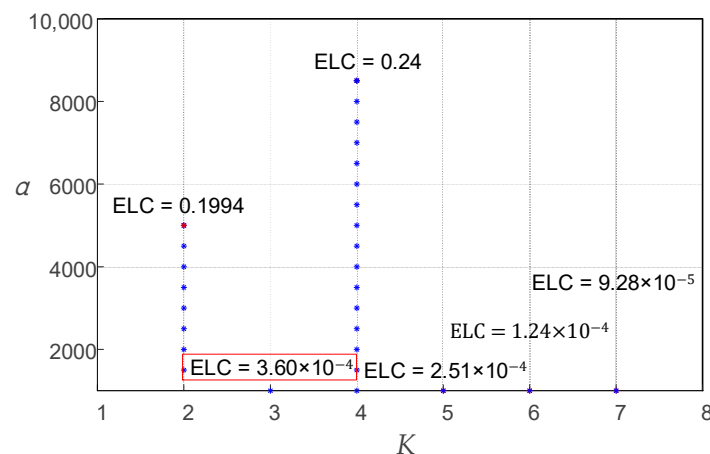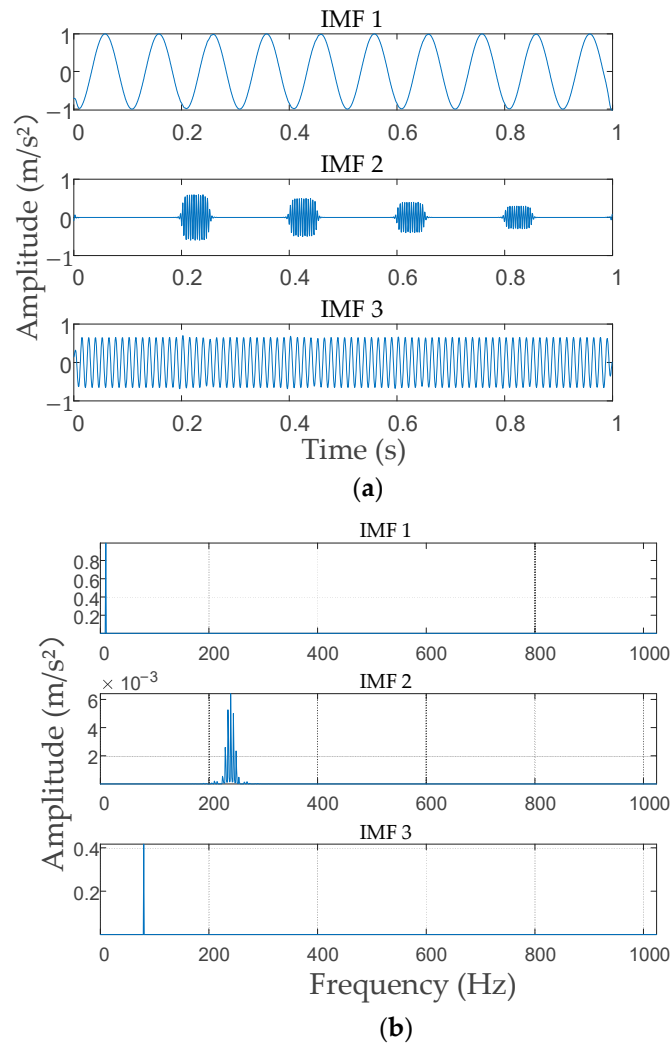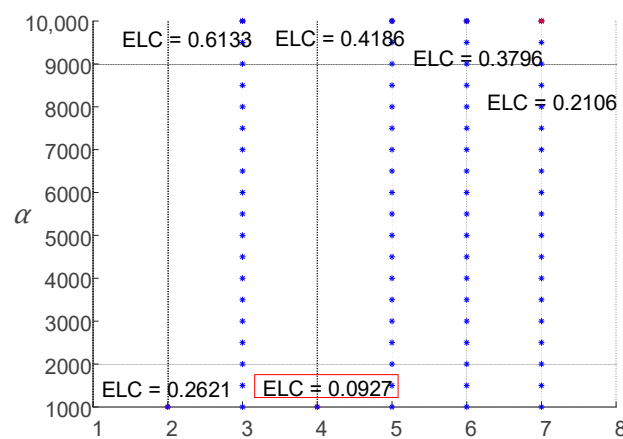
Figure 7. IVMD parameter optimization process.

From Figure 7, under the local optimal parameter combinations, with the increase in *K*, the ELC value decreases at the beginning and then increases, and then decreases continually. According to the trend in Figure 7, it can be predicted that when *K* continues to increase, the ELC will also decrease. Therefore, after the mode decomposition number *K* was 5 (including 5), the simulation signal showed over-decomposition and mode mixing. Based on the result (3, 1000), which is marked in the red box and selected as the optimal VMD decomposition parameter of the simulation signal, the signal was decomposed by VMD with that parameter. The decomposition results are presented in Figure 8. Figure 8a shows that after the simulation signal was decomposed with the parameter (3, 1000), each mode component signal obtained corresponded to one component of the original signal and there was no over or under-decomposition, which indicated that the simulation signal was completely decomposed in the time domain. Figure 8b shows that the simulation signal from low frequency to high frequency was clearly divided by IVMD and the mode components clearly reflected the original signal in the frequency spectrum of each component of the center frequency. Additionally, the modes showed clear differences in the frequency domain, and no mode mixing or false mode occurred, which illustrated the simulation signal in the frequency domain to obtain the complete decomposition. As a result, the simulation signal proved that the proposed IVMD method can effectively determine the decomposition parameters (*K*, *α*) of VMD.

In accordance with the above simulation signal analysis results, we used the IVMD algorithm to decompose gear signals under four conditions. Taking the tooth crack fault as an example, the optimization process of IVMD decomposition parameters is shown in Figure 9, where (4, 1000) marked in the red box is taken as the optimal VMD decomposition parameter. The EMD and IVMD algorithms were used to decompose signals with a tooth crack fault, separately, to validate that the proposed IVMD algorithm provides benefits in signal decomposition; the two methods' decomposition results are shown in Figure 10. EMD decomposed the signal into 17 groups of IMFs; due to space limitations, only the first four IMFs are expressed. From Figure 10a, between IMF1 and IMF2, there is an obvious mode mixing in the intermediate frequency stage. The same situation also occurs between IMF3 and IMF4. In addition, there are no obvious differences between the different components in the frequency domain, which proves that EMD cannot clearly decompose the signal from low frequency to high frequency. Figure 10b shows that all the decomposed mode components by IVMD. Each IMF shows good physical significance and the information overlapping in the full frequency band was decomposed in different frequency bands by IVMD. Additionally, Figure 10c shows the iteration curve of the central frequencies of each IMF during the IVMD decomposition; there is no crossing of curves and each curve moves away from each other, which means that no mode mixing occurred

in the process of IVMD decomposition. The effectiveness and advantages of the proposed IVMD algorithm in signal decomposition are verified through the above analysis.

(a)

(b)

Figure 8. Decomposition results of simulation signal. (**a**) Decomposition result in the time domain; (**b**) decomposition result in the frequency domain.


Figure 9. Parameter optimization process of IVMD.
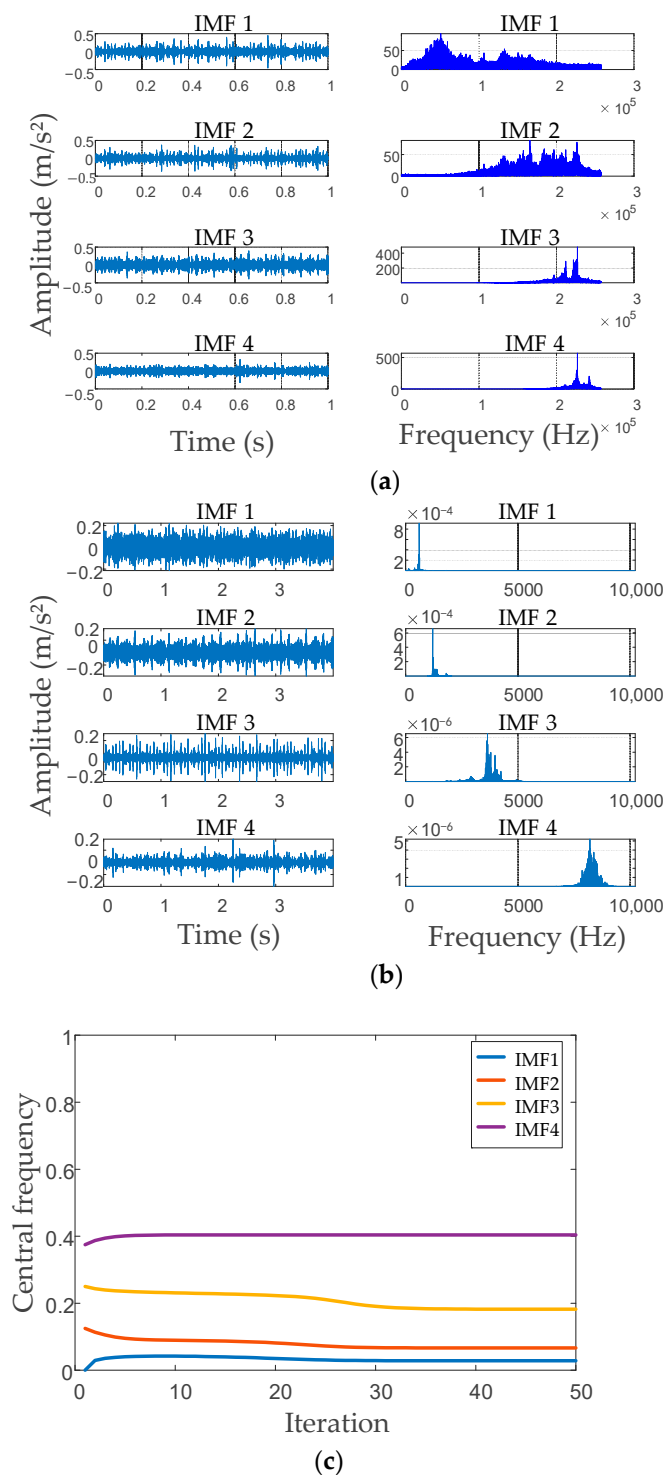
Figure 10. Decomposition results of a tooth crack fault. (**a**) First four decomposition results of EMD; (**b**) decomposition results of IVMD; (**c**) the iteration curve of the central frequencies.

To further verify the above findings, IMF3 was selected as the optimal IMF decomposed by IVMD and EMD according to the minimum envelope entropy criterion to perform envelope analysis, and the results are shown in Figure 11. From Figure 11a, the mode component fault characteristic frequency and the harmonics (2fr, 3fr, and 4fr) are clearly observed via IVMD decomposition. The results basically remove the noise interference compared with Figure 11b. Conversely, from Figure 11b, the mode component decomposed by EMD shows the mode-mixing problem. The signal fault characteristic frequency and

harmonics were submerged in noise, which resulted in difficulties in extracting the effective fault features. In summary, the above analysis proves that the IVMD algorithm proposed in this paper can effectively decompose signals.
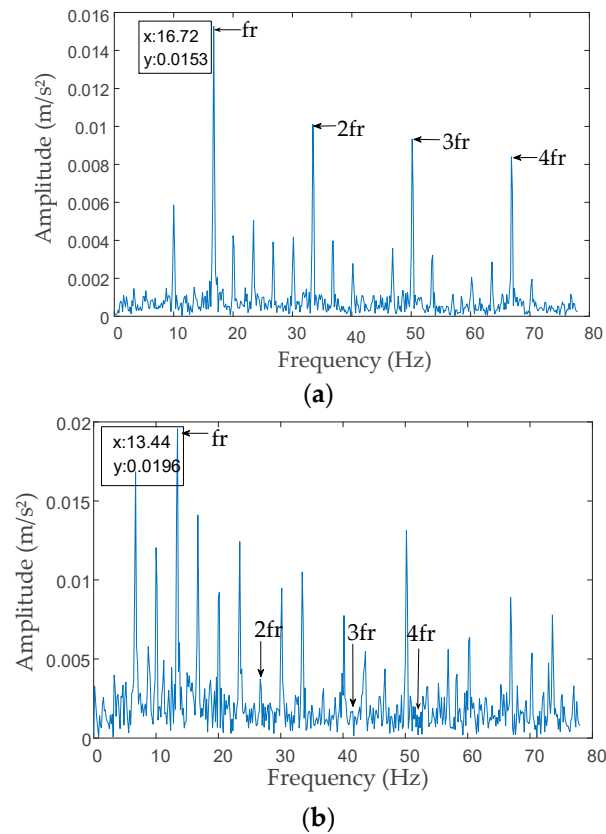
Figure 11. Envelope analysis results. (**a**) Envelope analysis results by IVMD; (**b**) envelope analysis results by EMD.

In a gearbox system, a local fault in the gear causes periodic vibration of the signal. The more obvious the periodicity of the fault signal, the smaller the envelope entropy value. Therefore, the optimal mode components in each work stage were selected according to the minimum envelope entropy value criterion for subsequent feature vector extraction, and time-domain waveform of the optimal IMFs is depicted in Figure 12.
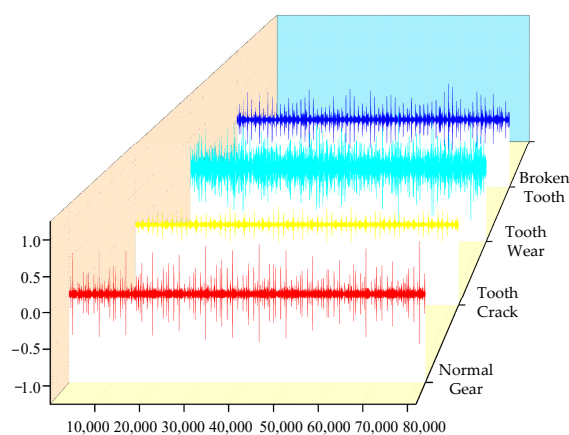

Figure 12. Time-domain waveform of the optimal IMFs.

6.2. Experimental Signal Feature Extraction Based on the TSMSE Algorithm

In this experiment, the TSMSE algorithm was used to extract the features of the optimal mode components selected in Section 6.1. The TSMSE and MSE algorithms were separately applied to extract the features of the signals of the four conditions to prove that the proposed TSMSE can effectively distinguish the gear signals in various conditions. The results are shown in Figure 13, where, compared with the TSMSE algorithm, the MSE algorithm cannot effectively distinguish the different conditions of gear. In Figure 13a, the feature entropy extracted by MSE shows serious mixing. Specifically, the normal gear and broken tooth signals in the first four scales can be effectively distinguished. With the increase in the scale, the two entropy curves nearly overlap. Tooth crack and tooth wear signals only overlap on the 6th, 11th, and 16th scale; on the other scales, the two signals can be identified. Considering the whole scale, the four features' entropy curves seriously cross, which means the four signals' conditions cannot be clearly distinguished. It also can be seen that MSE is unable to measure the complexity of the signal on higher scales due to the influence of coarse graining processing. The feature entropy obtained by the TSMSE algorithm shows a good degree of differentiation for each signal (Figure 13b), and there is almost no signal point crossing in the curves, which proves that the TSMSE algorithm can effectively distinguish gearbox signals under different conditions.

(**a**)

(**b**)
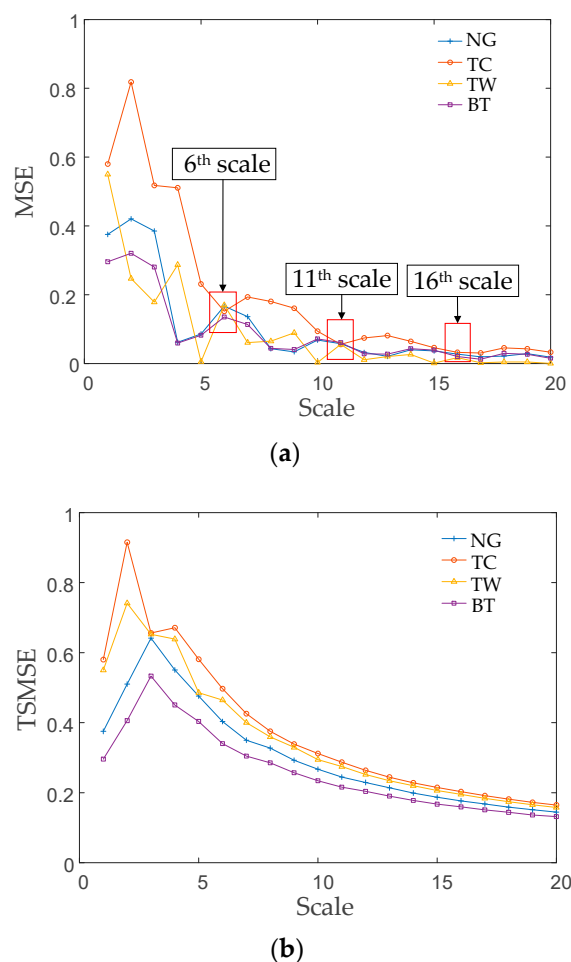
Figure 13. Feature entropy of signals of the four gearbox conditions. (**a**) MSE results of the signals of the four conditions; (**b**) TSMSE results of the signals of the four conditions.

6.3. Experimental Signal Fault Classification Based on SSA-SVM Algorithm

Based on the above experimental signal analysis results, the SSA was applied to perform adaptive optimization of the parameters (c, σ) of the SVM. The parameters of the algorithm are shown in Table 1.

Table 1. SSA parameter settings.

Parameters	Setting
Population size	50
iter_{max}	100
Percentage of producer	0.2
Range of parameter c	[1, 100]
Range of parameter σ	[1, 100]
ST	0.8

The feature entropy vectors from Section 6.1 were randomly and equally divided into a training sample and test sample, and the samples were used for training and testing of the SSA-SVM classification model. The optimal SVM parameters obtained from the search were [91.28, 1.78]. Figure 14 shows that the fault diagnosis model proposed in this paper has high identification accuracy (100%) and diagnosis capability of the vibration signals of wind turbine gearboxes.

The above analysis shows that the fault diagnosis model of mechanical equipment usually contains three parts, signal decomposition, fault feature extraction, and fault pattern identification. We optimized the above three parts separately and proved that the proposed optimization algorithm of each part (IVMD, TSMSE, SSA-SVM) is superior to the original algorithm (VMD, MSE, SVM) by comparative analysis. Furthermore, to prove that the final combination algorithm (IVMD-TSMSE-SSA-SVM) provides advantages in fault diagnosis of wind turbine gearbox, the proposed optimized combination algorithm model was compared with different combination algorithm models, and the results are shown in Table 2. The default VMD decomposition parameter of VMD was (3, 2000), and each method was repeated 5 times. From the table, three conclusions can be drawn:

(1) The average testing accuracy of the classification model using the IVMD algorithm (96.15%) is higher than that using the default VMD parameter (93.28%). Consequently, the above result proves that the proposed IVMD algorithm performs better than the default parameter VMD algorithm in signal decomposition.

(2) The average testing accuracy of 98.45% can be achieved by the classification model combined with the TSMSE feature vector, which outperforms the model with the feature vector selected by MSE (90.95%), as the TSMSE reflects the complexity and chaos of the signal on different scales and it is influenced by the length of the time series that is weakly compared with the MSE algorithm due to the time-shift technology. The above discussion verifies that the TSMSE algorithm can extract information with more abundant fault features and is superior in fault feature extraction.

(3) The SSA-SVM model has higher fault identification accuracy compared with the original SVM model; the testing accuracies of the two models are 97.05% and 92.35%, respectively, which indicate that the SSA can effectively determine the best SVM parameters. That is, the SSA-SVM model can effectively identify gear vibration signals under different conditions.

The above analysis shows that the performance of the proposed model in wind turbine gearbox fault diagnosis is more accurate than that of other methods. Therefore, this model shows great application potential in fault diagnosis.
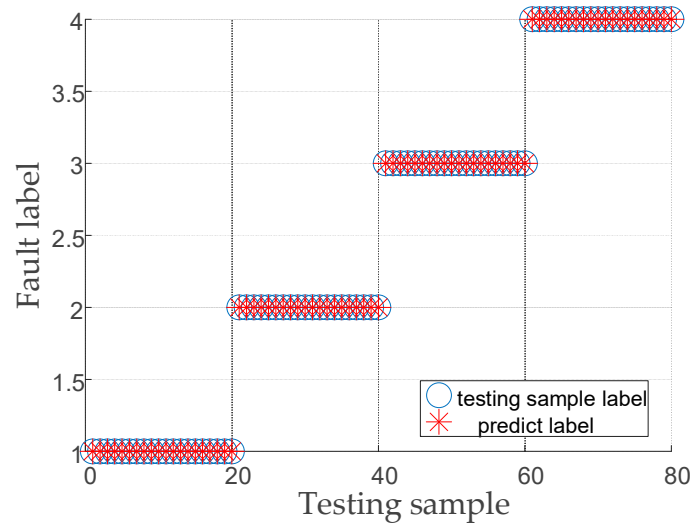
Figure 14. SSA-SVM classification results.

Table 2. Fault diagnosis accuracy of different combination algorithm models.

Signal Decomposition Algorithm	Fault Feature Extraction Algorithm	Classification Algorithm	Accuracy
IVMD	TSMSE	SSA-SVM	100%
IVMD	MSE	SSA-SVM	95.8%
VMD	TSMSE	SSA-SVM	98%
VMD	MSE	SSA-SVM	94.4%
IVMD	TSMSE	SVM	98.2%
IVMD	MSE	SVM	90.6%
VMD	TSMSE	SVM	97.6%
VMD	MSE	SVM	83.01%

7. Conclusions

This paper proposed a new wind turbine gearbox fault diagnosis model to accurately extract fault features and diagnose fault states. The effectiveness of this model in fault diagnosis was verified through experimentation; however, the proposed model has some limitations, which are listed as follows:

(1) At present, the rotating machinery equipment is developing toward being large-scale and intelligent, so the structure of the equipment is becoming increasingly complex, and the collected signals contain increasingly abundant information, which poses difficulties in effectively extracting the fault features of a signal. Although the IVMD algorithm effectively determines the parameters of the VMD adaptively, the time cost increases when it deals with complex signals. Therefore, improving the IVMD algorithm to reduce the time cost is a direction for future research.

(2) In the proposed method, SSA is used to optimize the parameters of the SVM model. Although SSA has some advantages compared with other algorithms, the parameters (percentage of producers and safety threshold) in the SSA are set artificially, which relies on the users' prior experience. This may lead to poor optimization results due to improper parameter selection. Therefore, future research should determine how to avoid the influence of human factors on the optimization effect of SSA.

Author Contributions: Conceptualization, F.Z., W.S., H.W. and T.X.; formal analysis, F.Z.; funding acquisition, W.S.; software, F.Z.; writing—original draft, F.Z.; writing—review and editing, W.S., H.W. and T.X. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study are all owned by the research groupand will not be transmitted.

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

References

1. Teng, W.; Ding, X.; Tang, S.; Xu, J.; Shi, B.; Liu, Y. Vibration Analysis for Fault Detection of Wind Turbine Drivetrains—A Comprehensive Investigation. *Sensors* **2021**, *21*, 1686. [CrossRef] [PubMed]
2. Noshirvani, G.; Askari, J.; Fekih, A. A robust fault detection and isolation filter for the pitch system of a variable speed wind turbine. *Int. Trans. Electr. Energy Syst.* **2018**, *28*, e2625. [CrossRef]
3. Casau, P.; Rosa, P.; Tabatabaeipour, S.M.; Silvestre, C.; Stoustrup, J. A Set-Valued Approach to FDI and FTC of Wind Turbines. *IEEE Trans. Control. Syst. Technol.* **2014**, *23*, 245–263. [CrossRef]
4. Odgaard, P.F.; Stoustrup, J.; Kinnaert, M. Fault-Tolerant Control of Wind Turbines: A Benchmark Model. *IEEE Trans. Control. Syst. Technol.* **2013**, *21*, 1168–1182. [CrossRef]
5. Badihi, H.; Zhang, Y.; Hong, H. Wind Turbine Fault Diagnosis and Fault-Tolerant Torque Load Control Against Actuator Faults. *IEEE Trans. Control. Syst. Technol.* **2015**, *23*, 1351–1372. [CrossRef]
6. Ma, Z.; Zhao, M.; Li, B.; Fan, H. A Novel Blind Deconvolution Based on Sparse Subspace Recoding for Condition Monitoring of Wind Turbine Gearbox. *Renew. Energy* **2021**, *170*, 141–162. [CrossRef]
7. Inturi, V.; Shreyas, N.; Chetti, K.; Sabareesh, G.R. Comprehensive fault diagnostics of wind turbine gearbox through adaptive condition monitoring scheme. *Appl. Acoust.* **2021**, *174*, 107738. [CrossRef]
8. Ziani, R.; Hammami, A.; Chaari, F.; Felkaoui, A.; Haddar, M. Gear fault diagnosis under non-stationary operating mode based on EMD, TKEO, and Shock Detector. *Comptes Rendus Mécanique* **2019**, *347*, 663–675. [CrossRef]
9. Minhas, A.S.; Singh, G.; Singh, J.; Kankar, P.; Singh, S. A novel method to classify bearing faults by integrating standard deviation to refined composite multi-scale fuzzy entropy. *Measurement* **2020**, *154*, 107441. [CrossRef]
10. Cheng, Y.; Wang, Z.; Chen, B.; Zhang, W.; Huang, G. An improved complementary ensemble empirical mode decomposition with adaptive noise and its application to rolling element bearing fault diagnosis. *ISA Trans.* **2019**, *91*, 218–234. [CrossRef]
11. Dragomiretskiy, K.; Zosso, D. Variational Mode Decomposition. *IEEE Trans. Signal Process.* **2014**, *62*, 531–544. [CrossRef]
12. Zhang, J.; Zhang, J.; Zhong, M.; Zheng, J.; Yao, L. A GOA-MSVM based strategy to achieve high fault identification accuracy for rotating machinery under different load conditions. *Measurement* **2020**, *163*, 108067. [CrossRef]
13. Pan, L.; Zhao, L.; Song, A.; She, S.; Wang, S. Research on gear fault diagnosis based on feature fusion optimization and improved two hidden layer extreme learning machine. *Measurement* **2021**, *177*, 109317. [CrossRef]
14. Dibaj, A.; Ettefagh, M.M.; Hassannejad, R.; Ehghaghi, M.B. Fine-tuned variational mode decomposition for fault diagnosis of rotary machinery. *Struct. Health Monit.* **2019**, *19*, 1453–1470. [CrossRef]
15. Fu, W.; Wang, K.; Li, C.; Tan, J. Multi-step short-term wind speed forecasting approach based on multi-scale dominant ingredient chaotic analysis, improved hybrid GWO-SCA optimization and ELM. *Energy Convers. Manag.* **2019**, *187*, 356–377. [CrossRef]
16. Shao, K.; Fu, W.; Tan, J.; Wang, K. Coordinated approach fusing time-shift multiscale dispersion entropy and vibrational Harris hawks optimization-based SVM for fault diagnosis of rolling bearing. *Measurement* **2021**, *173*, 108580. [CrossRef]
17. Kumar, A.; Zhou, Y.; Xiang, J. Optimization of VMD using kernel-based mutual information for the extraction of weak features to detect bearing defects. *Measurement* **2021**, *168*, 108402. [CrossRef]
18. Yan, X.; Jia, M. Application of CSA-VMD and optimal scale morphological slice bispectrum in enhancing outer race fault detection of rolling element bearings. *Mech. Syst. Signal Process.* **2019**, *122*, 56–86. [CrossRef]
19. Zheng, J.; Pan, H.; Liu, Q.; Ding, K. Refined time-shift multiscale normalised dispersion entropy and its application to fault diagnosis of rolling bearing. *Phys. A Stat. Mech. Appl.* **2020**, *545*, 123641. [CrossRef]
20. Yan, R.; Gao, R.X. Approximate Entropy as a diagnostic tool for machine health monitoring. *Mech. Syst. Signal Process.* **2007**, *21*, 824–839. [CrossRef]
21. Sharma, V. Gear fault detection based on instantaneous frequency estimation using variational mode decomposition and permutation entropy under real speed scenarios. *Wind Energy* **2021**, *24*, 246–259. [CrossRef]
22. Ding, J.; Xiao, D.; Huang, L.; Li, X. Gear Fault Diagnosis Based on VMD Sample Entropy and Discrete Hopfield Neural Network. *Math. Probl. Eng.* **2020**, *2020*, 8882653. [CrossRef]
23. Zhou, Q.; Zhang, J.; Zhou, T.; Qiu, Y.; Jia, H.; Lin, J. A parameter-adaptive variational mode decomposition approach based on weighted fuzzy-distribution entropy for noise source separation. *Meas. Sci. Technol.* **2020**, *31*, 125004. [CrossRef]
24. Yasir, M.N.; Koh, B.H. Data Decomposition Techniques with Multi-Scale Permutation Entropy Calculations for Bearing Fault Diagnosis. *Sensors* **2018**, *18*, 1278. [CrossRef]
25. Zhao, D.; Liu, S.; Cheng, S.; Sun, X.; Wang, L.; Wei, Y.; Zhang, H. Parallel multi-scale entropy and it's application in rolling bearing fault diagnosis. *Measurement* **2021**, *168*, 108333. [CrossRef]
26. Li, Y.X.; Jiao, S.B.; Gao, X. A novel signal feature extraction technology based on empirical wavelet transform and reverse dispersion entropy. *Def. Technol.* **2020**. [CrossRef]

27.　Xu, J.; Xu, P.; Wei, Z.; Ding, X.; Shi, L. DC-NNMN: Across Components Fault Diagnosis Based on Deep Few-Shot Learning. *Shock Vib.* **2020**, *2020*, 3152174.

28.　Pang, Y.; Jia, L.; Zhang, X.; Liu, Z.; Li, D. Design and implementation of automatic fault diagnosis system for wind turbine. *Comput. Electr. Eng.* **2020**, *87*, 106754. [CrossRef]

29.　Liang, P.; Deng, C.; Wu, J.; Yang, Z.; Zhu, J.; Zhang, Z. Compound Fault Diagnosis of Gearboxes via Multi-label Convolutional Neural Network and Wavelet Transform. *Comput. Ind.* **2019**, *113*, 103132. [CrossRef]

30.　Fei, S.-W. The Hybrid Method of VMD-PSR-SVD and Improved Binary PSO-KNN for Fault Diagnosis of Bearing. *Shock. Vib.* **2019**, *2019*, 4954920. [CrossRef]

31.　Yan, X.; Jia, M. A novel optimized SVM classification algorithm with multi-domain feature and its application to fault diagnosis of rolling bearing. *Neurocomputing* **2018**, *313*, 47–64. [CrossRef]

32.　Li, C.; Lin, S.; Xu, F.; Liu, D.; Liu, J. Short-term wind power prediction based on data mining technology and improved support vector machine method: A case study in Northwest China. *J. Clean. Prod.* **2018**, *205*, 909–922. [CrossRef]

33.　Luo, K.; Jiao, Y. Automatic fault detection of sensors in leather cutting control system under GWO-SVM algorithm. *PLoS ONE* **2021**, *16*, e0248515. [CrossRef] [PubMed]

34.　Xue, J.; Shen, B. A novel swarm intelligence optimization approach: Sparrow search algorithm. *Syst. Sci. Control Eng.* **2020**, *8*, 22–34. [CrossRef]

35.　Wang, Z.; Yao, L.; Chen, G.; Ding, J. Modified multiscale weighted permutation entropy and optimized support vector machine method for rolling bearing fault diagnosis with complex signals. *ISA Trans.* **2021**, *114*, 470–484. [CrossRef] [PubMed]

36.　Xie, X.; Chen, W.; Chen, B.; Cheng, J.; Tan, L. Comprehensive fatigue estimation and fault diagnosis based on Refined Generalized Multi-Scale Entropy method of centrifugal fan blades. *Measurement* **2020**, *166*, 108224. [CrossRef]

37.　Kou, Z.; Yang, F.; Wu, J.; Li, T. Application of ICEEMDAN Energy Entropy and AFSA-SVM for Fault Diagnosis of Hoist Sheave Bearing. *Entropy* **2020**, *22*, 1347. [CrossRef]

38.　Xu, Z.; Yue, M.; Li, C. Application of the proposed optimized recursive variational mode decomposition in nonlinear decomposition. *Acta Phys. Sin.* **2019**, *68*, 292–305.