



## Article

# Fault Diagnosis of Mine Ventilator Bearing Based on Improved Variational Mode Decomposition and Density Peak Clustering

Xi Zhang <sup>1</sup>, Hongju Wang <sup>1,\*</sup>, Xuehui Li <sup>2</sup>, Shoujun Gao <sup>1</sup>, Kui Guo <sup>1</sup> and Yingle Wei <sup>1</sup>

<sup>1</sup> School of Mechanical Electronic & Information Engineering, China University of Mining & Technology, Beijing 100083, China

<sup>2</sup> China North Vehicle Research Institute, Beijing 100083, China

\* Correspondence: bqt1900401005@student.cumtb.edu.cn

**Abstract:** The mine ventilator plays a role in protecting the life safety of underground workers, which is very significant to the production and development of coal mines. In total, 70% of ventilator failures are mechanical failures, and bearing failures are the most likely to occur in mechanical failures, which are also difficult to find. In order to identify fan bearing faults accurately, this paper proposes a fault diagnosis method based on improved variational mode decomposition and density peak clustering. First, the variational mode decomposition's modal number  $K$  and secondary penalty factor  $\alpha$  are chosen employing the improved sparrow optimization process. The bearing vibration signal is decomposed by the variational mode decomposition algorithm with optimized parameters. To create the characteristic vector, the multi-scale permutation entropy of the fourth order intrinsic mode function is determined. Then, the characteristic matrix is dimensionally reduced by kernel principal component analysis, and the two-dimensional matrix after dimensionality reduction is divided by density peak clustering method to find the clustering center of the training sample features. Lastly, the membership degree is assessed using the normalized clustering distance between the characteristic matrix of the test sample and the cluster center of the training sample. The accuracy of bearing fault identification on the self-constructed experimental platform can reach 100%, which verifies the effectiveness and potential of the proposed method.



**Citation:** Zhang, X.; Wang, H.; Li, X.; Gao, S.; Guo, K.; Wei, Y. Fault Diagnosis of Mine Ventilator Bearing Based on Improved Variational Mode Decomposition and Density Peak Clustering. *Machines* **2023**, *11*, 27.

<https://doi.org/10.3390/machines11010027>

Academic Editor: Davide Astolfi

Received: 10 November 2022

Revised: 18 December 2022

Accepted: 19 December 2022

Published: 26 December 2022



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** mine ventilator bearing; variational mode decomposition; multi-scale permutation entropy; density peak clustering

## 1. Introduction

As an important part of the mine ventilator, the running state of bearing directly affects its working performance. When the bearing fails, the damage point repeatedly collides with other parts in contact with it, resulting in impact vibration and non-stationary, nonlinear, and multifrequency signals [1]. Sudden failures, such as bearing looseness or damage can lead to uneven stress, increased resistance, or stall, which will cause misalignment, imbalance, surge, and other failures of the fan. The faults caused by bearings account for about half of the total faults of mine ventilators. Therefore, it is of great significance for the safety and reliability of the mine ventilator to be able to accurately identify bearing faults.

There are many factors that will adversely affect the motor bearing operation when the mine ventilator is running. When the air enters the impeller through the collector, if the air flow is uneven along the circumference, the axial thrust will be unbalanced. Long-term operation of bearings under alternating loads will reduce the service life of the bearings. When the motor is fixed in the air duct, the unbalanced rotation of the impeller will be caused due to the asymmetry of the support points and the distortion of the flow field at the inlet and outlet, so that the motor shaft will swing radially. The radial swing of the motor shaft not only has a large variable load on the bearing, but also makes the axis of the inner ring of the bearing incline to the axis of the outer ring, which is an important factor causing bearing fatigue failure.

The vibration analysis method shows the underlying characteristics of bearing defect; thus, it is frequently used to diagnose bearing faults. It is generally believed that vibration analysis methods mainly include three aspects: data preprocessing, extraction of fault feature, and fault mode classification [2]. Research in recent years has mostly concentrated on time–frequency analysis technology since the recorded vibration signals frequently exhibit nonlinear and nonstationary properties. The current time–frequency analysis technology is mainly divided into two categories. The first method does not need to set parameters before analyzing vibration signals. A typical example is local mean decomposition (LMD). LMD is an adaptive vibration analysis technology, which can decompose any complex signal into multiple product functions (PF) according to the inherent vibration in the vibration signal. Although the effectiveness of LMD in bearing fault diagnosis has been widely proved by many applications [3], it still has issues with endpoint effect and mode aliasing. The second needs to set some parameters, such as wavelet transform (WT). Although WT can decompose signals well, the setting of wavelet basis function and threshold requires a lot of prior experience, and different choices of wavelet basis function will have a significant impact on the final result. Wavelet transform lacks adaptive properties as a result.

Variational mode decomposition (VMD), a technique for identifying the frequency center and bandwidth of a variational model, was developed by Dragomiretshiy [4]. VMD has a firm mathematics theoretical base and can separate vibration signals reliably and effectively in contrast to LMD and WT. The VMD approach can adaptively split the vibration signal frequencies, but the outcomes of attenuation are still constrained by the choice of the modal number  $K$  and the secondary penalty parameter  $\alpha$  [5]. Researchers frequently integrate intelligent algorithms with the parameter optimization of VMD due to the growing applications of intelligent algorithms [6]. In order to boost  $K$ , Zijian Guo developed the cuckoo search algorithm [7]. Mengjiao Wang presented a sparrow search method that simultaneously optimizes  $K$  and  $\alpha$  [8]. Although using an intelligent algorithm to optimize the parameters of VMD takes a relatively long time, it has gained popularity in research because it considers how the coupling of the two components affects the decomposition result.

The following objective is to seek out how to extract the defect information from the acquired intrinsic mode function (IMF) weights once the vibration data signal has been broken down into a series of IMFs through VMD. Richman [9] made sample entropy (SE) his explicit suggestion. SE is of widespread significance since it is less susceptible to noise and data length. Bandt [10] proposed the use of permutation entropy (PE) to examine and evaluate the various mechanical systems. PE is straightforward and unaffected by noise since it evaluates complexity in terms of proximity that is comparatively near. PE and SE, on the other hand, only calculate complexity on a single scale, which has unfavorable effects when used to analyze data over various time periods. Due to this flaw, Costa [11] created a method employing a multi-scale sample entropy (MSE) approach for evaluating the complexity of unprocessed time series at several scales. MSE has the advantages of low requirements on sequence length and immunity to noise. However, it uses a step function to evaluate the similarity between two vectors, which is inconsistent with the actual signal characteristics. The multiscale permutation entropy (MPE) was developed by Aziz and Arif [12] to evaluate the complexity of time-series data. Additionally, the MPE's stability and toughness were confirmed. J. Zheng used MPE and SVM to find rolling bearing flaws, demonstrating MPE's advantage in feature extraction of these flaws [13]. MPE is thus chosen in this article as a particular tool for the density peak clustering (DPC) algorithm.

J. Tian [14] used spectral kurtosis and cross correlation to extract fault features that represent different faults, aiming to solve the difficulty of extracting fault features from vibration signals. It was found in the research that although the signal has been preprocessed, there is still some interference information around the spectral peak in the envelope spectrogram that may mislead the inspectors. If the inspectors set wrong parameters, it is very likely to cause errors in fault diagnosis, resulting in unnecessary consequences. This requires the inspectors to be careful and have a lot of prior experience, or use intelligent

methods to carry out fault diagnosis research on rolling bearings. In this paper, we focus on a lower-level but automated analysis based on machine learning (feature extraction and classification). In this case, industrial level acquisition systems and lower sampling frequencies are likely to be acceptable, and better for integration.

Machine learning techniques employ feature sets as training examples and test samples for pattern identification in defect detection. Dimension reduction, classification, regression, and clustering are the main topics in this section. K-means clustering was utilized by Yiakopolos to find rolling bearing defects in commercial settings. The fuzzy C-means (FCM) clustering technique was utilized by Hongwei Fan to separate signal feature sets with various levels of deterioration [15]. In order to decrease noise sensitivity and classification mistake, S.C. Shu suggested a novel tensor-based classifier that uses a fuzzy support tensor with marble loss [16]. A pattern recognition technique based on EEMD, PCA, and Gatt Geva (GG) clustering was proposed by Xiong Zhang [17]. Wu suggested a pattern recognition technique in which GG clustering organizes sample data while EEMD improves signal dimension and PE builds feature matrices [18]. The benefits of this approach are confirmed when compared to FCM clustering and Gustafson Kessel (GK) clustering.

Inspired by previous research, this paper describes a technique for generating bearing fault feature matrices using improved variational mode decomposition (IVMD) and MPE. Kernel principal component analysis (KPCA) is used to intuitively reduce the dimension of the feature matrices, and DPC clustering is then utilized to categorize the feature matrices. The generated feature matrices could have greater aggregation and higher distance between various categories due to the comprehensive theoretical IVMD model, the capability of signal subdivision, as well as the strong resilience of MPE indicators. Additionally, DPC has a clear segmentation border that may precisely split various data samples. The remaining chapters of this article are structured as follows: The algorithm basis, evaluation indicator, and technical processes are introduced in the second section. The effectiveness of this approach for bearing data of various fault kinds is examined in the third part. Concluding observations are included in the final part.

## 2. The Proposed Method

### 2.1. Improved Variational Mode Decomposition

As an increasingly widely used time–frequency analysis method, VMD has a good performance in signal processing. For the original signal  $y(t)$ , it can be decomposed into a series of IMFs  $u_k$  [19].

$$y(t) = \sum_k u_k \quad (1)$$

To minimize the summation of the bandwidth of each mode function, the procedure can be expressed as:

$$\min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[ (\delta(t) + \frac{j}{\pi t}) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \quad (2)$$

where  $\delta(t)$  is a pulse signal;  $j$  is an imaginary unit.

A sequence of decomposed IMFs is represented by  $\{u_k\} = \{u_1, \dots, u_k\}$ , and the central frequency corresponding to each IMF is represented by  $\{\omega_k\} = \{\omega_1, \dots, \omega_k\}$ . Secondary penalty factor  $\alpha$  and Lagrangian penalty factor  $L$  are included in Equation (2) to help discover the best answer. The expression is presented in Formula (3).

$$L(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_i \left\| \partial_t \left[ (\delta(t) + \frac{j}{\pi t}) * u_k(t) \right] e^{-j\omega_i t} \right\|_2^2 + \left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \langle \lambda(t), f(t) - \sum_k u_k(t) \rangle \quad (3)$$

where  $f(t)$  is the real part of the signal;  $\lambda(t)$  is Lagrange multiplier which changes with the number of iterations.

The decomposition level  $K$  and the secondary penalty factor  $\alpha$  impose a limit on the VMD algorithm's ability to decompose data. Under decomposition and false components

result from a  $K$  value that is too little, while over decomposition and false components result from a  $K$  value that is too big [20]. If  $\alpha$  is too big, each center frequency's bandwidth will be too tiny; if  $\alpha$  is too small, each center frequency's bandwidth will be too big. The ultimate accuracy of defect diagnosis will be impacted by the difficulty of feature extraction caused by improper parameter selection. Therefore, choosing appropriate parameters is essential to obtaining good decomposition results.

The predatory and anti-predatory behavior of sparrows in the natural world was the basis for the sparrow search algorithm (SSA). The Sparrow set matrix reads like this:

$$\begin{aligned} X &= [x_1, x_2 \cdots x_n]^T \\ X_i &= [x_{i,1}, x_{i,2} \cdots x_{i,d}] \end{aligned} \quad (4)$$

In Formula (4),  $n$  stands for the total number of sparrows,  $i$  equals "1, 2, ...,  $n$ ", and  $d$  refers to the number of dimensions.

The sparrow with better position in the population has priority in obtaining food and leads the whole population to the food source as the finder. The location of the finder is updated as follows:

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \cdot \exp\left(\frac{-i}{\alpha \cdot \text{iter}}\right) & R_2 < ST \\ X_{i,j}^t + Q \cdot L & R_2 \geq ST \end{cases} \quad (5)$$

where  $t$  represents the current number of iterations,  $j = (1, 2, \dots, d)$ ;  $X_{i,j}^t$  represents the position of the  $i$ th sparrow in the  $j$ th dimension; iter represents the maximum number of iterations [21];  $\alpha$  is a random number in the range of (0, 1);  $R_2$  is a random number ( $R_2 \in [0, 1]$ ); and  $ST$  is a constant number ( $ST \in [0.5, 1]$ ) which represent the danger value and security value, respectively;  $L$  stands for a 1D matrix, and each item in the matrix is 1.  $Q$  represents a random integer with a normal distribution of [0, 1]. All the others are followers except the finders. The formula for followers' location update is as follows:

$$X_{i,j}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{X_{\text{worst}}^t - X_{i,j}^t}{i^2}\right) & i > \frac{n}{2} \\ X_p^{t+1} + |X_{i,j}^t - X_p^{t+1}| \cdot A^+ \cdot L & i \leq \frac{n}{2} \end{cases} \quad (6)$$

where the overall worst position is represented by  $X_{\text{worst}}$ , while  $A$  represents  $1 \times D$ .  $A^+ = A^T (AA^T)^{-1}$  and 1 or  $-1$  are randomly allocated to each matrix element. When  $i > n/2$ , it indicates that the  $i$ th follower has a low fitness value [22], is not fed, and has a very low energy value. It must currently travel to other locations for food, i.e., energy intake.

ISSA is proposed to solve the engineering optimization problems of SSA, which is easy to fall into premature, resulting in low convergence accuracy and local convergence. Firstly, the algorithm employs Tent mapping to initialize the population and increase the homogeneity of the initial population. Chaos initialization has randomness, ergodicity, and initial value sensitivity, which can make the algorithm converge faster. The process of generating chaotic sequences based on Tent map is as follows:

$$T = \begin{cases} x(n+1) = \mu x(n), & 0 \leq x(n) \leq 0.5 \\ x(n+1) = \mu[1 - x(n)], & 0.5 < x(n) \leq 1 \end{cases} \quad (7)$$

where  $\mu$  is the mapping factor, which is set to 2 in this paper.

Secondly, in the basic SSA algorithm, as the number of iterations increases, each dimension of the individual sparrow becomes smaller, the search space gradually decreases, and the probability of falling into the local space increases. The discoverer location update approach incorporates the sine and cosine algorithm (SCA) to address this issue, and the nonlinear sine learning factor is also incorporated. In the early stages of the search, it is quite valuable and aids in worldwide exploration. It has a negligible value in the later

stages of the search, which helps to increase the precision and capability of local growth. The enhanced discoverer location formula and the learning factor formula are as follows:

$$\omega = \omega_{\min} + (\omega_{\max} - \omega_{\min}) \cdot \sin(t\pi / \text{iter}_{\max}) \quad (8)$$

$$X_{i,j}^{t+1} \begin{cases} (1 - \omega) \cdot X_{i,j}^t + \omega \cdot \sin(r_1) \cdot |r_2 \cdot X_{best} - X_{i,j}^t|, & R_2 < ST \\ (1 - \omega) \cdot X_{i,j}^t + \omega \cdot \cos(r_1) \cdot |r_2 \cdot X_{best} - X_{i,j}^t|, & R_2 \geq ST \end{cases} \quad (9)$$

In Formula (9),  $r_1$  is a random number in  $[0, 2\pi]$ , and  $r_2$  is a random number in  $[0, 2]$ .

Finally, in order to avoid the algorithm falling into local optimization, Lévy flight strategy is introduced into the follower update formula to improve the global search ability [23]. The improved formula is as follows:

$$X_{i,j}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{X_{worst}^t - X_{i,j}^t}{i^2}\right) & i > \frac{n}{2} \\ X_p^{t+1} + X_p^{t+1} \otimes \text{Levy}(d) & i \leq \frac{n}{2} \end{cases} \quad (10)$$

When using ISSA to determine the parameter setting of the VMD, it is necessary to select a proper fitness function. Since the low-frequency signal is the basic frequency of the component, while the high-frequency harmonic contains a lot of noise, this paper uses the mean envelope entropy (MEE) of the third to sixth intrinsic mode functions in the decomposition results as the fitness function.

$$\begin{aligned} Q_j &= a(j) / \sum_{j=1}^N a(j) \\ E_e &= -\sum_{j=1}^N Q_j \lg Q_j \\ \text{MEE} &= (E_{e1} + E_{e2} + E_{e3} + E_{e4}) / 4 \end{aligned} \quad (11)$$

where  $Q_j$  converts the envelope signal acquired following Hilbert's demodulation of the original  $Q_j$  signal into a series of probability distributions;  $a(j)$  is the envelope signal; and  $E_e$  is the envelope entropy value which has the ability to quantify the vibration signal's sparsity [24]. The fault impact and modulation phenomena caused by the fault will be buried in the signal when there are many noise components present, resulting in a weak signal sparsity and a higher value at this time [25]. The signal's sparsity is greater and the value at this point is lesser when the signal clearly exhibits fault impulse and modulation phenomena. The primary IMFs' sparsity may be thoroughly assessed using the mean envelope entropy. The VMD algorithm uses MME as the fitness function for parameter selection, and the optimization of the pertinent parameters may be finished by using its minimal value as the program's search objective.

## 2.2. Multiscale Permutation Entropy and Dimension Reduction

Referencing [26] illustrated the related ideas of multi-scale permutation entropy, and the following is a summary of its theory:

To establish a new time series, the existing time series of length  $n$  is coarsened:

$$y_j^{(s)} = \frac{1}{s} \sum_{i=(j-1)s+1}^{js} X_i, j = 1, 2, \dots, [N/s] \quad (12)$$

where  $s$  denotes the scale factor;  $N$  is the sequence length;  $[N/s]$  represents rounding. The following are the time reconstructions for each scale sequence:

$$Y_1^{(s)} = \{y_1^{(s)}, y_{1+\tau}^{(s)}, \dots, y_{1+(m-1)\tau}^{(s)}\} \quad (13)$$

where  $m$  is encapsulation size;  $\tau$  is time delay.

If the values of the reconstructed sequences are the same, they are ordered according to the footmark. For each scale sequence, a set of symbol sequences may be generated. The symbol sequence is one of the permutations, and the probability of each symbol sequence occurring is calculated ( $r = 1, 2, \dots, R$ ). The permutation entropy of various symbol sequences is defined in terms of information entropy.

$$H_p(m) = -\sum_{r=1}^R P_r \ln P_r \quad (14)$$

The highest value is attained when  $P_r = 1/m!$ . Normalization is generally accomplished for convenience.

$$H_p = H_p(m) / \ln(m!) \quad (15)$$

Prior to using MPE, four variables must be defined: encapsulation size ( $m$ ), series length ( $N$ ), scaling factor ( $s$ ), and delay ( $\tau$ ). Since  $m!$  is the maximum allowed number of permutations, it emerges that  $m$  is the crucial variable. PE is highly sensitive to the value of  $m$ , the encapsulation dimension. In addition, for robust statistics  $N$  should be greater than  $5 \times m$  factorial. Bandt remarked that this strategy works well if the encapsulation dimension is between three and seven. If the encapsulation dimension  $m$  is too low, there will not be enough distinct states for the method to work. However, it becomes impractical when the encapsulation dimension  $m$  is too high. The encapsulation dimension  $m$  is often selected with the trade-off between information content impairment and measurement complexity [27]. In this piece,  $m$  is 5. Since there is little effect of the delay on the result, we leave this  $\tau = 1$ . If  $N$  is too large, computation speed suffers. If  $N$  is too little, the condition  $N \geq 5 * m!$  cannot be satisfied. A data length of 2048 points is adequate to obtain a stable permutation entropy when this constraint is taken into consideration. Therefore, we put 2048 as  $N$ . The permutation entropy of each scale is calculated by setting the scaling factor  $s$  to 15.

Kernel principal component analysis (KPCA) is a simple and effective nonlinear principal component analysis method, whose nonlinear mapping is completed by kernel function. Initially, KPCA was developed by Schölkopf, and applied to fault diagnosis and process monitoring by W.U. [28]. The research on KPCA shows that it can effectively detect faults, but in KPCA-based monitoring, due to the difficulty or even impossibility of locating the inverse mapping function from the feature space to the original space [29], it is difficult to determine the contribution of the original process variables. In this article, we used a new contribution framework to provide the contribution of KPCA method. KPCA maps the input space to the feature space through nonlinear transformation and extracts the principal components in the feature space. Assuming a nonlinear mapping function  $\varphi(x)$  that converts the input space to the feature space, the covariance matrix in the feature space may be presented as:

$$\text{cov}(\Phi(\mathbf{X})) = \frac{1}{n-1} \sum_{j=1}^n \Phi(\mathbf{X}_j)^T \Phi(\mathbf{X}_j) \quad (16)$$

where  $x_j$  is the  $j$ th line of the  $\mathbf{X}$  vector. In order to complete the kernel principal component analysis, it is necessary to solve the eigenvalue problem in vector space.

$$\text{cov}(\Phi(\mathbf{X})) \mathbf{p}_i = \lambda_i \mathbf{p}_i \quad (17)$$

It is noticeable that the problem of Equation (17) is not easy to solve, because in most cases  $\varphi(x)$  is not available. However, there are coefficients  $\alpha_{ij}$  that can make the eigenvalue be approximately expressed as:

$$\mathbf{p}_i = \sum_{j=1}^n \alpha_{ij} \Phi(\mathbf{X}_j)^T \quad (18)$$

By combining Equations (16)–(18), the following equation can be obtained:

$$(n-1)\lambda_i \alpha_i = \mathbf{K} \alpha_i \quad (19)$$

Here,  $K$  is defined as  $[K]_{ij} = k(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$ . KPCA is applied to reduce the dimension of the extracted multi-scale permutation entropy feature matrix to two dimensions.

### 2.3. Density Peak Clustering

The clustering approach described in “Clustering by quick search and discovery of density peaks”, a work by Alex Rodriguez and Alessandro Laio [30] published in *Science* in June 2014, is both simple and sophisticated. Different shaped clusters may be easily identified using this approach, which is based on density and whose characteristics can be easily determined. It fixes the issue with generic density-based clustering algorithms, whereby the density of various classes varies substantially, and the neighborhood range is hard to determine.

The density peak clustering (DPC) method relies on the following two tenets: (1) for a given dataset, the cluster center is surrounded by several data points with low local density; and (2) these low local density points are relatively far away from other points with high local density.

Local density  $\rho_i$  is defined as:

$$\rho_i = \sum_j \chi(d_{ij} - d_c) \tag{20}$$

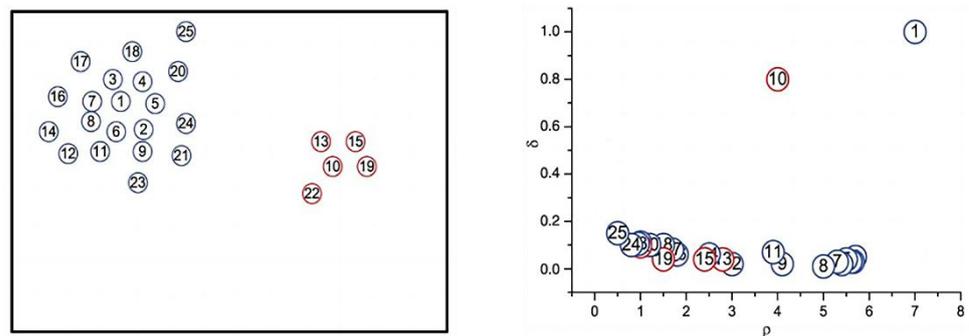
where  $\chi(x) = \begin{cases} 1, & x < 0 \\ 0, & x \geq 0 \end{cases}$  ;

$d_c$  is the cutoff distance.  $\rho_i$  means the number of objects whose distance to the  $i$ th object is less than  $d_c$ . The selection of  $d_c$  is critical [31], because the algorithm is sensitive to the relative value of  $\rho_i$ .

High local density distance  $\delta_i$  is defined as:

$$\delta_i = \min_{j: \rho_j > \rho_i} (d_{ij}) \tag{21}$$

The points with high local density  $\rho_i$  and high local density distance  $\delta_i$  are considered as the center of the cluster. The points with high local density distance  $\delta_i$  and low local density  $\rho_i$  are the abnormal points. After determining the cluster center, other points are classified according to the nearest distance from the known cluster center. As shown in Figure 1, point 1 and point 10 are selected as the cluster center. Point 8 and point 7 are excluded because their  $\rho$  values are large while  $\delta$  values are low.

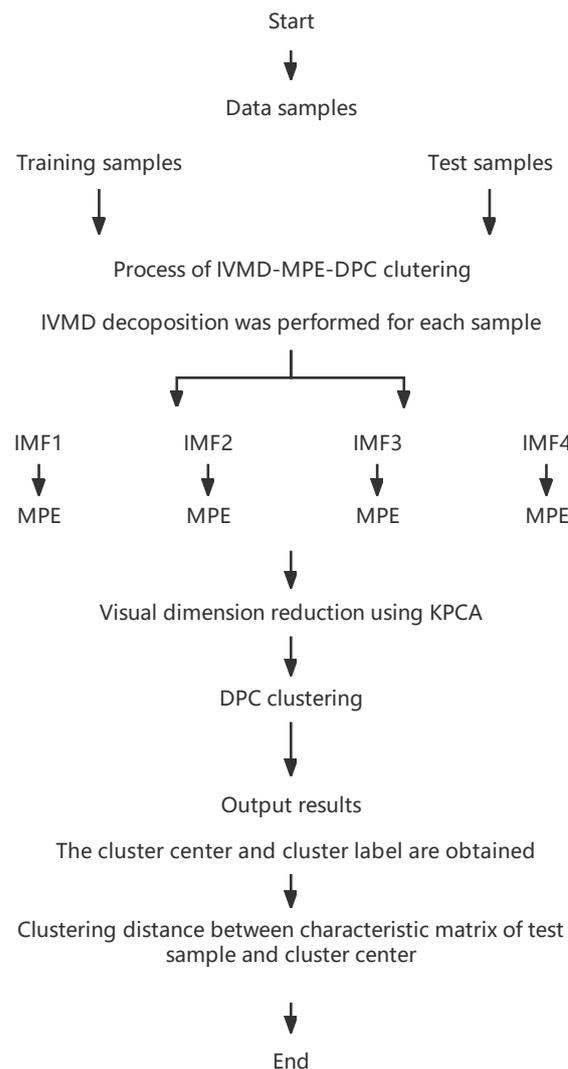


**Figure 1.** Determination of cluster center of density peak clustering.

The flowchart of the bearing fault diagnosis method proposed in this paper is shown in Figure 2. The specific process is described as follows:

- (1) Establish characteristic matrix. Select the original vibration signal to form time series  $x = (x_1, x_2, \dots, x_n)$ , and each time series  $x$  is processed with IVMD. The third to sixth intrinsic mode functions form a matrix. The characteristic matrix is constructed by computing the multi-scale permutation entropy of each IMF.

- (2) The characteristic matrix's dimension is decreased with the help of KPCA. Projecting the feature matrix onto a two-dimensional space, we then use the two primary components with the greatest contribution rate to build a feature matrix in the two-dimensional space.
- (3) Training evaluation model. The training set composed of principal components is input into the DPC classifier to obtain the cluster category and cluster center.
- (4) Use a test set to confirm. Replicate steps 1 and 2, then feed the trained DPC classifier the test set's main component matrix. The clustering distance between the test samples and the cluster center of the training set is used to categorize test samples.



**Figure 2.** Flowchart of IVMD-DPC.

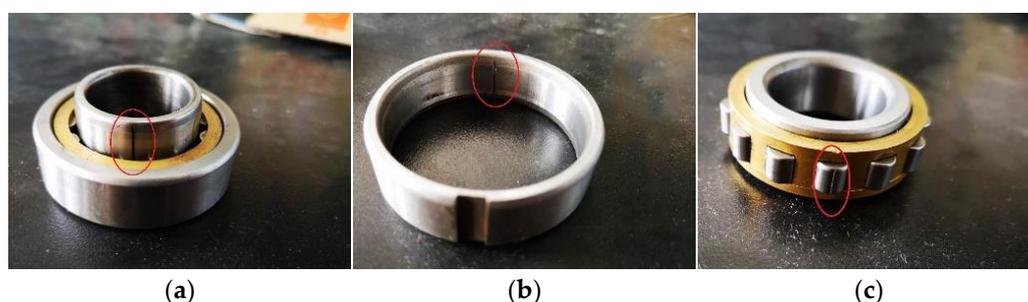
### 3. Experiment and Results

#### 3.1. Experimental Platform

The experimental data were collected from the bearing simulation failure part of the rotating machinery test bench in the laboratory of the China University of Mining and Technology (Beijing, China). The test bench is shown in Figure 3. Bearing fault types are inner race fault, outer race fault, and rolling element fault. The inner race, outer race and rolling element were wired by EDM 0.2 mm to simulate fatigue wear after long-time operation, as shown in Figure 4a–c.



**Figure 3.** Bearing Simulation Fault Platform.



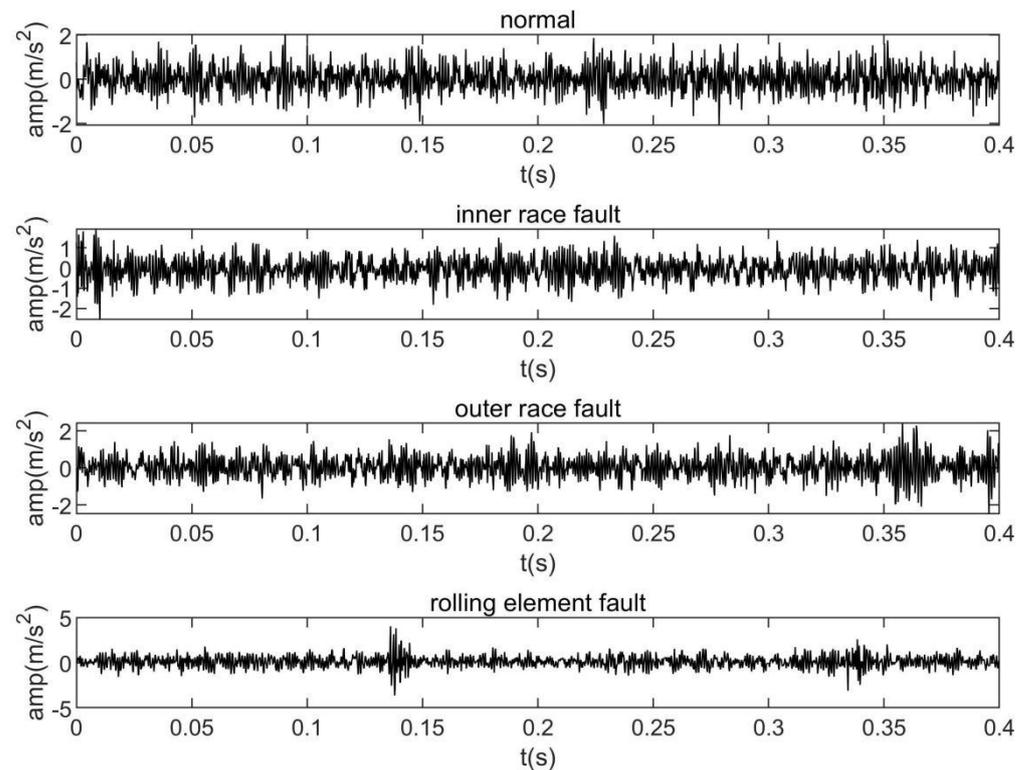
**Figure 4.** Different faults of rolling bearings: (a) rolling element (b) inner race (c) outer race.

First, the vibration acceleration signals of normal bearings were collected, and then the three kinds of faulty bearings were installed to the test bench for fault signal collection. The INV3062C acquisition instrument was used to connect to the computer through the adapter cable. INV3062C data acquisition instrument is an industrial high-precision distributed acquisition instrument, which is suitable for data acquisition of rotating machinery and modal testing. We opened the DASP digital signal acquisition software to refresh and connected the acquisition instrument equipment with the software. Then, we connected one end of the adapter wire to the acquisition instrument and the other end to the INV9822 acceleration sensor. The sensitivity of INV9822 acceleration sensor was 100 mv/g and the frequency range was 0.5 Hz~8 kHz. We installed the magnetic base of the acceleration sensor and confirmed that the connection was not loose. The acceleration sensor was attached to the upper part of the bearing bracket, that is, the red circle in Figure 3. The sampling frequency was set to 5.12 kHz according to Shannon sampling theorem. Meanwhile, 5.12 kHz was behind the middle of the maximum range of the acceleration sensor, which belongs to the proper working range of the instrument. We collected 40 groups of samples containing 2048 points. The sampling time was set to exactly 16 s. After multiple zero-point calibrations, sampling was performed and the collected vibration signal data were saved.

Table 1 shows that there are one group of normal data and three groups of fault data of bearing vibration acceleration signal data collected in the laboratory. The original waveform of signal data is shown in Figure 5. A total of 50 samples were collected for each type, of which 30 samples were used as training set and 20 samples as test set. The length of each sample was 2048.

**Table 1.** Signal Data Type Description.

Signal Type	Motor Load (HP)	Motor Speed (rpm)
NOR		
IF		
OF	0	1500
RF		

**Figure 5.** The time domain wave forms of original signals under four bearing states.

### 3.2. Results and Discussion

IVMD algorithm was used to decompose the vibration signal. Taking rolling element fault as an example, the parameters of VMD algorithm were optimized by using the improved sparrow algorithm. The minimal mean envelope entropy value change curve is shown in Figure 6 as the number of iterations in the ISSA optimization process increases. The graphic shows that the minimal mean envelope entropy is 3.7268 at the seventh iteration. Because the improved sparrow optimization algorithm introduces chaos initialization and sine cosine ideas, it has a smaller initial value and can jump out of the local optimal solution. The optimization procedure is finished and the optimization parameters  $[K, \alpha]$  are  $[11, 2257]$  when the number of iterations approaches 10. The ideal value  $K$  was used to reset VMD, and the optimized VMD was then used to generate 11 modal components.

The ideal values  $K$  and  $\alpha$  were found by using ISSA to optimize the VMD algorithm under various failure circumstances of rolling bearings. Set the VMD algorithm's settings using the recommendations in Table 2. The VMD method was used to dissect the vibration signals at various rolling bearing damage areas once the parameters were optimized.

In order to quantify the fault information contained in the IMFs, MPE values of the third to sixth intrinsic mode functions of vibration signals in different states after VMD were calculated. Figure 7 shows the MPE of the first IMF of different samples of rolling element fault state. It can be seen that the trend of multi-scale arrangement entropy value of different samples changing with the change of scale factor  $s$  is basically the same.

Under the condition of the same scale factor, the difference of arrangement entropy of different samples is very small. This shows that multi-scale permutation entropy has good robustness as a characteristic index of quantifying fault information and can stably extract fault information contained in vibration signals.

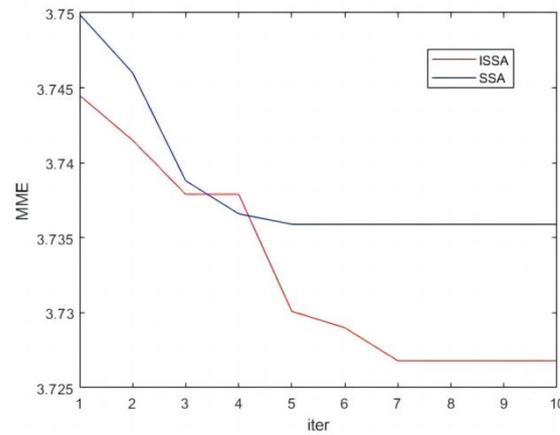


Figure 6. Minimum mean envelope entropy iteration curve.

Table 2. Optimization parameters obtained by using the ISSA.

	Normal	Inner race	Outer Race	Rolling Element
K	11	11	11	11
$\alpha$	2761	2822	3995	2257

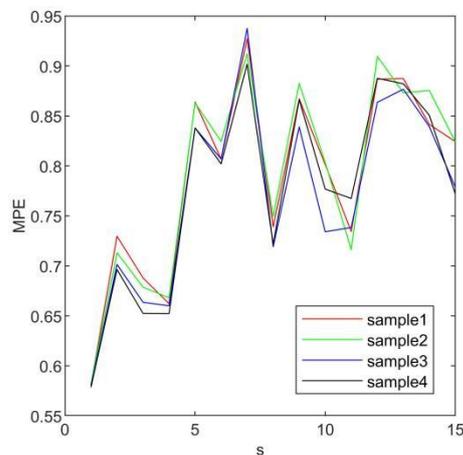
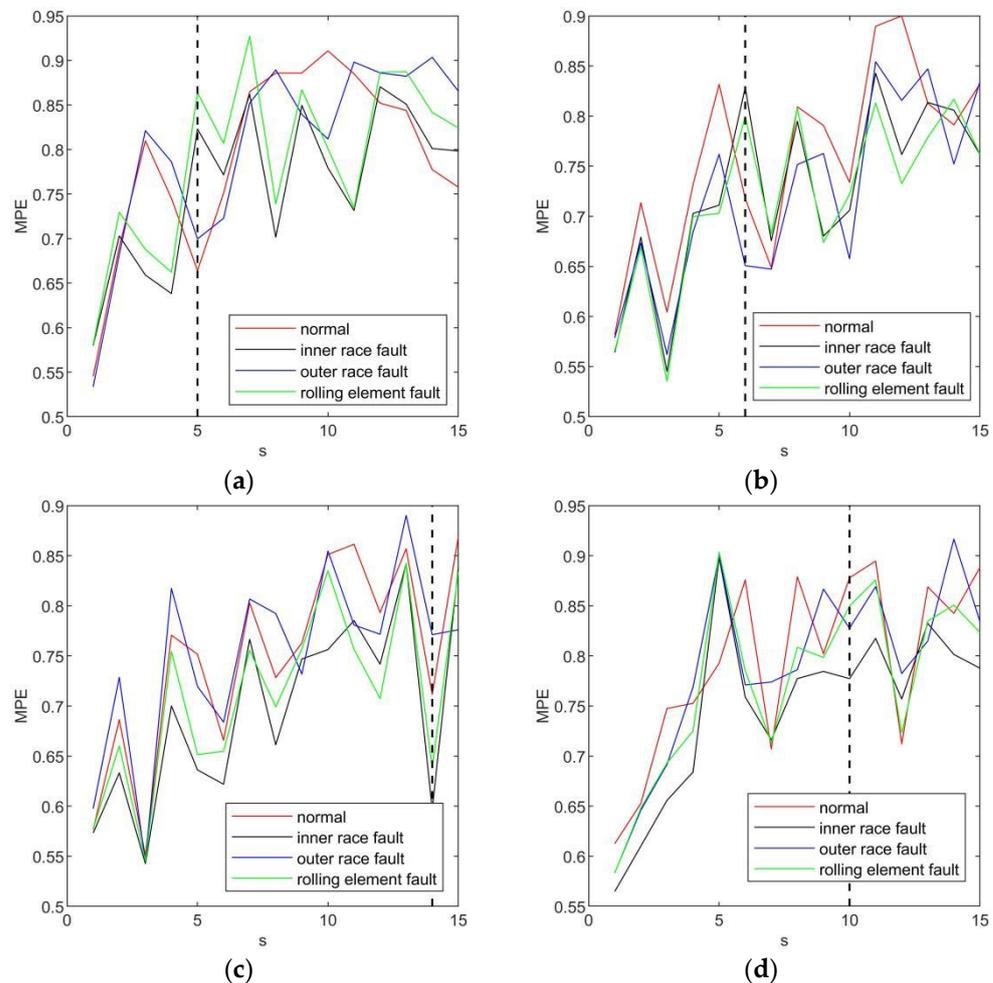


Figure 7. MPE of the first intrinsic mode function of different samples.

In order to facilitate the subsequent accurate judgment of the fault type of bearings, the multi-scale permutation entropy difference extracted from bearings in different states should be obvious enough. Figure 8a shows the multi-scale permutation entropy of different scale factors of the first intrinsic mode function. It can be seen from the figure that when  $s = 1$ , the permutation entropy values of the normal state and the inner race fault are very close, and the permutation entropy values of the outer race fault and the rolling element fault are also very close, which is easy to cause confusion in the later classification. Through many groups of experiments, we found that both the maximum difference and minimum difference of permutation entropy have an impact on the final classification results, and the minimum difference has a greater impact. Setting the weight of the minimum difference to 0.7 and the weight of the maximum difference to 0.3 is a reasonable weight distribution. In this way, we set the scale factor  $s$  to 5 for calculating the permutation

entropy of the first IMF. According to the same principle, the scale factors  $s$  of the second IMF, the third IMF and the fourth IMF were set to 6, 14, and 10, respectively. Then, the extracted four-dimensional permutation entropy matrix was reduced to two-dimensional by kernel principal component analysis.



**Figure 8.** MPE of four intrinsic mode functions: (a) imf 1, (b) imf 2, (c) imf 3, (d) imf 4.

The selection of cutoff distance  $d_c$  greatly affects the results of density peak clustering, while the current research on the selection of cutoff distance  $d_c$  still largely depends on human subjective experience. Generally, the optimal  $d_c$  value is 0.1–0.2 times that of the maximum distance between two points. The Davies–Bouldin index (DBI) was selected here as the index to evaluate the clustering results, which guided the selection of the truncation distance  $d_c$ .

According to Table 3, when the cutoff distance  $d_c$  is 0.07, the DBI reaches the minimum value of 0.2174, so the cutoff distance  $d_c$  was set to 0.07. The result of clustering the training set with peak density is shown in Figure 9.

**Table 3.** DBI of different cutoff distances.

$d_c$	0.05	0.06	0.07	0.08	0.09
DBI	0.2196	0.2176	0.2174	0.2189	0.2237

It can be seen from Figure 9 that the coordinate points representing normal state and rolling element fault have good aggregation. Although the aggregation of coordinate points

representing inner race fault and outer race fault is slightly poor, it does not affect the final result. The coordinates of the cluster center are shown in Table 4.

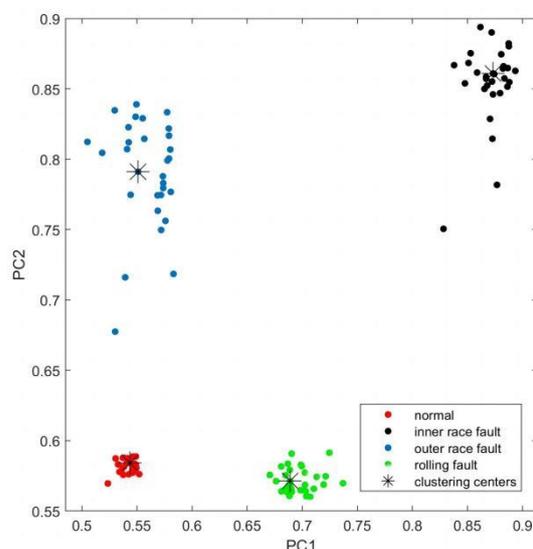


Figure 9. Result of density peak clustering.

Table 4. Coordinates of four cluster centers.

	Cluster Center 1	Cluster Center 2	Cluster Center 3	Cluster Center 4
PC1	0.5435	0.6889	0.5510	0.8732
PC2	0.5842	0.5713	0.7911	0.8609

Then, 80 groups of test set samples were processed in the same way to construct the principal component feature matrix. The cluster distance between the coordinate point representing the test set and the cluster center of the training set was calculated to determine the classification of the unknown state samples of the test set. The results are shown in Figure 10 (the abscissa 1–20 interval represents the clustering distance between each sample and cluster center 1, and the next represents the clustering distance between each sample and cluster center 2, 3, and 4 in turn). It can be seen that the distance between different test samples and the four cluster centers is obviously different. Therefore, the state types of the four test samples can be accurately determined. Figure 11 shows the confusion matrix results obtained by using IVMD-DPC method under four different fault states. All 80 groups of test samples were classified correctly, and the final recognition accuracy was 100%. The results show that the fault diagnosis method can accurately identify bearing faults in different states.

To verify the effectiveness of the proposed method, SSA-VMD was used for signal processing, and then density peak clustering was used for classification. The parameters of VMD were set according to the reference. The results are shown in Figure 12. The final classification accuracy of this method is 96.25%. It verifies that the ISSA-VMD proposed in this paper has excellent signal dissection capability. The ISSA-VMD was used for signal processing, and then fuzzy mean clustering (FCM) was used for clustering. The results are shown in Figure 13. The final classification accuracy of this method is 97.5%. It verifies that DPC has better clustering performance than FCM.

In order to improve the confidence in the results, the experiments were repeated ten times to take the average identification accuracy for comparison. Every time, the training set and test set were randomly divided. The results are shown in Table 5.

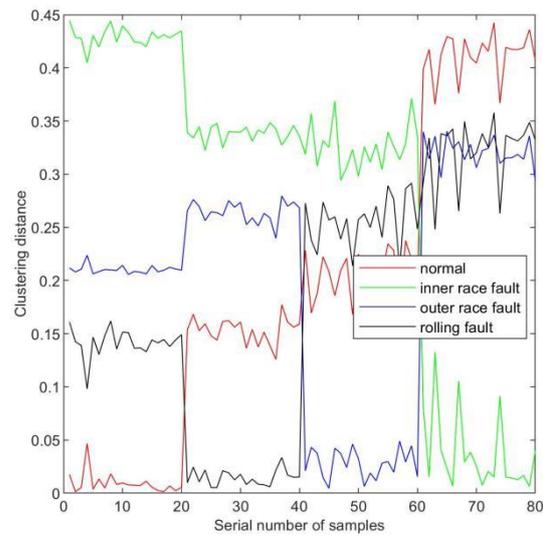


Figure 10. Distance between test set feature matrix and cluster center.

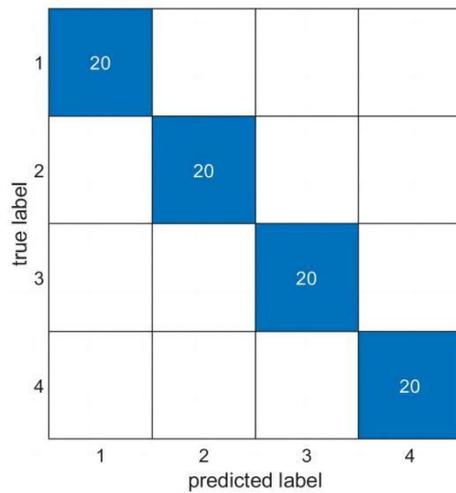


Figure 11. Confusion matrix of ISSA-VMD-DPC.

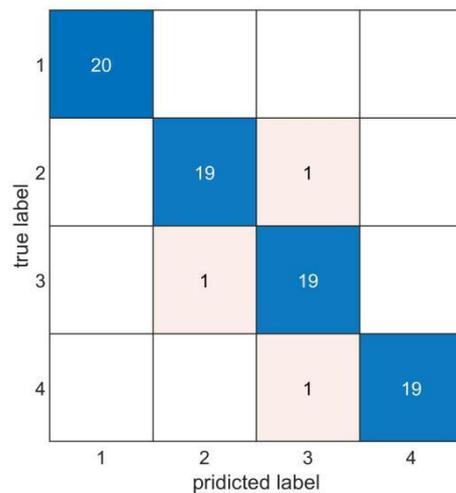
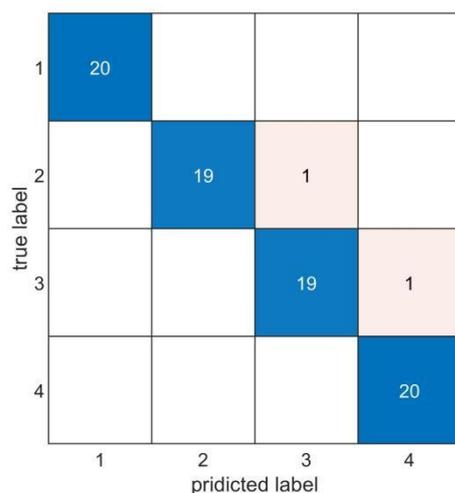


Figure 12. Confusion matrix of SSA-VMD-DPC.



**Figure 13.** Confusion matrix of ISSA-VMD-FCM.

**Table 5.** Average accuracy of different classifiers.

Classifier	DPC	SVM	ELM	KNN
Average accuracy	99.25%	97.75%	96.25%	97.37%

As shown in Table 5, we compared the average accuracy of the algorithm in this paper with other methods. The methods of feature extraction remain the same, and only different classifiers are applied at the end. The average accuracy of other methods cannot reach above 98%. Comprehensively analyzing the experimental results of different classifiers, the bearing fault diagnosis method based on the MPE feature extraction and DPC classifier proposed in this paper has achieved satisfactory results.

#### 4. Conclusions

This work proposes a technique combining ISSA-VMD, MPE, and DPC for ventilator bearing problem diagnostics. Data preprocessing, fault feature extraction, and fault feature identification were used to diagnose and analyze bearing faults, and the experimental verification was carried out on the self-built experimental platform.

For VMD parameter selection, the optimal VMD parameters were found using the improved sparrow search algorithm, and information on fault feature extraction was subsequently obtained using the optimized VMD approach. The results show that ISSA-VMD can effectively extract bearing fault information. The cutoff distance  $d_c$  in the DPC fault diagnosis model was set according to the DBI index. The findings demonstrate the DPC method's strong classification efficiency, and the diagnostic accuracy can reach 100%. It can be observed from the findings that this approach can accurately identify various forms of bearing damage and has better identification accuracy when compared to the outcomes of other methods.

In future work, we will be committed to collecting the on-site signal of the coal mine ventilator bearing to verify the feasibility of the proposed method in practical production.

**Author Contributions:** Conceptualization, X.Z. and H.W.; methodology, H.W.; software, H.W. and X.L.; validation, S.G. and K.G.; formal analysis, S.G.; investigation, Y.W.; resources, Y.W.; data curation, X.L.; writing—original draft preparation, H.W.; writing—review and editing, X.Z.; visualization, H.W.; supervision, X.L.; project administration, S.G.; funding acquisition, X.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by funds by the National Natural Science Foundation of China (Grant No. 52121003).

**Data Availability Statement:** Reasonable data use request will be approved, subject to restrictions (e.g., privacy or ethical restrictions).

**Acknowledgments:** The authors would like to thank the editor and the anonymous reviewers for their constructive comments and suggestions.

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

## References

1. Bian, L.; Sun, H.; He, H.; Liu, C.; Guan, Z. Research on Fault Diagnosis of Mine Ventilator Bearing based on Cross Entropy Algorithm. In Proceedings of the 2020 35th Youth Academic Annual Conference of Chinese Association of Automation (YAC), Zhanjiang, China, 16–18 October 2020.
2. Li, J.; Chen, W.; Han, K.; Wang, Q. Fault Diagnosis of Rolling Bearing Based on GA-VMD and Improved WOA-LSSVM. *IEEE Access* **2020**, *8*, 166753–166767. [[CrossRef](#)]
3. Li, Y.; Xu, M.; Wei, Y.; Huang, W. A new rolling bearing fault diagnosis method based on multiscale permutation entropy and improved support vector machine based binary tree. *Measurement* **2015**, *77*, 80–94. [[CrossRef](#)]
4. Zhang, X.; Wang, H.; Ren, M.; He, M.; Jin, L. Rolling Bearing Fault Diagnosis Based on Multiscale Permutation Entropy and SOA-SVM. *Machines* **2022**, *10*, 485. [[CrossRef](#)]
5. Zhang, X.; Miao, Q.; Zhang, H.; Wang, L. A parameter-adaptive VMD method based on grasshopper optimization algorithm to analyze vibration signals from rotating machinery. *Mech. Syst. Signal Process.* **2018**, *108*, 58–72. [[CrossRef](#)]
6. Ma, H.; Tong, Q.; Zhang, Y. Applications of Optimization Parameters VMD to Fault Diagnosis of Rolling Bearings. *Zhongguo Jixie Gongcheng China Mech. Eng.* **2018**, *29*, 390–397.
7. Guo, Z.; Liu, M.; Wang, Y.; Qin, H. A New Fault Diagnosis Classifier for Rolling Bearing United Multi-Scale Permutation Entropy Optimize VMD and Cuckoo Search SVM. *IEEE Access* **2020**, *8*, 153610–153629. [[CrossRef](#)]
8. Wang, M.; Wang, W.; Zeng, J.; Zhang, Y. An Integrated Method Based on Sparrow Search Algorithm Improved Variational Mode Decomposition and Support Vector Machine for Fault Diagnosis of Rolling Bearing. *J. Vib. Eng. Technol.* **2022**, *10*, 2893–2904. [[CrossRef](#)]
9. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol.* **2000**, *278*, H2039–H2049. [[CrossRef](#)]
10. Bandt, C.; Pompe, B. Permutation Entropy: A Natural Complexity Measure for Time Series. *Phys. Rev. Lett.* **2002**, *88*, 174102. [[CrossRef](#)]
11. Costa, M.; Goldberger, A.L.; Peng, C.K. Multiscale Entropy Analysis of Complex Physiologic Time Series. *Phys. Rev. Lett.* **2007**, *89*, 705–708. [[CrossRef](#)]
12. Aziz, W.; Arif, M. Multiscale Permutation Entropy of Physiological Time Series. In Proceedings of the International Multitopic Conference, Karachi, Pakistan, 23–24 December 2005.
13. An, G.; Tong, Q.; Zhang, Y.; Liu, R.; Li, W.; Cao, J.; Lin, Y. An Improved Variational Mode Decomposition and Its Application on Fault Feature Extraction of Rolling Element Bearing. *Energies* **2021**, *14*, 1079. [[CrossRef](#)]
14. Tian, J.; Morillo, C.; Azarian, M.H.; Pecht, M. Motor Bearing Fault Detection Using Spectral Kurtosis-Based Feature Extraction Coupled With K-Nearest Neighbor Distance Analysis. *IEEE Trans. Ind. Electron.* **2016**, *63*, 1793–1803. [[CrossRef](#)]
15. Ting-Ting, X.; Yan, Z.; Zong, M.; Xiao-Lin, G. A fault diagnosis method of rolling bearing based on VMD Tsallis entropy and FCM clustering. *Multimed. Tools Appl.* **2020**, *79*, 30069–30085. [[CrossRef](#)]
16. Shu, S.C.; Han, D. Approach for a hydraulic pump's feature extraction based on multiscale optimal fuzzy entropy. *J. Vib. Shock.* **2016**.
17. Zhang, X.; Zhang, M.; Wan, S.T.; He, Y.L.; Wang, X.L. A bearing fault diagnosis method based on multiscale dispersion entropy and GG clustering. *Measurement* **2021**, *185*, 110023. (In English) [[CrossRef](#)]
18. Wu, S.D.; Wu, P.H.; Wu, C.W.; Ding, J.J.; Wang, C.C. Bearing Fault Diagnosis Based on Multiscale Permutation Entropy and Support Vector Machine. *Entropy* **2012**, *14*, 1343–1356. [[CrossRef](#)]
19. Yan, X.; Jia, M.; Xiang, L. Compound fault diagnosis of rotating machinery based on OVMD and a 1.5-dimension envelope spectrum. *Meas. Sci. Technol* **2016**, *27*, 075002. [[CrossRef](#)]
20. Zhongliang, L.; Baoping, T.; Yi, Z.; Chuande, Z. A Novel Method for Mechanical Fault Diagnosis Based on Variational Mode Decomposition and Multikernel Support Vector Machine. *Shock. Vib.* **2016**, *2016*, 3196465.
21. Wu, C.; Fu, X.; Pei, J.; Dong, Z. A Novel Sparrow Search Algorithm for the Traveling Salesman Problem. *IEEE Access* **2021**, *9*, 153456–153471. [[CrossRef](#)]
22. Yu, Y.; Cheng, J.; Kang, Z. An ensemble local means decomposition method and its application to local rub-impact fault diagnosis of the rotor systems. *Measurement* **2012**, *45*, 561–570.
23. Campeau, W.; Simons, A.M.; Stevens, B. The evolutionary maintenance of Lévy flight foraging: A numerical simulation. *PLOS Comput. Biol.* **2021**, *18*, e1009490.
24. Gao, S.; Wang, Q.; Zhang, Y. Rolling Bearing Fault Diagnosis Based on CEEMDAN and Refined Composite Multi-Scale Fuzzy Entropy. *IEEE Trans. Instrum. Meas.* **2021**, *23*, 259.
25. Dragomiretskiy, K.; Zosso, D. Variational Mode Decomposition. *IEEE Trans. Signal Process.* **2014**, *62*, 531–544. [[CrossRef](#)]

26. He, C.; Wu, T.; Gu, R.; Jin, Z.; Ma, R.; Qu, H. Rolling bearing fault diagnosis based on composite multiscale permutation entropy and reverse cognitive fruit fly optimization algorithm—Extreme learning machine. *Measurement* **2021**, *173*, 108636. [[CrossRef](#)]
27. Hou, J.; Wu, Y.; Gong, H.; Ahmad, A.S.; Liu, L. A Novel Intelligent Method for Bearing Fault Diagnosis Based on EEMD Permutation Entropy and GG Clustering. *Appl. Sci.* **2020**, *10*, 386. [[CrossRef](#)]
28. Shengqiang, W.U.; Jiang, W.; Zhao, L. Research on KPCA Fault Diagnosis Method Based on Sound Signal. *Mach. Tool Hydraul.* **2016**.
29. Dai, C.; Liu, Z.; Hu, K.; Huang, K. Fault diagnosis approach of traction transformers in high-speed railway combining kernel principal component analysis with random forest. *IET Electr. Syst. Trans.* **2016**, *6*, 202–206. [[CrossRef](#)]
30. Ziwen, G.U.; Peng, L.I.; Xun, L.; Yixuan, Y.U.; Xin, S.; Min, C.A.O. A Multi-Granularity Density Peak Clustering Algorithm Based on Variational Mode Decomposition. *Chin. J. Electron.* **2021**, *30*, 658–668. [[CrossRef](#)]
31. Zhang, W.; Zhou, W.; Luo, J. Mining And Application of User Behavior Pattern Based on Operation And Maintenance Data. In Proceedings of the 2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM), Washington, DC, USA, 8–12 April 2019.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.