Hindawi
Shock and Vibration
Volume 2022, Article ID 7294795, 11 pages
https://doi.org/10.1155/2022/7294795

Research Article

Application of Variational Mode Decomposition and Multiscale Permutation Entropy in Rolling Bearing Failure Analysis

Haorui Liu, Haijun Li, Rongyan Wang, Hengwei Zhu, and Jianchen Zhang

College of Computer and Information, Dezhou University, Dezhou, Shandong 253500, China

Correspondence should be addressed to Haorui Liu; liuhaorui@dzu.edu.cn

Received 14 May 2022; Revised 1 September 2022; Accepted 27 September 2022; Published 26 October 2022

Academic Editor: Chengwei Fei

The rolling bearing fault test signal has nonstationary and nonlinear characteristics. The feature extraction method based on variational mode decomposition (VMD) and permutation entropy can effectively measure the regularity of the signal and detect weak changes. Since the center frequency of the intrinsic mode function (IMF) of each fault test signal contains more details, this paper further extracts the multiscale permutation entropy feature for each IMF. The training samples and test samples of each IMF are constructed, and then the support vector machine (SVM) and the *K*-nearest neighbor algorithm (KNN) are used to identify the faults. The test results of the IMF components are used to determine the classification results combined with the maximum attribution index. Compared with the relevant feature extraction, the experimental results show that the method achieves a certain improvement in the accuracy of fault identification. The research results of rolling bearing fault data show that the multiscale permutation entropy and SVM/KNN can more accurately diagnose different fault modes, different fault sizes, and different operating states of rolling bearings.

1. Introduction

In the diagnosis field of rolling bearings fault, machine learning-based fault diagnosis methods have attracted much attention, mainly including two procedures: feature extraction and pattern recognition.

In terms of feature extraction of faults, many statistical parameter estimation methods have been applied. In recent decades, entropy has commonly been used as the evaluation criteria of classification or clustering results, of which approximate entropy, permutation entropy, sample entropy, and fuzzy entropy have been introduced by many scholars in rolling bearing fault diagnosis.

In 1948, Shannon [1] proposed to take Shannon entropy as a measure of uncertainty for stochastic process results. As described in Shannon's seminal paper, information entropy provides a measure of information content. High entropy levels imply a high degree of "unpredictability"in the system. Conversely, if each subsequent state of a system can be easily predicted from the previous state, the system is said to have low entropy. Subsequently, Kolmogorov [2] defined the concept of entropy for a new class of dynamic systems. In 1991, Pincus et al. [3, 4] modified Kolmogorov–Shannon (KS) entropy and created an approximate entropy (ApEn), which can be used to measure the complexity of real-life time series without coarse-grained procedures. In 2000, Richman and Moorman [5, 6] further modified ApEn and created sample entropy (SampEn), which has two advantages over ApEn: data length independence and higher relative agreement under multiple parameters. In 2002, Costa et al. proposed multiscale entropy (MSE) [7, 8] to address the problem of classifying healthy individuals, patients with congestive heart failure, and patients with unstable arrhythmias on a single time scale. The concept of a coarse-grained process was proposed in the entropy calculation process. In addition, Richman and Moorman [5] proposed cross-SampEn for the measurement of asynchrony and phase anisotropy of two different time series.

Guo et al. used dynamic time warping (DTW) and symplectic geometry mode decomposition(SGMD) to recognize the fault type of the test sample [9]. Kumar et al. employed an entropy-based state space model for estimating

the remaining useful life (RUL) and carrying out the dynamic degradation monitoring of rolling element bearings [10]. Liu et al. solved the problem of failure samples being difficult to obtain by constructing FEM models with faults to obtain simulation signals [11]. Li et al. proposed a new complexity measure for analyzing time series and termed it as reverse dispersion entropy (RDE) [12].

In terms of uncertainty and randomness, sample entropy is superior to approximate entropy [9], and permutation entropy (PE) is superior to sample entropy [10]. PE has high sensitivity to signal change, strong noise elimination ability, and fast operation speed, as proposed by Bandt and Pompe, it is also available for the detection of the randomness of time series [11]. Considering these features, it can be used to quantify the fault sign gearbox signal present in vibration [12]. When any gear fault occurs during operation, it shows obvious signal characteristics. Therefore, the vibration signals obtained from such a system result in varied PE values in different health conditions. The smaller the entropy values, the lower the uncertainties present in the signal, and vice versa [13]. Due to the better predictability of the faults, the entropy will increase as the faults increase. As a consequence, PE can be used to distinguish bearing faults.

Among them, permutation entropy, as one of the most commonly used entropy methods, can effectively measure the regularity of signals and detect weak changes. It has the advantages of antinoise and invariance to nonlinear monotonic transformations. In the case of rolling bearing failure, the nonlinear dynamic complexity changes accordingly. Therefore, PE is well suited for feature extraction of rolling bearings. However, since PE analyzes time series on only one scale, the implementation of multiscale methods [14] has led to multiscale permutation entropy (MPE) techniques [15]. This enables permutation entropy to search for information contained in long-term trends in the signal, which can be lost if its original data points are analyzed directly. Similar to traditional single scale nonlinear metrics, PE can only measure the complexity of a signal (time series) on a single scale. In order to make up for the shortcomings, Aziz and Arif developed a multiscale permutation entropy (MPE) with better robustness than PE on the basis of PE [16], which was successfully applied to estimate the complexity and randomness of time series at different scales. Hsieh et al. used MPE to quantitatively analyze the vibration signals of rotating machinery at different scales and construct the original feature set [17]. A multichannel fused convolutional neural network was constructed to extract features from PE at multiple scales for fault identification of rolling bearings. Wang et al. proposed a data-driven method combining multiscale permutation entropy and linear local tangent spatial alignment to diagnose faults in vehicle suspension systems [18].

Furthermore, PE and MPE rely on pattern counts inside the signal, which itself is simply a sample series from a wider phenomenon. Thus, the measured entropy content is not a measure of certainty but an estimator [19, 20].

Considering the above theory and application results, after decomposing the original time series by VMD, we used multiscale permutation entropy (MPE) to perform feature extraction of intrinsic modal function components again, which were classified by the SVM method. Finally, the features were comprehensively judged using the calculation method of attribution degree. Multiscale permutation entropy of the original time series and VMD-based permutation entropy were applied for feature extraction. Moreover, two classification methods, KNN and SVM, were used for comparison.

2. Multiscale Permutation Entropy

Permutation entropy is used to measure the complexity of time series. The size of the permutation entropy value corresponding to different time series varies. The more complex the time series, the greater the permutation entropy value; the more regular the time series, the smaller the permutation entropy value [21–23]. Due to the different harmonic content of current signals of different types of faults, the complexity is quite different. Therefore, the fault signal feature extraction can be realized by calculating the arrangement entropy of fault current signals. Compared with permutation entropy, multiscale permutation entropy can measure the complexity of time series from different time scales by overcoming the problem of a single feature of time series in a single scale permutation entropy metric. The steps to calculate multiscale permutation entropy are as follows:

Step 1. Coarse-grain the IMF components obtained after decomposition at the s-scale and obtain a new time series at the s-scale, which is formulated as follows:

$$y_j(s) = \frac{1}{s} \sum_{i=(j-1)s+1}^{js} x_i. \tag{1}$$

In (1) $s = 1, 2, \cdots, n$ is the scale factor; $j = 1, 2, \cdots, N/s$; N is the sequence length; and $\cdot$ indicates rounding down.

Step 2. Perform phase space reconstruction with embedding dimension m and delay time t on the new time series after coarse-graining processing and incrementally arrange the interior of each subsequence $X(i)$. At this time, map each m-dimensional subsequence $X(i)$ to one of the $m!$ arrangements.

Step 3. Through the above steps, represent an m-dimensional subspace as a symbol sequence and denote the probability distribution of all symbols by $P_1, P_2, \cdots P_K$, where $K \leq m!$. Calculate the probability of the occurrence of each set of sequences at different time scales. The multiscale permutation entropy expression for each IMF component can be obtained as follows:

$$H_p(m) = -\sum_{i=1}^{k} P_i \ln P_i. \tag{2}$$

When $P_i = 1/m!$, take the maximum value $\ln(m!)$ of $H_p(m)$. Normalize $H_p(m)$, then

$$H_p(m) = \frac{H_p(m)}{\ln(m!)}. \quad (3)$$

In (3): $0 \le H_p(m) \le 1$, which reflects the degree of randomness of the time series, of which the degree becomes stronger as the entropy value increases.

The multiscale permutation entropy algorithm is mainly affected by three parameters: scale factor, embedding dimension, and delay time. Therefore, it is very important to select three parameters of the embedding dimension m, delay time τ, and scale factor s of multiscale permutation entropy. Furthermore, the changes of these three parameters will affect our calculation results to a great extent. Zheng [24] et al. experimentally analyzed the change of the entropy value of the multiscale arrangement of the signal when the value of delay τ was different and verified that when τ was changed from 1–6, the entropy value of multiscale arrangement changed little, indicating that the effect of delay τ on entropy value of the multiscale arrangement was small, so $\tau = 1$ was selected to calculate the entropy of the multiscale arrangement. Aziz and Arif [14] suggested that when the value of embedding dimension m was generally 3–7, the operation result reached the optimal state. When the value was 1–2, the number of reconstruction vector states was small, and the change of signal dynamics mutation could not accurately detect the change of signal. When the value was 6–7, the reconstruction of signal space homogenizes the time series, and the calculation amount of multiscale arrangement entropy increased dramatically and could not well reflect the subtle change in the time-series. He et al. [25] showed that the value of multiscale permutation entropy reaches an optimal state when the embedding dimension $m = 5$, $m = 6$, or $m = 7$. Therefore, in this paper, the embedding dimension $m = 6$ was selected.

In this paper, the vibration data of roller bearings were taken as the research object. The data source was the bearing vibration experimental data of the Electrical Engineering Laboratory of Case Western Reserve University. The driving end bearing model was deep groove ball bearing SKF6205, with 9 rollers, fault diameter of 0.1778 mm, motor load of 2 horsepower, motor speed of 1750 r/min, and sampling frequency of 12 KHz. The fault location was the inner ring, roller, and outer ring, so the data labels were divided into 4 types (normal state, inner ring fault, roller fault, and outer ring fault}. Figure 1 is the data display in one of the time series windows, and the number of sampling points is 2048.

It can be seen from the data diagram that the occurrence of a fault leads to the occurrence of a vibration shock, resulting in shock fluctuations in the amplitude of the three vibration signals. The shock frequency of the inner ring fault was large, the amplitude was small, and the shock frequency of the outer ring fault was relatively small, but the amplitude was large.

The results of the calculation of permutation entropy for the original time series of the four states are shown in Figure 2. The number of sampling points per frame was 2048, the head of the next frame was adjacent to the tail of the previous frame, and a total of 50 frames were obtained,

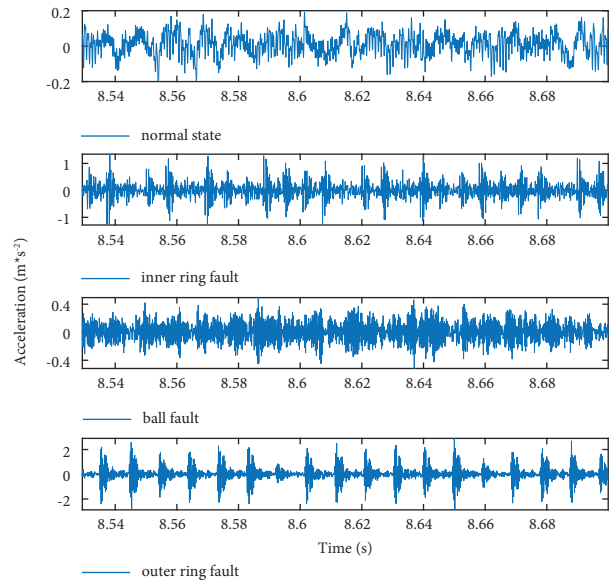


Figure 1: Original time series of four bearing states.

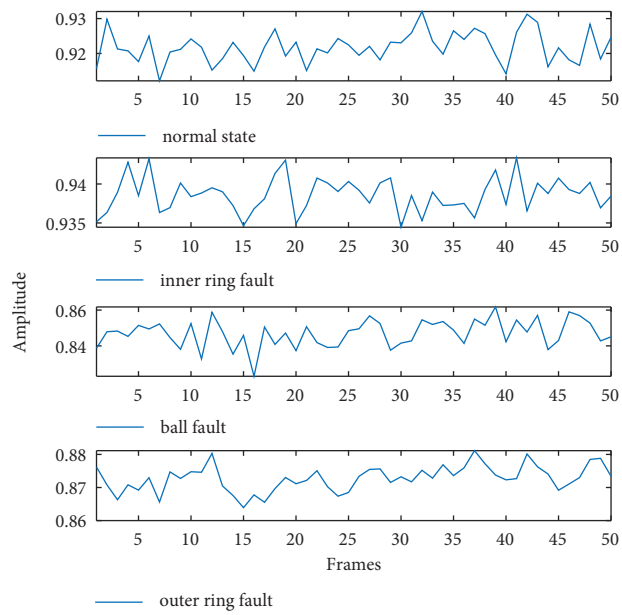


Figure 2: Permutation entropy calculation of the original time series.

with an average of 1.9247, 2.1105, 2.4871, and 1.5809 for the four states. Different thresholds were selected for comparison, and the results are shown in Table 1.

The normal state vibration data can be regarded as the superposition of noise and a certain frequency signal. According to the permutation entropy calculation rule, the appearance of a single periodic signal will make the entropy value reduce, and the amplitude and the number of a single signal will increase this reduction. So, the appearance of a periodic signal and its amplitude reduce the entropy value by 0.3438. For the roller, the irregular signal increases due to the

Table 1: Comparison of calculation results of different thresholds.

	$R=0.1$	$R=0.15$	$R=0.2$	$R=0.25$
Normal status	0.5914	0.7515	0.8424	0.8488
Inner ring fault	0.6892	0.8767	0.9222	0.9068
Roller fault	0.6896	0.8851	0.9218	0.9089
Outer ring fault	0.6844	0.8905	0.9241	0.9169

irregular rotation of the roller itself, and then the entropy value increases.

The appearance of fault signals makes it difficult to analyze the entropy features of samples, so it is necessary to perform modal decomposition of the original time series.

3. VMD

3.1. VMD Rationale. VMD is a signal decomposition estimation method with better time-frequency distribution than EMD and LMD (local mean decomposition), and its overall framework is a variational problem, that is, it decomposes the signal according to the number of preset modal components. The original signal $f(x)$ is decomposed into K modal function u_k with central frequencies of ω_k, where K is the number of preset modal components. In the VMD algorithm, the intrinsic mode function (IMF) is redefined as an AMP-FM signal.

$$u_k(t) = A_k(t)\cos\left(\phi_k(t)\right), \tag{4}$$

where, phase $\phi_k(t)$ is a nondecreasing function, $\phi_k'(t) \geq 0$; the envelope is non-negative, $A_k(t) \geq 0$; and the envelope $A_k(t)$ and the instantaneous frequency $\omega_k(t) = \phi_k'(t)$ are retarded for the phase $\phi_k(t)$.

In order to obtain K modal components with certain bandwidth frequencies, first of all, for each modal function u_k, the marginal spectrum is obtained by Hilbert transform; then, for each modal analytical signal, a central frequency is mixed and estimated; the spectrum of each modal is modulated to the corresponding base frequency band; then the square L^2 norm of the analytical signal gradient is calculated to estimate the bandwidth of each modal signal. The constrained variable problem is as follows:

$$\begin{aligned} &\min_{\{u_k\},\{\omega_k\}} \left\{ \sum_k \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t}{}_2^2 \right\}, \\ &\text{subject.to.} \sum_{k=1}^{K} u_k = f(t), \end{aligned} \tag{5}$$

where, $\{u_k\} = \{u_1, u_2, \cdots, u_k\}$ is number (K) of modal components obtained by decomposition; $\{\omega_k\} = \{\omega_1, \omega_2, \cdots, \omega_k\}$ is the frequency center of each component; and $\delta(t)$ is the pulse function.

In order to obtain the optimal solution to the above constrained variable problem, the quadratic penalty factor and Lagrange multiplication operator are introduced, and the alternate direction method of multipliers is used to solve the above variable problem. The saddle point of the extended Lagrange expression is gained by alternately updating u_k^{n+1}, ω_k^{n+1} and λ^{n+1}.

The value of the modal function u_k^{n+1} can be expressed as follows:

$$u_k^{n+1} = \underset{u_k \in X}{\operatorname{argmin}} \left\{ \alpha \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t}{}_2^2 +, \right. \tag{6}$$

where, a is the penalty parameter; λ is the Lagrange multiplier.

By using the Parseval/Plancherel Fourier isometric transform, the (6) is transited to the frequency domain.

$$\begin{aligned} \hat{u}_k^{n+1} = \underset{\hat{u}_k, u_k \in X}{\operatorname{argmin}} \Bigg\{ & \alpha j(\omega - \omega_k) \Big[(1 + \operatorname{sgn}(\omega) \hat{u}_k(\omega) \Big]_2^2 \\ & + \hat{f}(\omega) - \sum_i \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}{}_2^2 \Bigg\}. \end{aligned} \tag{7}$$

The solution of the quadratic optimization is obtained after further transformation as follows:

$$\begin{aligned} \hat{u}_k^{n+1}(\omega) = & \left(\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2} \right), \\ & \frac{1}{1 + 2a(\omega - \omega_k)^2}, \end{aligned} \tag{8}$$

where, ω_k is the core of the power spectrum of the current modal function, and it can be seen that the Wiener filter is embedded in the VMD algorithm, and the algorithm has better noise robustness.

The value of the center frequency ω_k can be expressed as follows:

$$\begin{aligned} & f(t) \sum_i u_i(t) + \frac{\lambda(t)^2}{2}{}_2^2 \Bigg\}, \\ & \omega_k^{n+1} = \underset{\omega_k}{\operatorname{argmin}} \left\{ \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} * u_k(t) \right) \right] e^{-j\omega_k t}{}_2^2 \right\}. \end{aligned} \tag{9}$$

According to the same process, the value taking of the center frequency is first transformed into the frequency domain.

$$\omega_k^{n+1} = \underset{\omega_k}{\operatorname{argmin}} \left\{ \int_0^\infty (\omega - \omega_k)^2 \left| \hat{u}_k(\omega) \right|^2 \mathrm{d}\omega \right\}. \tag{10}$$

The solution to the quadratic optimization of the center frequency is as follows:

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega \left| \hat{u}_k(\omega) \right| \mathrm{d}\omega}{\int_0^\infty \left| \hat{u}_k(\omega) \right|^2 \mathrm{d}\omega}. \tag{11}$$

The VMD algorithm steps are as follows:

Step 1. Initialize $\left\{ \hat{u}_k^1 \right\}, \left\{ \hat{\omega}_k^1 \right\}, \hat{\lambda}^1$ and n;

Step 2. According to the formula (11) update u_k and ω_k;

Step 3. Update λ;

Table 2: Comparison of different thresholds ($K = 3$).

	Center frequency (Hz)		
Normal condition	108	1046	2101
Inner ring fault	684	2753	3559
Roller fault	677	2918	3372
Outer ring fault	779	2836	3422

$$\hat{\lambda}^{n+1}(\omega) \longleftarrow \hat{\lambda}^{n}(\omega) + \tau\left(\hat{f}(\omega) - \sum_{k}\hat{u}_k^{n+1}(\omega)\right). \quad (12)$$

Step 4. Repeat Steps 2 and 3, the cycle is ended until the iteration stop condition $\sum_k \hat{u}_k^{n+1} - \hat{u}_k^{n}{}_2^2 / \hat{u}_k^{n}{}_2^2 < \varepsilon$ is met, and the result is output to obtain a modal component.

4. Data Analysis and Discussion

VMD was used to decompose the original time series. It was necessary to determine the number of modes K and the penalty factor. In this paper, three cases of $K = 3$, 4, and 5 were analyzed, and the results are shown in Table 2–4. The center frequency aliasing was small in the case of $K = 4$, so $K = 4$ was determined. By referring to the references, the penalty factor was finally determined to be = 2000.

The original time series of various states are selected for VMD decomposition. The decomposition results of the normal state are selected, and the results are shown in Figure 3.

For the four selected bearing operation states (normal state, loss diameter 0.1778 inner ring fault, roller fault, and outer ring fault), the time windows of each running state were decomposed by VMD. The time window length selected was 2048, and the overlap between adjacent windows is determined. A total of 50 time windows were used for permutation entropy extraction, and the permutation entropy mean value of 50 time windows was also calculated. The results are shown in Figure 4.

One of the time windows was selected for the extraction of entropy from multiscale samples, with a scale factor selection of 20 and nonoverlapping adjacent scales. The results are shown in Figure 5.

For each modal component of the four states, 100 time windows were selected and extracted for multiscale permutation entropy to obtain 100 ∗ 20 training samples, and finally, 400 ∗ 20 samples of the four states were labeled and used as input for SVM and KNN classifiers, with effects shown in Figure 6. The remaining IMF classification results were 100%.

To compare the accuracy of multiscale permutation entropy, the permutation entropy of modal components of each state was constructed as 100 ∗ 5 training samples, and finally 400 ∗ 5 samples of the four states were labeled and used as inputs for SVM and KNN classifiers. The effects are shown in Figure 7. It can be seen from the figure that four normal test samples are misjudged as roller faults when performing SVM classification.

To further compare the accuracy of multiscale permutation entropy, the time window of each state was used to extract the multiscale permutation entropy, which was constructed as 100 ∗ 20 training samples. Finally, 400 ∗ 20 samples of the four states were labeled and used as the input of SVM and KNN classifiers. From Figure 8, it can be seen that both classification methods reached 100% when performing classification.

Table 3: Comparison of different thresholds ($K = 4$).

	Center frequency (Hz)			
Normal condition	108	1046	2101	4884
Inner ring fault	608	1385	2757	3560
Roller fault	577	1053	2921	3373
Outer ring fault	756	2830	3370	3614

Table 4: Comparison of different thresholds ($K = 5$).

	Center frequency (Hz)				
Normal condition	106	1040	1289	2102	4888
Inner ring fault	608	1384	2745	3336	3636
Roller fault	576	1050	2867	3268	3428
Outer ring fault	644	1105	2836	3421	5183

For the selected four bearing operation states (normal state, inner ring fault with a loss diameter of 0.1778, roller fault, and outer ring fault), the test of data combination 1 was performed. The comprehensive test results are shown in Table Table 3. In the VMD-comprehensive test, the calculation method was to perform a membership degree calculation for the prediction results of four components and the residual error. That is, if one result belongs to one state with the most times, the result is judged to belong to the state. It can be seen from the table that the comprehensive judgment method can reach 100% judgment (Table 5).

To further validate the method proposed in this paper, the test of data combination 2 was performed. The running states of the four bearings were under load of 2 horsepower, the loss diameters of the inner ring fault were 0.1778 mm, 0.3556 mm, 0.5334 mm, and 0.7112 mm, respectively, which are shown in Table 6. It can be seen that in this test, except for misinformation in a small number of samples in IMF4 and residual errors, the VMD-synthesis method can reach 100% judgment.

The test of data combination 3 was performed. The four states were bearing normal operation states under different loads. From Table 7, it can be seen that the use of SVM for multiscale permutation entropy of the original time series has a better test effect; the use of KNN for permutation entropy obtained after VMD decomposition of the original time series has a better test effect; the use of VMD-multiscale permutation entropy combined with KNN has the best test effect, and its SVM classification effect is intermediate between the other two methods.

Table 5: Classification results for data combination 1.

		Multiscale MSE		Single scale SE	
		KNN (%)	SVM (%)	KNN (%)	SVM (%)
VMD	IMF1	100	100	100	99
	IMF2	100	100		
	IMF3	100	100		
	IMF4	100	99		
	Residual error	100	100		
	Integrative	100	100	100	99
Original time series (OTS)		100	100		

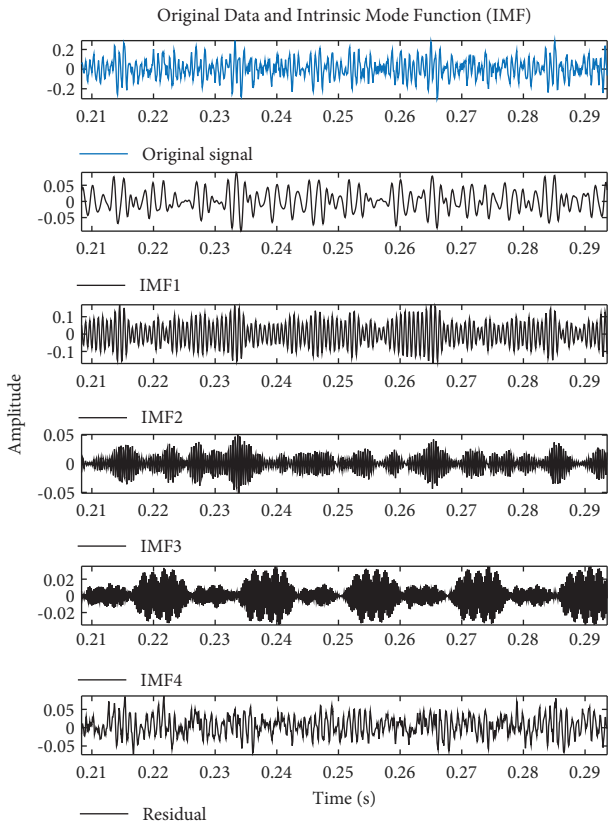
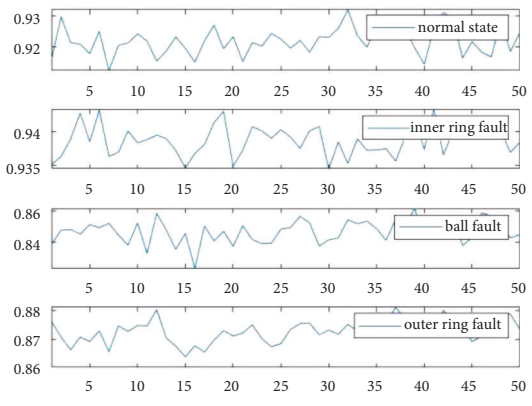


Figure 3: Multiscale permutation entropy for each IMF.



(a)

Figure 4: Continued.

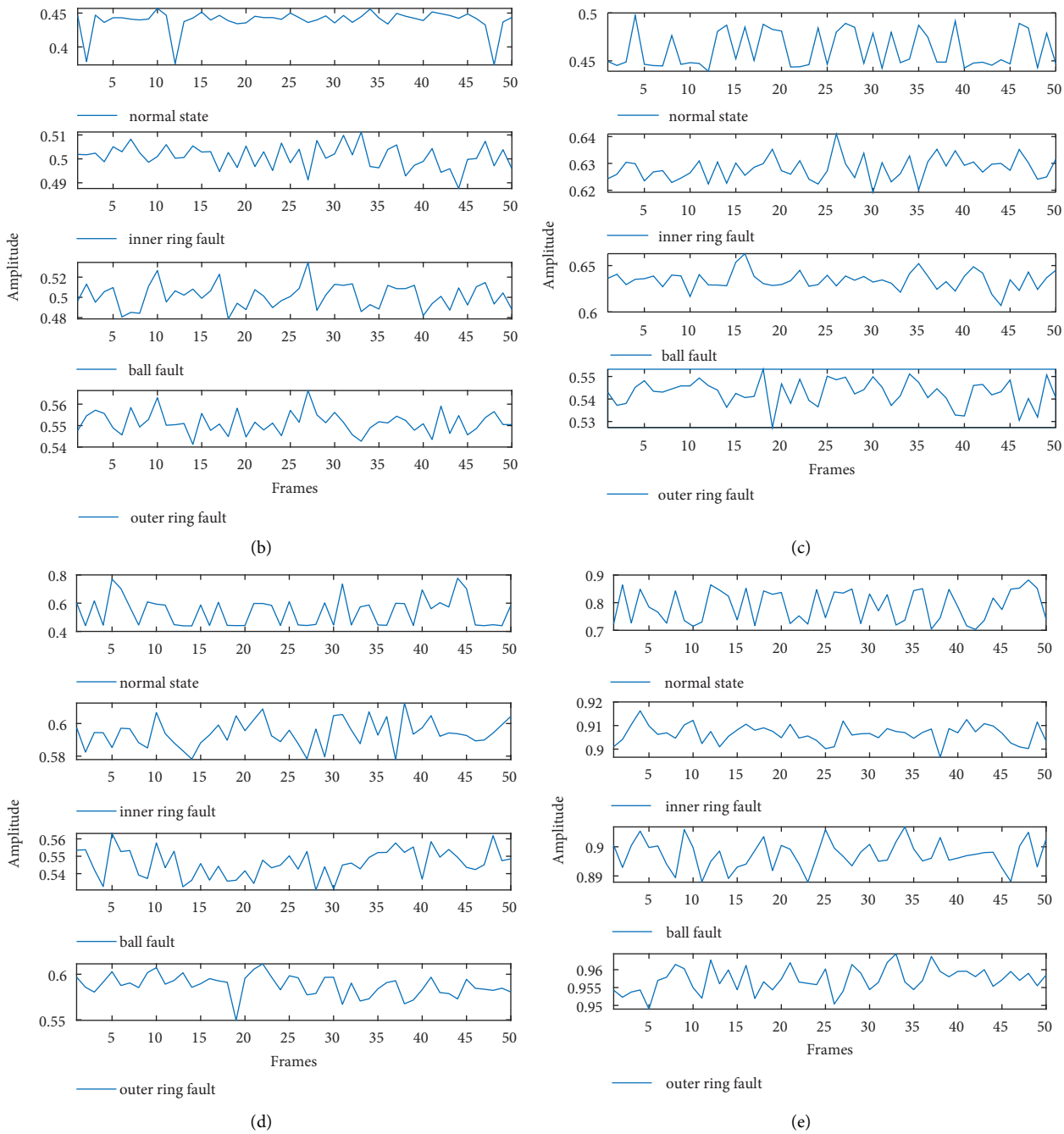


Figure 4: Permutation entropy calculation of the original time series. (a) Mean values of IMF1: 0.2710, 0.3767, 0.3960, and 0.4611, (b) mean values of IMF2: 0.4428, 0.5097, 0.4970, and 0.5571, (c) mean values of IMF3: 0.4708, 0.6383, 0.6432, and 0.5447, (d) mean values of IMF4: 0.5982, 0.5996, 0.5562, and 0.5847, and (e) mean values of residual: 0.7986, 0.9045, 0.9071, and 0.9573.

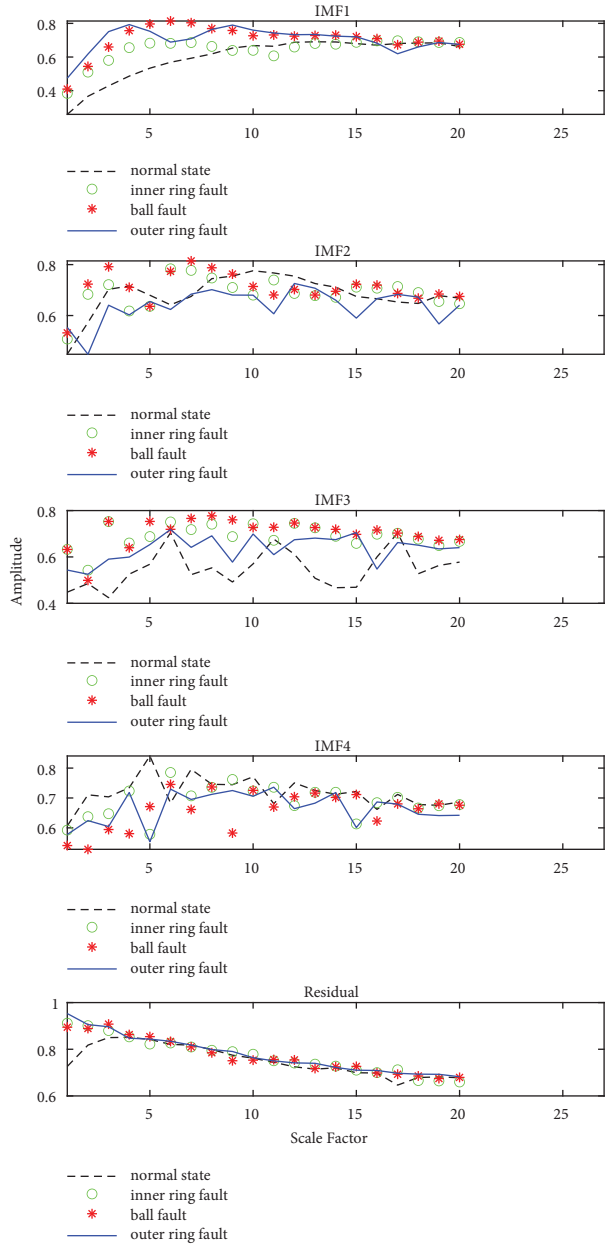


Figure 5: Multiscale permutation entropy for each IMF.

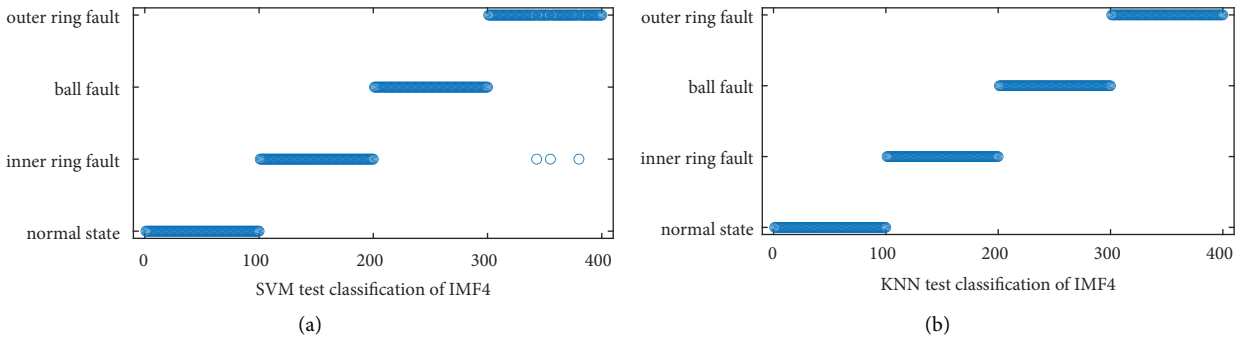


Figure 6: KNN and SVM classification comparison of IMF4 multiscale permutation entropy. (a) SVM test classification and (b) KNN test classification.

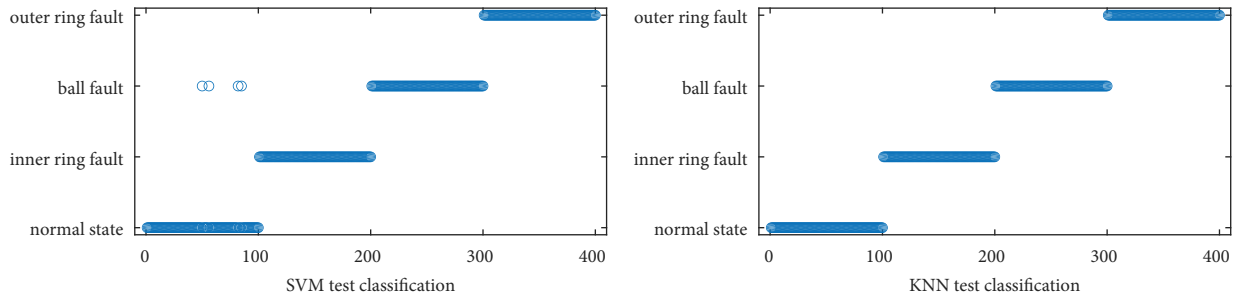


Figure 7: Comparison of permutation entropy classification after the original time series VMD decomposition. (a) SVM test classification and (b) KNN test classification.

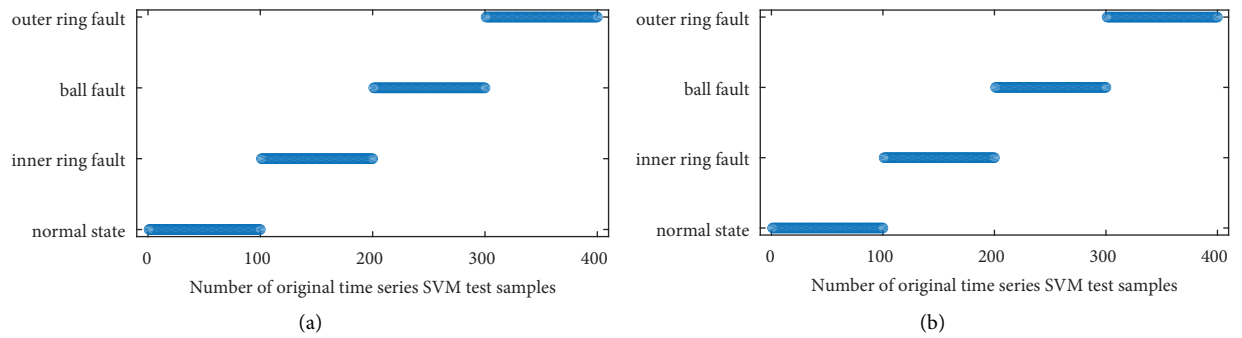


Figure 8: Comparison of multiscale permutation entropy classification for original time series. (a) SVM test classification and (b) KNN test classification.

Table 6: Classification results for data combination 2.

		Multiscale MSE		Single scale SE	
		KNN (%)	SVM (%)	KNN (%)	SVM (%)
VMD	IMF1	100	100	100	99
	IMF2	100	100		
	IMF3	100	100		
	IMF4	100	99		
	Residual error	92.25	82.75		
	Integrative	100	100		
Original time series (OTS)		100	100		

Table 7: Classification results for data combination 3.

		Multiscale MSE		Single scale SE	
		KNN (%)	SVM%	KNN (%)	SVM (%)
VMD	IMF1	37	43.25%	87.50	66.25
	IMF2	77.50	85.25%		
	IMF3	87.25	80.75		
	IMF4	46.5	53.75%		
	Residual error	59.75	65.75%		
	Integrative	87.75	84.25%		
Original time series (OTS)		76.75	87.25%		

5. Conclusion

In view of the fact that the central frequency of the intrinsic mode function (IMF) after test signal VMD decomposition contains more details, this paper further extracted the features of multiscale permutation entropy for each IMF. Subsequently, the training and test samples of each IMF were constructed, followed by the fault identification using support vector machines and KNN. The test results of the five modal components were applied to determine the classification results through the maximum assignment index. The feature extraction used at the same time includes the permutation entropy after VMD decomposition of the original time series as well as the multiscale permutation entropy of the original time series. The combinations of data used include varied fault modes, fault sizes, and operating states of rolling bearings. On account of obvious characteristics of fault mode and fault size, the above feature extraction mode could reach 99% extraction at the lowest. Compared with other feature extractions, this method could improve the accuracy of fault identification in different load test data combinations under normal operations. The research results of fault data of rolling bearings showed that the diagnosis of varied fault modes, fault sizes, and operating states of rolling bearings could be accurately performed based on multiscale permutation entropy and SVM/KNN.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research received financial supports from the Key Laboratory of Intelligent Industrial Equipment Technology of Hebei Province (Hebei University of Engineering) (NO. 202203) and China University Innovation Fund (NO. 2021DZ026).

References

[1] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[2] A. N. Kolmogorov, "New metric invariant of transitive dynamical systems and automorphisms of Lebesgue spaces," *Doklady of Russian Academy of Sciences*, vol. 119, no. 5, pp. 861–864, 1958.

[3] Y. G. Sinai, "On the notion of entropy of a dynamical systems," *Doklady of Russian Academy of Sciences*, vol. 124, pp. 768–771, 1959.

[4] S. M. Pincus, I. M. Gladstone, and R. A. Ehrenkranz, "A regularity statistic for medical data analysis," *Journal of clinical monitoring*, vol. 7, no. 4, pp. 335–345, 1991.

[5] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology - Heart and Circulatory Physiology*, vol. 278, no. 6, pp. H2039–H2049, 2000.

[6] D. E. Lake, J. S. Richman, M. P. Griffin, and J. R. Moorman, "Sample entropy analysis of neonatal heart rate variability," *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, vol. 283, no. 3, pp. R789–R797, 2002.

[7] M. Costa, A. L. Goldberger, and C. K. Peng, "Multiscale entropy analysis of complex physiologic time series," *Physical Review Letters*, vol. 89, no. 6, p. 068102, Article ID 068102, 2002.

[8] M. Costa, A. L. Goldberger, and C. K. Peng, "Multiscale entropy analysis of biological signals," *Physical Review*, vol. 71, no. 2, Article ID 021906, 2005.

[9] J. Guo, Z. Si, Y. Liu, J. Li, Y. Li, and J. Xiang, "Dynamic time warping using graph similarity guided symplectic geometry mode decomposition to detect bearing faults," *Reliability Engineering & System Safety*, vol. 224, Article ID 108533, 2022.

[10] A. Kumar, C. Parkash, G. Vashishtha, H. Tang, P. Kundu, and J. Xiang, "State-space modeling and novel entropy-based health indicator for dynamic degradation monitoring of rolling element bearing," *Reliability Engineering & System Safety*, vol. 221, Article ID 108356, 2022.

[11] X. Liu, H. Huang, and J. Xiang, "A personalized diagnosis method to detect faults in a bearing based on acceleration sensors and an FEM simulation driving support vector machine," *Sensors*, vol. 20, no. 2, p. 420, 2020.

[12] Y. Li, X. Gao, and L. Wang, "Reverse dispersion entropy: a new complexity measure for sensor signal," *Sensors*, vol. 19, no. 23, p. 5203, 2019.

[13] Y. Deng, W. Wang, C. Qian, Z. Wang, and D. Dai, "Boundary-processing-technique in EMD method and Hilbert transform," *Chinese Science Bulletin*, vol. 46, no. 11, pp. 954–960, 2001.

[14] W. Aziz and M. Arif, "Multiscale permutation entropy of physiological time series," in *Proceedings of the 2005 Pakistan Section Multitopic Conference*, pp. 1–6, Karachi, Pakistan, December 2005.

[15] L. Zhang, G. Xiong, H. Liu, H. Zou, and W. Guo, "Bearing fault diagnosis using multi-scale entropy and adaptive neuro-fuzzy inference," *Expert Systems with Applications*, vol. 37, no. 8, pp. 6077–6085, 2010.

[16] Y. Li and J. Jiang, "Fault diagnosis of bearing based on LMD and MSE," in *Proceedings of the 2017 Prognostics and System Health Management Conference (PHM-Harbin)*, pp. 1–4, Harbin, China, July 2017.

[17] N. K. Hsieh, W. Y. Lin, and H. T. Young, "High-speed spindle fault diagnosis with the empirical mode decomposition and multiscale entropy method," *Entropy*, vol. 17, no. 4, pp. 2170–2183, 2015.

[18] Z. Wang, L. Yao, and Y. Cai, "Rolling bearing fault diagnosis using generalized refined composite multiscale sample entropy and optimized support vector machine," *Measurement*, vol. 156, Article ID 107574, 2020.

[19] A. Dávalos, M. Jabloun, P. Ravier, and O. Buttelli, "Improvement of statistical performance of ordinal multiscale entropy techniques using refined composite downsampling permutation entropy," *Entropy*, vol. 23, no. 1, p. 30, 2020.

[20] H. Shaobo, S. Kehui, and W. Huihai, "Modified multiscale permutation entropy algorithm and its application for multiscroll chaotic systems," *Complexity*, vol. 21, no. 5, pp. 52–58, 2016.

[21] H. Li, J. Huang, X. Yang, J. Luo, L. Zhang, and Y. Pang, "Fault diagnosis for rotating machinery using multiscale

permutation entropy and convolutional neural networks," *Entropy*, vol. 22, no. 8, p. 851, 2020.

[22] Y. Ye, Y. Zhang, Q. Wang, Z. Wang, Z. Teng, and H. Zhang, "Fault diagnosis of high-speed train suspension systems using multiscale permutation entropy and linear local tangent space alignment," *Mechanical Systems and Signal Processing*, vol. 138, Article ID 106565, 2020.

[23] Y. Hu, S. Zhang, A. Jiang, L. Zhang, W. Jiang, and J. Li, "A new method of wind turbine bearing fault diagnosis based on multi-masking empirical mode decomposition and fuzzy C-means clustering," *Chinese Journal of Mechanical Engineering*, vol. 32, no. 1, p. 46, 2019.

[24] X. Zheng, G. Zhou, D. Li, H. Ren, and H. Ren, "Application of variational mode decomposition and permutation entropy for rolling bearing fault diagnosis," *International Journal of Acoustics and Vibrations*, vol. 24, no. 2, pp. 303–311, 2019.

[25] T. He, R. Zhao, Y. Wu, and C. Yang, "Fault identification of rolling bearing using variational mode decomposition multiscale permutation entropy and adaptive GG clustering," *Shock and Vibration*, vol. 2021, 13 pages, Article ID 9212759, 2021.