

Article

A Hydraulic Pump Fault Diagnosis Method Based on the Modified Ensemble Empirical Mode Decomposition and Wavelet Kernel Extreme Learning Machine Methods

Zhenbao Li ^{1,2}, Wanlu Jiang ^{1,2,*}, Sheng Zhang ^{1,2}, Yu Sun ^{1,2} and Shuqing Zhang ³

¹ Hebei Provincial Key Laboratory of Heavy Machinery Fluid Power Transmission and Control, Yanshan University, Qinhuangdao 066004, China; lizhenbao@stumail.ysu.edu.cn (Z.L.); zsqhd@ysu.edu.cn (S.Z.); zhang_py@stumail.ysu.edu.cn (Y.S.)

² Key Laboratory of Advanced Forging & Stamping Technology and Science, Yanshan University, Ministry of Education of China, Qinhuangdao 066004, China

³ School of Electrical Engineering, Yanshan University, Qinhuangdao 066004, China; zhshq-yd@163.com

* Correspondence: wljiang@ysu.edu.cn

Abstract: To address the problem that the faults in axial piston pumps are complex and difficult to effectively diagnose, an integrated hydraulic pump fault diagnosis method based on the modified ensemble empirical mode decomposition (MEEMD), autoregressive (AR) spectrum energy, and wavelet kernel extreme learning machine (WKELM) methods is presented in this paper. First, the non-linear and non-stationary hydraulic pump vibration signals are decomposed into several intrinsic mode function (IMF) components by the MEEMD method. Next, AR spectrum analysis is performed for each IMF component, in order to extract the AR spectrum energy of each component as fault characteristics. Then, a hydraulic pump fault diagnosis model based on WKELM is built, in order to extract the features and diagnose faults of hydraulic pump vibration signals, for which the recognition accuracy reached 100%. Finally, the fault diagnosis effect of the hydraulic pump fault diagnosis method proposed in this paper is compared with BP neural network, support vector machine (SVM), and extreme learning machine (ELM) methods. The hydraulic pump fault diagnosis method presented in this paper can diagnose faults of single slipper wear, single slipper loosing and center spring wear type with 100% accuracy, and the fault diagnosis time is only 0.002 s. The results demonstrate that the integrated hydraulic pump fault diagnosis method based on MEEMD, AR spectrum, and WKELM methods has higher fault recognition accuracy and faster speed than existing alternatives.

Keywords: hydraulic pump; fault diagnosis; modified ensemble empirical mode decomposition (MEEMD); wavelet kernel extreme learning machine (WKELM)



Citation: Li, Z.; Jiang, W.; Zhang, S.; Sun, Y.; Zhang, S. A Hydraulic Pump Fault Diagnosis Method Based on the Modified Ensemble Empirical Mode Decomposition and Wavelet Kernel Extreme Learning Machine Methods. *Sensors* **2021**, *21*, 2599. <https://doi.org/10.3390/s21082599>

Academic Editor: Chris Karayannis

Received: 11 March 2021

Accepted: 5 April 2021

Published: 7 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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/>).

1. Introduction

In recent years, modern industrial equipment has developed in the direction of complexity, high speed, and intelligence. This not only ensures the quality of the products, but also greatly reduces the associated production costs and significantly improves production efficiency. However, in the actual production process, due to the continuous expansion of system scale and the increasing complexity of the structures involved, the probability of equipment failure has been gradually increasing. Therefore, it is very important to improve the safety and reliability of such equipment. In this regard, it is necessary not only to improve the reliability of the equipment in the process of equipment design and manufacturing, but also to carry out real-time status monitoring and fault diagnosis during the operation of the equipment. In this way, we can detect the fault initially, then predict the location and cause of the fault, in order to provide accurate information for predictive maintenance. Therefore, more and more attention has been paid to fault diagnosis and prediction technology, for which a lot of research work has been carried out [1–4].

Hydraulic systems have the advantages of small size, light weight, high power, smooth transmission, and fast response. Hydraulic systems have been widely used in the aerospace, shipbuilding, metallurgical, machine tool, agricultural machinery, and engineering machinery industries, among other fields. The wide application of hydraulic systems has greatly promoted the development of industrial production [5]. Since entering the 21st century, hydraulic systems and hydraulic equipment have become more complicated and large-scale. Once a failure occurs, it not only can damage the equipment, but may even cause casualties [6]. Hydraulic pumps have the characteristics of complex structure, high speed, high pressure, and continuous operations, and are the most prone to failure. Therefore, fault diagnosis technology for hydraulic systems has become a research hotspot in recent years.

In fault diagnosis, the first step is to extract fault features from the acquired data [7,8]. Feature extraction is the process of obtaining fault features through signal processing [9,10]. Considering that the vibration signal of a hydraulic pump has non-linear and non-stationary characteristics, it is difficult to extract the fault features accurately by using traditional time- and frequency-domain analyses. The key information of the hydraulic pump vibration signal can be extracted more accurately by using time-frequency characteristic analysis methods. Time-frequency analysis methods include non-adaptive analysis and adaptive analysis. Adaptive analysis refers to the modal decomposition and reconstruction of the signal, according to the characteristics of the signal, in the process of signal processing, in order to achieve good analysis results.

Considering the mode mixing phenomenon of the empirical mode decomposition (EMD) method [11,12], Wu et al. proposed the ensemble empirical mode decomposition (EEMD) method, which effectively suppresses modal mixing by adding Gaussian white noise to the original signal [13]. In view of the shortcomings of the EEMD method, scholars have carried out a lot of research. Finally, Zheng et al. proposed the modified ensemble empirical mode decomposition (MEEMD) method [14], which adds a pair of Gaussian white noises with the same amplitude and opposite signs, such that they cancel each other and the influence of the noise on the signal is reduced. Compared with the EEMD method, this method has a better decomposition effect. Jiang et al. applied the MEEMD method to the fault diagnosis of the main shaft bearing of a wind turbine [15]. Their method showed good results, in terms of training accuracy and training speed, which verifies the feasibility of this method. Zheng proposed a method combining MEEMD and generalized adaptive S transform (AGST), which was successfully applied to the vibration signal analysis of an internal combustion engine [16]. Shi et al. applied the MEEMD method to fault feature extraction of rolling bearings, which not only effectively suppressed the phenomenon of modal mixing but also greatly reduced the number of pseudo-components and accurately extracted the weak fault information of rolling bearings [17]. Although the MEEMD method has been widely used in the field of fault diagnosis, it has seldom been reported in hydraulic system fault diagnosis. Therefore, in this paper, we applied it to the vibration signal feature extraction of an axial piston pump, in order to improve the accuracy of fault identification.

In 2004, Professor Huang from Nanyang Technological University in Singapore proposed a new single-hidden layer feedforward neural network, called the extreme learning machine (ELM) [18–20]. In 2018, Zhang et al. proposed a method combining evidence theory and ELM, which effectively solved the classification problem of incomplete data [21]. However, it should be noted that the optimal number of neurons in the hidden layer obtained by the incremental extreme learning machine and its improved method may be very large, thus wasting a lot of unnecessary training time.

As the ELM method needs to set the number of hidden layer neurons and randomly select input weights and hidden layer biases, ELM instability is inevitable. Therefore, Huang et al. introduced a kernel function into the ELM to propose the kernel-ELM (KELM) method. This method avoids the problems of artificially setting the number of hidden layer neurons and randomly selecting the input weights and hidden layer biases in ELM

by means of kernel function mapping. KELM can achieve very high training accuracy by using the powerful mapping ability of kernel function [22]. To date, KELM has been widely used in various fields. One study [23] has proposed a parallel KELM, which effectively improved the computing speed of KELM. Cho et al. applied a K-ELM to correctly classify the stress levels to reflect five different levels of stress situations [24]. Zhang et al. proposed a soft sensing algorithm that can detect alumina concentration by the electrical signals such as voltages and currents of the anode rods [25]. Yang et al. combined the multi-scale feature energy with the KELM to effectively diagnose the airflow jamming fault of a fluidized bed [26]. Zhang et al. applied the KELM to the fault diagnosis of a coal mill and optimized the kernel parameters with particle swarm optimization (PSO) [27]. Yang et al. proposed a deep kernel extreme learning machine (DK-ELM), which has been successfully applied to aero-engine fault diagnosis [28]. Considering that wavelet functions have the advantages of time–frequency local analysis and multi-resolution analysis, a wavelet kernel function has been introduced into KELM to construct the wavelet kernel extreme learning machine (WKELM).

In this paper, a fault diagnosis method for a hydraulic pump based on the integration of the MEEMD, autoregressive (AR) spectrum energy, and WKELM methods is proposed and studied. First, the vibration signals of an axial piston pump under different working states were acquired from a hydraulic pump fault simulation test bed. Secondly, signal decomposition and feature extraction were carried out for the acquired original vibration signals, using the fault feature extraction method based on MEEMD and AR spectrum. Finally, the fault diagnosis method based on WKELM was used to diagnose the working states of the hydraulic pump. The results show that the diagnosis accuracy rate can reach 100%. Compared with BP, SVM, and ELM, the feasibility and superiority of the method proposed in this paper are demonstrated.

2. Basic Principles of Modified Ensemble Empirical Mode Decomposition (MEEMD) and AutoRegressive (AR) Spectrum Method

2.1. Basic Principles of MEEMD Method

MEEMD is an improvement on EEMD [29]. First, considering that the set of Gaussian white noise added in EEMD cannot be completely canceled out, MEEMD adds a pair of Gaussian white noise signals instead, where the amplitude of the white noise is the same but the sign is opposite. The purpose of this is to allow the added Gaussian white noise signals to cancel each other out, thereby reducing the influence of noise on the signal [30]. Secondly, considering that EEMD will decompose more pseudo-components, the concept of PE is introduced to MEEMD, in order to detect whether the signal is abnormal. The abnormal signals are decomposed again and eliminated, such that the pseudo-components can be effectively reduced. At the same time, the MEEMD method can also suppress the phenomenon of mode mixing in the EMD and EEMD methods. Therefore, the MEEMD method has a better decomposition effect than EMD or EEMD.

The Specific steps of MEEMD are as follows [14]:

(1) Two groups of Gaussian white noises, $noise_l(i)$ and $-noise_l(i)$, with equal absolute values, are added to the original signal $\{x(i), i = 1, 2, \dots, n\}$:

$$\begin{aligned} x^+(i) &= x(i) + a_l noise_l(i) \\ x^-(i) &= x(i) - a_l noise_l(i) \end{aligned} \quad (1)$$

where $noise_l(i)$ is the added Gaussian white noise signal, a_l is the amplitude of the added noise signal, and $l = 1, 2, \dots, Ne$, where Ne is the number of white noise additions.

The signals $x^+(i)$ and $x^-(i)$, after adding Gaussian white noise, are decomposed by EMD. The first-order Intrinsic Mode Function (IMF) component sequences are respectively

expressed as $\{c_{l,1}^+(i)\}$ and $\{c_{l,1}^-(i)\}$, where $l = 1, 2, \dots, Ne$. The overall average of the above components can eliminate the residual white noise to the greatest extent, that is:

$$c_1(i) = \frac{1}{2Ne} \sum_{l=1}^{Ne} [c_{l,1}^+(i) + c_{l,1}^-(i)]. \quad (2)$$

The PE of the IMF components $c_1(i)$ is calculated after the overall average. If the PE is greater than θ , it is considered to be an abnormal signal; otherwise, it is approximately considered to be a stationary signal. The general range of θ is 0.55–0.6; θ is taken in this paper.

(2) Determine whether $c_1(i)$ is an abnormal signal. If $c_1(i)$ is an abnormal signal, the first IMF component $c_1(i)$ will be separated from the original signal $x(i)$, where the new signal $r_1(i)$ is:

$$r_1(i) = x(i) - c_1(i). \quad (3)$$

Taking the signal $r_1(i)$ as the original signal, Step (1) is repeated until the IMF component $c_q(i)$ obtained after the overall average is not an abnormal signal.

(3) The first $q - 1$ decomposed abnormal signal components whose PE is greater than θ are separated from the original signal, namely:

$$r_{q-1}(i) = x(i) - \sum_{j=1}^{q-1} c_j(i), \quad (4)$$

where $j = 1, 2, \dots, q - 1$, $q - 1$ is the number of abnormal components, and $r_{q-1}(i)$ is the residual signal after removing the abnormal components.

(4) After the remaining signal, $r_{q-1}(i)$, is decomposed by EMD, the decomposed IMF components are arranged, in order from high frequency to low frequency:

$$r_{q-1}(i) \xrightarrow{EMD} \sum_{k=1}^z [c_k'(i)] + r'(i). \quad (5)$$

Finally, MEEMD method can be expressed as:

$$x(i) \xrightarrow{MEEMD} \sum_{j=1}^{q-1} c_j(i) + \sum_{k=1}^z [c_k'(i)] + r'(i), \quad (6)$$

where $c_k'(i)$ are the IMF components finally obtained, $k = 1, 2, \dots, z$, z is the number of IMF components finally obtained, and $r'(i)$ is the residual component finally obtained.

2.2. Permutation Entropy

Permutation entropy (PE) [31] is a method for detecting the sudden change of dynamic behavior and signal regularity, which has the advantages of fast calculation speed and strong anti-interference ability. It is especially suitable for non-linear and non-stationary signal analysis and has good robustness. The steps of the permutation entropy method are as follows:

Given a period of time-series: $\{x(i), i = 1, 2, \dots, n\}$, the reconstruction matrix B can be obtained by phase space reconstruction:

$$B = \begin{bmatrix} x(1) & x(1 + \lambda) & \cdots & x(1 + (m - 1)\lambda) \\ x(2) & x(2 + \lambda) & \cdots & x(2 + (m - 1)\lambda) \\ \vdots & \vdots & \ddots & \vdots \\ x(k) & x(k + \lambda) & \cdots & x(k + (m - 1)\lambda) \end{bmatrix}, \quad (7)$$

where λ is the time delay, m is the embedding dimension, and $k = n - (m - 1)\lambda$. Each row $b(i)$ in the reconstruction matrix B can be regarded as a reconstruction component.

The elements of the i^{th} reconstruction component $b(i)$ in the reconstruction matrix B are arranged in ascending order, according to the size of the value:

$$x(i + (j_1 - 1)\lambda) \leq x(i + (j_2 - 1)\lambda) \leq \dots \leq x(i + (j_m - 1)\lambda), \quad (8)$$

where j_1, j_2, \dots, j_m are the indices of the column of the reconstructed component element. If $x(i + (j_1 - 1)\lambda) = x(i + (j_2 - 1)\lambda)$, they are sorted by the sizes of j_1, j_2 ; that is, when $j_1 < j_2$, $x(i + (j_1 - 1)\lambda) \leq x(i + (j_2 - 1)\lambda)$.

Therefore, for any discrete sequence $\{x(i), i = 1, 2, \dots, n\}$, a set of symbolic sequences can be obtained for each row of its reconstruction matrix B :

$$S(l) = (j_1, j_2, \dots, j_m), \quad (9)$$

where $l = 1, 2, \dots, k, k \leq m!$, as there are $m!$ permutations of m symbols $\{j_1, j_2, \dots, j_m\}$ and, so, the m -dimensional phase space can map $m!$ kinds of symbol sequences, where $S(l)$ is just one of them. The probability $\{P_1, P_2, \dots, P_k\}$ of each symbol sequence is calculated, following which the PE of the discrete sequence $\{x(i), i = 1, 2, \dots, n\}$ can be defined as [32]:

$$H_P(m) = -\sum_{j=1}^k P_j \ln P_j. \quad (10)$$

When $P_j = 1/m!$, then $H_P(m)$ reaches its maximum $\ln(m!)$. For convenience, $\ln(m!)$ is often used to standardize $H_P(m)$. Thus, the PE obtained in Equation (4) can be standardized as follows:

$$PE_{norm} = \frac{H_P(m)}{\ln(m!)}, \quad (11)$$

where $PE_{norm} \in [0, 1]$. The smaller the value of PE_{norm} , the more regular the time-series; thus, PE_{norm} can reflect the randomness of the time-series [33].

2.3. AR Spectrum Method

The autoregressive (AR) model is a time-series analysis method. It not only can fully reveal the characteristics of a signal, but also can improve the resolution of the signal. By analyzing the signal using an AR model, the characteristic frequency contained in the signal can be effectively extracted, such that feature extraction can be carried out accurately. The AR spectrum method can overcome the limitations of the Hilbert method, as there is neither estimation error nor a windowing effect. Therefore, the AR spectrum has great advantages in fault feature extraction. The general expression of the AR model is [34]:

$$x(n) = u(n) - \sum_{k=1}^p a_k x(n - k), \quad (12)$$

where $a_k (k = 1, 2, \dots, p)$ are the parameters of the AR model, p is the order of the model, and $u(n)$ is a stationary white noise sequence with a mean value of zero and a variance of σ^2 .

The AR(p) process $\{x(n)\}$ defined by Equation (12) can be regarded as a white noise sequence $u(n)$ generated by an all-pole filter, whose transfer function is:

$$H(z) = \frac{1}{1 + \sum_{k=1}^p a_k z^{-k}}. \quad (13)$$

Therefore, the power spectral density, $p_x(e^{j\omega})$, of $\{x(n)\}$ can be expressed as:

$$p_x(e^{j\omega}) = \sigma^2 |H(e^{j\omega})|^2 = \frac{\sigma^2}{\left|1 + \sum_{k=1}^p a_k e^{-jk\omega}\right|^2}. \quad (14)$$

The AR spectrum has the advantages of high resolution, smoothness, ease of locating, and so on. By analyzing a signal through the AR model, the characteristic frequency contained in the signal can be effectively extracted and the feature extraction can be carried out accurately.

3. Wavelet Kernel Extreme Learning Machine

WKELM has been proposed as an improvement on the basis of ELM. Considering that a wavelet kernel function has the characteristics of multi-scale approximation by wavelet analysis, it has a better effect for non-linear classification. Therefore, a wavelet kernel function was used to replace the matrix multiplication of the output matrix, H , of the ELM hidden layer neurons. This not only can improve the fitting ability of the method for non-linear samples but can also improve the stability of the method and shorten the training time.

3.1. Extreme Learning Machine Theory

The key characteristic of ELM is that the input weights and hidden layer biases of the network are randomly selected [35], and there is no need for iterative adjustment in the process of operating the method. Therefore, the calculation speed of ELM is extremely fast, which can greatly save time and effectively overcome the shortcomings of traditional single-hidden layer feed-forward neural networks, which have slow convergence speed and can easily become stuck in local minima. ELM can minimize the training error in the process of training and ensure that the norm of the output weights is very small, such that it has higher function approximation ability and generalization performance than traditional neural networks.

ELM [18] is a single-hidden layer feed-forward neural network, which is composed of an input layer, a hidden layer, and an output layer. The ELM network structure is similar to that of a single-hidden layer feed-forward neural network, as shown in Figure 1.

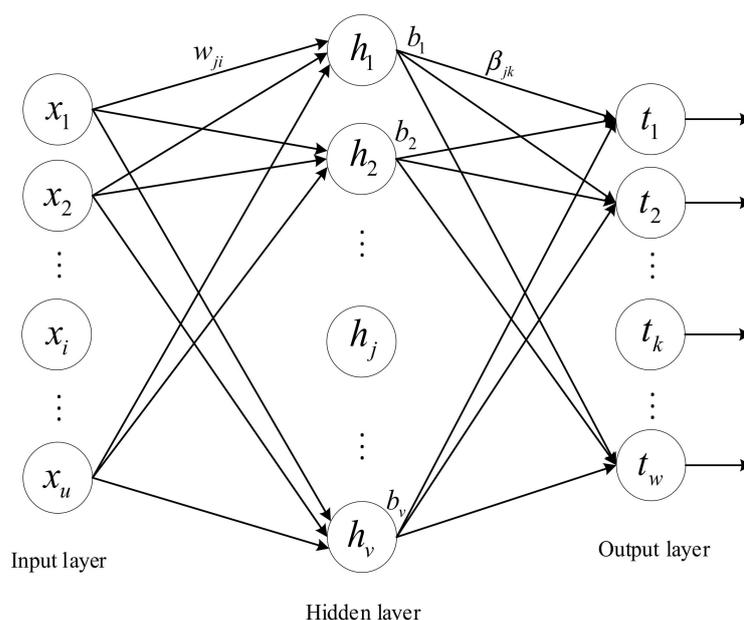


Figure 1. Structure of a single-hidden layer feed-forward neural network.

Suppose we have an ELM with u neurons in the input layer, v neurons in the hidden layer, and w neurons in the output layer. Given a training sample set $(X, T) = \{(x_j, t_j) | 1 \leq j \leq M\}$ of M groups, where the input sample set is $x_j = [x_{j1}, x_{j2}, \dots, x_{ju}]^T$ and the expected output sample set is $t_j = [t_{j1}, t_{j2}, \dots, t_{jw}]^T$, the actual output of the ELM is:

$$\sum_{i=1}^v \beta_i g(w_i \cdot x_j + b_i) = o_j, j = 1, 2, \dots, M, \quad (15)$$

where $w_i = [w_{i1}, w_{i2}, \dots, w_{iu}]^T$ are the weights between the input neurons and the i^{th} hidden layer node; β_i are the weights between the i^{th} hidden layer node and the output layer neurons, $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{iw}]^T$; b_i are the biases of the i^{th} hidden layer node; and $g(\cdot)$ is the activation function of the ELM hidden layer.

When the number of hidden layer neurons, v , and the number of training set samples, M , of single-hidden layer feedforward neural network are equal, the neural network can approach the training samples with zero error, that is:

$$\sum_{j=1}^M \|o_j - t_j\|_2 = 0. \quad (16)$$

Then, we have β_i , w_i , and b_i , such that:

$$\sum_{i=1}^v \beta_i g(w_i \cdot x_j + b_i) = t_j, j = 1, 2, \dots, M. \quad (17)$$

Define:

$$H = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_M) \end{bmatrix}_{M \times v} = \begin{bmatrix} g(w_1 \cdot x_1 + b_1) & \cdots & g(w_v \cdot x_1 + b_v) \\ \vdots & \vdots & \vdots \\ g(w_1 \cdot x_M + b_1) & \cdots & g(w_v \cdot x_M + b_v) \end{bmatrix}_{M \times v}, \quad (18)$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_v^T \end{bmatrix}_{v \times w}, \quad (19)$$

$$T = \begin{bmatrix} t_1^T \\ \vdots \\ t_v^T \end{bmatrix}_{M \times w}. \quad (20)$$

Therefore, Equation (17) can be written as:

$$H\beta = T, \quad (21)$$

where H is the hidden layer output matrix of ELM.

When the activation function, $g(\cdot)$, of the hidden layer is infinitely differentiable, the weight matrix W of the input and hidden layers and the bias b of the hidden layer can be randomly determined before training and remain unchanged during training. In this case, the hidden layer output matrix H is a constant matrix. In this way, the solution of the weight matrix β between the hidden layer and the output layer can be transformed into the solution of the least-squares solution $\hat{\beta}$ of the linear equations $H\beta = T$, namely:

$$\hat{\beta} = H^+T, \quad (22)$$

where H^+ is the Moore–Penrose generalized inverse of the hidden layer output matrix H .

According to Karush–Kuhn–Tucker (KKT) theory, the solution of Equation (22) can be transformed into the solution of the Lagrangian function [36]:

$$\min L_{\text{ELM}} = \frac{1}{2} \|\beta\|_1^2 + \frac{C}{2} \sum_{i=1}^M \|\zeta_i\|_2^2 - \sum_{i=1}^M \alpha_i \cdot [h(x_i)\beta - t_i + \zeta_i], \quad (23)$$

where β is the output weight vector, C is a penalty factor, ζ_i is the training error, α_i are Lagrange operators, $h(x_i)$ is the feature mapping of the hidden layer, t_i are the target output values, and M is the number of samples.

Calculating the partial derivative of Equation (23), the output weights, $\hat{\beta}$, of the ELM can be solved as:

$$\hat{\beta} = H^T \left(HH^T + \frac{I}{C} \right)^{-1} T, \quad (24)$$

where I is the identity matrix of order v .

Finally, the output function of the ELM can be expressed as [37]:

$$f(x) = h(x)\hat{\beta} = h(x)H^T \left(HH^T + \frac{I}{C} \right)^{-1} T. \quad (25)$$

3.2. Theory of Kernel Extreme Learning Machine

As the weights between the input layer and the hidden layer and hidden layer biases of the ELM are randomly selected during training, a good training accuracy cannot be guaranteed every time; even the same training sample can have different results in two training processes. In order to solve this problem, Huang et al. proposed an improved ELM method in 2012; namely, kernel-ELM (KELM). The KELM method introduces a kernel function into the ELM method, which replaces the method of randomly selecting the input weights and hidden layer biases in original ELM method.

In Equation (21), the hidden layer output matrix, H , of ELM can be written as:

$$H = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_M) \end{bmatrix}_{M \times v}, \quad (26)$$

where $h(x_i)$ is the feature mapping of the hidden layer. HH^T in Equation (24) can be replaced by constructing a kernel function without knowing the specific form of $h(x_i)$:

$$HH^T(i, j) = K(x_i, x_j), \quad (27)$$

where $K(x_i, x_j)$ is the kernel function and $i, j \in (1, 2, \dots, M)$.

The ELM kernel matrix is defined as:

$$\Omega_{\text{ELM}} = HH^T = \begin{bmatrix} K(x_1, x_1) & \cdots & K(x_1, x_M) \\ \vdots & \vdots & \vdots \\ K(x_M, x_1) & \cdots & K(x_M, x_M) \end{bmatrix}. \quad (28)$$

Then, Equation (25) for the ELM can be changed into the form of KELM:

$$f(x) = \begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_M) \end{bmatrix} \left(\Omega_{\text{ELM}} + \frac{I}{C} \right)^{-1} T. \quad (29)$$

The KELM method was obtained by introducing a kernel method into the ELM method. Compared with basic ELM, the KELM method has better non-linear function approximation ability and superior data classification ability. In addition, KELM does

not need to randomly select input weights and hidden layer biases during training, only requiring the selection of appropriate kernel functions to overcome the shortcomings of extreme learning machines. In addition, KELM not only can overcome the dimensional disaster problem of ELM, but also has very high recognition accuracy.

3.3. Wavelet Kernel Extreme Learning Machine Theory

The basic use of a wavelet function is to fit any function with wavelets of different scales. As wavelet functions have strong mapping ability, a wavelet kernel can be used as the kernel function in the KELM method [38]. Therefore, a wavelet kernel function is constructed from a wavelet function in this paper, which is used as the kernel function of KELM. Thus, a WKELM is constructed.

Given a mother wavelet function, $\psi(x)$, whose scale factor and translation factor are a and b respectively, the wavelet basis function can be expressed as:

$$\psi_{a,b}(x) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{x-b}{a}\right). \quad (30)$$

According to tensor product theory, a multidimensional wavelet function can be expressed by tensor products of several one-dimensional wavelet functions:

$$\psi(x) = \prod_{i=1}^s \psi(x_i), \quad (31)$$

where s is the number of one-dimensional wavelet functions, when the multidimensional wavelet function is written in tensor product form; and x_i is the independent variable of the i^{th} one-dimensional wavelet function.

According to Equation (31), a translation-invariant kernel function can be constructed:

$$K(x, x') = \prod_{i=1}^s \psi\left(\frac{x_i - x'_i}{a}\right). \quad (32)$$

In this paper, the Morlet wavelet is used to construct the wavelet kernel function. The Morlet wavelet function is given as:

$$\psi(x) = \cos(1.75x) \exp\left(-\frac{x^2}{2}\right). \quad (33)$$

The wavelet kernel function constructed from the Morlet wavelet function is:

$$K(x, x') = \prod_{i=1}^s \left[\cos\left(1.75 \times \frac{x_i - x'_i}{a}\right) \exp\left(-\frac{(x_i - x'_i)^2}{2a^2}\right) \right], \quad (34)$$

where x_i are the i^{th} components of x .

As the wavelet kernel function is constructed using the translation-invariant kernel of the wavelet function, the wavelet kernel function inherits the multi-scale approximation characteristics of the wavelet function. Therefore, taking a wavelet kernel function as the kernel function in KELM will have a better effect on classification. The method flow is as follows:

- (1) The number of input neurons and output neurons are determined, according to the number of features and categories, and the WKELM classifier model is established;
- (2) Set the value of the penalty factor C and the parameter of the wavelet kernel a ;
- (3) The input characteristics of the training samples and their corresponding output categories are used to train the WKELM;
- (4) The features of the test sample are used as the input for the WKELM and, then, the trained WKELM classifier is used for the category decision.

4. Fault Feature Extraction Method of Axial Piston Pump Based on MEEMD and AR Spectrum

4.1. Test System and Data Acquisition

A hydraulic pump fault simulation test bed was used for the test. The vibration sensor used in this paper is piezoelectric and charge output type vibration sensor. The charge amplifier is used to amplify and convert the output charge of the vibration sensor, which is used to sample the vibration signal of the hydraulic pump shell in x , y and z directions. The pressure sensor is a piezoresistive pressure sensor, which is used to sample the pressure pulsation signal of the hydraulic oil at the outlet of the hydraulic pump. The main components used in the test are described in Table 1.

Table 1. Types and performance parameters of test components.

Num.	Name	Model	Performance Parameters
1	Motor	Y132M-4	Rated speed: 1480 rpm
2	Axial piston pump	MCY14-1B	Theoretical displacement: 10 mL/r rated pressure: 31.5 MPa; 7 pistons rated speed: 1500 rpm
3	Data acquisition card	USB-6221	Maximum sampling rate: 250 kS/s
4	Vibration sensor	YD-72D	Frequency range: 0.3 Hz–18 kHz Charge sensitivity: 0.35 pC/(m/s ²) Linear range of amplitude: 1000 m/s ²
5	Charge amplifier	DHF-10	Maximum output voltage: ± 10 V Gain: 0.1m V/pC–1 V/pC; Precision: < 1.5%
6	Pressure Sensor	SYB-351	Measuring range: 0–25 MPa Precision: 0.2%; Output range: 0–5 V
7	Rotational speed measurer	LT-XSMP	Measuring range: 6–45000 rpm

The installation of an experimental installation is shown in Figure 2. The vibration sensors are used to collect the vibration signals of the hydraulic pump in different states. The schematic diagram of the test system is shown in Figure 3. During the test, the different fault states of the hydraulic pump were simulated by replacing the plunger and slipper components of different fault types, as well as through use of a worn center spring. The outlet pressure of the hydraulic pump could be controlled by adjusting the pilot-operated proportional overflow valve 18. Vane pump 3 was used as a supplementary pump, in order to make up for the poor self-priming capacity of the piston pump.

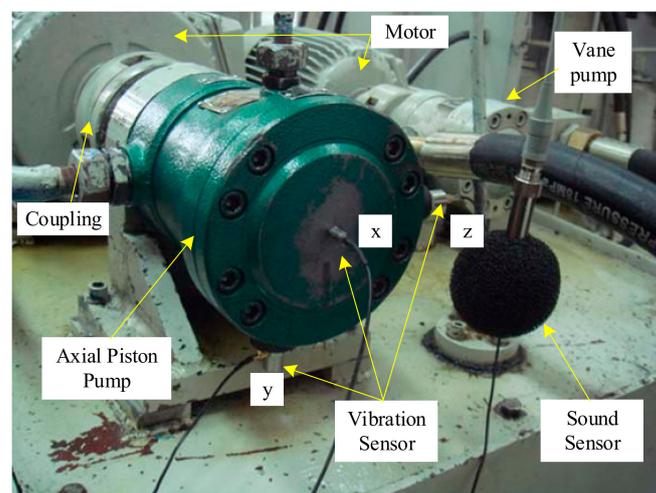
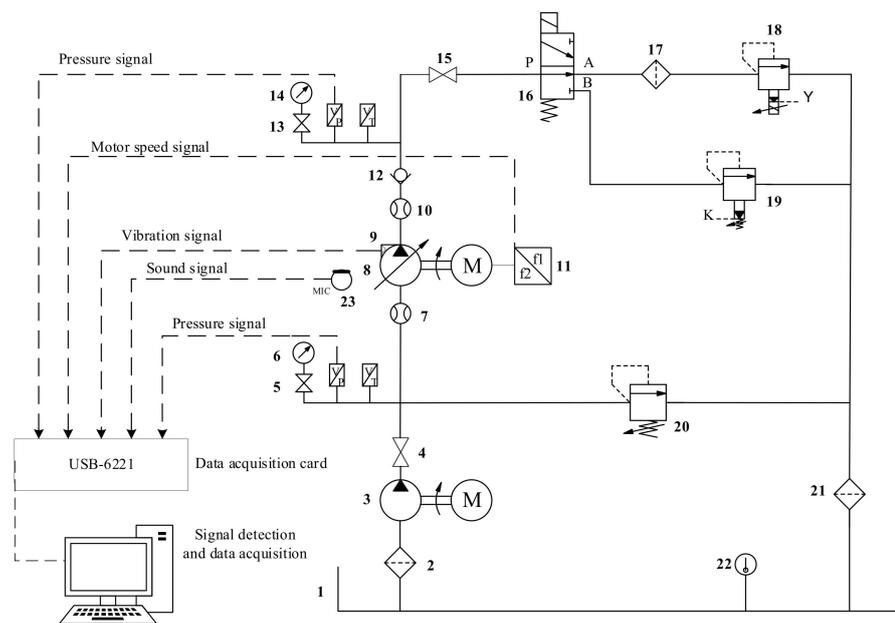


Figure 2. Installation of an experimental installation.



1-Oil tank; 2,21,17-filter; 3-vane pump; 4,15-stop valve; 5,13- pressure gauge switch; 6,14-pressure gauge; 7,10- flow meter; 8-axial piston pump; 9-vibration sensor; 11-frequency converter; 12-check valve; 16- two-position three-way solenoid valve; 18-pilot-operated proportional overflow valve; 19-pilot-operated overflow valve; 20-overflow valve; 22-temperature meter; 23-impulse sound level meter;

Figure 3. Schematic diagram of the test system.

During the test, a data acquisition system was required to display and store the vibration signal collected by the vibration sensor in real-time [39]. The data acquisition system used in this test was based on the LabVIEW software (National Instruments (NI), Austin, Texas, USA). As shown in Figure 4, the sampling frequency, sampling time, and controls for starting and stopping the data collection could be set using the front panel. At the same time, the collected pump speed, pressure, vibration, and other information could be displayed in real-time.

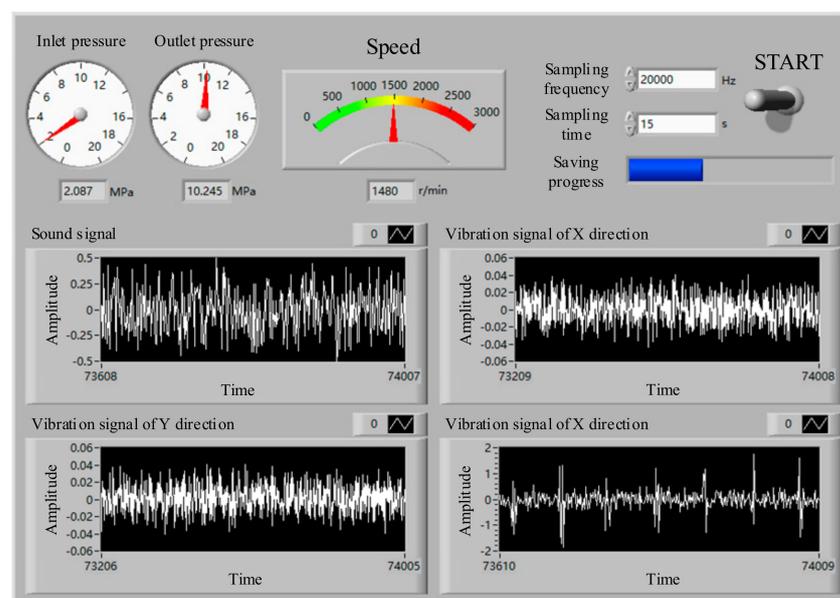


Figure 4. Front panel of the LabVIEW program.

The fault setting methods of the axial piston pump during the test are shown in Table 2. The process of the test was as follows: First, the three vibration sensors were fixed to the end cap and shell of the hydraulic pump, respectively, as shown in Figure 3. When the hydraulic system was unloaded, the test equipment was started. When the system reached normal speed, the system pressure was set to 10 MPa through the pilot-operated proportional overflow valve 18. The speed of the motor was set to 1480 r/min. Then, after the sampling frequency was set to 20 kHz in the data acquisition system, the data acquisition system was started, in order to collect and store the speed and other signals of the hydraulic pump in the normal state. Finally, the normal components were replaced by the fault components of the axial piston pump, and the vibration signal of the hydraulic pump in the fault state was acquired. After the data acquisition stage, the original signals of the hydraulic pump in normal state and three fault states were obtained. Some pictures of the fault components are shown in Figure 5.

Table 2. Fault setting method of axial piston pump.

Num.	States	Fault Setting Method
1	Normal working	—
2	Single slipper wear fault	Grind off a rounded corner of the slipper
3	Single slipper loosening fault	Replace the normal components with a slipper loosening fault components
4	Center spring worn fault	Grind off the center spring by 1.2 mm

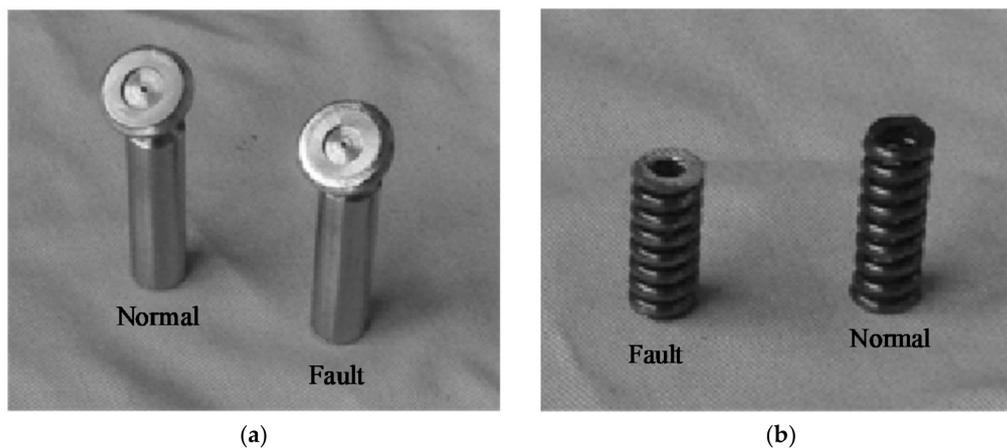


Figure 5. Some pictures of the fault elements: (a) Single slipper wear fault; and (b) Center spring worn fault.

4.2. Simulation Signal Analysis of MEEMD Method

A simulation signal, $x(t)$, with a sampling frequency of 1000 Hz and a sampling time of 2s was constructed:

$$x(t) = x_1(t) + x_2(t) + n(t), \quad (35)$$

where $x_1(t)$ is a sine signal, $x_2(t)$ is an amplitude modulation signal, and $n(t)$ are two intermittent random signals:

$$x_1(t) = 2 \sin\left(2\pi \cdot 30t + \frac{\pi}{2}\right), \quad (36)$$

$$x_2(t) = (t + 1)2 \sin\left(2\pi \cdot 8t + \frac{\pi}{3}\right). \quad (37)$$

Figure 6 shows the time-domain diagram of the synthesized signal and its components. The synthesized signal was decomposed by EEMD, and the decomposition results obtained are shown in Figure 7.

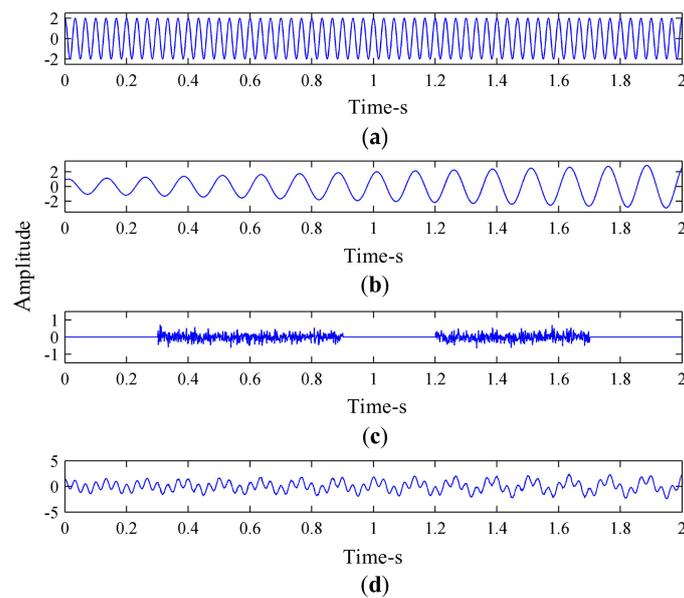


Figure 6. Time-domain diagram of the synthesized signal and its components: (a) Sine signal $x_1(t)$; (b) Amplitude modulation signal $x_2(t)$; (c) Intermittent random signals $n(t)$; and (d) Synthetic simulation signal $x(t)$.

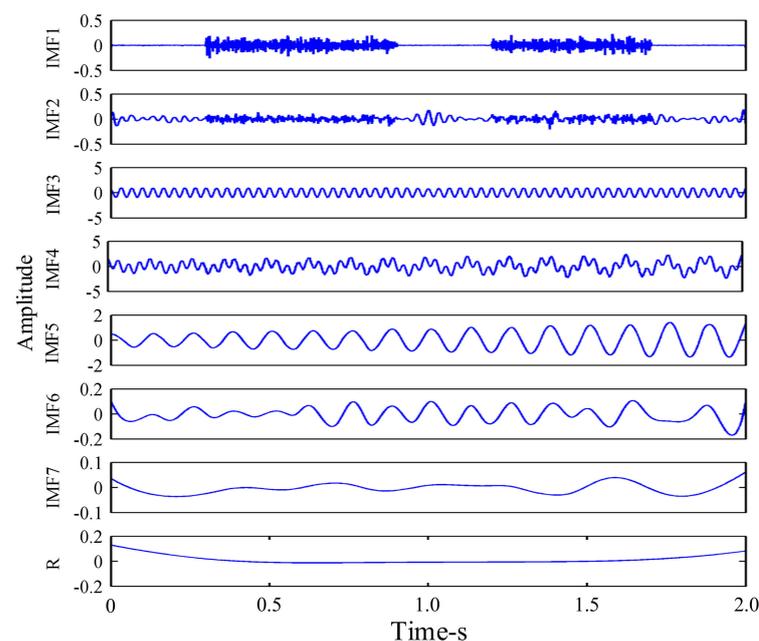


Figure 7. Ensemble empirical mode decomposition (EEMD) decomposition of the synthesized signal.

The amplitude of the added Gaussian white noise was $a_1 = 0.2$, and the total number of noises added was $N_e = 100$. The results obtained by EEMD decomposition of the signal are shown in Figure 7. It can be seen, from Figure 7, that EEMD decomposed many pseudo-components. After EEMD decomposition of the simulation signal $x(t)$, seven IMF components and one residual component were decomposed, among which IMF3 and IMF5, respectively, represent the sine signal $x_1(t)$ and the amplitude modulation signal $x_2(t)$ in the synthetic simulation signal $x(t)$. The simulation signal $x(t)$ was also decomposed by MEEMD, as shown in Figure 8. It can be seen, from Figure 3, that three IMF components and one residual component were decomposed, among which IMF2 and IMF3, respectively, represent $x_1(t)$ and $x_2(t)$ in the synthetic simulation signal $x(t)$. Compared with EEMD, MEEMD can effectively eliminate the pseudo-components.

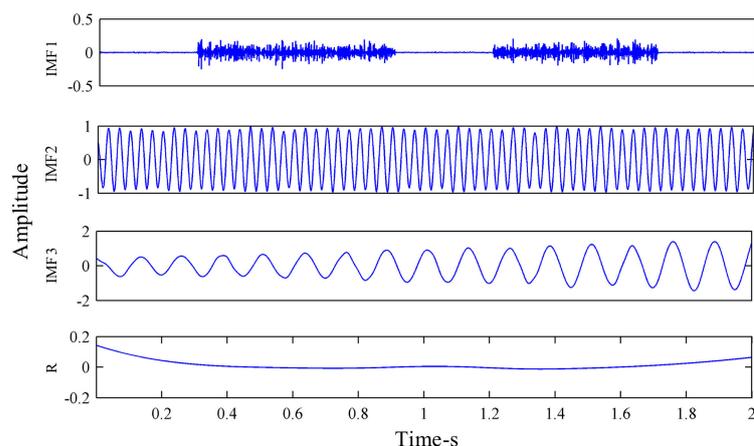


Figure 8. Modified ensemble empirical mode decomposition (MEEMD) decomposition of the synthesized signal.

4.3. Fault Feature Extraction Process

In this paper, a fault feature extraction method based on MEEMD and AR spectrum is proposed. First, the non-stationary vibration signals are decomposed by MEEMD, in order to obtain some IMF components. Secondly, the effective IMF components are screened by the Pearson correlation coefficient method, then analyzed by the AR spectrum method. Finally, the calculated AR spectrum energy is taken as the fault feature of the signal. Compared with the Hilbert-Huang transform (HHT) method, the feature extraction method based on MEEMD and AR spectrum proposed in this paper not only can solve the mode mixing phenomenon of EMD decomposition by the HHT method but can also overcome the limitation of Hilbert analysis in HHT method. The specific steps of this method are as follows:

- (1) The signal is decomposed by the MEEMD method, from which several IMF components are obtained.
- (2) As there may still be pseudo-components in the IMF components obtained through MEEMD decomposition, it is necessary to eliminate such components. In this paper, the pseudo-components are eliminated by the Pearson correlation coefficient method. First, the Pearson correlation coefficient r between all IMF components and the original signal is calculated. Then, t IMF components with large correlation coefficients are taken as the effective components. The Pearson correlation coefficient r is calculated as follows:

$$r = \frac{\sum_{i=1}^n [x(i) - \bar{x}][y(i) - \bar{y}]}{\sqrt{\sum_{i=1}^n [x(i) - \bar{x}]^2} \sqrt{\sum_{i=1}^n [y(i) - \bar{y}]^2}}, \quad (38)$$

where $x(i)$ is the original signal, $y(i)$ is each order of IMF components, $i = 1, 2, \dots, n$, n is the signal length, \bar{x} is the average of $x(i)$, and \bar{y} is the average of $y(i)$.

- (3) AR spectrum estimation of the t IMF components is carried out and the AR spectrum energies of the t IMF components are calculated. After normalization, the characteristic matrix, P , of the signal is composed as follows:

$$P = [p_1, p_2, \dots, p_M]_{t \times M}, \quad (39)$$

where $p_i = [p_{1i}, p_{2i}, \dots, p_{ti}]^T$, $i = 1, 2, \dots, M$, and p_{ki} is the AR spectrum energy of the i^{th} IMF component of the k^{th} sample.

4.4. Analysis of Acquired Hydraulic Pump Signals

The x -direction vibration signal of a single slipper loosening fault was selected for analysis. Figure 9 shows the time-domain diagram of the vibration signal in the state of the single slipper loosening fault of the axial piston pump.

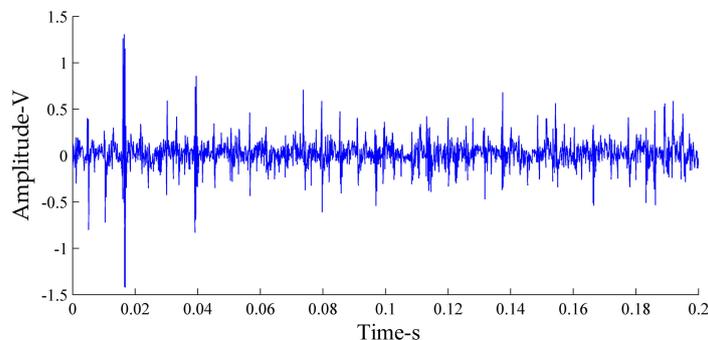


Figure 9. Time-domain diagram of the vibration signal in the single slipper loosening fault state.

In order to show that the MEEMD method had a better effect than the EEMD method in the signal processing of the hydraulic pump, the number of IMF components and the index of orthogonality (IO) decomposed by the two methods were compared. Figure 10 shows the vibration signal decompositions of hydraulic pump in the single slipper loosening fault state by the two methods. Table 3 gives the IO of EEMD and MEEMD methods. Orthogonality refers to the correlation between different components. The orthogonality of the signal can be measured by calculating the IO , after which the degree of modal mixing between the components can be judged. The larger the IO value, the worse the orthogonality; that is, the more serious the degree of modal mixing. The Equation for IO of IMF components is [11]:

$$IO = \sum_{i=0}^n \left[\sum_{j=1}^I \sum_{k=1}^I \frac{c_j(i)c_k(i)}{x^2(i)} \right], \quad (40)$$

where $j \neq k$, $c_j(i)$ is the j^{th} IMF component, $c_k(i)$ is the k^{th} IMF component, $x(i)$ is the original sequence, n is the length of the time-series, I is the number of IMF components, $j = 1, 2, \dots, I$, and $k = 1, 2, \dots, I$.

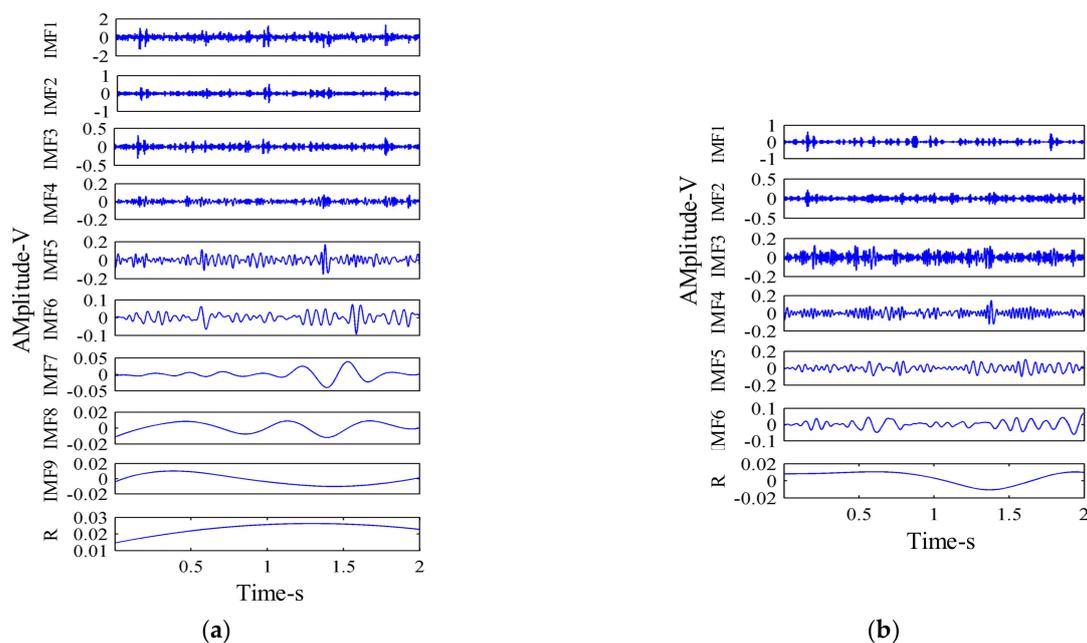


Figure 10. Vibration signal decomposition of hydraulic pump: (a) EEMD; and (b) MEEMD.

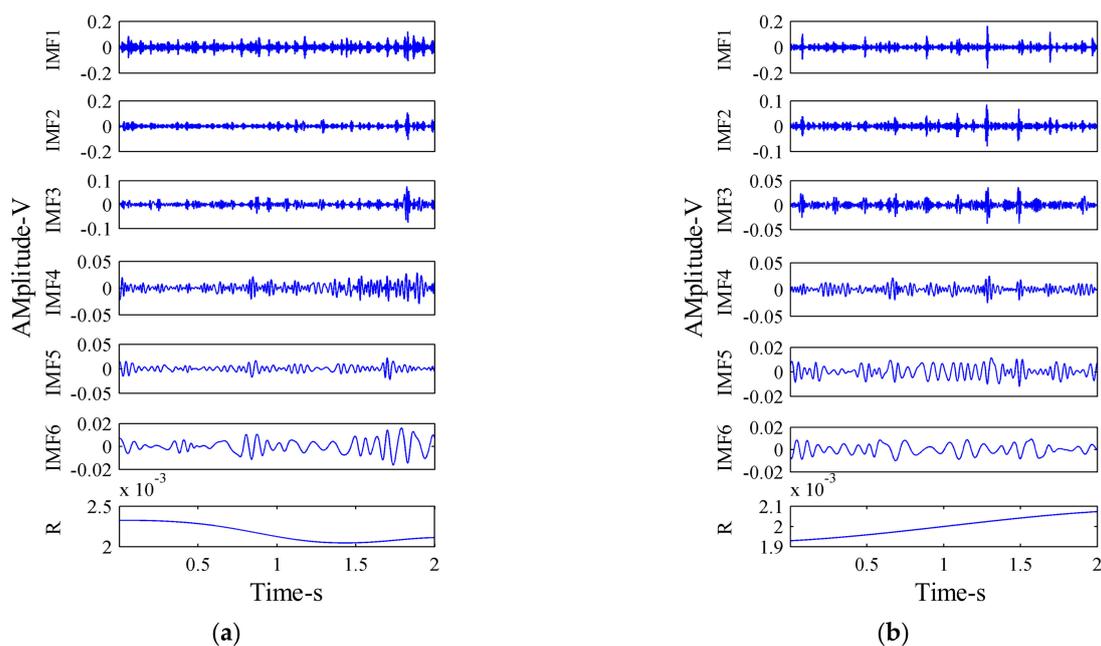
Table 3. Comparison of Ensemble empirical mode decomposition (EEMD) and Modified ensemble empirical mode decomposition (MEEMD) indicators.

Method	IO	Number of IMF Components
EEMD	0.2069	9
MEEMD	0.1131	6

It can be seen, from Figure 10, that nine IMF components were decomposed by the EEMD method, while six IMF components were decomposed by the MEEMD method. It is obvious that the MEEMD method addressed the problem of excessive false modal components, as was obvious in the EEMD method. It can be seen, from Table 3, that the IO value for the MEEMD method was smaller than that of the EEMD method. This indicates that the MEEMD method is more effective than the EEMD method in suppressing the phenomenon of modal mixing.

4.5. Fault Feature Extraction of Hydraulic Pump Based on MEEMD-AR Spectrum

First, the x -direction vibration signals of the hydraulic pump in four states were decomposed by the MEEMD method, and several IMF components were obtained. The vibration signals of the hydraulic pump in normal working state and single slipper wear fault state were decomposed by MEEMD, respectively. The decomposition results are shown in Figure 11.

**Figure 11.** MEEMD decomposition results of vibration signal: (a) Normal working state; and (b) Single slipper wear fault state.

The Pearson correlation coefficient method was used to calculate the correlation coefficient between all IMF components decomposed by MEEMD and the original signal. Figure 12 shows the mean value of correlation coefficients between each IMF component decomposed by the MEEMD method and the original signal of 50 samples in each of the four hydraulic pump states. It can be seen, from Figure 12, that the correlation coefficient between the first IMF component and the original signal was more than 0.75, the correlation coefficient between the second IMF component and the original signal was basically more than 0.4, the correlation coefficient between the third IMF component and the original signal was more than 0.34, and the correlation coefficient between the fourth IMF component and the original signal is more than 0.3. The correlation coefficients between the other IMF components and the original signal were all less than 0.3.

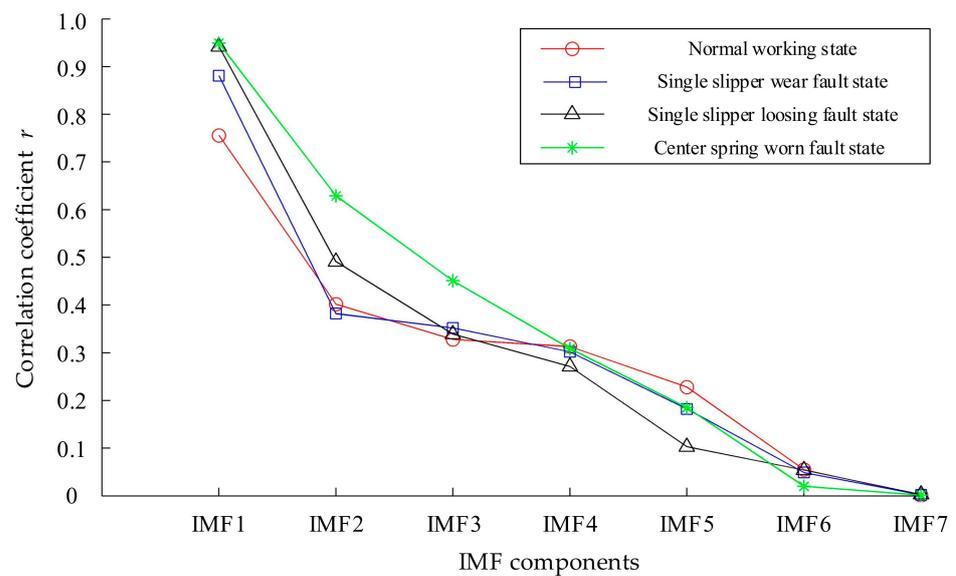


Figure 12. Correlation coefficients between intrinsic mode function (IMF) components of hydraulic pump and original signal.

Considering the complexity and accuracy of the method, the components whose correlation coefficient between the IMF component and original signal was greater than 0.3 were selected as the effective components in this paper. It can be seen, from the analysis of Figure 12, that the correlation coefficients between the first four IMF components and the original signal in the four states of the hydraulic pump were greater than 0.3. Therefore, AR spectrum analysis was only performed for the first four IMF components decomposed by MEEMD in various states, as shown in Figure 13.

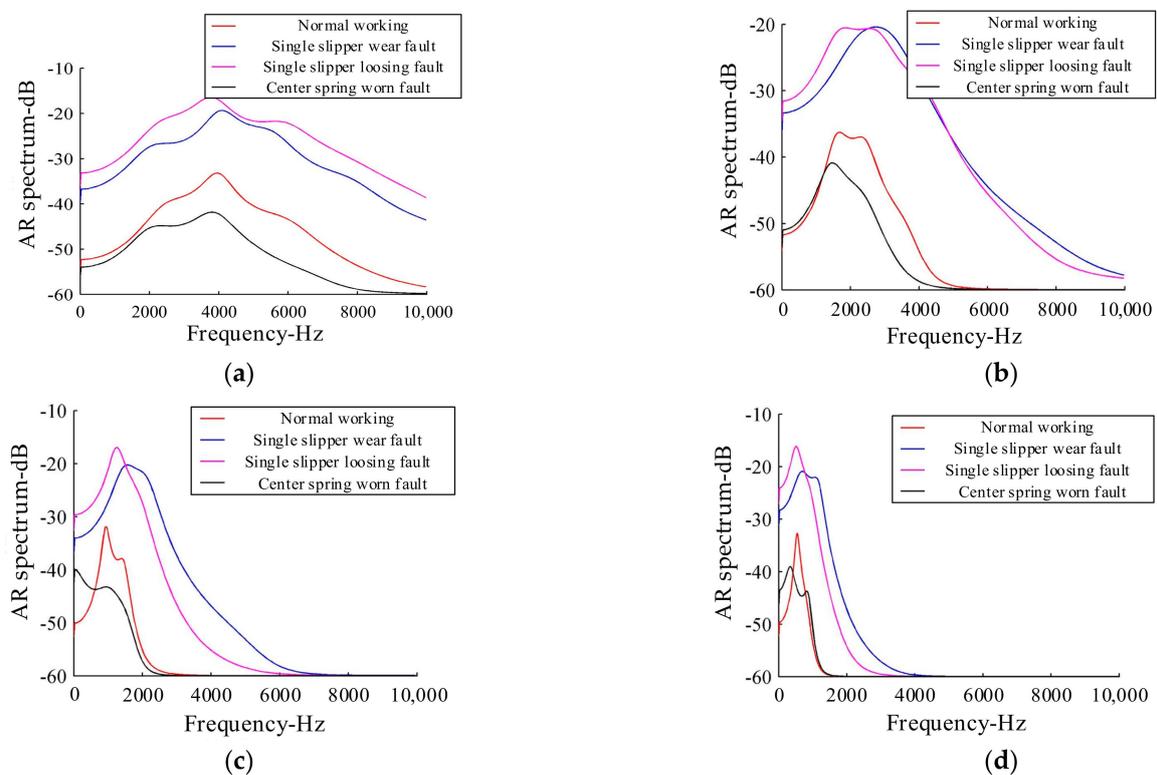


Figure 13. Autoregressive (AR) power spectra of the first four order IMF components: (a) IMF 1; (b) IMF 2; (c) IMF 3; and (d) IMF 4.

It can be seen from Figure 13 that the AR spectra of the same-order IMF components of hydraulic pumps in the four states were different, where there were different characteristics among the different IMF components. In addition, the maps obtained by AR spectrum analysis were not only smooth, but also easy to distinguish. The AR spectrum energies of different states (i.e., corresponding to each IMF component in Figure 12) were calculated, and the results are shown in Table 4.

Table 4. Autoregressive (AR) spectrum energies of intrinsic mode function (IMF) components in different states.

State	IMF1	IMF2	IMF3	IMF4
Normal working	0.2102	0.0712	0.0374	0.0210
Single slipper wear fault	0.1156	0.0316	0.0151	0.0106
Single slipper losing fault	4.7201	1.6651	1.3350	0.9085
Center spring worn fault	8.9050	3.5852	2.8181	1.6852

It can be intuitively seen, from Table 4, that the feature values extracted by different states had different distributions, demonstrating the effectiveness of the feature extraction method, to a certain extent.

5. Fault Diagnosis Method of Hydraulic Pump Based on WKELM

Equipment will inevitably fail in the process of operation. In order to identify the equipment state information from the vibration signal, a fault diagnosis method based on the integration of MEEMD, AR spectrum, and WKELM methods is proposed in this paper, in order to identify the equipment state. Figure 14 shows the fault diagnosis flow chart, based on the MEEMD–AR–WKELM integration method.

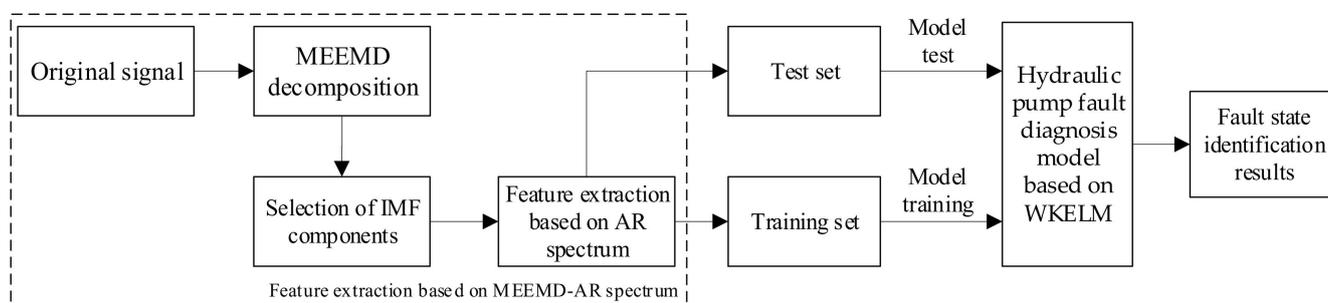


Figure 14. Fault diagnosis flow chart based on the MEEMD–AR–WKELM integration method.

5.1. Hydraulic Pump Fault Diagnosis Based on WKELM

The z-direction vibration signal of the axial piston pump was selected for analysis. A sample was intercepted every 2000 data points of the hydraulic pump vibration signal in four states, where 50 samples were taken in each state. The training sample set was composed of 30 samples selected from 50 samples in each state, while the test sample set was composed of the remaining 20 samples. Two groups of AR spectrum energies in four different states of the hydraulic pump were taken to draw the AR spectrum energy distribution diagram of IMF components in different states, as shown in Figure 15.

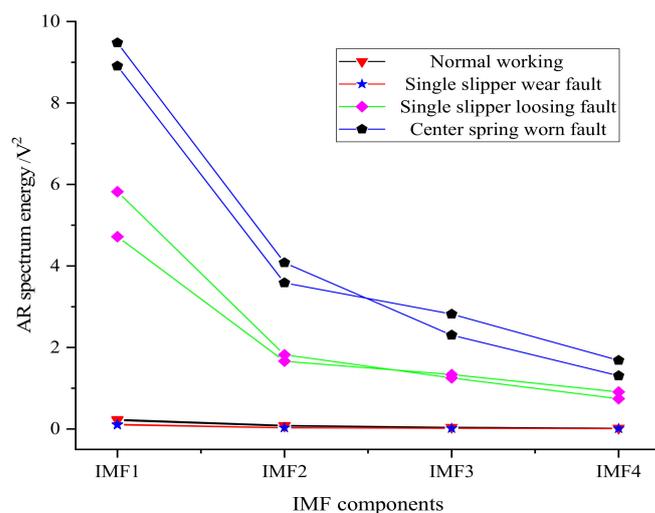


Figure 15. AR spectrum energy distribution of IMF components in different states.

It can be seen from Figure 15 that the features extracted from the different hydraulic pump working states had different distributions, while the features taken from the same working state had a similar distribution, thus indicating that the feature extraction method had a good effect.

The features of 200 samples in the four states of the axial piston pump were extracted, and the 4-dimensional fault features of each sample were extracted. A total of 30 samples were selected from 50 samples in each state, in order to form the training sample matrix, $P_{120 \times 4}^{train}$, while the remaining 20 samples were used to form the test sample matrix, $P_{80 \times 4}^{test}$. After the extraction of feature vectors, the WKELM method was used to identify the state of the hydraulic pump. The number of input neurons was $u = 4$ and the number of output neurons was $w = 4$ for the WKELM method. In consideration of the influence of the WKELM parameters on diagnosis accuracy, the penalty factor C and wavelet kernel coefficient a were analyzed. The influences of C and a on the recognition accuracy of WKELM are shown in Figure 16.

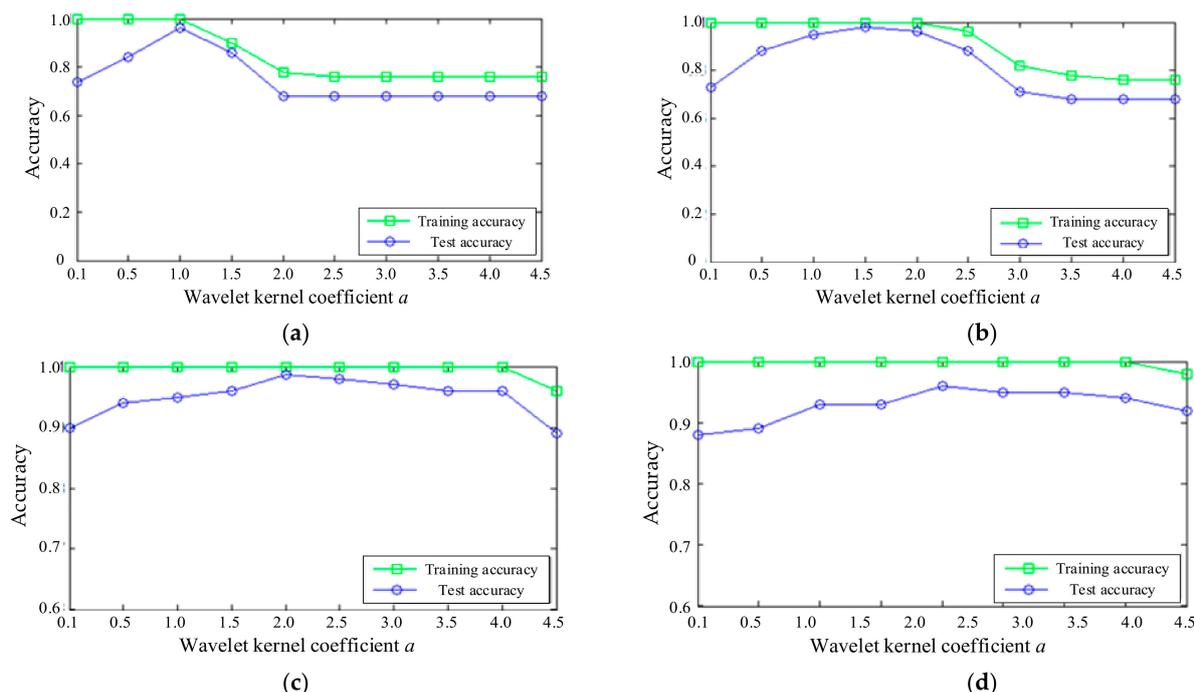


Figure 16. The influence of parameters on accuracy: (a) $C = 10$; (b) $C = 100$; (c) $C = 1000$; and (d) $C = 10,000$.

It can be seen from Figure 16, when $C=10$, the training accuracy reaches 100%. When $a \leq 1$, and it gradually decreases and then reaches a plateau. When $a > 1$, the test accuracy first increases and then decreases, reaching the maximum 98% at $a = 1$. When $C = 100$, the training accuracy reaches 100%. When $a \leq 2$, and it gradually decreases. When $a > 2$, the test accuracy first increases and then decreases, reaching the maximum 99% at $a = 1.5$. When $C = 1000$, the training accuracy is 100%; when $a \leq 4.5$, the test accuracy first increases and then decreases, reaching the maximum 98% at $a = 1.5$. When $C = 10000$, although the training accuracy is generally high, the test accuracy is very low. In general, when the value range of a is between 1 and 2.5, the training accuracy and the test accuracy are relatively high. After taking values within this range many times, we finally found that when $a = 1.9$ and $C = 500$, both the training accuracy and the test accuracy reached 100%. Therefore, the parameters of WKELM are determined as $a = 1.9$ and $C = 500$.

After determining the parameters of WKELM, the WKELM model was trained. The characteristic matrix, $P_{120 \times 4}^{train}$, of the training samples was taken as the input of WKELM, while the labels corresponding to each state were taken as the output (the label for normal working was 1, the label for single slipper wear fault was 2, the label for single slipper loosening fault was 3, and the label for center spring worn fault was 4). Finally, the characteristic matrix, $P_{80 \times 4}^{test}$, of the test samples was tested. The obtained recognition results are shown in Figure 17.

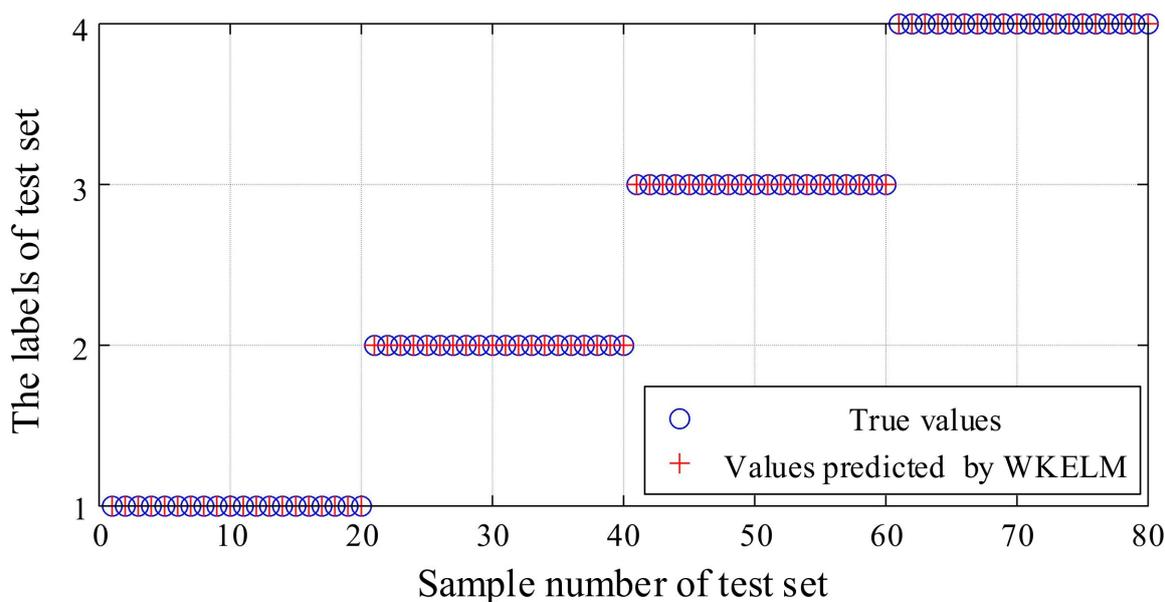


Figure 17. Fault diagnosis of hydraulic pump based on wavelet kernel extreme learning machine (WKELM).

It can be seen, from Figure 17, that all the predicted values identified by WKELM were the same as the true values; that is, the accuracy of the recognition results was 100%. This demonstrates that the proposed fault diagnosis method, based on the integration of MEEMD, AR spectrum, and WKELM methods, is effective.

5.2. Comparison and Analysis of Different Hydraulic Pump Fault Diagnosis Methods

In order to prove that WKELM has more advantages than the BP neural network, SVM, and ELM methods, hydraulic pump fault diagnosis was carried out using these methods under the same conditions. Among them, the number of neurons in the input, hidden, and output layers of BP neural network were four, 15, and four, respectively, and the activation function was the Sigmoid function. The LibSVM toolkit, provided by the Taiwanese scholar Lin Chih-Jen, was selected for SVM. The number of hidden layer neurons for the ELM was 14 and the activation function was the sigmoid function. The diagnosis accuracies of the four fault diagnosis methods are shown in Table 5. The training and testing times of the four methods were also compared, as shown in Table 6.

Table 5. Comparison of diagnostic accuracy of four fault diagnosis methods.

Condition of Hydraulic Pump		Normal Working	Single Slipper Wear Fault	Single Slipper Loosing Fault	Center Spring Worn Fault	Total
MEEMD-BP	Number of samples correctly diagnosed	20	20	20	19	79
	Recognition accuracy	100%	100%	100%	95%	98.75%
MEEMD-SVM	Number of samples correctly diagnosed	20	19	19	20	78
	Recognition accuracy	100%	95%	95%	100%	97.5%
MEEMD-ELM	Number of samples correctly diagnosed	20	19	20	20	79
	Recognition accuracy	100%	95%	100%	100%	98.75%
MEEMD-WKELM	Number of samples correctly diagnosed	20	20	20	20	80
	Recognition accuracy	100%	100%	100%	100%	100%

Table 6. Comparison of training time and testing time of the four fault diagnosis methods.

Fault Diagnosis Method	MEEMD-WKELM	MEEMD-ELM	MEEMD-SVM	MEEMD-BP
Training time (s)	0.0020	0.0054	0.058	0.749
Testing time (s)	0.0011	0.0019	0.0067	0.0042

It can be seen from Tables 5 and 6 that in terms of fault recognition accuracy, WKELM had the highest fault recognition accuracy, followed by the BP neural network and ELM, while the SVM had the lowest fault recognition accuracy. In terms of time, WKELM had the shortest time, followed by ELM, while the BP neural network took the longest time. It can be seen that WKELM performed better than the other three methods, in terms of fault recognition accuracy and time consumption.

The Case Western Reserve University dataset was used to demonstrate the fault diagnosis method proposed in this paper has higher fault recognition accuracy. The sampling frequency of the data is 12 kHz, the model of the bearing is SKF6205, the damage diameter is 0.36 mm, the motor loads is 2 horsepower. Take the normal working state, rolling ball fault state, inner race fault state, and outer race fault data, 0.2 s for each state data as a sample, take 50 samples under each state, of which 30 as the training sample, 20 as test samples, the results as shown in Table 7. It can be seen that the fault diagnosis method proposed in this paper has a high recognition accuracy.

Table 7. Fault diagnosis of rolling bearings based on wavelet kernel extreme learning machine (WKELM).

Condition of Rolling Bearing	Normal Working	Inner Race Fault	Rolling Ball Fault	Outer Race Fault	Total
Number of samples correctly diagnosed	20	20	20	20	80
Recognition accuracy	100%	100%	100%	100%	100%

In order to prove that the advantages of the fault diagnosis method proposed in this paper, hydraulic pump fault diagnosis was carried out using these methods under the same conditions. The diagnosis accuracies of the fault diagnosis methods [40,41] are shown in Table 8. Therefore, the advantages of the fault diagnosis method proposed in this paper, based on the integration of MEEMD, AR spectrum, and WKELM, were proved.

Table 8. Comparison of diagnostic accuracy of fault diagnosis methods.

Fault Diagnosis Method	MEEMD–WKELM	IMSE-KFCM [40]	RQA-KFCM [41]
Recognition accuracy	100%	98.75%	98%

6. Conclusions

In this paper, an axial piston pump fault diagnosis method based on MEEMD, AR spectrum, and WKELM integration was proposed and studied. Our main conclusions are as follows:

- (1) Compared with the EEMD method, the MEEMD method can better suppress the phenomenon of mode mixing in hydraulic pump vibration signal decomposition. At the same time, the MEEMD method has better orthogonality and fewer pseudo-components;
- (2) The hydraulic pump fault feature extraction method based on the MEEMD and AR spectrum methods proposed in this paper can effectively solve the problem of low efficiency inherent to traditional feature extraction methods. This method can effectively extract the fault features from the vibration signals of a hydraulic pump;
- (3) The WKELM method was improved and introduced for the fault diagnosis of a hydraulic pump, and a full hydraulic pump fault diagnosis method was proposed based on MEEMD–AR–WKELM integration. The diagnosis accuracy of different hydraulic pump fault states using this method was 100%. Compared with the BP, SVM, and ELM methods, the fault diagnosis method proposed in this paper was shown to have higher fault identification accuracy and faster identification speed.

Author Contributions: Conceptualization and methodology, Z.L. and W.J.; software, Y.S.; validation, S.Z. (Sheng Zhang) and W.J.; formal analysis, Z.L. and S.Z. (Shuqing Zhang); writing—original draft preparation, Z.L. and Y.S.; writing—review and editing, Z.L.; visualization, S.Z. (Sheng Zhang); supervision, S.Z. (Shuqing Zhang) and W.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China (Grant No. 51875498), Key Project of Natural Science Foundation of Hebei Province, China (Grant No. E2018203339, F2020203058).

Institutional Review Board Statement: Not applicable, this study not involving humans or animals.

Informed Consent Statement: Not applicable, this study not involving humans.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

1. Li, C.; Sánchez, R.V.; Zurita, G.; Cerrada, M.; Cabrera, D. Fault Diagnosis for Rotating Machinery Using Vibration Measurement Deep Statistical Feature Learning. *Sensors* **2016**, *16*, 895. [\[CrossRef\]](#)
2. Liu, T.; Luo, Z.; Huang, J.; Yan, S. A Comparative Study of Four Kinds of Adaptive Decomposition Algorithms and Their Applications. *Sensors* **2018**, *18*, 2120. [\[CrossRef\]](#)
3. Jiang, W.; Li, Z.; Lei, Y.; Zhang, S.; Tong, X. Deep learning based rolling bearing fault diagnosis and performance degradation degree recognition method. *J. Yanshan Univ.* **2020**, *44*, 526–536. [\[CrossRef\]](#)
4. Chen, J.; Wang, J.; Zhu, J.; Lee, T.; Silva, C. Unsupervised Cross-domain Fault Diagnosis Using Feature Representation Alignment Networks for Rotating Machinery. *IEEE/ASME Trans. Mechatron.* **2020**, *99*, 1. [\[CrossRef\]](#)
5. Peng, X.; Chen, J. The Future Trends of Hydraulics. *China Hydraul. Pneum.* **2007**, *3*, 1–5.
6. Ge, X.; Lin, Y. Current Status and Development Tendency of Hydraulic System Fault Diagnosis Technology. *China Hydraul. Pneum.* **2008**, *7*, 1–3.
7. Chen, J.; Li, T.; Wang, J.; Silva, C. WSN Sampling Optimization for Signal Reconstruction Using Spatiotemporal Autoencoder. *IEEE Sens. J.* **2020**, *99*, 14290–14301. [\[CrossRef\]](#)

8. Chen, J.; Li, T.; Wang, J.; Silva, C. Optimization of Wireless Sensor Network Deployment for Spatiotemporal Reconstruction and Prediction. *Electr. Eng. Syst. Sci.* **2019**, *1*, 1–11.
9. Jiang, W.; Wang, Z.; Zhu, Y.; Hu, H.; Zhang, B. Isomap and intrinsic mode function feature energy-based multi-fault recognition method of rolling bearing. *ICIC Express Lett. Part B Appl.* **2016**, *10*, 279–285.
10. Zhu, Y.; Jiang, W.; Kong, X.; Zheng, Z.; Hu, H. An Accurate Integral Method for Vibration Signal Based on Feature Information Extraction. *Shock Vib.* **2015**, *2015*. [[CrossRef](#)]
11. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.C.; Tung, C.C.; Liu, H.H. The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-stationary Time Series Analysis. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **1998**, *454*, 903–995. [[CrossRef](#)]
12. Ahn, J.; Kwak, D.; Koh, B. Fault Detection of a Roller-Bearing System through the EMD of a Wavelet Denoised Signal. *Sensors* **2014**, *14*, 15022–15038. [[CrossRef](#)] [[PubMed](#)]
13. Wu, Z.H.; Huang, N.E. Ensemble Empirical Mode Decomposition: A Noise-assisted Data Analysis Method. *Adv. Adapt. Data Anal.* **2009**, *1*, 1–41. [[CrossRef](#)]
14. Zheng, J.; Cheng, J.; Yang, Y. Modified EEMD algorithm and its applications. *J. Vib. Shock* **2013**, *32*, 21–26, 46. [[CrossRef](#)]
15. Jiang, S. Fault Diagnosis Analysis of Fan Spindle Bearing Based on CPSO-BBO Optimization SVM. Master's Thesis, Northeast Electric Power University, Jilin, China, 2019.
16. Zheng, X. Research on Blind Separation and Noise Source Identification for the Vibro-acoustic Signals of Vehicle and Engine. Ph.D. Thesis, Zhejiang University, Hangzhou, China, 2012.
17. Shi, Y.; Wang, Y.; Jiang, Y.; Gao, H.; Xiang, J. Rolling Bearing Fault Diagnosis Method Based on MEEMD. *Coal Mine Mach.* **2014**, *35*, 266–268. [[CrossRef](#)]
18. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks. In Proceedings of the 2004 IEEE International Joint Conference on Neural Networks—Proceedings, Budapest, Hungary, 25–29 July 2004.
19. Huang, G.B.; Chen, L.; Siew, C.K. Universal Approximation Using Incremental Constructive Feedforward Networks with Random Hidden Nodes. *IEEE Trans. Neural Netw. Learn. Syst.* **2006**, *17*, 879–892. [[CrossRef](#)]
20. Huang, G.B.; Chen, L. Convex Incremental Extreme Learning Machine. *Neurocomputing* **2007**, *70*, 3056–3062. [[CrossRef](#)]
21. Zhang, Y.; Liu, Y.; Chao, H.; Zhang, Z.; Zhang, Z. Classification of Incomplete Data Based on Evidence Theory and an Extreme Learning Machine in Wireless Sensor Networks. *Sensors* **2018**, *18*, 1046. [[CrossRef](#)]
22. José, M.M.; Pablo, E.; Emilio, S.; José, D.M.; Rafael, M.; Juan, G. Regularized Extreme Learning Machine for Regression Problems. *Neurocomputing* **2011**, *74*, 3716–3721. [[CrossRef](#)]
23. Bi, X.; Zhao, X.; Wang, G.; Zhang, P.; Wang, C. Distributed Extreme Learning Machine with Kernels Based on MapReduce. *Neurocomputing* **2015**, *149*, 456–463. [[CrossRef](#)]
24. Cho, D.; Ham, J.; Oh, J.; Park, J.; Kim, S.; Lee, N.; Lee, B. Detection of Stress Levels from Biosignals Measured in Virtual Reality Environments Using a Kernel-Based Extreme Learning Machine. *Sensors* **2017**, *17*, 2435. [[CrossRef](#)]
25. Zhang, S.; Zhang, T.; Yin, Y.; Xiao, W. Alumina Concentration Detection Based on the Kernel Extreme Learning Machine. *Sensors* **2017**, *17*, 2002. [[CrossRef](#)]
26. Yang, X.; Ma, Z.; Shen, H.; Chen, H. Fault diagnosis of airflow jamming fault in double circulating fluidized bed based on multi-scale feature energy and KELM. *Ciesc J.* **2019**, *70*, 2616–2625. [[CrossRef](#)]
27. Zhang, W.; Ma, Y.; Dong, Z.; Du, J. Fault diagnosis of coal mill based on PSO-KELM. *Electr. Power Sci. Eng.* **2018**, *34*, 54–58. [[CrossRef](#)]
28. Pang, S.; Yang, X.; Zhang, Y.; Wei, X. Application of Deep Kernel Extreme Learning Machine in Aero Engine Components Fault Diagnosis. *J. Propuls. Technol.* **2017**, *38*, 2613–2621. [[CrossRef](#)]
29. Jiang, W.; Zheng, Z.; Zhu, Y.; Li, Y. Demodulation for Hydraulic Pump Fault Signals Based on Local Mean Decomposition and Improved Adaptive Multiscale Morphology Analysis. *Mech. Syst. Signal Process.* **2015**, *58–59*, 179–205. [[CrossRef](#)]
30. Zheng, X.; Hao, Z.; Lu, Z.; Yang, J. Separation of piston-slap and combustion shock excitations via MEEMD method. *J. Vib. Shock* **2012**, *31*, 109–113. [[CrossRef](#)]
31. Massimiliano, Z.; Luciano, Z.; Osvaldo, A.R.; David, P. Permutation Entropy and Its Main Biomedical and Econophysics Applications: A Review. *Entropy* **2012**, *14*, 1553–1577. [[CrossRef](#)]
32. Kuai, M.; Cheng, G.; Pang, Y.; Li, Y. Research of Planetary Gear Fault Diagnosis Based on Permutation Entropy of CEEMDAN and ANFIS. *Sensors* **2018**, *18*, 782. [[CrossRef](#)]
33. Bandt, C.; Pompe, B. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.* **2002**, *88*, 174102. [[CrossRef](#)]
34. Sun, J.; Xu, X.; Liu, Y.; Zhang, T.; Li, Y. FOG Random Drift Signal Denoising Based on the Improved AR Model and Modified Sage-Husa Adaptive Kalman Filter. *Sensors* **2016**, *16*, 1073. [[CrossRef](#)]
35. Jiang, W.; Li, Z.; Li, J.; Zhu, Y.; Zhang, P. Study on a Fault Identification Method of the Hydraulic Pump Based on a Combination of Voiceprint Characteristics and Extreme Learning Machine. *Processes* **2019**, *7*, 894. [[CrossRef](#)]
36. He, C.; Liu, C.; Wu, Y.; Wu, T.; Chen, T. Improved Grey Wolf Optimization Algorithm for Life Prediction of Medical Lithium Batteries. *J. Chongqing Norm. Univ. (Nat. Sci.)* **2019**, *36*, 21–28. [[CrossRef](#)]
37. Li, Q.; Yu, Y.; Zheng, S.; Ma, T. Short-term Power Load Forecasting Based on EWT-WKELM. *Proc. CSU-EPSA* **2018**, *30*, 83–89. [[CrossRef](#)]

-
38. Zhang, W.; Guo, W.; Zhang, C.; Zhao, S. An Online Calibration Method for a Galvanometric System Based on Wavelet Kernel ELM. *Sensors* **2019**, *19*, 1353. [[CrossRef](#)]
 39. Chen, J.; Shu, T.; Li, T.; Silva, C. Deep Reinforced Learning Tree for Spatiotemporal Monitoring With Mobile Robotic Wireless Sensor Networks. *IEEE Trans. Syst. Man Cybern. Syst.* **2019**, 1–15. [[CrossRef](#)]
 40. Dong, K. Improved Multi-Scale Entropy and Apply to Fault Feature Extraction and Diagnosis of Rotating Machinery. Master's Thesis, Yanshan University, Qinhuangdao, China, 2016; pp. 60–67.
 41. Jiang, W.; Li, Z.; Zhang, S.; Lei, Y.; Wang, H. Fault Recognition Method Based on Recurrence Quantitation Analysis for Hydraulic Pump. *Chin. Hydraul. Pneum.* **2019**, *02*, 18–23. [[CrossRef](#)]