Prediction of particulate content in oil based on SPA vibration feature selection

Liu Ge[1], Chen Bin*[2]

1 School of Environmental Engineering, North China Institute of Science and Technology, Hebei, 065201, China, lycy@ncist.edu.cn

2 School of Mechanical & Electrical, Hebei Key Laboratory of Safety Monitoring of Mining Equipment, North China Institute of Science and Technology, Hebei, 065201, China, hustchb@ncist.edu.cn.

Corresponding authors: Chen Bin, Tel: +86-10-61591483; E-mail: hustchb@ncist.edu.cn.

Abstract: Aiming at the non-stationary characteristics of oil pressure vibration signals containing particulate, a method for predicting particulate content in oil was proposed based on vibration characteristic frequency extraction by vibrational mode decomposition (VMD), variable selection using successive projections algorithm (SPA) and T_S fuzzy identification combined. Firstly, the pressure vibration signal was decomposed by VMD and a series of narrow-band characteristic frequency matrices were obtained. Then, variables were selected using SPA to construct the feature vector matrix. Finally, the feature vector matrix was used as the input of T_S fuzzy identification to identify the content of particulate in oil. The results showed that the VMD reconstruction of the original oil sample pressure signals could well characterize the main variation of the original signal; the 19 variables were selected from the characteristic frequency of the vibration signal from the oil pressure using SPA, the 19 pressure vibration characteristic frequency of 11 sample sets SPA selected was taken as the input variable of T_S identification model; for each set of sample, the predicted output of the content of particulate in oil was obtained, model prediction decision coefficient is 0.8637, the root mean square error is 0.1979, a reasonable prediction effect was obtained.

Key words: vibrational feature; particulate in oil; content predict; SPA; T_S fuzzy identification ;

1 Introduction

The mixing of particulate pollutants into the oil will not only cause carbon deposition, blockage, and corrosion of the system, but also degrade the physical and chemical properties of the oil. In addition, the particulate in oil has a certain role in promoting the aging reaction of oil, which accelerates the aging of oil. Therefore, it is very important to take effective measures to reduce the particulate content in oil to improve the performance stability of oil and the safety of system operation. So far, the studies on particulate in oil have mainly focused on the monitoring and control of oil pollution[1-3], the measurement and characterization of pollution degree[4,5], etc., however, the studies on the interaction between particulate and oil have been less reported.

Scholars have made in-depth studies on the influence of particulate on the turbulence characteristics of fluids (oil). Liu et al.[6] analyzed the distribution law of particulate in the fluid, and believed that the dispersion of particulate in the fluid was a long-term fluid-induced diffusion process of accumulated particles. The dispersion of particulate was related to the flow velocity of the flow field, and a basic equation was proposed to calculate the particle concentration and velocity. Tiger et al.[7] experimental study shows that even if the mass load of particles in the fluid is only 10^{-4} orders of magnitude, the turbulent properties of the fluid can be significantly changed. Such as the friction velocity of the wall surface increases by about 7% effectively, and the normal turbulence pulsation and Reynolds shear stress of the outer boundary layer increase by about 8-10%; there is a lag in the average flow velocity of particles in the fluid, but the local slip flow velocity near the particles is almost negligible. On the contrary, the average normal velocity of the particles is basically the same as that of the fluid, but there is a local slip velocity in the outer region of the flow, which is about 40% of the particle settling velocity. Wu [8] considers that changes in the structure of flow are the turbulent changes main mechanism under the low particle load, particles contributed to the large scale flow structure to the flow structure in the small scale of decomposition, and along with the increase of the energy transfer rate, in order to maintain turbulent state, so large scale structure acquire more turbulent kinetic energy from the average flow, causing the average flow weakens. Li et al.[9] claimed that the coarse wall surface formed by particle gravity subsidence enhanced the turbulent burst behavior near the wall surface, resulting in the gas-phase normal pulsation velocity and Reynolds shear stress significantly increase in the viscous substrate. The collision between particles and the wall surface strengthens the up-throwing of low-speed fluid, weakens the down-sweeping of high-speed fluid, and enhances the cross-track effect, thus inhibiting the development of turbulent coherent structure, and significantly reducing the normal pulsation velocity and Reynolds shear stress in the area above the viscous bottom layer.

In the flow field analysis of the containing particulate, mode decomposition technique is one of the effective methods to the study of complex flow states such as Proper orthogonal decomposition(POD)[10,11], Global eigenfrequency decomposition (GED)[12], and Dynamic modal decomposition (DMD)[13] methods, for example, Ma et al.[12] put forward a modal analysis method based on fast Fourier transform, the full-field power spectrum density is employed as a foundation to select the characteristic frequency, the full-field amplitude spectrum, phase spectrum, and mode corresponding to the characteristic frequency can be obtained, and the complex flow field around the cylinder under the control of the composite jet can be analyzed efficiently and accurately. Wang et al. [14]used the eigen orthogonal decomposition to separate the scale from the flow field, and found that the velocity distribution corresponding to the large scale structure was close to the original flow field, while the vorticity distribution and spectral characteristics of the small scale flow field were highly consistent with the original vorticity distribution. Chen et al.[15] established the dynamic model of oil with suspended particles by using the continuum medium model, and solved it numerically by using the characteristic line method, and obtained the variation trend of each phase parameter under different initial velocity conditions. The results show that the pulsation period of oil pressure and flow velocity at the beginning, middle, and end of the pipeline decreases with the increase of the initial velocity under different initial velocity conditions, and the concentration distribution of suspended particles is contrary to the changing trend of oil pressure.

So, the presence of particulate in oil, the turbulence characteristics of oil show different macroscopic characteristics. Therefore, the outline of the rest of this paper is as follows. In Section 2, the experiment was set up for oil with different particulate. In Section 3.1, to obtain the transient pressure field of oil by the experiment of oil with different particulate, and analyze the pressure vibration characteristics of the oil flow field with particulate through VMD method firstly. In Section 3.2, SPA variables were selected for the pressure vibration characteristics to obtain the minimum variables characterizing the oil turbulence characteristics. Then, the selected variable of SPA is used as the input of the T_S fuzzy identification model to establish the internal relationship between the oil pressure vibration characteristics and the particulate content in the oil, to predict the content of the particulate in the oil, in section 3.3. In Section 4, we present results to provide the preliminary basis for obtaining the influence rule of the particulate on the stability of the oil. Finally, we conclude in Section 5.

2 Test set up

The oil pressure vibration data used in this paper was collected from an experimental apparatus as shown in Fig.1, in order to analyze the intrinsic relationship between the turbulence characteristics of the oil with different particulate and the particulate content.

Fig.1 shows the schematic of the test rig, including the power source, experimental pipeline and LabVIEW oil data acquisition, processing, and analysis system.

$25^{\#}$ transformer oil was used as the experimental oil, and Cu, Fe, and SiO_2 powder with a medium diameter of 5μm, 15μm, 25μm, and 50μm were selected for particles in the oil. The 0.02g Cu, Fe, and SiO_2 powder of different medium diameters were weighed and mixed with 1L $25^{\#}$ transformer oil, respectively, as the initial oil sample after 8h (temperature 30 ~ 60℃) of ultrasonic oscillator oscillation. Then, $25^{\#}$ transformer oil was added to the initial oil sample according to a certain volume proportion, and after 8h (temperature 30 ~ 60℃), the ultrasonic oscillator was used to oscillate so as to these samples uniformly. The particle counter was used to measure the results, and 53 graded oil samples with different particle sizes and quantities were obtained, respectively.

The oil samples were thoroughly and evenly mixed and placed in the oil tank through the peristaltic pump into the experimental pipeline. The initial velocity of each experimental condition was V_0=0.0362m/s, and the Reynolds number of the oil was Re = 138.3.

A pressure sensor was used during the experiment, which of YP-01S type sensor with a sensitivity of 0.20% and measuring range of -0.10 to 2.00MPa. The YP-01S type sensor was installed on the inlet of the experimental pipeline, which is mainly used to collect pressure signals of the experimental pipeline and convert them into electrical signals for transmission to the data acquisition card. The data acquisition card used is the PXI-6133 acquisition system of NI company production; the acquisition card is a synchronous sampling multi-function data acquisition equipment, can provide 8 analog input channels (14 bits, the rate of each channel is as high as 2.50 MS/s), article 8 digital I/O lines, two dozen counter and trigger, mainly converts electrical signals to digital signals, then transmit to the LabVIEW software in a computer. LabVIEW software was used to compile the data acquisition program of the experimental system, and some parameters such as the connection mode of the NI-DAQmx task pressure signal were selected as RSE, the sampling rate was 1000Hz, the sampling points were 1000Hz, and the continuous sampling mode was set. Then, a total of 53×1000 pressure vibration data can be obtained with 53 graded oil samples experiments in 1s sampling time, providing original data for the acquisition of pressure vibration characteristics of oil samples.

3 The modeling methods

An architectural scheme of the predict model method between the oil pressure vibration characteristics and the particulate content in the oil is presented in Fig.2. Firstly, the pressure vibration signal was decomposed by VMD and a series of narrow-band characteristic frequency matrices were obtained. Then, variables were selected using SPA to construct the feature vector matrix. Finally, the feature vector matrix was used as the input of T_S fuzzy identification to identify the content of particulate in oil ; and detailed steps of the algorithm are presented in Sections 3.1–3.3.

3.1 Characterization of pressure vibration characteristics based on VMD

Oil flow has the non-stationary signal of the multi-scale, quasi-periodic characteristics due to containing different particle sizes and number. Common wavelet decomposition, DFFT decomposition method can appear problems such as frequency aliasing and endpoint effect for processing these style signal, and the Variational Mode Decomposition (VMD)[16] were a multivariate decomposition method for a non-stationary signal, its mechanism is to assume that the vast majority of each of the modes are closely around a center frequency, By transforming the decomposition problem of multivariable non-stationary signals, namely the modal bandwidth problem, into a variational constraint problem, each mode around the central frequency is decomposed, so as to adaptively divide the frequency domain of signals and effectively separate each component. Li et al.[17] presented a milling chatter detection methodology based on VMD and difference of power spectral entropy, Joshuva et al.[18] used variational mode decomposition (VMD) for signal preprocessing, to the raw vibration signal to retrieve the required descriptive statistical features.

About the acquisition of oil pressure vibration signal to analysis and process using VMD, the number of VMD decomposition components of each sample is determined according to a criterion for comparison the slope of a center frequency of the VMD decomposition component with the dominant frequency of the component together, so as to realize extraction and analysis of the main VMD component of the pressure vibration signal of oil containing different particle size and content.

The particle size and content of pollutants of $3^{\#}$ oil sample are 5μm 827 particles per ml, 15 μm 51.333 particles per ml, 25μm 8.935 particles per ml and 50μm 1.283 particles per ml, the corresponding ISO4406 pollution degree is 17/13. For the pressure vibration signal of the oil sample, the central frequency slope criterion is used to determine the number of components decomposed by the VMD is 11, and perform various Mode component resulting from the decomposition of VMD is shown in Fig.3.

As shown in Fig.3, as the number of decomposed components increases, the component contains information - amplitude modulation (AM) and frequency modulation (FM) characteristics are more obvious. The first order component is the trend signal of the oil sample pressure signal, the pressure signal is characterized by cosine variation. The second order component can see the signal characteristics with an intermittent pulse; the components of the 3-7th order have obvious AM signal characteristics, while the components of the 8-11th order not only have AM modulation characteristics, but also have FM modulation characteristics. It can be seen that $3^{\#}$ oil sample can decompose each component of the signal effectively after the decomposition of VMD, which plays an important role in extracting the characteristics of the pressure signal.

Fig.4 is the spectrum diagram of each component of the oil sample after VMD decomposition. It can be seen that the frequency bands for each component basically have no aliasing phenomenon after VMD decomposition, and the spectrum bandwidth of the components of the

1-7th order is within 40Hz, and the bandwidth of the components of the 8-11th order is within 50Hz. This is due to the fact that each component of the 8-11th order has the characteristics of frequency modulation, and its corresponding bandwidth is composed of the sum or difference of the central frequency and the multipliers of the modulation frequency. Thereby the spectrum of each component of the pressure signal decomposed by the VMD can represent the characteristics of the signal through the frequency.

Fig.5 shows the comparison of the original pressure signal and their corresponding frequency spectrum of $3^{\#}$ oil sample with the reconstructed signal after VMD decomposition. The components of the pressure signal decomposed by VMD were reconstructed and compared with the original signal, as shown in fig.5 (a). The reconstructed signal can better represent the main change characteristics of the original signal, the error between the two is shown in fig.5 (b), and the error range is between -19.14-17.77%. The frequency spectrum comparison and error between the reconstructed signal and the original signal are shown in Fig.5 (c) and (d), and the error range is between -0.2-16.76%. The spectrum difference between the reconstructed signal and the original signal lies in the interval part of the bandwidth of each component of the original signal after VMD decomposition. It can be seen that the frequency of each component after VMD decomposition is the characteristic frequency of the oil sample. In the same way, each oil sample was decomposed by VMD to extract its corresponding component frequency, which was taken as the characteristic frequency of the corresponding oil sample and provided a prerequisite for constructing the nonlinear relationship between the pressure vibration characteristic signal of the oil sample and the content of particle pollutants in the oil.

3.2 SPA selection of effective variables for vibration characteristics

Although the frequency spectrum of the reconstructed signal after VMD decomposition can represent the characteristic frequency of the original signal, there are 512 characteristic frequencies in each oil sample. As can be seen from Fig.1 and Fig.2, each component of VMD decomposition contributes differently to the characteristic of the original signal, that is, the component decomposed by VMD to each oil sample is redundant. For constructing the nonlinear relationship between the pressure vibration characteristic signal of the oil sample and the particle pollutant content in the oil, these characteristic frequencies will affect the accuracy and validity of the final results.

Successive projection algorithm (SPA)[19], a multivariate calibration process dimension reduction method is presented by the Araujo, which converts the selection of input variables to a combination of the constraint optimization problem, tests and comparison the variable subset using the performance of obtained model, then variable subset optimization search, finally, according to a series of matrix projection operations involving the oil sample indicators form a subset of selected variables, a specific algorithm can be found in the literature[20].

The characteristic frequency of each oil sample was selected by SPA as an effective variable, and according to the SPA algorithm, by means of calculating projection a certain frequency to other frequencies within the characteristic frequency matrix constructed by 53 oil samples, the biggest projection quantity frequency was selected as the frequency of each frequency sequence, each frequency in the sequence has the least correlation with the previous frequency, so as to achieve the greatest degree to eliminate collinearity's influence on the model, reducing model complexity. The number of variables and their variables selected by SPA for the characteristic frequency of 53 oil samples is shown in Table 1.

Since the rotation frequency of the spindle of the oil sample driven pump was 30Hz, and the fundamental vibration frequency of the oil sample was 6.8359Hz during the experiment, each characteristic frequency selected by SPA was the multiplication frequency under the interaction of the rotation frequency of the spindle and the fundamental frequency. The chosen first frequency is 53.7109Hz in Table 1, which is 5 times the difference between the 2 times rotation frequency of the spindle (30Hz) and the 7 times frequency of the fundamental frequency. Secondly, 236.3281Hz is 22 times the difference between the 2 times rotation frequency of the spindle (30Hz) and the 7 times frequency of the fundamental frequency. Next, 8.7891Hz is the difference between the spindle rotation frequency 30Hz and the 3 times fundamental frequency; 11.7188hz is the difference between the spindle rotation frequency 30Hz and the 6 times fundamental frequency. 117.1875Hz is 10 times of 11.7188Hz. 13.6719Hz is 2 times the fundamental frequency. 12.6953Hz is the difference between the 3 times frequency of 8.7891Hz and 13.6719Hz. 19.5313Hz is the sum of 12.6953Hz and fundamental frequency. 48.8281Hz is the sum of the spindle rotation frequency 30Hz and 19.5313Hz. 49.8047Hz is the sum of the rotation frequency of the spindle and the 3 times fundamental frequency. 55.6641Hz is 3 times the sum of 11.7188Hz and the fundamental frequency. 92.7734Hz is 5 times the sum of 11.7188Hz and fundamental frequency. 56.6406Hz is 2 times the sum of 19.5313Hz and 8.7891Hz. 169.9219Hz is 3 times of 56.6406Hz. 179.6875Hz is 2 times the sum of 48.8281Hz and 6 times the fundamental frequency. 209.9609Hz is 5 times the sum of 19.5313Hz and the rotation frequency of the spindle. 327.1484Hz is 5 times the sum of 56.6406Hz and 8.7891Hz. 355.4688Hz is 52 times the fundamental frequency.

It can be seen that SPA selected variables are arranged according to the projected root mean square scale of each variable, in which the optimal variable is placed in the first position. For example, 53.7109Hz is the optimal characteristic frequency of the pressure signal of the experimental oil sample selected by SPA in Table 1.

The variables selected by SPA are shown in Fig.6. It can be seen that of the 19 variables selected from the characteristic frequencies of the oil sample pressure vibration signal, 11 of them are mainly distributed in the lower frequency region of 0-50Hz, which is consistent with the case that the larger frequency amplitude of the first two components decomposed by VMD in Fig.3. Moreover, it can be seen that the selected variable is the position where the frequency spectrum of the pressure vibration signal varies greatly, indicating that the chosen SPA variable can better represent the frequency characteristics of the original signal and has a certain physical significance.

3.3 Establish T_S identification model

Since the relationship between the content of particulate pollutants in the oil and the pressure vibration signal of the oil is nonlinear and implicit, it is difficult to be characterized by an accurate mathematical expression. The fuzzy T_S identification system is a kind of identification method developed for nonlinear problems widely used[21-23]. The relationship between the oil pressure vibration signal and the content of particle pollutants in the oil is the MIMO system. The 19 pressure vibration characteristic frequencies selected by SPA are used as the input data, and x (i, k) is used to represent the parameter set of the k-th pressure vibration characteristic frequency for the input of the i-th oil sample, the output variables are the particle pollutants with different particle sizes and contents in the oil, represented by y (i, j). 53 oil samples can be divided into Ⅰ and Ⅱ group used for training and validation model respectively, among them, the Ⅰ group is used to examine model fitting precision and generalization ability, a total of 42 sets; Ⅱ used for testing the validity of the model, there are 11 sets of data. A MIMO T_S model of oil pressure vibration system containing particle pollutants can be represented by a fuzzy rule set R, then the fuzzy rule i-th can be described as:

$$\begin{aligned} R^i: \text{if} \quad & x_1 = A_1^i \cup x_2 = A_2^i \cup \cdots \cup x_m = A_m^i \\ \text{then} \quad & y^i(k) = p_0^i + p_1^i x_1^i + p_2^i x_2^i + \cdots + p_m^i x_m^i, \quad i = 1, 2. \cdots, r \end{aligned} \tag{1}$$

Where R^i is the i-th rule in the fuzzy set; A_j^i is a fuzzy subset of the j-th characteristic frequency of pressure vibration, and the parameters in its membership function are called antecedent parameters. y^i is the output variable of i-th rule; p_j^i is the coefficient of identification parameter; $x_1(\cdot), x_2(\cdot), \cdots x_m(\cdot)$ is the input pressure vibration characteristic frequency variable; $y(\cdot)$ is the output variable of particle pollutant with different particle size and content in the oil;

Then, β_i is defined as the effective function of the weighted average rule of the i-th fuzzy rule, and the expression is $\beta_i = \mu^i \Big/ \sum_{j=1}^{r} \mu^i$, where, $\mu^i(x)$ is the activation degree of the i-th fuzzy rule in formula (1), that is $\mu^i(x) = \prod_{j=1}^{m} A_j^i(x_j)$, $\prod$ is a fuzzy operator and takes a small operation.

Then the output of the T_S fuzzy identification system for the oil pressure vibration signal containing particle pollutants is:

$$y = \sum_{i=1}^{r} \mu^i y^i \Big/ \sum_{i=1}^{r} \mu^i = \sum_{i=1}^{r} \beta_i y^i = \sum_{i=1}^{r} \beta_i \left(p_0^i + p_1^i x_1 + p_2^i x_2 + \cdots + p_m^i x_m \right) \tag{2}$$

(1) Antecedent parameters identified by fuzzy clustering of C-means

The Fuzzy C-means Algorithm (FCM) was proposed by Bezdek, a clustering algorithm based on objective function optimization. The core problem is to determine the clustering center according to the reasonable clustering index (equation (3)), so as to the partition of fuzzy input space is optimized.

$$J(Z, U, V) = \sum_{i=1}^{c} \sum_{k=1}^{N} (\mu_{ik})^m \left\| z_k - v_i \right\|_A^2 \tag{3}$$

Where, Z is the set that needs clustering, $U = [\mu_{ik}]$ is the fuzzy division of Z, V=(v_1, v_2,..., v_c)^r is the clustering center vector. c is the number of clusters, N is the number of oil samples,

μ_{ik} is the membership degree of the pressure vibration characteristic frequency z_k relative to the cluster center v_i, and meets $\mu_{ik} \in [0,1]$ and $\sum_{i=1}^{c} \mu_{ik} = 1$, $m \in [1, \infty]$ is the fuzzy index.

Then, the objective function equation (3) is optimized according to the Lagrange operator, and the conditions that make equation (3) have a minimum point (U, V) are as follows:

$$\mu_{ik} = 1 \Big/ \sum_{j=1}^{c} \left(D_{ik} / D_{jk} \right)^{2/(m-1)} \tag{4}$$

$$v_i = \sum_{k=1}^{N} (\mu_{ik})^m z_k \Big/ \sum_{k=1}^{N} (\mu_{ik})^m \tag{5}$$

Where i = 1, 2, ... , c; D_{ikA}^2 is the distance norm of the square inner product, $D_{ikA}^2 = \|z_k - v_i\|_A^2 = (z_k - v_i)^T A (z_k - v_i)$, and A determines the shape of clustering, in FCM,A is generally taken as the unit vector I.

FCM clustering algorithm is:

① Firstly select the number of clusters c=3 and the fuzzy index m=2, and assign the initial value to the fuzzy partition matrix U, then the number of iterations I=1, and the iteration termination condition $\varepsilon > 0$;

② According to equation (4) update the cluster center v_i and distance norm D_{ikA}^2;

③ According to equation (3) update the fuzzy partition matrix U ;

④ If $\left\| U^{(l+1)} - U^{(l)} \right\| \prec \varepsilon$, then stop the iteration, if not I=I+1, go to step ②.

Then, according to the FCM algorithm, the oil-containing particle pollutants can be fuzzy identified, and the three clustering centers obtained are:

$$\begin{bmatrix} \text{cluster center 1} \\ \text{cluster center 2} \\ \text{cluster center 3} \end{bmatrix} = \begin{bmatrix} 0.1628 & 0.1361 & 0.1139 & 0.1470 & 0.1618 & 0.0992 & 0.1216 & 0.1762 & 0.1062 & 0.1378 & 0.0690 & 0.1883 & 0.1151 & 0.2388 & 0.3742 & 0.0967 & 0.1049 & 0.1498 & 0.0760 \\ 0.1047 & 0.2419 & 0.0997 & 0.1171 & 0.0789 & 0.0478 & 0.0744 & 0.1585 & 0.1008 & 0.0962 & 0.0541 & 0.8029 & 0.0688 & 0.1139 & 0.1253 & 0.2067 & 0.0835 & 0.1155 & 0.0649 \\ 0.1239 & 0.0933 & 0.0534 & 0.6763 & 0.1057 & 0.0725 & 0.0918 & 0.1050 & 0.0542 & 0.1145 & 0.0679 & 0.1031 & 0.1375 & 0.1238 & 0.0991 & 0.0937 & 0.2370 & 0.1225 & 0.0556 \end{bmatrix} \tag{6}$$

(2) Latter parameters identified used recursive least square method

According to the input data of pressure vibration characteristic frequency of oil sample selected by SPA and the output data of particle pollutants with different particle size and content in the oil, the training data pair $x_{1i}, x_{2i}, x_{3i}, x_{4i} \to y_i \ (i = 1, 2, \cdots, 42)$ of T_S fuzzy identification model was constructed.

When, the recursive least square method is used to obtain the latter parameters of the model. The algorithm is as follows:

Let X be an N×r (m+l) matrix, Y be an N-dimensional vector, and P be an r(m+l) dimensional vector:

$$
\begin{aligned}
&X=\begin{bmatrix}
\lambda_{11},\cdots,\lambda_{r1},\lambda_{11}x_{11},\cdots,\lambda_{r1}x_{11},\cdots,\lambda_{11}x_{m1},\cdots,\lambda_{r1}x_{m1} \\
\vdots \qquad\qquad\qquad\qquad \vdots \\
\lambda_{1N},\cdots,\lambda_{rN},\lambda_{1N}x_{1N},\cdots,\lambda_{rN}x_{1N},\cdots,\lambda_{1N}x_{mN},\cdots,\lambda_{rN}x_{mN}
\end{bmatrix} \\
&Y=\left[y_1,\cdots,y_N\right] \\
&P=\left[\,p_0^1,\cdots,p_0^r,p_1^1,\cdots,p_1^r,\cdots,p_m^1,\cdots,p_m^r,\right]
\end{aligned}
\tag{7}
$$

Then $Y = XP$. So we get the least-squares estimate of P is $P^* = \left(X^T X\right)^{-1} X^T Y$.

The recursive least squares algorithm can avoid matrix inversion and carry out iterative optimization of parameter matrix P. This is, the k-th row vector of X is x_k and the k-th component of Y is y_k, then the recursive method is:

$$
\begin{aligned}
&P_{k+1}=P_k+K_k\left(y_{k+1}-x_{k+1}P_k\right) \\
&K_k=S_k x_{k+1}^T\left(1+x_{k+1}S_k x_{k+1}^T\right)^{-1} \\
&S_{k+1}=S_k-K_k x_{k+1}S_k \qquad k=0,1,\cdots,N-1
\end{aligned}
\tag{8}
$$

The initial conditions are P=0, S_0=aI. A is always going to be more than 10,000 real numbers. I is the unit matrix of L×L(L=r(m+l)). The purpose of parameter identification is to minimize the following performance indicators:

$$
J=\left(\sum_{i=1}^{N} y_i-\hat{y}_i\right)^2 \Big/ N \tag{9}
$$

Where, y_i is the actual output, and $\hat{y}_i$ is the model output.

Thus, the rules of the T_S fuzzy identification algorithm for oil pressure vibration signal containing particle pollutants are as follows:

$$
\begin{aligned}
&R^{1\text{-}3}:\text{if} \quad \left[x_1,x_2,\ \cdots,\ x_{,18},x_{19}\right] \in Z_1 \\
&\text{then}\begin{bmatrix} y(1,j) \\ y(2,j) \\ y(3,j)\end{bmatrix}=\Psi\cdot\begin{bmatrix} x(1,k) \\ x(2,k) \\ \vdots \\ x(19,k)\end{bmatrix}+\begin{bmatrix} y_0(1,j) \\ y_0(2,j) \\ y_0(3,j)\end{bmatrix}
\end{aligned}
\tag{10}
$$

Where:

$$
\Psi=\begin{bmatrix}
-0.5908 & -0.7937 & 0.4109 & 0.7966 & -0.8709 & -0.4687 & -0.2716 & -1.1646 & -0.4349 & -1.1463 & -0.8593 & 2.2614 & 0.1352 & -0.6321 & -0.5328 & -0.4135 & -0.9798 & 0.2554 & -1.1796 \\
0.3859 & 1.4993 & -0.0670 & -0.7178 & 0.4938 & -1.2313 & -2.0832 & 0.6670 & 0.4935 & 1.6253 & -0.2725 & -0.7844 & 0.2097 & 1.6636 & -0.3512 & -0.2118 & 0.7676 & 0.1731 & 2.5360 \\
-0.4331 & 0.0222 & -0.3616 & 0.2078 & 0.6448 & 0.1877 & 0.4200 & 0.1964 & -0.0687 & -0.4771 & -0.5024 & -1.4509 & -0.6498 & -0.5333 & 0.5182 & -0.8874 & -0.4405 & -1.5051 & 0.6356
\end{bmatrix}
$$

$$
\begin{bmatrix} y_0(1,j) \\ y_0(2,j) \\ y_0(3,j)\end{bmatrix}=\begin{bmatrix} 1.1600 \\ 0.2225 \\ 0.0965\end{bmatrix}。
$$

,

Then T_S identification model was established for the pressure vibration signal of oil samples containing particulate pollutants.

4. Result analysis

According to the established T_S identification model of pressure vibration signals of oil samples containing particulate pollutants, the content of particulate pollutants in oil was predicted with the determination coefficient (R^2), and root mean square error (RMSE) as the evaluation indexes.

The 19 pressure vibration characteristic frequencies selected by SPA of 11 groups of oil samples were used as the input variables of the T_S identification model (M1), and the predicted output of particle pollutant content of each oil sample was obtained. The evaluation index values of the model are shown in table 2. Among them, the determination coefficient predicted by the model is 0.8637, and the root-mean-square error is 0.1979, indicating that the 19 characteristic frequencies selected by SPA can better represent the characteristic spectrum of each oil sample, and the prediction effect is better.

Fig.7 shows the comparison between the M1 predicted value and the measured value of 11 verification sets and their correlation graphs. As can be seen from Fig. 7 (a), the predicted output of the constructed T_S identification model of pressure vibration signal of oil samples containing particulate pollutants can well track the change of the content of particulate pollutants in the oil, with the error range between -9.507-0.3819%. It can be seen from the correlation diagram in Fig.7(b) that the error between the predicted value and the measured value of data $4^{\#}$ in the verification set is large, which affects the overall accuracy of the predicted output. The difference between the predicted value and the measured value of other data is relatively small. It can also be seen that except the 3 data of $5^{\#}$, $8^{\#}$, and $9^{\#}$ below the center line, the other 8 data are all distributed on the top of the center line, indicating that the predicted value of the pressure output of the constructed T_S model is slightly larger than the measured value. The reason may be caused by oil sample uneven or sedimentation phenomenon when measuring the content of particle pollutants in the oil sample. Therefore, these errors can be reduced by taking the mean value of multiple measurements.

In order to compare the effectiveness of the SPA selected variables, the spectrum of the signal by VMD reconstruct of the 11 sets of sample pressure vibration signal data in the verification set, a total of 512 data as the input variable of the constructed T_S identification model (M2) of the pressure vibration signal of the oil sample containing particulate pollutants, get 11 sets predicted output of sample containing particle pollution, the model evaluation indexes are shown in table 2. The full frequency spectrum of the pressure vibration signal is used as the input of the identification model, the pressure output prediction determination coefficient of the model is 0.7369, and the root mean square error is 0.1513. Compared with the characteristic frequency selected by SPA as the input of the model, the prediction determination coefficient is more minor.

Fig.8 shows the comparison between the M2 predicted value and the measured value of 11 verification sets and their correlation graphs. As shown in Fig. 8 (a), the predicted value of the content of particulate pollutants in oil can basically maintain the same trend of change as the measured value, but the error between the two is relatively large, varying within the range of -4.32-45.1257%. In Fig.8 (b), it can be clearly seen from the correlation diagram that the 3 data of $7^{\#}$, $9^{\#}$ and $10^{\#}$ are far away from the center line, which are large data affecting the error of the predicted value and the measured value. In addition, the 11 groups of data in the verification set are relatively concentrated in two regions, and the uneven distribution is also a factor causing large errors.

It is shown that the established T_S identification model offers a better prediction performance. It can be used for on-line monitoring of the key indicators of the oil in use of the machine. Through intelligent data processing of the system, it can reflect the changing trend of the state of deterioration, pollution, and mechanical wear of the oil in use of the machine in real time, prevent the major lubrication accidents of the machine in time, and provide technical support for enterprises to make reasonable oil-changing cycle and maintenance decisions. For example, when applied to the compressor oil quality analysis as shown in Figure 9, there are many impurities in the compressor oil, mainly composed of copper, iron, carbonized particles, and a small amount of calcium, sulfur, and other components. These powder copper and iron particles are brought into the pipeline system, which will cause the compressor to block or form a system block. The proposed monitoring method for particulate matter content in oil can continuously and continuously monitor the particulate pollutants index of oil products in real time, and provide a strong guarantee for the reliable operation of equipment.

Usually, in order to obtain the failure state of the equipment, the vibration signal of the equipment is used for online monitoring[24], and the oil in the equipment will be mixed with particles during operation. Scholars at home and abroad have put forward some views on the interaction characteristics between these particles and the oil. Such as changes in the structure of flow is the turbulent changes main mechanism under the low particle load[8]. It can be seen that the pressure vibration signal of the oil can be used to predict the content of particulate matter in the oil, so as to obtain the pollution status of the oil, and then evaluate the failure degree of the equipment, to ensure the safe operation of the equipment. Therefore, this paper puts forward that the prediction of particulate content in oil by means of oil pressure vibration signal is different from the traditional vibration monitoring method of equipment, which has a certain theoretical basis and innovation.

In terms of the analysis and processing of vibration signals, researchers have put forward many mature and effective processing methods[25], but these methods are basically for the equipment vibration signals, such as using FFT, MSAF-12, and mean of vector sum[24] to obtain feature vectors, proposed a parsimonious network based on a fuzzy inference system[25], etc., have received good results. This paper is a fluid vibration signal, the flow characteristics of different particle sizes and amounts of oil in the non-stationary signal of the multi-scale, quasi periodic characteristics, the common method such as wavelet decomposition, DFFT decomposition can appear problems such as frequency aliasing and endpoint effect. For non-stationary signal, VMD is a method of multivariate decomposition and the quantity of VMD decomposition component of each sample is according to a criterion for comparison the slope of a center frequency of the VMD decomposition component with the dominant frequency of the component together, to determine the main frequency; so as to realize extraction and analysis of the main component of VMD of the oil pressure vibration signal, It is also one of the advantages and innovations of this paper.

Because the contribution of the VMD decomposition component to the original signal is different, that is, the decomposed component of each oil sample has redundancy. SPA variable selection to each component of VMD decomposition, there are 19 variables selected, and 11 of them are mainly distributed in the lower frequency region of 0-50Hz, the selected variable is the position where the frequency spectrum of the pressure vibration signal varies greatly, indicating that the chosen SPA variable can better represent the frequency characteristics of the original signal and has a certain physical significance. It can guarantee the accuracy and validity of the relationship between the characteristic signal of pressure vibration and the content of particulate pollutants in oil.

Based on feature vectors of the vibration signals, models of feature vectors and failure forms are constructed for different failure forms, such as ANFIS, eTS, Simp_eTS and fuzzy inference system[25] , Nearest Neighbor (NN), Linear Discriminant Analysis (LDA), and Linear Support Vector Machine (LSVM)[24]. According to the relationship of the particle pollution content in oil and the oil pressure vibration signals is the MIMO system, this paper is different from the common vibration modeling method, will use the SPA selection pressure vibration feature vectors as the input data, pollutants content for the output variable, to build T_S fuzzy identification model, has received the good prediction effect. Because more samples are needed for the learning and training of the T_S fuzzy identification model, the method proposed in this paper requires a larger amount of data than other vibration analysis methods.

In order to apply the proposed method for real-time prediction of particulate content in oil, a feasibility study to embed the predicted method into real-time measurement, is also expected to be conducted in the future.

5 Conclusions

This paper describes the prediction of particulate content in oil based on SPA vibration feature selection. The proposed techniques were based on oil pressure vibration signals containing different particle sizes and number. The authors extracted vibration characteristic frequency by VMD. The authors selected variables using SPA to construct the feature vector matrix, the feature vector matrix was used as the input of T_S fuzzy identification to identify the content of particulate in oil.

(1) The VMD reconstruction signal of the original pressure signal of the oil sample can well represent the main change characteristics of the original signal, and the frequency of each component decomposed by VMD is the characteristic frequency of the oil sample.

(2) Among the 19 variables selected by SPA from the characteristic frequencies of the oil sample pressure vibration signal, 11 variables are mainly distributed in the low frequency region of 0-50Hz. The selected variables are the locations where the frequency spectrum of the pressure vibration signal varies greatly, which can better represent the frequency characteristics of the original signal, and have certain physical significance.

(3) The 19 pressure vibration characteristic frequencies selected by SPA of 11 sample set used for input variables of the T_S identification model, to obtain the predicted output of particle pollution content of each sample, the determination coefficient predicted by the models is 0.8637, the root mean square error is 0.1979. Compared with the full frequency spectrum of model prediction, it shows that the 19 characteristic frequency selected by SPA can better represent the spectrum characteristics of each sample, the predicted effect is better.

Funding:

This study was funded by Local science and technology development fund projects guided by the central government (206Z5201G), Chongqing Basic Research and Frontier Exploration Project (cstc2018jcyjAX0121), Fundamental Research Funds for the Central Universities (3142019001, 3142019055).

Declaration of Competing Interest:

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authorship contributions:

Liu Ge: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing - original draft, Visualization. **Chen Bin**: Validation, Resources, Writing – review & editing, Supervision, Funding acquisition.

References:

[1] Wen, Z., Yin, X., Jiang, Z. "Applications of electrostatic sensor for wear debris detecting in the lubricating oil", *Journal of The Institution of Engineers (India): Series C*, **94**(3),pp. 281-286 (2013).

[2] Levi, O., Eliaz, N. "Failure analysis and condition monitoring of an open-loop oil system using ferrography", *Tribol. Lett.*, **36**(1), pp. 17-29 (2009).

[3] Wu, T., Peng, Y., Sheng, C., et al. "Intelligent identification of wear mechanism via on-line ferrograph images", *Chinese Journal of Mechanical Engineering*, **27**(2), pp. 411-417 (2014).

[4] Wang, Y., Tan, R., Hu, K., et al. "Experimental study on the viscoelastic behaviors of debris flow slurry", *J. Mt. Sci.-Engl.*, **9**(4), pp. 501-510 (2012).

[5] Chen, B., Liu, G.,Zhang, X. "Progress in on-line monitoring of polluted lubricant", *Applied Chemical Industry*, **41**(7),pp. 1248-1252 (2012).

[6] Liu, S. "Particle dispersion for suspension flow", *Chem. Eng. Sci.*, **54**(54), pp. 873-891 (1999).

[7] Kiger, K. T., Pan, C. "Suspension and turbulence modification effects of solid particulates on a horizontal turbulent channel flow", *J. Turbul.*, **3**(3), pp. 27-29 (2002).

[8] Wu, Y. "Experimental investigation on turbulence modification in a horizontal channel at relatively low mass loadings". Doctor, *Huazhong University of Science and Technology*(2007).

[9] Li, J., Jiang, G.,Wang, H. "Modifications of coherent structures in turbulent boundary layer of dilute gas-particle horizontal channel flow", *Journal of chemical industry*, **63**(11),pp. 3407-3417 (2012).

[10] Greig, D., Siddiqui, K., Karava, P., et al. "Investigation of fundamental flow mechanisms over a corrugated waveform using proper orthogonal decomposition and spectral analyses", *Int. J. Therm. Sci.*, **96**(160-172 (2015).

[11] Chen, B.,Liu, G. "Anslysis of kinetic energy dissipation rate and POD reconstruction of a channel flow", *Chinese Journal of Computational physics*, **35**(2),pp. 169-177 (2018).

[12] Ma, L., Feng, L.,Pan, C. et al. "Fourier mode decomposition of PIV data", *Science China Technological Sciences*, **58**(11),pp. 1935-1948 (2015).

[13] Narasingam, A.,Kwon, S. I. "Development of local dynamic mode decomposition with control: Application to model predictive control of hydraulic fracturing", *Comput. Chem. Eng.*, **106**(2),pp. 501-511 (2017).

[14] Wang, H., Gao, Q.,Wu, J. "The statistical study of vortex structure in turbulent boundary layer flow based on Tomographic PIV", *Sci Sin-Phys Mech Astron*, **45**(12),pp. 73-86 (2015).

[15] Chen, B.,Liu, G. "Influence of oil velocity on characteristics of suspension flow", *Journal of China University of Petroleum*, **41**(4),pp. 167-173 (2017).

[16] Upadhyay, A.,Pachori, R. B. "Instantaneous voiced/non-voiced detection in speech signals based on variational mode decomposition", *Journal of the Franklin Institute*, **352**(7),pp. 2679-2707 (2015).

[17] Li, X., Wan, S.,Huang, X., et al. "Milling chatter detection based on VMD and difference of power spectral entropy", *The International Journal of Advanced Manufacturing Technology*, **111**(7-8),pp. 2051-2063 (2020).

[18] Joshuva, A., Kumar, R. S.,Sivakumar, S., et al. "An insight on VMD for diagnosing wind turbine blade faults using C4.5 as feature selection and discriminating through multilayer perceptron", *Alexandria Engineering Journal*, **59**(5),pp. 3863-3879 (2020).

[19] Jiang, W., Lu, C.,Zhang, Y., et al. "Moving-window-improved monte carlo uninformative variable elimination combining successive projections algorithm for near-infrared spectroscopy (NIRS)", *J. Spectrosc.*, **2020**(1-12 (2020).

[20] Chen, B., Liu, G.,Zhang, X. "Analysis on near infrared spectroscopy of water content in lubricating oil using successive projections algorithm", *Infrared and Laser Engineering*, **42**(12),pp. 3168-3174 (2013).

[21] Chen, B.,Liu, G. "Analysis on Near Infrared Spectroscopy of water content in oil using T-S Fuzzy Identifying Model", *Acta photonica sinica*, **43**(2),pp. 230001 (2014).

[22] Li, W., Qiao, J.,Zeng, X., et al. "Identification and simplification of T-S fuzzy neural networks based on incremental structure learning and similarity analysis", *Fuzzy Set. Syst.*, **394**(65-86 (2020).

[23] Yan, J., Yang, Guang-H., Li, X. "Adaptive fault-tolerant compensation control for T–S fuzzy systems with mismatched parameter uncertainties", *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, **50**(9),pp. 3412-3423 (2020).

[24] Glowacz, A., Glowacz, W.,Kozik, J., et al. "Detection of deterioration of three-phase induction motor using vibration signals", *Meas. Sci. Rev.*, **19**(6),pp. 241-249 (2019).

[25] Caesarendra, W., Pratama, M.,Kosasih, B., et al. "Parsimonious network based on a fuzzy inference system (PANFIS) for time series feature prediction of low speed slew bearing prognosis", *Applied Sciences*, **8**(12),pp. 2656 (2018).

List of captions

Fig.1 The experimental apparatus of the pressure vibration signal of the oil sample

Fig.2 An architectural scheme of the predict model method

Fig.3 Components of the VMD decomposition of the 3# oil sample

Fig.4 The spectrum of each component of VMD decomposition of $3^{\#}$ oil sample

Fig.5 Comparison of original pressure signal and spectrum with the reconstructed signal of $3^{\#}$ oil sample

Fig.6 The SPA selected variable

Fig. 7 The comparison of model predicted value and measured value of SPA selection variables

Fig. 8 The comparison of model predicted value and measured value of full frequency spectrum

Fig.9 The compressor of application of the proposed research

Table 1 SPA selected variables and their variables

Table 2 The model evaluation index

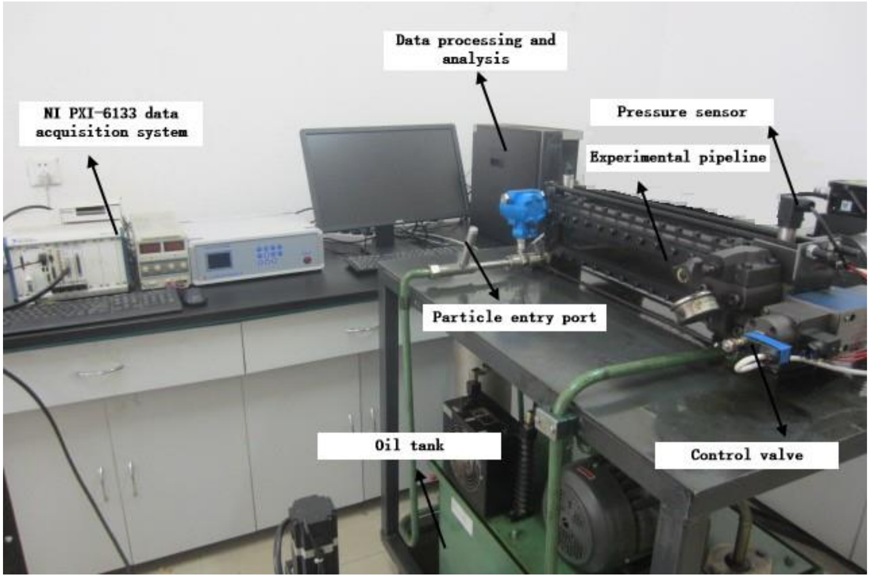

Fig.1 The experimental apparatus of the pressure vibration signal of the oil sample

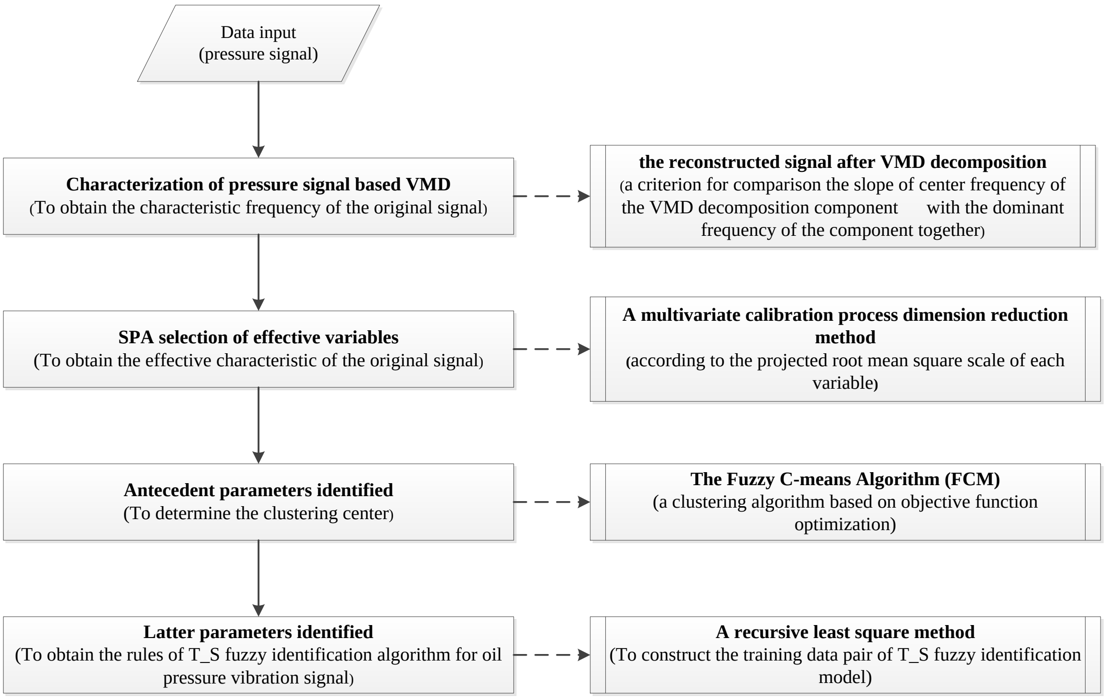

Fig.2 An architectural scheme of the predict model method

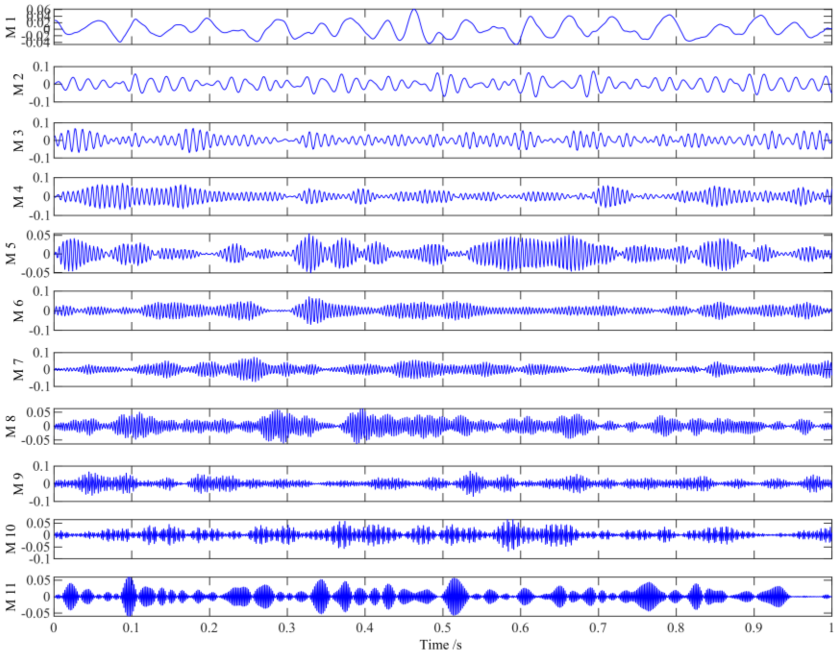

Fig.3 Components of the VMD decomposition of the 3# oil sample

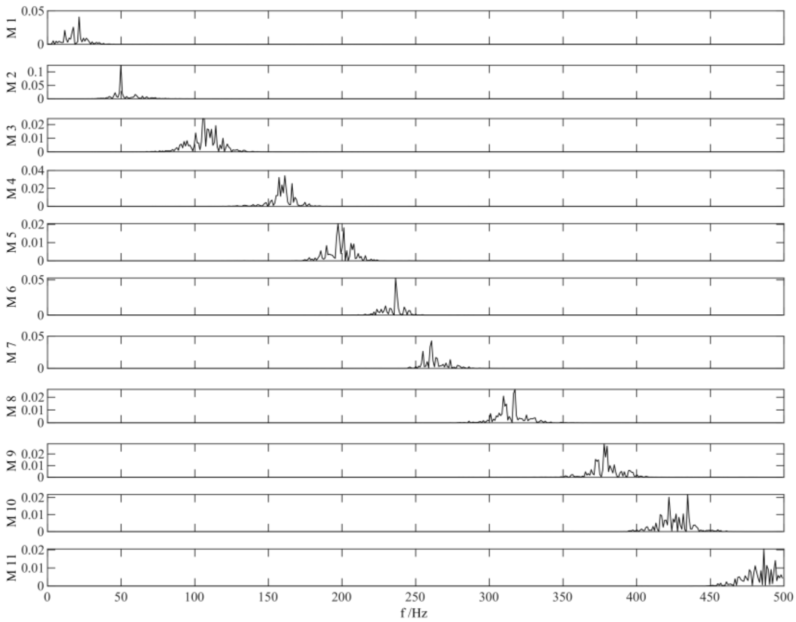

Fig.4 The spectrum of each component of VMD decomposition of $3^{\#}$ oil sample

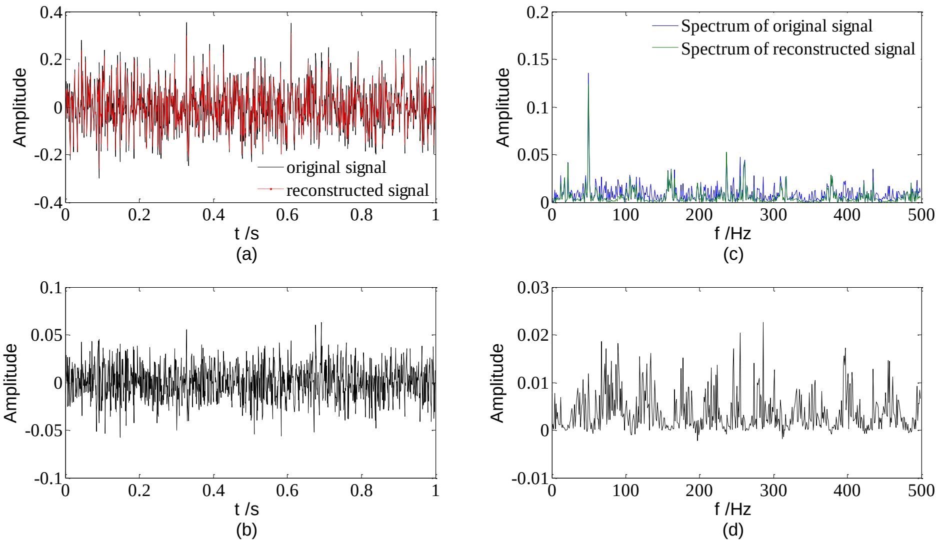

Fig.5 Comparison of original pressure signal and spectrum with the reconstructed signal of 3# oil sample

(a) Comparison of the original signal with the reconstructed signal. (b)The error of original signal with the reconstructed signal. (c) Spectrum comparison of the original signal with the reconstructed signal. (d) Spectrum error of original signal with the reconstructed signal

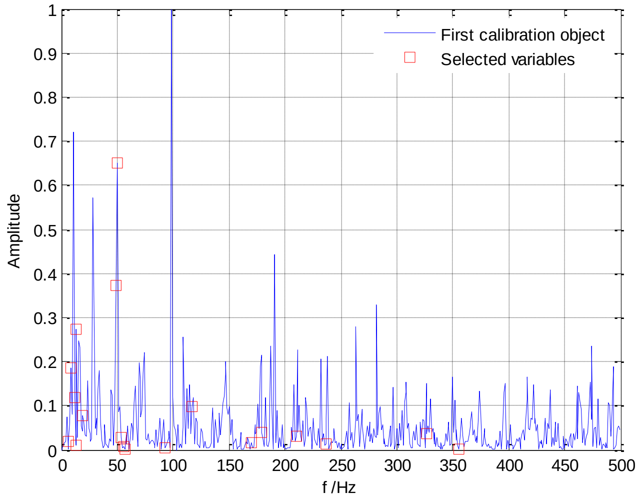

Fig.6 The SPA selected variable

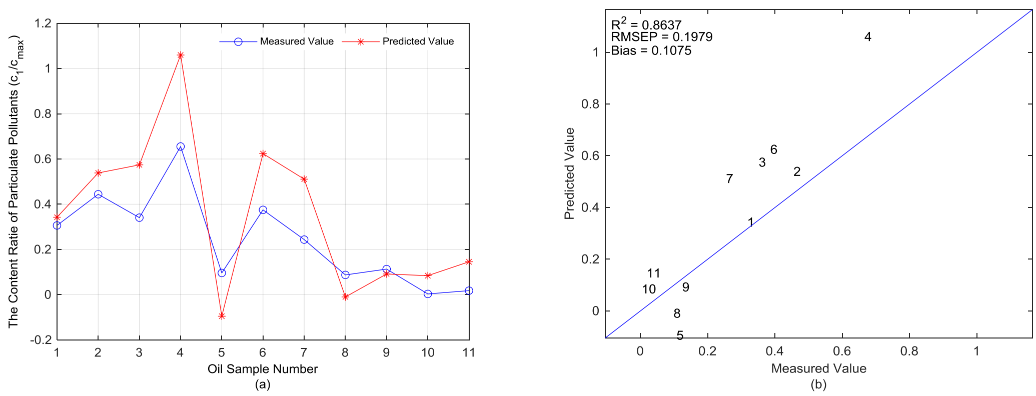

Fig. 7 The comparison of model predicted value and measured value of SPA selection variables

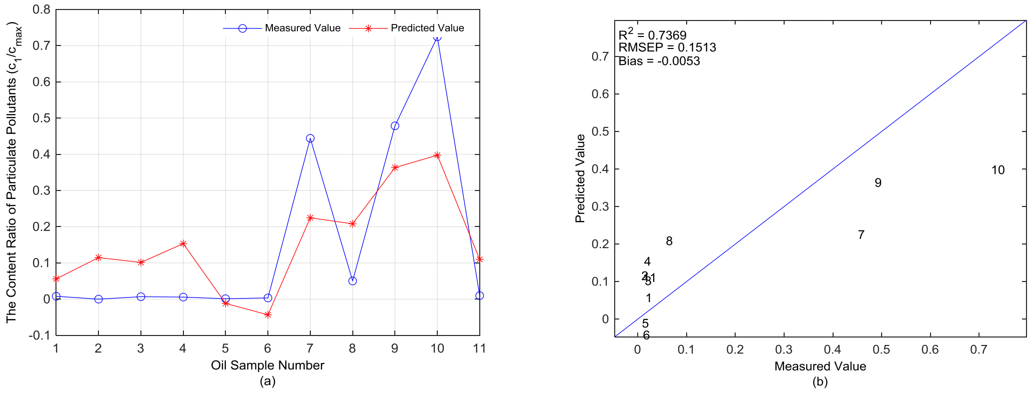

Fig. 8 The comparison of model predicted value and measured value of full frequency spectrum

Fig.9 The compressor of application of the proposed research

Table 1 SPA selected variables and their variables

method	number	variable
SPA	19	53.7109 , 55.6641 , 12.6953 , 49.8047 , 117.1875 , 179.6875 , 355.4688 ,8.7891 ,92.7734 ,13.6719 ,209.9609 ,56.6406 ,236.3281 , 19.5313 , 6.8359 , 169.9219 , 48.8281 , 11.7188 , 327.1484

Table 2 The model evaluation index

Input variable	Variable number	evaluation indexes	
		R^2	RMSE
full frequency spectrum	512	0. 7369	0. 1513
SPA selected	19	0. 8637	0. 1979

A brief technical biography of each author:

Liu Ge received the M.Eng. degree in Chenical Engineering from Kunming University of Science and Technology, Yunnan, China, in 2004. She is currently an associate Professor in chemical machinery, North China Institute of Science and Technology, Hebei, China. Her research interests include Chemical Mechatronics, signal processing, and measurement.

Chen Bin received the B.Eng. degree in Mechanical Engineering from Huazhong University of Science and Technology, Wuhan, China, in 2007. He is currently a Professor in Mechatronics, North China Institute of Science and Technology, Hebei, China. His research interests include Mechatronics, signal processing, control, and measurement.