

# Fault detection and diagnosis method of diesel engine by combining rule-based algorithm and Bayesian/neural networks

Baoping Cai (✉ [caibaoping@upc.edu.cn](mailto:caibaoping@upc.edu.cn))

China University of Petroleum <https://orcid.org/0000-0002-4499-492X>

**Xiutao Sun**

China University of Petroleum

**Jiaying Wang**

China University of Petroleum

**Yonghong Liu**

China University of Petroleum

**Chao Yang**

China University of Petroleum

**Xiangdi Kong**

China University of Petroleum

**Zengkai Liu**

China University of Petroleum

---

## Original Article

**Keywords:** Bayesian networks, BP neural networks, Diesel engine, fault diagnosis, rule-based algorithm

**Posted Date:** May 29th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-31637/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Fault detection and diagnosis method of diesel engine by combining rule-based algorithm and Bayesian/neural networks

Baoping Cai<sup>1,2\*</sup>, Xiutao Sun<sup>2,3</sup>, Jiaying Wang<sup>3</sup>, Yonghong Liu<sup>1,2</sup>, Chao Yang<sup>1,2</sup>, Xiangdi Kong<sup>1,2</sup>, Zengkai Liu<sup>1,2</sup>

**Abstract** The stable operation of diesel engine is critical to the normal production of the industry, and the prevention, monitoring and identification of faults are of great significance. At present, the fault research on diesel engine still has some defects, such as only few types of faults diagnosis are identified, the accuracy of fault diagnosis is still low, and fault identification is located at a constant speed. Therefore, a rule-based algorithm for fault diagnosis is proposed. Bayesian networks (BNs) and BP neural networks are used to identify seven faults at different speeds. Changchai EV80 diesel engine is taken as an example, and the feature values are extracted from the vibration signals measured from the cylinder head. The signals are processed by wavelet threshold de-noising and Ensemble Empirical Mode Decomposition (EEMD). The signal-sensitive feature values extracted from the decomposed Intrinsic Mode Function are used to distinguish different faults. After obtaining the feature values, a rule-based algorithm using IF... THEN's logic statement is proposed. BNs and BP neural networks established by parameter learning method are used for fault identification. Furthermore, this paper considers the uncertain factors and the interference of the external environment. Gaussian white noise is added to the raw signal and external excitation interference is applied to the diesel engine when it is running under normal operation condition. The results show that the proposed fault diagnostic method can accurately identify the faults.

**Keywords** Bayesian networks, BP neural networks, Diesel engine, fault diagnosis, rule-based algorithm

---

Correspondence: caibaoping@upc.edu.cn

<sup>1</sup>National Engineering Laboratory of Offshore Geophysical and Exploration Equipment, China University of Petroleum, Qingdao, Shandong 266580, China

<sup>2</sup>College of Mechanical and Electronic Engineering, China University of Petroleum (East China), Qingdao, Shandong 266580, China

<sup>3</sup>CRRC Qingdao Sifang Rolling Sock Research Institute CO., LTD., Qingdao, Shandong 266033, China

## 1 Introduction

Fault diagnostic accuracy of diesel engine is still low because of large number of components, rough work and large vibration, making it difficult to identify faults. In general, the fault diagnosis method of diesel engine can be classified into three types, including model-based method, signal-based method and knowledge-based method [1].

The earliest development and the most mature research is the model-based fault diagnosis method among them. The method has a profound theoretical foundation and can penetrate the essence of the dynamic system to achieve the identification of faults. For example, Salehi et al. [2] established a mathematical model of the engine exhaust pipe and identified accurately the faults of leakage and blockage. Nahim et al. [3] proposed a corresponding model for the defects of the diesel engine and pointed out the impact of the fault on the overall system. However, this method is only useful for the system with known mathematical models. It is difficult to obtain precise mathematical models of the system in actual situations. The uncertainty and time-variation of the system structure and parameters are difficult to predict, so the method has great limitations.

Signal-based method is also widely used for the fault diagnosis of diesel engine. The real-time operation of the diesel engine can be obtained through the sensors installed on the diesel engine and the online real-time monitoring system of the diesel engine without destroying the diesel engine is established. Jing et al. [4] identified the faults of different valve conditions by using the vibration acceleration signal accurately. Dykas et al. [5] measured the acoustic emission signal of the single-cylinder diesel engine and accurately identified the injector fault. However, sometimes the signal patterns extracted for the same type of fault each time are different, which interferes the identification of the fault.

The knowledge-based method is also called data-driven method. The method does not require prior mathematical model of the structure or component, and is ideal for the fault diagnosis of complex systems. Haghani et al. [6] used data-driven methods to detect experiment deviations, optimize engine parameters and improve

operation performance of diesel engine. Li et al. [7] monitored the pressure signals of diesel engines using the data-driven method to identify faults of valve trains and fuel systems. Data-driven method requires enough data to train the diagnostic model. In the current work, the faults are set artificially and enough data are obtained. This method requires a large amount of historical data. In order to obtain enough historical data, the experiment of fault diagnosis is repeated for many times.

The vibration sensors are used to collect the fault signals in this paper. The movement of the diesel engine belongs to the reciprocating mechanical motion which has a strong periodicity. One motion cycle consists of four strokes and the running curves of each stroke are completely different. Therefore, when the internal components of the diesel engine fail, the fault is reflected in the corresponding stroke process and the running curve also changes, so it is very convenient for the fault diagnosis of the diesel engine. Wang et al. [8] successfully identified the faults of the fuel failure and the fuel supply advance angle increase by decomposing the vibration signal. Wang et al. [9] calculated the cone-shaped nuclear distribution of the vibration signal under different faults of the diesel engine and identified the fault of the diesel engine valve mechanism. Yuan et al. [10] effectively monitored and evaluated the health and reliability of the engine based on the vibration signal method. Janssens et al. [11] proposed a multi-sensors system for fault diagnosis of rotating machinery by combining vibration sensor with other sensors. It turns out that multiple faults can be diagnosed effectively using multi-sensors.

Fault identification is a very important step in the fault diagnosis process of data-driven method. The faults cannot be accurately identified without the proper fault identification method when the fault signals are collected and decomposed. Proper identification methods can reduce the interference during signal acquisition and the errors during signal decomposition. The current popular fault identification methods include mixed causality and fuzzy rule based methods, blind source separation, support vector machines, BNs and neural networks. BNs have strong uncertainty problem processing ability because they can learn and reason under incomplete and uncertain

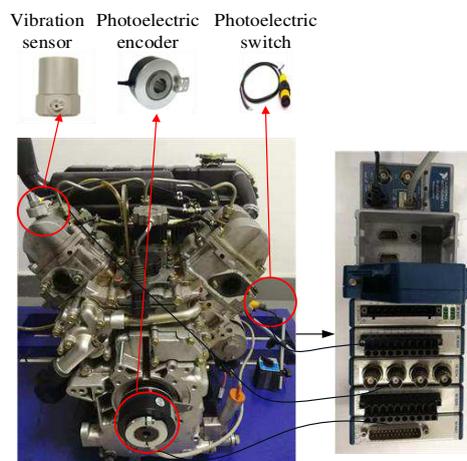
information conditions. BP neural networks have the arbitrarily complex pattern classification ability and excellent multidimensional function mapping ability. For example, Cai et al. [12] proposed an object-oriented BN modeling method for fault diagnosis of complex systems. Subrahmanya et al. [13] proposed a new algorithm of selecting sensors based on the Bayesian formula, and applied it to the fault diagnosis of diesel engines. Morgan et al. [14] used the feature values obtained from oil analysis to establish a BN to identify diesel engine faults. Zhao et al. [15] defined fault diagnosis disciplines using the BNs to identify the faults of air handling devices. Lu et al. [16] input the feature values extracted from the vibration signal into the BP neural network for identifying the turbine fault. Ahmed et al. [17] trained the network with BP neural networks and stacked self-encoder to achieve high-precision diagnosis of faults in rolling bearings. Wang et al. [18] trained the fault data of the electronic equipment system with BP neural networks to establish a fault diagnostic system for complex electronic equipment.

So far, many researchers have done a lot of work on the fault diagnosis of diesel engines. However, there are still many problems, for instance, only few types of faults are identified, the accuracy of fault diagnosis is still low, and fault identification is located at a constant speed. This paper proposes a diesel engine fault diagnostic method which combines the rule-based algorithm and BNs or BP neural networks. The method can accurately identify diesel engine faults under different speeds. The impact of white noise and external excitation on the diagnostic accuracy of diesel engines has also been verified. The structure of this paper is as follows: Section 2 describes the construction of the test bench and the simulation of the fault; Section 3 proposes the overall process of fault diagnosis, including signal acquisition, signal preprocessing, feature value extraction and fault identification; Section 4 gives the fault diagnostic results; Section 5 summarizes the paper.

## **2 System description and fault analysis**

The fault diagnosis system of diesel engine is shown in Figure 1. The vibration sensors, photoelectric switch, photoelectric encoder are installed on the diesel engine and the measured signals are delivered to the PC through

the corresponding data acquisition card. Both of the vibration sensors are placed on the same cylinder head, and they are located above the intake and exhaust valves of the cylinder head, respectively. The photoelectric encoder is placed at the output shaft of the diesel engine to monitor the real-time speed. A reflective sticker is attached to the combustion top dead position of the flywheel. The top dead position information is monitored using the photoelectric switch, and all signals are measured from the top dead position. It is convenient for the comparison of different faults signals.



**Figure 1** Diesel engine fault diagnosis platform

The engine used in the experiment is EV80 diesel engine produced by Changchai Company, and it is a two-cylinder, in-line V-type diesel engine with a speed range from 1800 rpm to 3600 rpm and a power range from 10.8 kw to 14 kw. The basic parameters are shown in Table 1. The vibration signals are measured by LC0103 accelerometers from Beijing Xiangyu Company, and its parameters are as follows: sensitivity 50 mV/g, resonant frequency 32kHz and resolution 0.0004g. The photoelectric switch adopts the Shanghai Kete KTE18-T30BP1 photoelectric switch, and its working distance varies from 10 to 500 cm. The speed is measured using the photoelectric incremental encoder with a resolution of 1024 pulses/r and three input terminals of A/B/Z from Shanghai Kaidi Electronics Corporation. The fault diagnostic system uses NI Labview software and NI acquisition

board. The vibration sensor is matched with the NI 9234 board, the photoelectric switch is matched with the NI 9421 board and the photoelectric encoder is matched with the NI 9401 board. The card slot uses CDAQ-9188.

**Table 1** Basic parameters of the EV80 diesel engine

Characteristics	Parameters
Cylinder number	2
Stroke number	4
Cylinder diameter×stroke	80mm×79mm
Displacement	0.794L
Cooling system	Water cooled
Rated power	14kW
Rated speed	3600rpm
Fuel system	Eddy current combustion
Lubrication method	Pressure and splash
Engine size	616mm×486mm×528mm

After a long period of operation, the diesel engine has various failure problems because of the hard working environment and large number of components. In the past literatures, the faults are mainly confined to the faults of valve clearance, such as the increasing or reducing of the valve clearance. In actual environment, the failure problems of the diesel engine contain a wide variety of faults. Therefore, the research is strengthened on the fault type of diesel engine. The faults are mainly divided into the following types, as shown in Table 2.

**Table 2** Setted fault types

Fault type	Setted details
Intake valve clearance with 0.1mm (IS)	The clearance between the valve and the transmission is adjusted to 0.1mm
Exhaust valve clearance with 0.8mm (EL)	The clearance between the valve and the transmission is adjusted to 0.8mm
The bending of push rod at exhaust valve (EC)	The putter is manually bent to a certain curvature
The leakage of high pressure oil pipe (OL)	The bolt is loosen at the high pressure inlet pipe to achieve slow oil leakage
The wear of intake valve base (ISE)	The side of the intake valve base is scratched so that the seal is not tight
The wear of cylinder pad (WE)	The black pad for the seal is scratched
The spring wear at the intake valve (ISP)	The spring is cut for a length by wire cutting machine

In addition, when the normal operating condition (NOR) is counted, there are totally 8 types and the valve clearance in the normal operating station is 0.2 mm.

### 3 Fault diagnosis methodology

#### 3.1 The proposed fault diagnosis methodology

The flow chart of the intelligent fault diagnosis method is shown in Figure 2. The different faults signals of the diesel engine are measured using the signal acquisition system. The signals are de-noised by the wavelet threshold method, and then are decomposed by the Ensemble Empirical Mode Decomposition (EEMD) method. Different layers of Intrinsic Mode Function (IMF) are selected to extract the feature values. Finally, the faults are identified using BP neural networks or BNs. There are many external interferences during the process of fault diagnosis, including the external environmental interference and the bias of sensors. The effects of the uncertainty factors for the accuracy should be considered. Gaussian white noise and external excitation are added to the signal and the diesel engine, respectively, and the effects are researched.

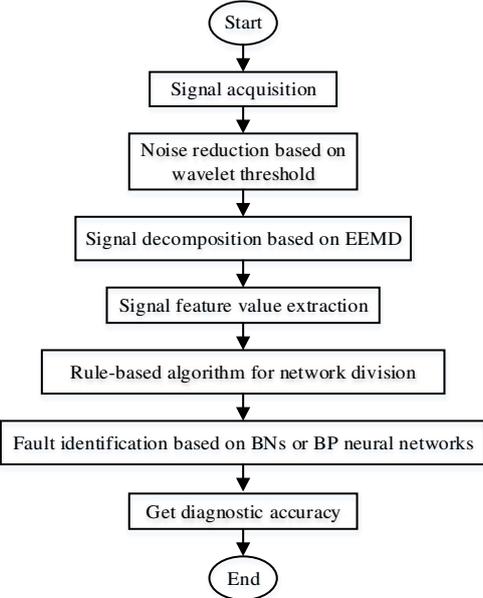


Figure 2 Overall fault diagnosis flowchart

### 3.2 Signal noise reduction and feature value extraction

Decomposing various frequency components of the signal into mutual non-overlapping frequency bands provides an effective way for signal filtering, signal-to-noise separation and feature extraction. It has a unique advantage in signal de-noising. The principle of wavelet threshold de-noising can be expressed as follows:

$$x(n) = s(n) + u(n) \quad (1)$$

where  $s(n)$  is useful signal,  $u(n)$  is sequence signal of noise. Assume that  $u(n)$  is zero-mean, and obeys Gaussian distribution of random sequences and the distribution of  $N:(0, \sigma_u^2)$ , the above Eq. (1) can be carried out by the wavelet transform, defined as:

$$WT_x(a,b) = WT_s(a,b) + WT_u(a,b) \quad (2)$$

Since  $u(n)$  is a zero mean and stationary random signal with independent distribution. Note that  $u = (u(0) u(1) \dots u(N-1))^T$ , then the formula can be defined as follows:

$$E\{uu^T\} = \sigma_N^2 I \triangleq Q \quad (3)$$

Here,  $E\{\bullet\}$  represents the mean operation,  $Q$  is the covariance matrix of  $u$ ,  $I$  is a unit matrix.

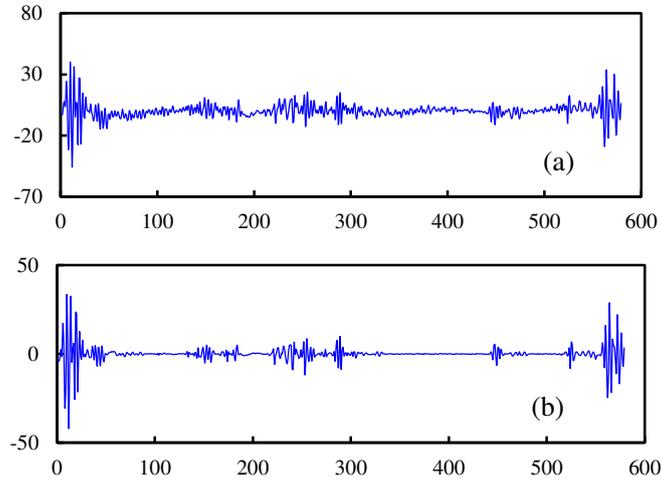
Assume that  $W$  is the wavelet transform matrix,  $x$  and  $s$  are the corresponding vectors of  $x(n)$  and  $s(n)$ . The vectors  $X$ ,  $S$ , and  $U$  is the wavelet transform of  $x(n)$ ,  $s(n)$ ,  $u(n)$ , respectively.  $P$  is the covariance matrix of  $U$ , then the formula can be defined as follows:

$$P = E\{UU^T\} = E\{Wuu^TW^T\} = WQW^T \quad (4)$$

It can be seen from Eq. (4) that after experiencing orthogonal wavelet transform, the energy of the signal is concentrated on a few sparse, large amplitude wavelet coefficients, and the noise is distributed on all time axes with various time scales. The amplitude is small and the noise interference is suppressed.

The raw signal not only contains the fault information but also contains the signal interference generated by the interaction of many other components and the external environment. Therefore, the raw signal needs to be de-

noised by the wavelet threshold method to decrease high frequency interference. The normal operating condition is taken as an example to compare the different signals before and after signal de-noising. The Figure 3 shows the comparison results.



**Figure 3** Comparison of time domain signals before and after signal de-noising (a) Original signal, (b) Noise reduction signal.

The EEMD method is an adaptive time-frequency analysis method, which basically gets rid of the limitations of the Fourier transform method, and is suitable for the processing of non-stationary signals. The method utilizes the statistical characteristics of Gaussian white noise with uniform frequency distribution. The signal added to white noise has a uniformly distributed decomposition scale. This method can smooth abnormal events such as pulse interference, and effectively solve the problem of pattern aliasing and false components. The basic algorithm is as follows:

$y_i(t)$  can be obtained by adding normal distributed white noise  $n_i(t)$  with equal length to the raw signal  $y(t)$  for multiple times.

$$y_i(t) = y(t) + n_i(t) \quad (5)$$

where  $y_i(t)$  is the signal that is added to white noise for the  $i$ th time.

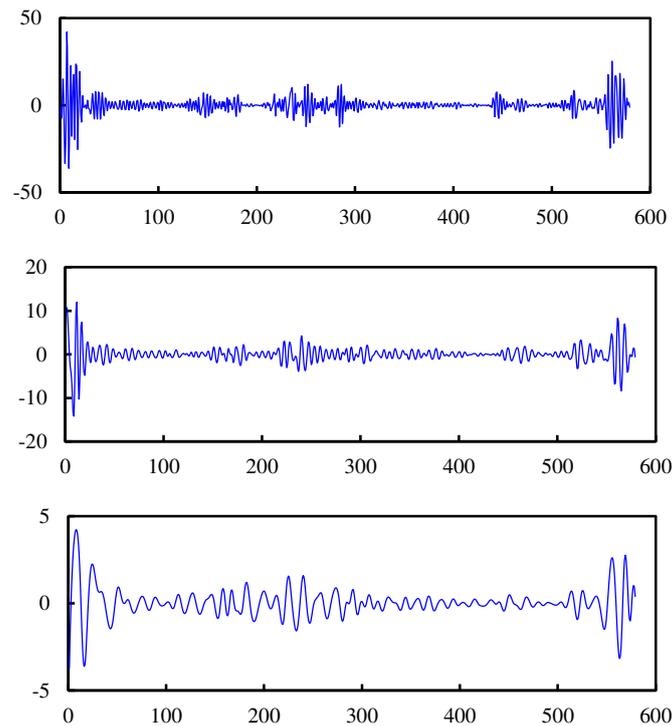
$y_i(t)$  is decomposed by the EEMD method to get component  $c_{ij}(t)$  and residual  $r_i(t)$ . The principle that the statistical mean of the uncorrelated random sequence is zero is utilized, then the components of  $c_{ij}(t)$  are averaged

to compensate the effect of white noise on the real IMF for multiple times. The final result of EEMD decomposition is shown as follows:

$$c_j(t) = \frac{1}{N} \sum_{i=1}^N c_{ij}(t) \quad (6)$$

where  $N$  refers to the number of added white noise sequences,  $c_j(t)$  is the averaged IMF component.

The normal operating condition is decomposed by EEMD method. The result is presented in Figure 4. Figure 4 shows the first three layers of signals after signal decomposition. It can be seen that the Intrinsic Mode Function in each layer contain the signals with the same frequency component. The problems of modal aliasing and endpoint effect are eliminated. The EEMD algorithm can decompose the signals based on different time scales, and satisfy the experimental requirements.



**Figure 4** The first three layers of IMFs after decomposition by EEMD

The decomposed fault signals are used to extract the feature values by some certain mathematical method. The feature values are used as the input parameters in the subsequent fault identification. The peak of the marginal spectrum, the root mean square value, the mean, the integrated area of the spectrum, the peak of the time domain

image and the integrated area of the envelope spectrum are selected as feature values. A large number of feature values can be obtained but not all that are helpful for the fault diagnosis. All feature values extracted from the signals need to be sorted and filtered. When the feature values are sorted in ascending or descending order, some feature values mix many faults together. These feature values have no effect on fault identification, and need to be deleted. Some feature values have the same effect on distinguishing faults, so only one feature value needs to be selected, and the other redundant feature values need to be removed.

### **3.3 Rule-based algorithm for model division**

The intelligent fault diagnostic system is able to identify all kinds faults at different speeds. The measured speed range varies from 1800 rpm to 2600 rpm. In the current work, it is unrealistic that only a model of BNs achieves the goal of seven faults identification within such a large speed span. Therefore, the model is established at 1900 rpm, 2200 rpm and 2500 rpm, respectively. The maximum speed span is 100 rpm in each model at central speed. The measured marginal speeds contain 1800 rpm, 2000 rpm, 2100 rpm, 2300 rpm, 2400 rpm and 2600 rpm. Therefore, all faults at all speeds can be diagnosed and identified.

In each model, seven faults are divided into three sub-models. The principle of dividing seven faults into three sub-models is to utilize the rule-based algorithm. Some feature values can clearly distinguish one or two types of faults. The faults are divided into different models in sequence based on the feature values, so the difference of different feature values in each model is significant. The number of faults in each model is less and the accuracy of diagnosis is improved. When the magnitude of the feature value meets the threshold setting of rule algorithm, the rules are triggered and the program is executed. The signals are automatically divided into the specified sub-model to generate an intelligent diagnostic classifier using IF...THEN logical statement. By selecting and dividing the faults identified by the feature values, it can be seen that the best diagnosis effect can be obtained using three models at each center speed. The number of models does not make the diagnosis system too complicated. The

operating procedure of fault diagnosis is as follows: the fault signals are measured by two vibration sensors placed on the cylinder head, which are de-noised and decomposed by Matlab software. The feature values are extracted from the processed signals to obtain a large number of feature value data streams. The fault signals are classified into different sub-models based on the magnitude of feature values. Then the faults in each sub-model are distinguished in detail through the magnitude judgment of multiple feature values. Finally, the fault identification results are obtained.

All fault signals at the central speed of 2500 rpm are taken as an example. The rule-based algorithm is used to divide the faults into three sub-models. The meaning of feature value node is shown in Table 3.

**Table 3** Physical meaning of feature value node

Nodes	Physical meaning
TDP	the peak of the time domain signal
TDA 1-4	the integral area of time domain signal with different frequency
RMS 1-4	root mean square value of signal with different frequency
AVE	mean value of the signal
FRE 1-4	the integrated area of spectrum with different frequency

The feature values that can obviously distinguish the fault are sorted in ascending order. It can be seen from Table IV that when the feature values satisfy the condition of  $TDA4 < 1111$  &&  $RMS2 < 152$ , the normal operating condition (NOR) can be distinguished. When the feature values satisfy the condition of  $TDA4 < 1220$  &&  $RMS3 < 130$  &&  $AVE < 114$ , the leak of high pressure fuel pipe (OL) can be distinguished. The threshold interval of different fault feature values is different. The magnitude of feature value that can accurately distinguish the faults can be chosen by sorting, and the feature value is defined as the interval threshold parameter. The program is shown in Table 4.

**Table 4** Rule-based program algorithm for model division

---

---

```
1: Given  $X=(X_1, X_2, \dots, X_n)$  data matrix
TDP1-4( $X$ ), TDA1-4( $X$ ), RMS1-4( $X$ ), AVE ( $X$ ), FRE1-4( $X$ ), function
2: for  $i=1:n$  do
3: if (TDA4<1111&&RMS2<152)  $Y(i)=\text{network1}$ 
4: else if (TDA4<1220&&RMS3<130&&AVE<114)  $Y(i)=\text{network1}$ 
5: else if (TDP1<15.3&& AVE<127)  $Y(i)=\text{network1}$ 
6: else if (TDP1>24.4&& TDA3>1660&&RMS2>198)  $Y(i)=\text{network3}$ 
7: else  $Y(i)=\text{network2}$ 
8: end if
9: end for
10: output classification result matrix  $Y=(Y_1, Y_2 \dots Y_n)$ 
```

---

---

### 3.4 Fault identification using BNs

BNs are probabilistic networks and the specific occurring probability of an event is inferred by establishing structure and parameter relationship between the parent and child nodes. It is a graphical network based on probabilistic reasoning. The relationship between fault type nodes and fault feature value nodes can be obtained through the parameter learning method. The BNs can be used to solve the problems of uncertainty and incompleteness. They have great advantages for identifying faults caused by the uncertainty and correlation of complex equipment [19].

The main algorithm in parameter learning is the Bayesian estimation method based on the estimation of complete data parameters. The samples are independent, and satisfy the principle of independent and identical distribution. All possible values of parameter  $t$  in this topology  $D$  are given, and the parameter value of maximum posterior probability given  $D$  using prior knowledge can be obtained.

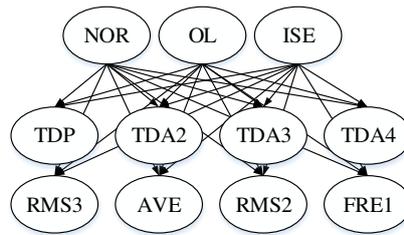
$$P(\theta|D,S) = \frac{P(\theta|D,S)P(\theta|S)}{P(D|S)} \quad (7)$$

where  $P(\theta|D,S)$  is the prior probability of parameter  $\theta$  in state  $S$ , and satisfies the Dirichlet distribution.  $\theta$  meets the conditions of  $\sum_{i=1}^k \theta_i = 1$ . Dirichlet distribution is composed of  $\alpha_1 \dots \alpha_k$ . When  $P(\theta|S)$  satisfies the

Dirichlet distribution, the corresponding posterior probability is as follows:

$$P(\theta|D,S) = \text{Dirichlet}(\theta|\alpha_1 \cdots \alpha_k) = \frac{T(\alpha)}{\prod_i T(\alpha_i)} \prod_i \theta^{\alpha_i-1} \quad (8)$$

The first model of BNs at the central speed of 2500 rpm is established to obtain the fault diagnostic results. The faults include normal operating condition (NOR), the leak of high pressure fuel pipe (OL) and the wear of intake valve base (ISE). The measured signals contain 30 groups, 80% of them are used for training models and 20% are used for result validation. The measured fault signals whose speed is near the central speed contain 8 groups, and all of them are used for result validation. The first model of BNs is shown in Figure 5, three parent nodes represent three fault states and eight child nodes represent eight feature values.



**Figure 5** First model of BNs at 2500 rpm

The second model and the third model at 2500 rpm speed are established using the same method of the fault model. Similarly, three fault sub-models at 1900 rpm speed and 2200 rpm speed are established. All fault signals are divided into corresponding sub-models using the IF...THEN logical statement. A complete intelligent fault diagnostic system of diesel engine based on BNs is built. The overall fault diagnostic flow chart is shown in Figure 6. Taking the fault of oil leakage with the central speed of 2500 rpm as an example, the fault is firstly divided into the first sub-model of BNs by judging the magnitude of the feature values, that is, TDA4, RMS3 and AVE, then the fault of oil leakage is accurately identified using the feature values of TDP, TDA, RMS, AVE and FRE in the first sub-model of BNs.

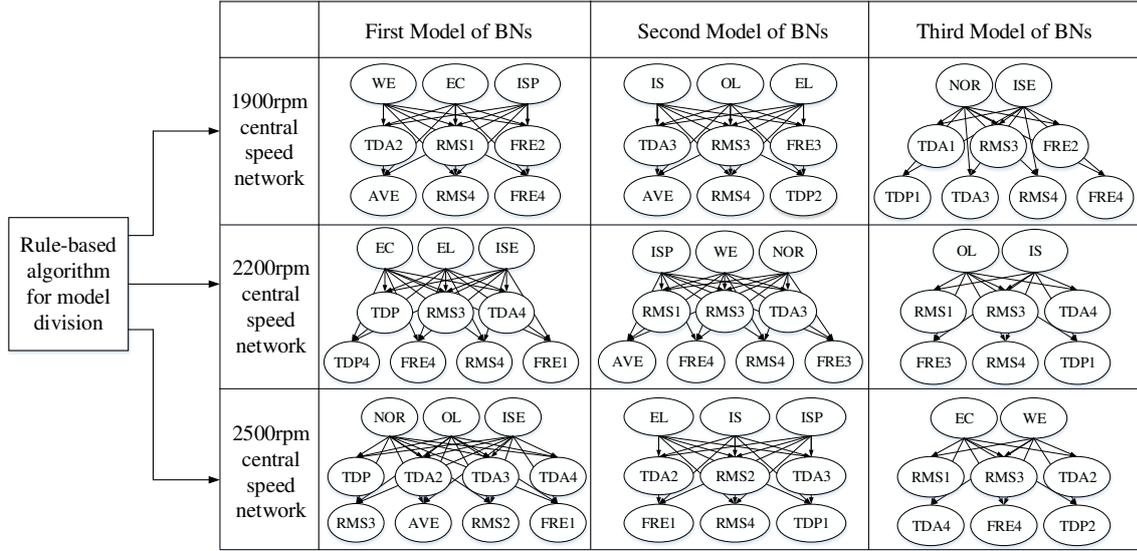


Figure 6 Complete BNs-based fault diagnosis model of diesel engine

### 3.5 Fault identification using BP neural networks

Neural networks have the capabilities of self-learning, nonlinear mapping, approximation for arbitrary functions and parallel computing. The capabilities provide a powerful guarantee for constructing new fault diagnostic model. The learning process of BP neural networks is divided into two stages. The first stage is to input the known learning samples and calculate the output of each neuron through the network structure, weights and thresholds. The second stage is to modify the weight and threshold to minimize the total error  $E$ .

The weight and the threshold between the input layer and the hidden layer is  $w_{ij}$  and  $\theta_j$ , the weight and the threshold value between the hidden layer and the output layer is  $w_{jk}$  and  $\theta_k$ , then the output of each layer neuron is as follows:

$$\begin{aligned}
 x'_j &= f(u_j) = f\left(\sum_{i=0}^n w_{ij}x_i\right) \\
 y_k &= f(u_k) = f\left(\sum_{j=0}^m w_{jk}x'_j\right)
 \end{aligned} \tag{9}$$

where  $x'_j$  is the output of hidden layer neuron,  $y_k$  is the output of the output layer neuron and  $x_i$  is the input of hidden layer neuron.

The sample  $p$  is input into the network, then the output  $y_k^{p1}$  ( $k=0,1,\dots,m-1$ ) is obtained. The error is the sum of

the output units errors.  $t_k^p$  is the expected output for the corresponding sample. The total error for  $p$  learning samples can be defined as:

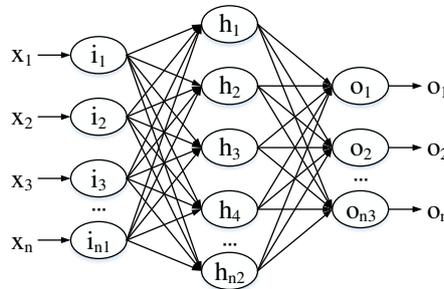
$$E_{sum} = \frac{1}{2} \sum_{p=1}^p \sum_{k=0}^{m-1} (t_k^{p1} - y_k^{p1})^2 \quad (10)$$

Assume that  $w_{sq}$  is the connection weight between any two neurons in the network, and the correction value using the gradient method to correct each  $w_{sq}$  element is defined as:  $\Delta w_{sq} = -\sum_{p=1}^p \eta \frac{\partial \varepsilon}{\partial w_{sq}}$ ,  $\eta$  is the step size, for which

$$\Delta E_{sum} = -\eta \sum_{p=1}^p \sum_{sq} \left( \frac{\partial \varepsilon}{\partial w_{sq}} \right)^2 \leq 0 \quad (11)$$

The total error changes with the direction of decreasing through the gradient method, which is stopped until  $\Delta E_{sum}=0$ .

The basic structure of BP neural networks is shown in Figure 7.  $x_n$  is the input layer of the network,  $h_n$  is the hidden layer of the network, and  $o_n$  is the output layer of the network.



**Figure 7** The structure of BP neural networks

Similarly, the rule-based algorithm is used to classify faults into different sub-models before the fault is identified by BP neural networks. The overall diagnostic flow chart is shown in Figure 8. Taking the wear of cylinder pad with the central speed of 2500 rpm as an example, the fault is firstly divided into the third sub-model of BNs by judging the magnitude of the feature values, that is, TDP1, TDA3 and RMS2, then the fault of the wear of cylinder pad is accurately identified using the feature values of TDP, TDA, RMS and FRE in the third sub-model of BNs.

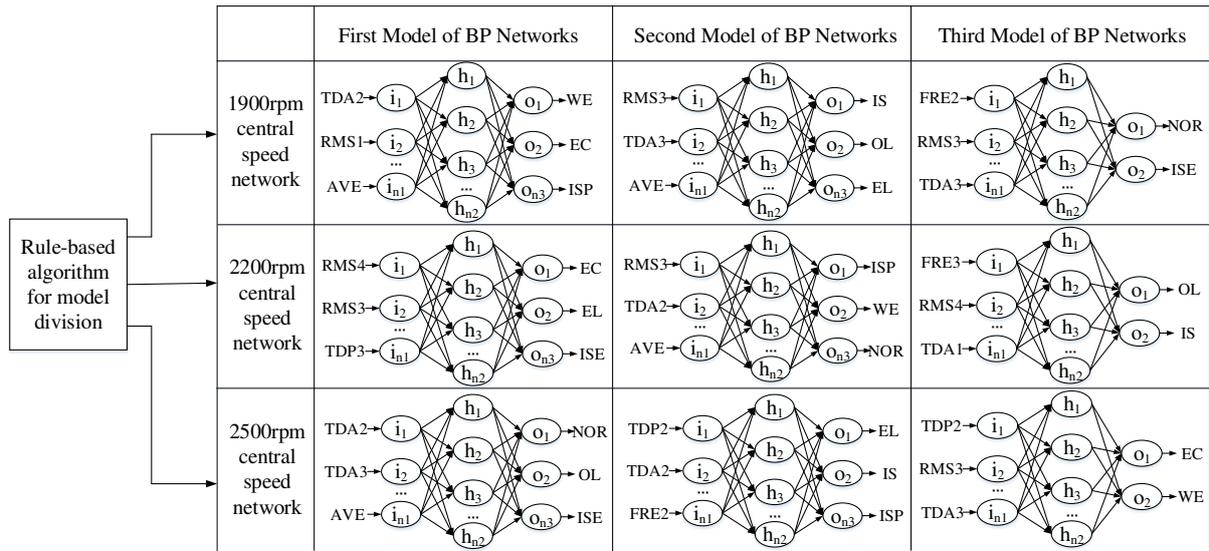


Figure 8 Complete BP-neural-networks-based fault diagnosis model of diesel engine

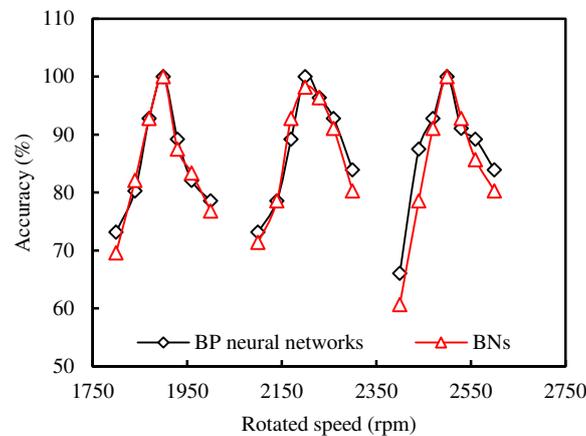
## 4 Results and discussion

### 4.1 Fault identification using BNs and BP neural networks

All fault signals can be obtained through the above-mentioned test bench of fault diagnosis, and the fault signals are de-noised by wavelet threshold method and decomposed by EEMD method. Eight fault signals are divided into three sub-models by the rule-based algorithm after the feature values are extracted. Finally, the faults are identified by BP neural networks or BNs. The diagnostic results are shown in Figure 9.

It can be seen that the diagnostic accuracy using BP neural networks is similar to that using BNs in each model established at each central speed. The accuracies of fault diagnosis using different identification method are high in the middle and low on both sides. Furthermore, the diagnostic accuracy with higher speed than the central one is completely higher than that with lower speed than the central one. When the speed exceeds the central speed and the speed difference reaches 100 rpm, the accuracy of fault diagnosis is basically higher than 80%, while the speed is lower than the central speed and the speed difference reaches 60 rpm, the accuracy of fault diagnosis is higher than 80%, and the diagnostic accuracy can reach more than 60% when the speed difference reaches 100

rpm. Only when the central speed is 2500 rpm, the diagnostic accuracy using BP neural network is slightly better than that using BNs. When the speed is 2440 rpm, the accuracy of diagnosis using BP neural networks is 8.9% higher than that using BNs. This result shows that BP neural networks have good compatibility for high speed and high intensity vibration. The BNs interval division is artificially defined, and the goal of perfect interval division cannot be achieved with limited training data. When the diesel engine is running at high speed, once the engine has abnormal fluctuation, the extracted feature values will exceed the upper limit of the interval threshold, which results in the decrease of diagnostic accuracy.

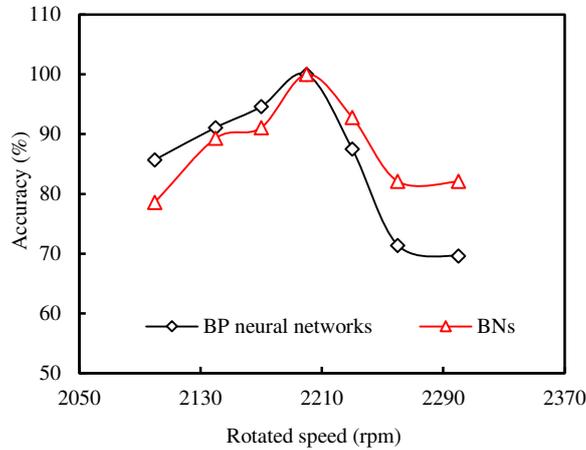


**Figure 9** Diagnostic results using BP neural networks and BNs

#### 4.2 Impact of different layers of signals on fault diagnosis

In the above mentioned fault diagnostic process of diesel engine, the processed signal is the first layer signal decomposed by EEMD method. The frequency range of each fault is different. For example, the vibration response caused by the pressure of combustion gas is a low-frequency response, and it is mainly concentrated at 1 kHz. While the valve seat impact is a high-frequency response, and it is mainly concentrated at 2-4 kHz. It is uncertain to determine which layer or layers signals are the best for fault diagnosis, since all first few layers of the decomposed signals contain the fault characteristics of the original signal, and the signal composition needs to be analyzed. The synthesized signal consisting of the second, third and fourth layers of decomposed signals is used

to identify the faults. The subsequent methods of signal decomposition and feature value extraction are the same as that adopted by the first layer decomposed signal. Taking the fault signals with the central speed at 2200 rpm as an example, the result of fault diagnosis is shown in Figure 10.



**Figure 10** Synthesized signal diagnostic results using BP neural networks and BNs

It can be inferred from Figure 10 that when the speed is lower than the central speed, the diagnostic accuracy using BP neural network is slightly better than that using BNs, and the maximum difference of accuracy reaches 7.1% at 2100 rpm speed. However, when the speed exceeds the central speed, the BNs are obviously superior to the BP neural networks. Especially the higher the speed is, the more obvious the difference of accuracy is. The maximum difference can reach 12.5% at the 2300 rpm speed. On the whole, when the synthesized signal consisting of the second, third and fourth layers of decomposed signals is used to identify the faults, the method using BNs is superior to the method using BP neural networks. The synthesized signal is composed of low frequency signals, which naturally contain a little high frequency information. When the measured speed is lower than the central speed, the vibration intensity weakens and the extracted feature values are lower than the low limit of the interval threshold, which can cause a lower accuracy of fault diagnosis. On the contrary, when the measured speed exceeds the central speed, the vibration intensity increases and the high-frequency information of the measured signals increases, which can compensate the shortcomings of little high-frequency information in the synthesized signal. So the interval of the feature values is stable and the fault diagnostic accuracy using BNs is high.

The accuracy of fault diagnosis with the first layer signal and synthesized signal at different speeds using BP neural networks and BNs can be obtained from Figure 9 and Figure 10. It can be observed that the trend lines of accuracy using different fault diagnosis methods are similar. The maximum difference of accuracy using BP neural networks at the both ends of the trend line is 12.5% and 21.4%, respectively. When the BNs are used to identify faults, the maximum difference of accuracy is 10.7% and 9%, respectively. The accuracy using the first layer signal is relatively high at high speeds, and the accuracy using the synthesized signal is high at low speeds. Since the EEMD algorithm decomposes the original signal according to the magnitude of the time feature scale, the first layer decomposed signal contains more high frequency information and the synthesized signal contains more low frequency information. When the speed of diesel engine increases, the vibration frequency increases and the high-frequency signal contains more fault information, so the diagnostic accuracy using the first layer signal is high. When the speed of diesel engine reduces, the vibration frequency slows down and the low-frequency signal contains more fault information, the effect of fault diagnosis using the synthesized signal is better.

#### **4.3 Influence of Gaussian white noise**

In this section, the effect of Gaussian white noise on the accuracy of fault diagnosis is studied. Gaussian white noise is used to simulate the unknown real noise. In the real environment, the noise is often not caused by a single source, and is a noise compound of many different sources. The noise source in most electronic systems such as radar and communication systems is thermal noise, and thermal noise is typical Gaussian white noise.

Figure 11 shows the accuracy of fault diagnosis when Gaussian white noise is added to the signal before signal denoising. The Gaussian white noise from 0 dB to 10 dB is added to the measured signal. The diagnostic accuracy of signal with white noise is basically the same as that without white noise in Figure 11(a). The fault diagnostic accuracy of the signal with signal-to-noise ratio (SNR) 0 dB is only reduced by 3.6% at 2230 rpm speed. It can be seen from the Figure 11(b) that the fault diagnostic accuracy of the signal with SNR 0 dB is reduced by 5.3%

at 2230 rpm speed. It can be concluded from Figure 11(a) and Figure 11(b) that the method using BNs is superior to the method using BP neural networks. The signal de-noising method can eliminate the added Gaussian white noise, the noise interference of the environment and the precision interference of the sensor can be eliminated.

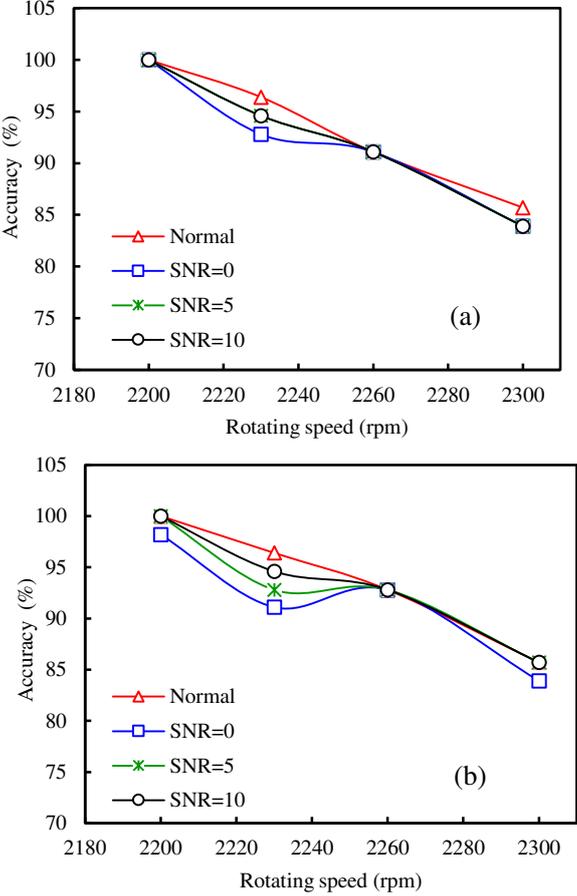
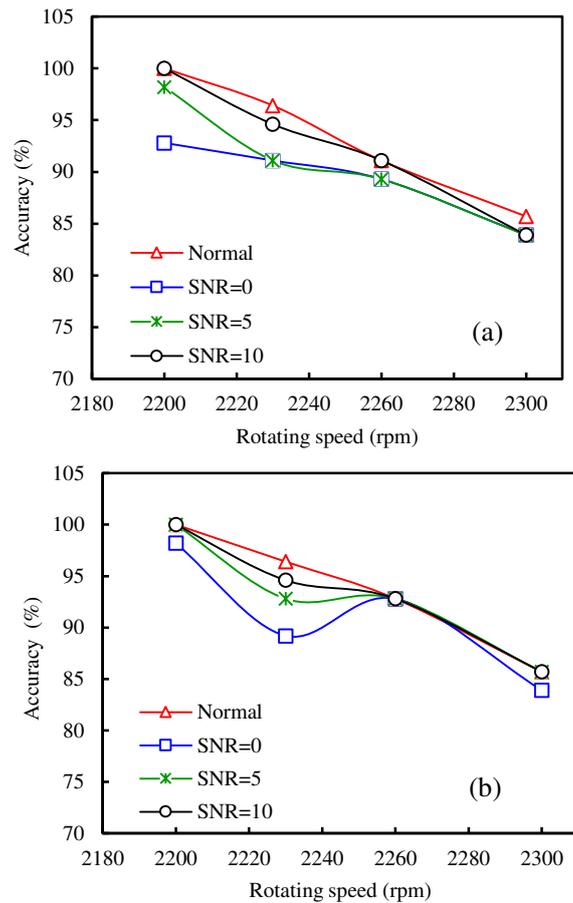


Figure 11 Fault diagnostic results when white noise is added before signal de-noising (a) BNs, (b) BP neural networks

Figure 12 shows the fault diagnostic accuracy when Gaussian white noise is added to the signal after signal de-noising. Similarly, the Gaussian white noise from 0dB to 10dB is added to the measured signal. It can be seen from the Figure 12(a) that the fault diagnostic accuracy of the signal with SNR 0 dB is reduced by 5.3% at 2230 rpm speed, and is reduced by 7.2% at 2200 rpm speed. As shown in Figure 12(b), the fault diagnostic accuracy of the signal with SNR 0 dB is reduced by 7.2% at 2230 rpm speed. It can be concluded from Figure 12(a) and Figure 12(b) that the method using BP neural networks is superior to the method using BNs. Combining Figure 11 and Figure 12, the signal de-noising method and the fault diagnostic method can accurately identify the faults.



**Figure 12** Fault diagnostic results when white noise is added after signal de-noising (a) BNs, (b) BP neural networks

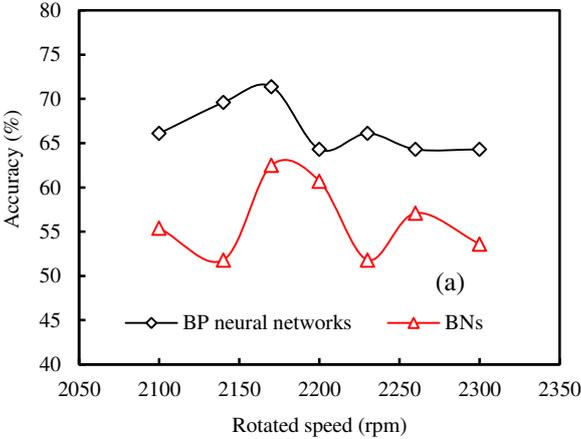
#### 4.4 Influence of external excitation interference

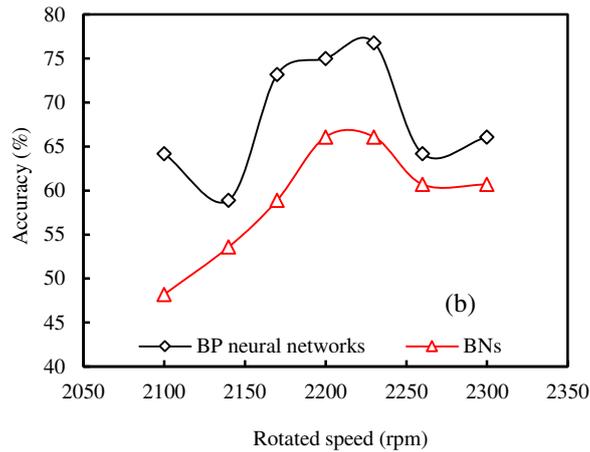
In this section, the external excitation is applied to the diesel engine, including continuous excitation (the gear driven by a motor wears the lower end of the cylinder head) and intermittent excitation (the screw nut drew by a motor strikes the diesel engine intake pipe). The excitation is not directly applied to the cylinder head, and is transferred to the cylinder head through some certain conduction.

Figure 13 shows the fault diagnostic accuracy using different fault identification methods when external excitation is applied to diesel engine. It can be seen that the trend lines of accuracy have great fluctuation and have no obvious regularity, The accuracy of fault diagnosis is greatly affected by external excitation interference. The result of fault diagnosis using BP neural networks is significantly better than that using BNs, no matter excitation is caused by continuous excitation or intermittent excitation. It can be inferred that BP neural networks have more

strong toleration than BNs for external interference. The maximum difference of diagnostic accuracy can reach 17.8% when continuous excitation is applied to diesel engine. When intermittent excitation is applied to diesel engine, the maximum difference of diagnostic accuracy can reach 16%. The reason is that when BNs are used to identify faults, the interval division of feature value is fixed and cannot be changed arbitrarily. Some differences between different interval values are small. Therefore, when the diesel engine is disturbed by external excitation, the magnitude of extracted feature value changes, which causes the change of original interval position. Then the posterior probability of the fault accordingly changes.

The same feature value extracted at different speeds or even the same speed has a relatively large variation range when the same excitation is applied to the diesel engine. The reason is as follows: when the continuous excitation is applied to diesel engine, the contact tightness between the gear and the diesel engine changes with the vibration intensity of the diesel engine. The produced frictional excitation also changes accordingly. The generated excitation signal is unstable, and the variation range of feature value varies greatly. Due to the discontinuity of the intermittent excitation, when the vibration signal is measured within one operation cycle, the intermittent excitation may not be collected. Therefore, the variation range of the feature value will increase, and a situation that a certain feature value suddenly increases at some speed will show up, which reduces the accuracy of the fault diagnosis.





**Figure 13** Fault diagnostic results under different excitations (a) Continuous excitation (b) Intermittent excitation

## 5 Conclusion

An intelligent method of fault diagnosis for diesel engine by combining rule-based algorithm and Bayesian/neural networks is proposed. The innovation of the method is that using a small amount of data trained at a speed realizes the fault diagnosis with a wide speed range. The fault diagnosis application on Changchai EV80 two-cylinder diesel engine is researched and some important conclusions are summarized as follows.

(1) The diagnostic accuracy with higher speed than the central one is completely higher than that with lower speed than the central one. When the speed exceeds the central speed and the speed difference reaches 100 rpm, the accuracy of fault diagnosis is basically higher than 80%, while the speed is lower than the central speed and the speed difference reaches 60 rpm, the accuracy of fault diagnosis is higher than 80%, and the diagnostic accuracy can reach more than 60% when the speed difference reaches 100 rpm. When synthesized signal is used to identify faults, the method using BNs is superior to the method using BP neural networks and the maximum difference of accuracy can reach 12.5% at 2300 rpm speed.

(2) When white noise is added to the signal before signal de-noising, the method using BNs is better than the method using BP neural networks, and the diagnostic accuracy is increased by 1.7% at 2230rpm speed. When white noise is added to the signal after signal de-noising, the method using BP neural networks is better than the

method using BNs, and the diagnostic accuracy is increased by 5.3% at 2300 rpm speed.

(3) The accuracy of fault diagnosis is greatly affected by external excitation interference, and the accuracy using BP neural networks is higher than the accuracy using BNs. When BP neural networks are used to identify the faults, the accuracy of fault diagnosis for continuous excitation can reach more than 60%, and the highest accuracy reaches more than 70%. While the intermittent excitation is applied to the diesel engine, the fluctuation of the accuracy is relatively large. The highest accuracy can reach 76.8% and the lowest accuracy can reach 58.9%.

### **Acknowledgements**

The authors sincerely thank Professor Xiaopeng Li of *China University of Petroleum* for his critical discussion and reading during manuscript preparation.

### **Authors' Contributions**

BC provided fundamental ideas and all support conditions of this paper. XS was in charge of the trial and wrote the manuscript. JW, YL and YC conducted proof reading and made some critical revisions. XK and ZL assisted the trial and experiment. All authors read and approved the final manuscript.

### **Funding**

This work was supported by the National Natural Science Foundation of China (No. 51779267), the Taishan Scholars Project (tsqn201909063), the Fundamental Research Funds for the Central Universities and the Opening Fund of National Engineering Laboratory of Offshore Geophysical and Exploration Equipment (20CX02301A), and the Science and Technology Support Plan for Youth Innovation of Universities in Shandong Province (No. 2019KJB016).

## Competing Interests

The authors declare no competing financial interests.

## References

- [1] L Wen, X Y Li, L Gao, et al. A New Convolutional Neural Network Based Data-Driven Fault Diagnosis Method. *IEEE Transactions on Industrial Electronics*, 2017, 65(7): 5990-5998.
- [2] R Salehi, A Alasty, M Shahbakhti and G Vossoughi. Detection and isolation of faults in the exhaust path of turbocharged automotive engines. *International Journal of Automotive Technology*, 2015, 16(1), 127-138.
- [3] H M Nahim, R Younes, H Shraim, and M Ouladsine. Oriented review to potential simulator for faults modeling in diesel engine. *Journal of Marine Science and Technology*, 2016, 21(3), 533-551.
- [4] Y B Jing, C W Liu, F R Bi, X Bi, Y Wang and K Shao. Diesel engine valve clearance fault diagnosis based on features extraction techniques and FastICA-SVM. *Chinese Journal of Mechanical Engineering*, 2017, 30(4), 991-1007.
- [5] B Dykas and J Harris. Acoustic emission characteristics of a single cylinder diesel generator at various loads and with a failing injector. *Mechanical Systems and Signal Processing*, 2017, 93, 397-414.
- [6] A Haghani, T Jeinsch, M Roepke, S X Ding and N Weinhold. Data-driven monitoring and validation of experiments on automotive engine test beds. *Control Engineering Practice*, 2016, 54, 27-33.
- [7] Y Li, W Zhou and Y Zi. A graphic pattern feature-mapping-based data-driven condition monitoring method for diesel engine malfunction identification and classification. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 2019, 233(1), 202-212.
- [8] X Wang, C Liu, F Bi, X Bi and K Shao. Fault diagnosis of diesel engine based on adaptive wavelet packets and EEMD-fractal dimension. *Mechanical Systems and Signal Processing*, 2013, 41(1-2), 581-597.
- [9] Q H Wang, Y Y Zhang, L Cai and Y S Zhu. Fault diagnosis for diesel valve trains based on non-negative matrix factorization and neural

- network ensemble. *Mechanical Systems and Signal Processing*, 2009, 23(5), 1683-1695.
- [10] N Yuan, W Yang, B Kang, S Xu and C Li. Signal fusion-based deep fast random forest method for machine health assessment. *Journal of Manufacturing Systems*, 2018, 48, 1-8.
- [11] O Janssens, M Loccupier and S Hoecke. Thermal Imaging and Vibration-Based Multisensor Fault Detection for Rotating Machinery. *IEEE Transactions on Industrial Informatics*, 2018, 15(1), 434-444.
- [12] B P Cai, H Liu and M Xie. A real-time fault diagnosis methodology of complex systems using object-oriented Bayesian networks. *Mechanical Systems and Signal Processing*, 2016, 80, 31-44.
- [13] N Subrahmanya, Y C Shin and P H Meckl. A Bayesian machine learning method for sensor selection and fusion with application to on-board fault diagnostics. *Mechanical systems and signal processing*, 2010, 24(1), 182-192.
- [14] I Morgan, H Liu, B Tormos and A Sala. Detection and diagnosis of incipient faults in heavy-duty diesel engines. *IEEE Transactions on Industrial Electronics*, 2009, 57(10), 3522-3532.
- [15] Y Zhao, J Wen and S Wang. Diagnostic Bayesian networks for diagnosing air handling units faults–Part II: Faults in coils and sensors. *Applied Thermal Engineering*, 2015, 90, 145-157.
- [16] S Lu, J Wang and Y Xue. Study on multi-fractal fault diagnosis based on EMD fusion in hydraulic engineering. *Applied Thermal Engineering*, 2016, 103, 798-806.
- [17] H O A Ahmed, M L D Wong and A K Nandi. Intelligent condition monitoring method for bearing faults from highly compressed measurements using sparse over-complete features. *Mechanical Systems and Signal Processing*, 2018, 99, 459-477.
- [18] X Wang, J Wang and M Privault. Artificial intelligent fault diagnosis system of complex electronic equipment. *Journal of Intelligent & Fuzzy Systems*, 2018, 35(4), 4141-4151.
- [19] B P Cai, Y Zhao, H Liu and M Xie. A data-driven fault diagnosis methodology in three-phase inverters for PMSM drive systems. *IEEE Transactions on Power Electronics*, 2016, 32(7), 5590-5600.

# Figures

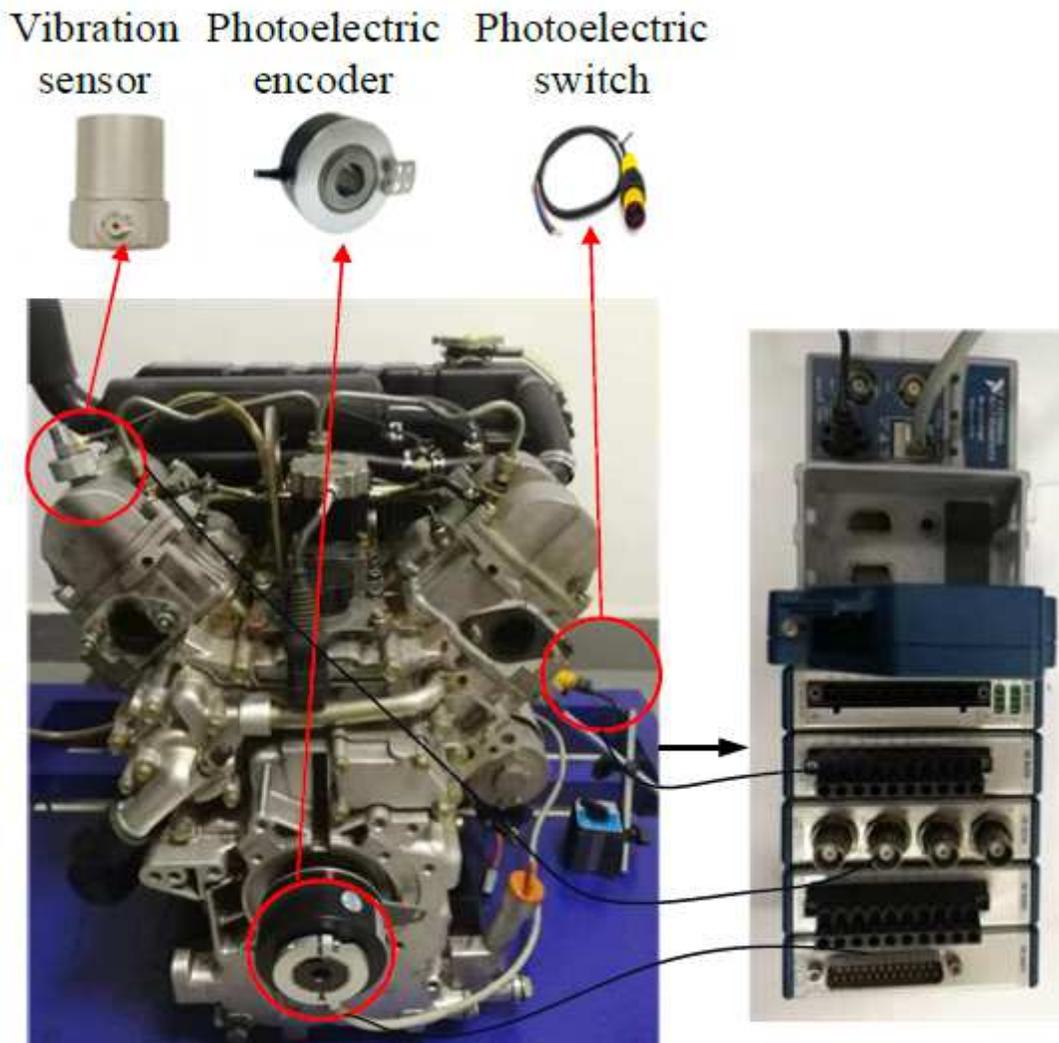


Figure 1

Diesel engine fault diagnosis platform

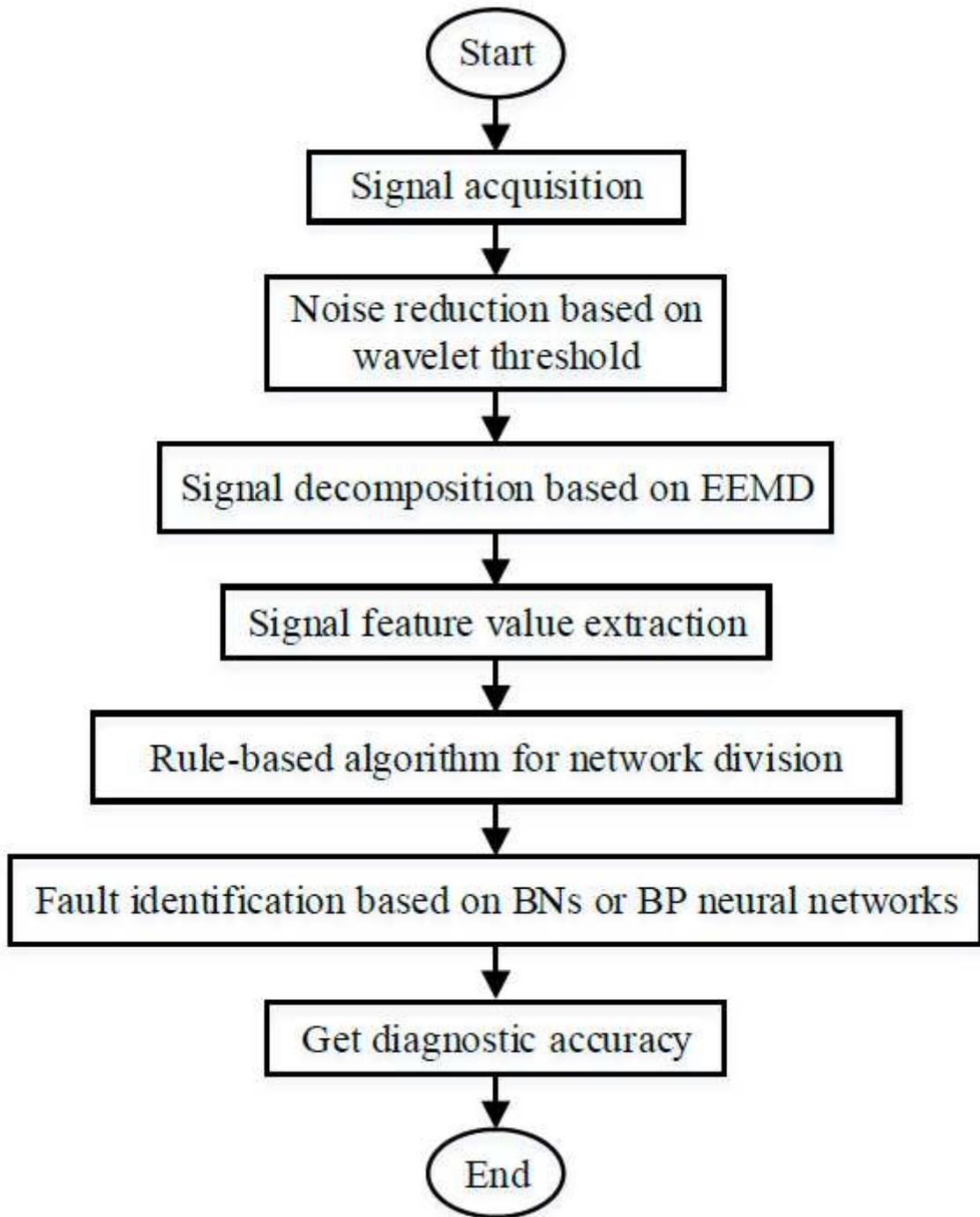
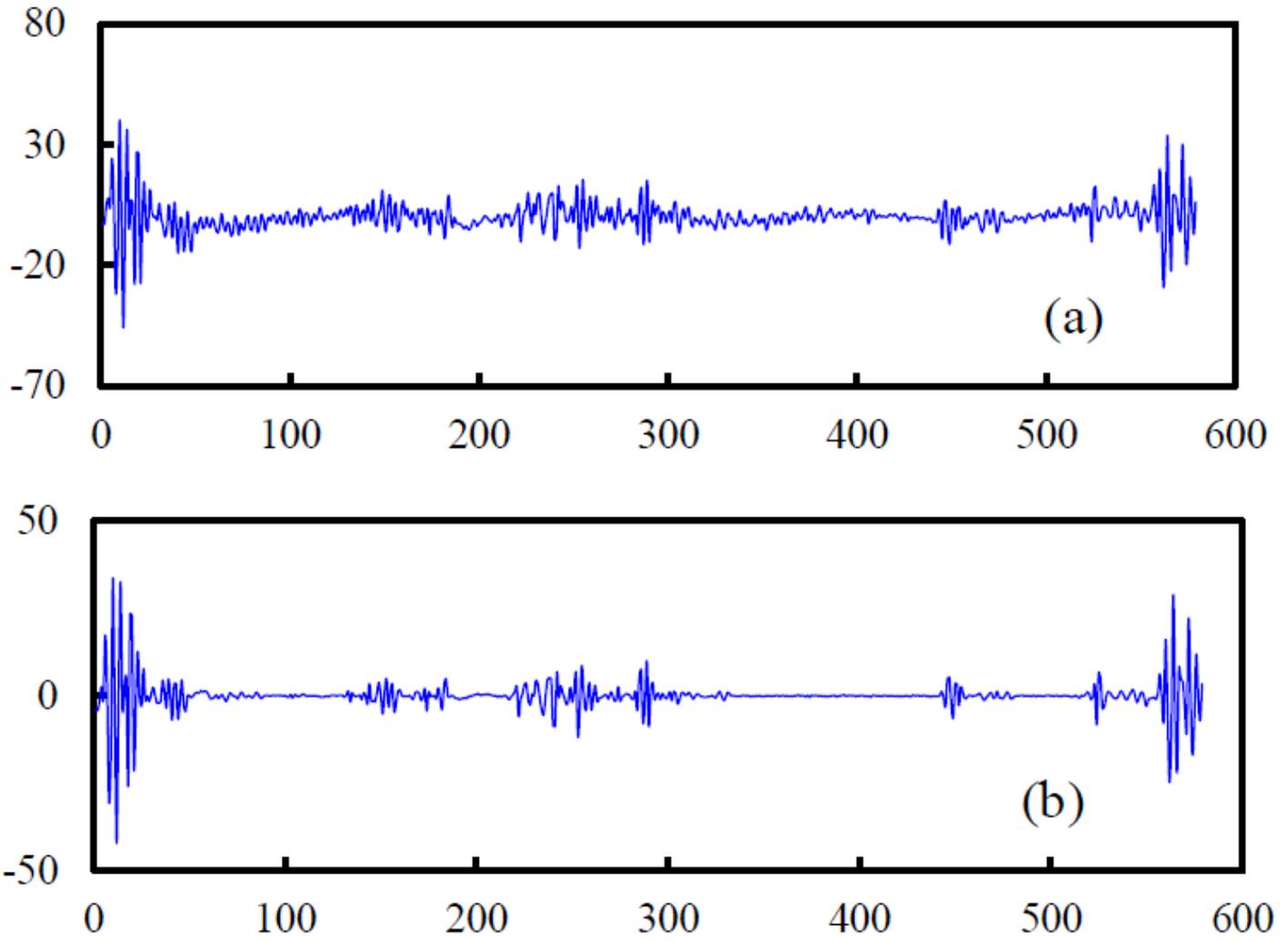


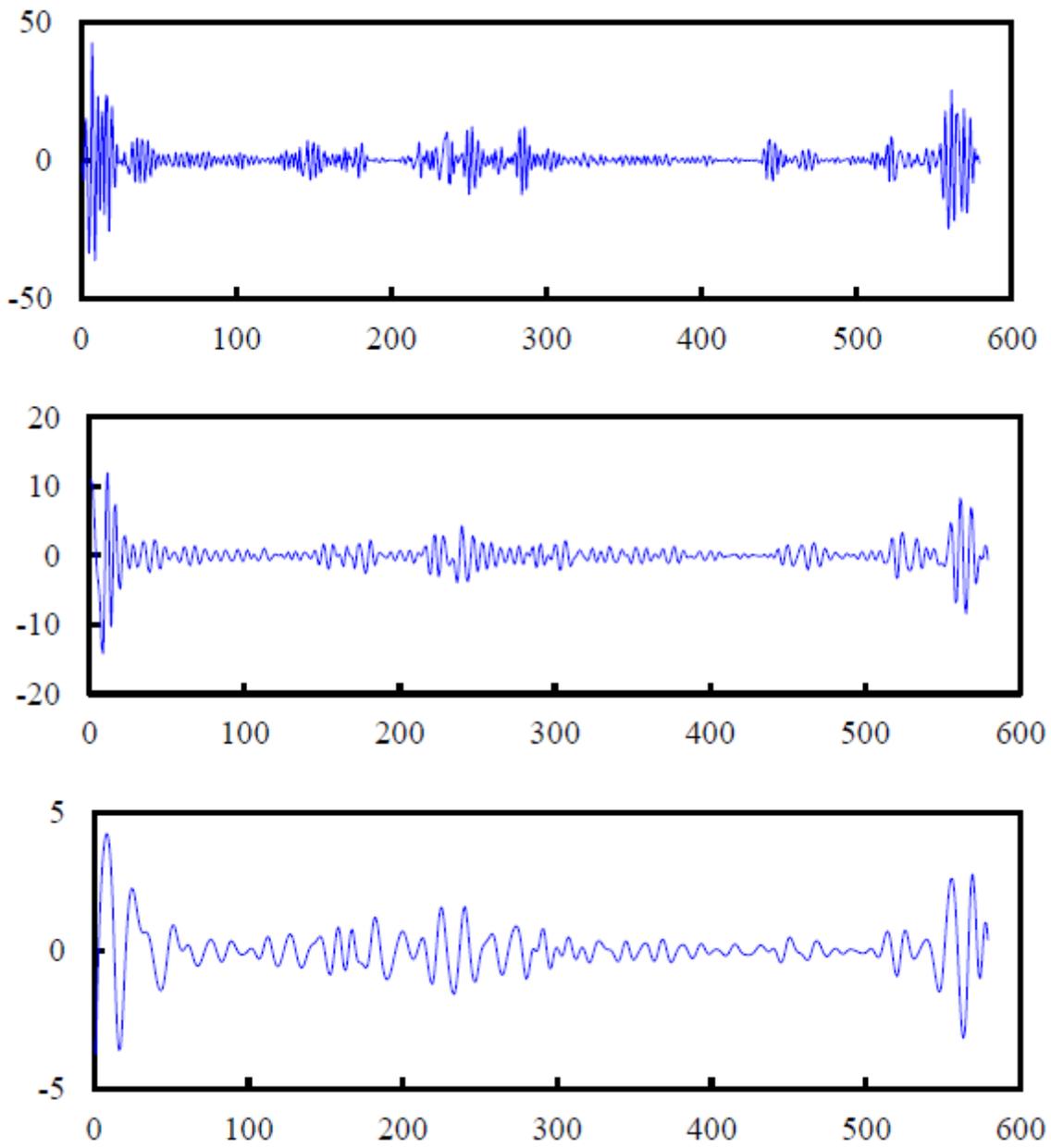
Figure 2

Overall fault diagnosis flowchart



**Figure 3**

Comparison of time domain signals before and after signal de-noising (a) Original signal, (b) Noise reduction signal.



**Figure 4**

The first three layers of IMFs after decomposition by EEMD

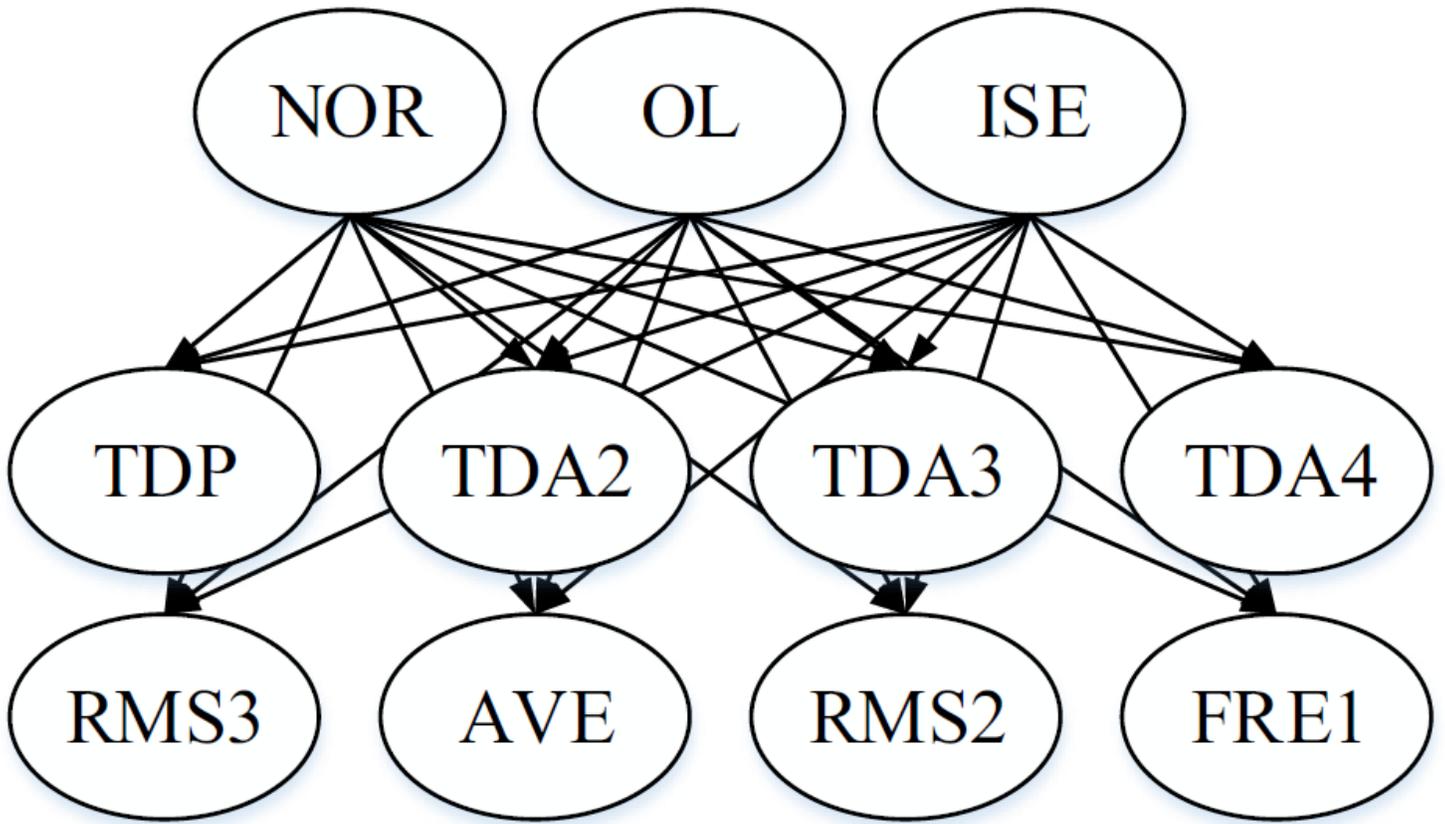


Figure 5

First model of BNs at 2500 rpm

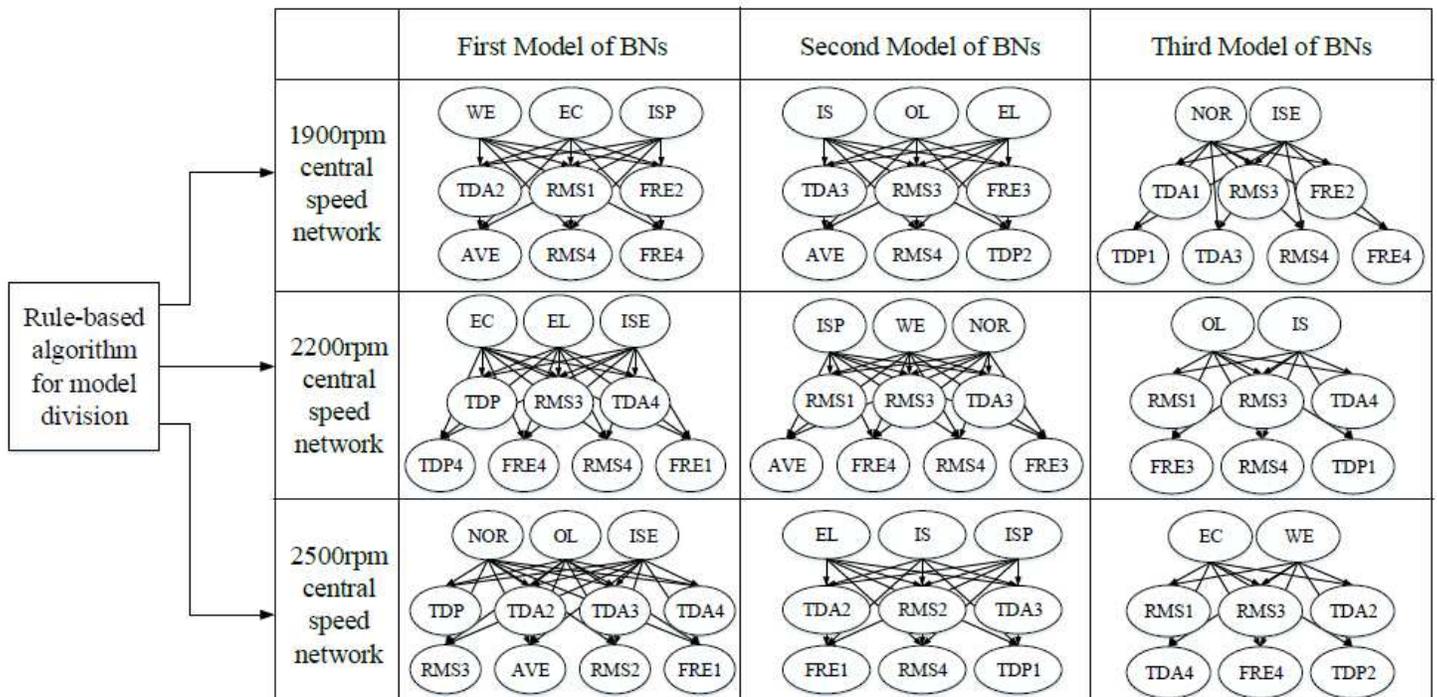


Figure 6

Complete BNs-based fault diagnosis model of diesel engine

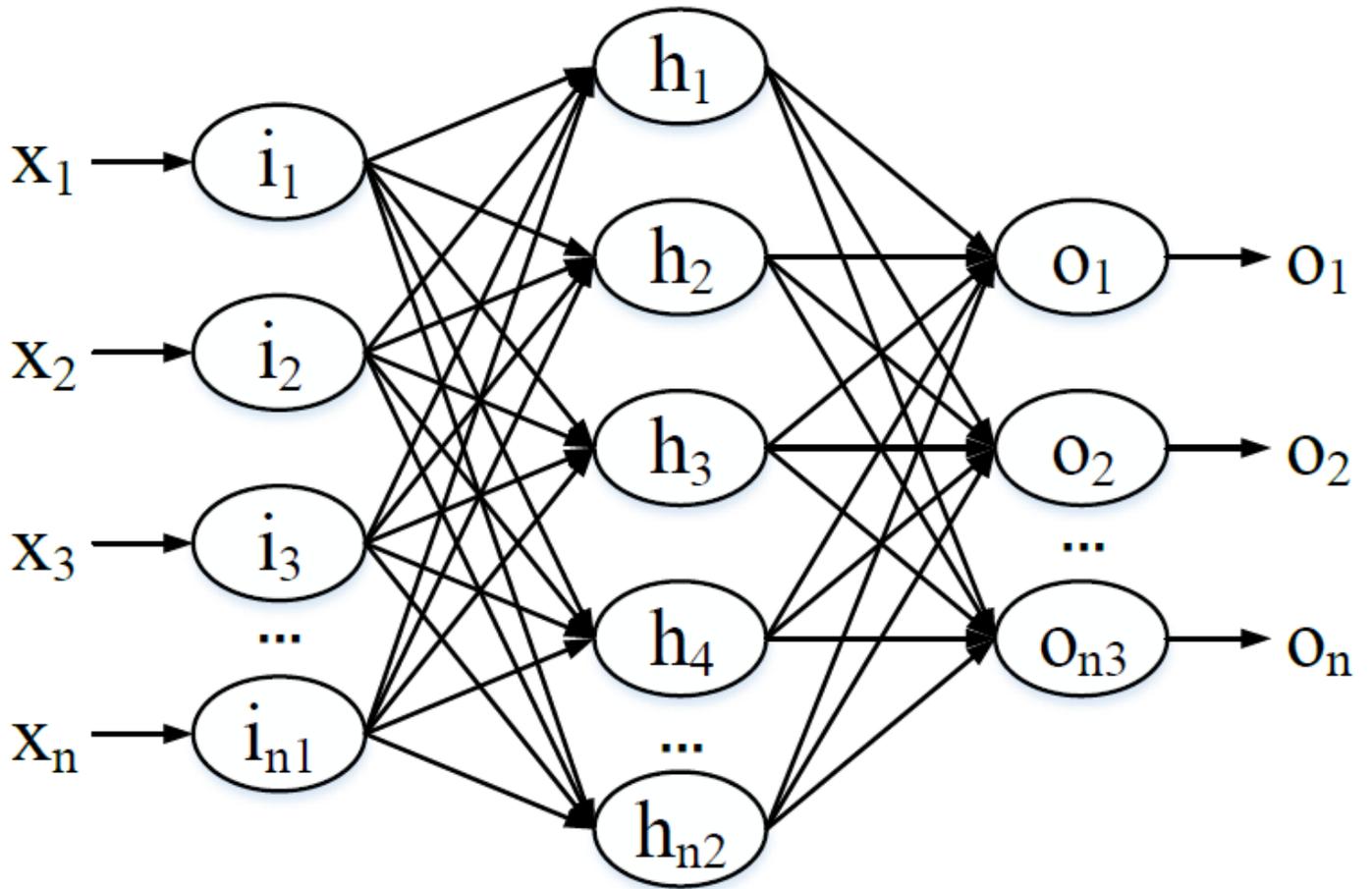


Figure 7

The structure of BP neural networks

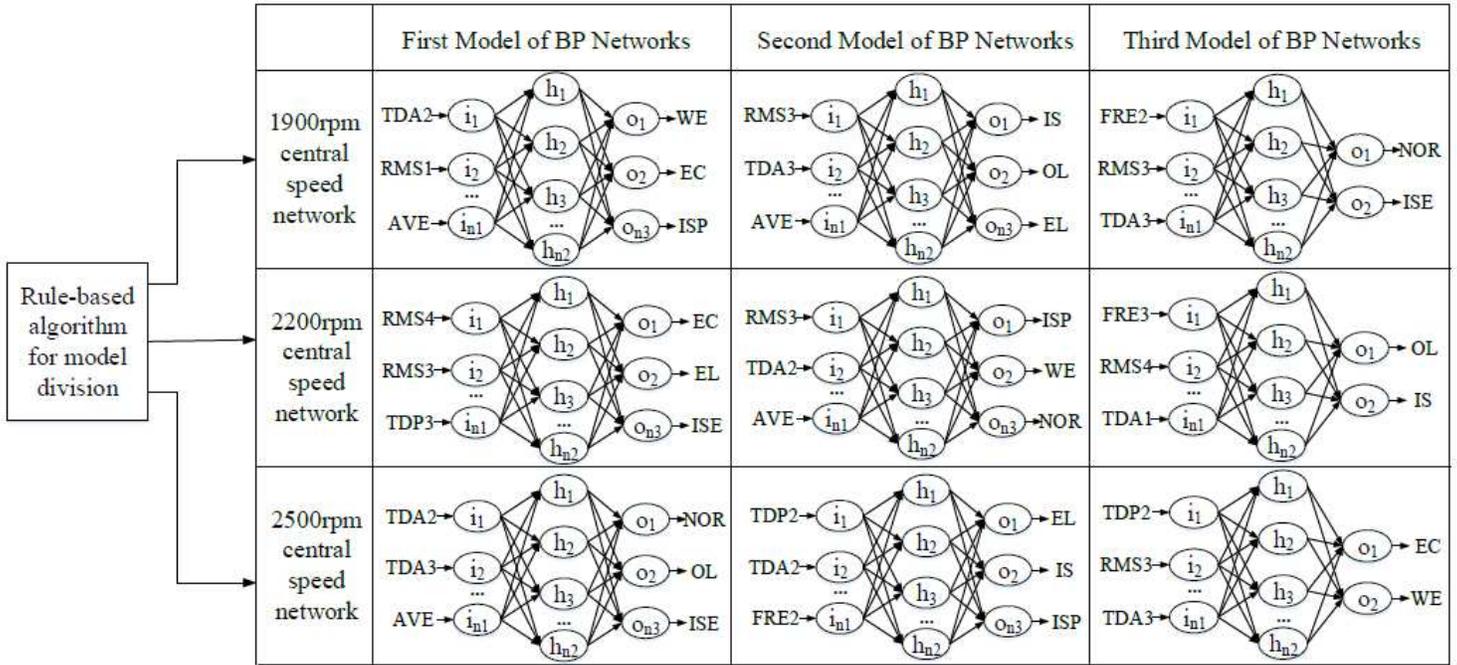


Figure 8

Complete BP-neural-networks-based fault diagnosis model of diesel engine

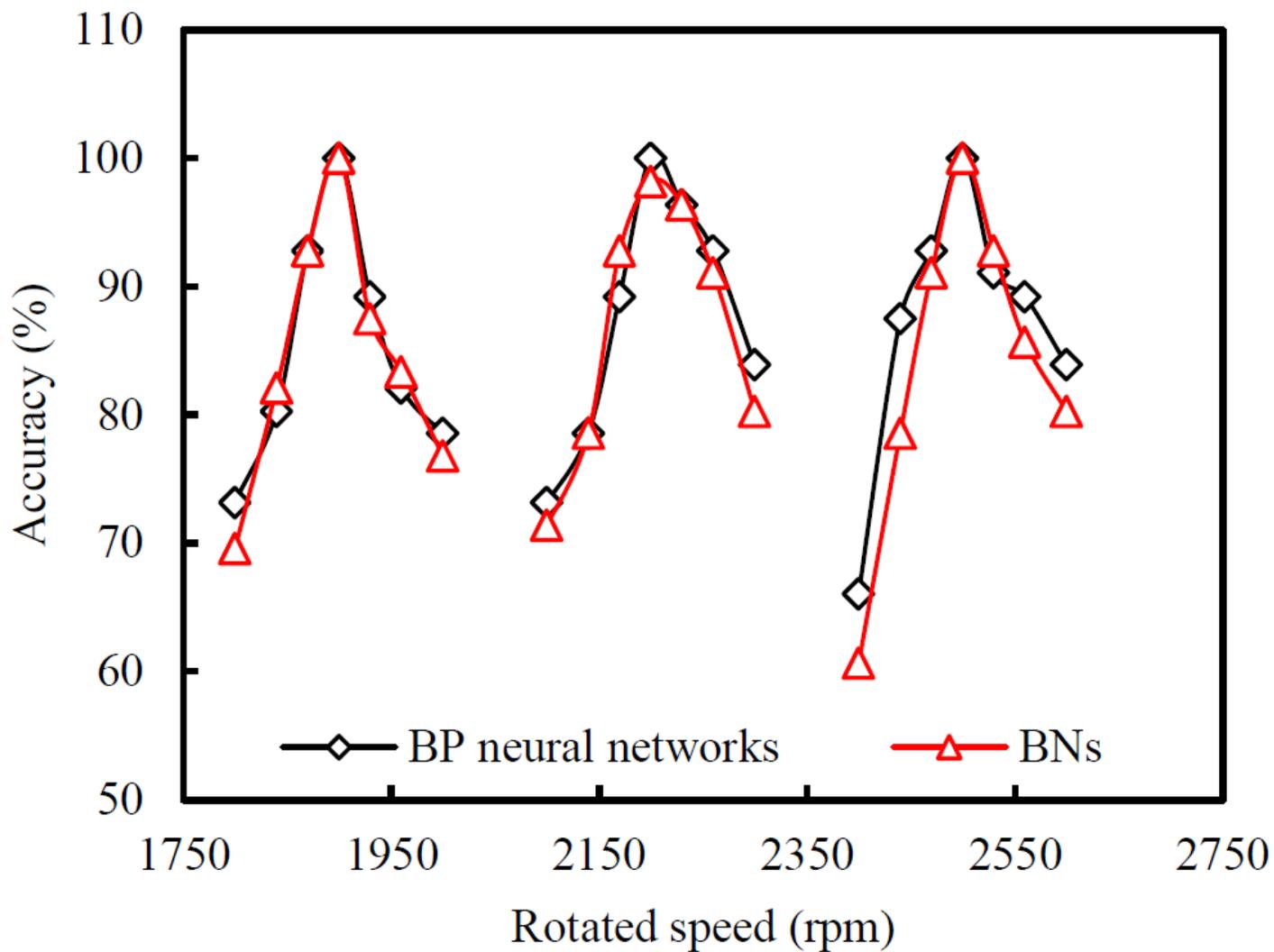


Figure 9

Diagnostic results using BP neural networks and BNs

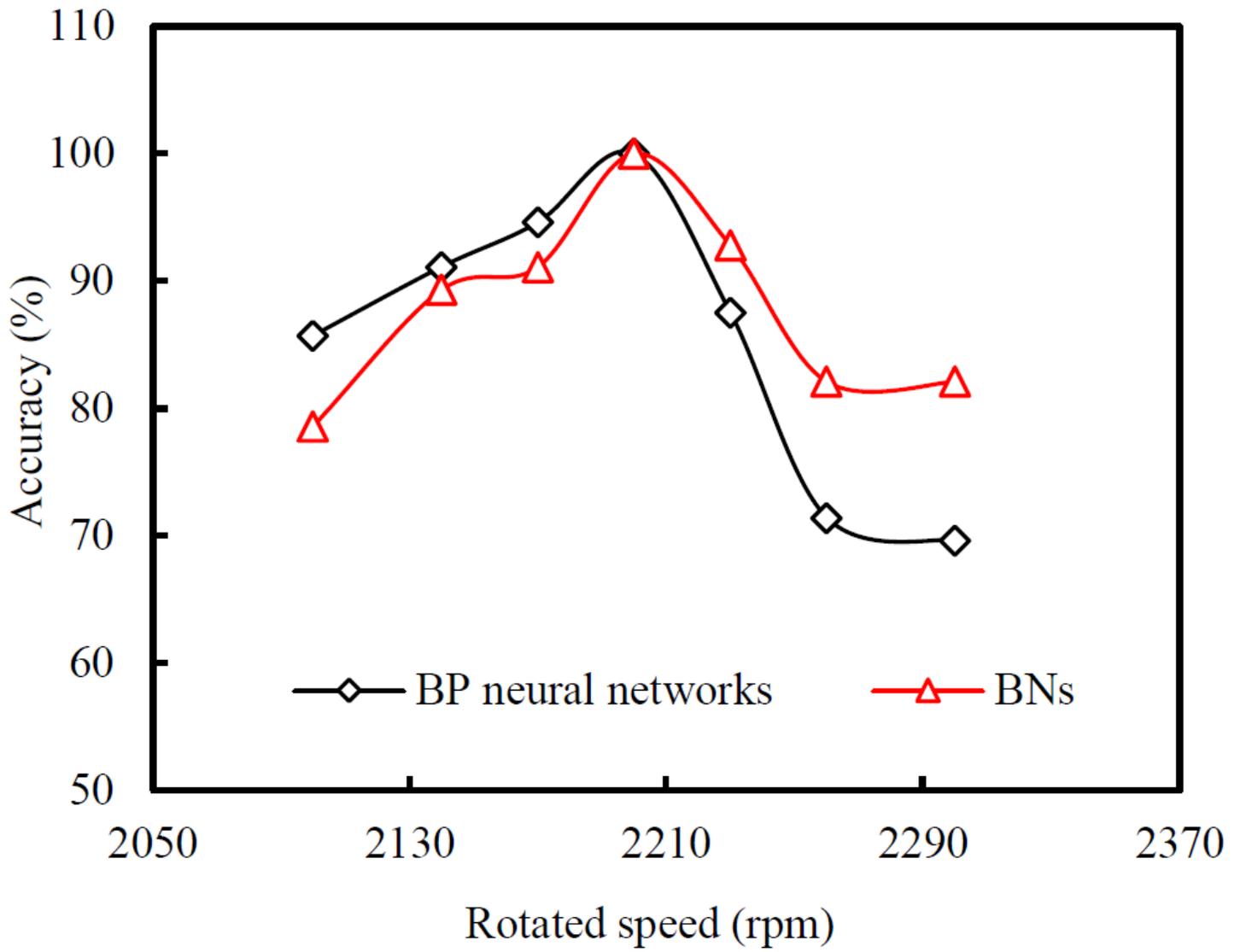
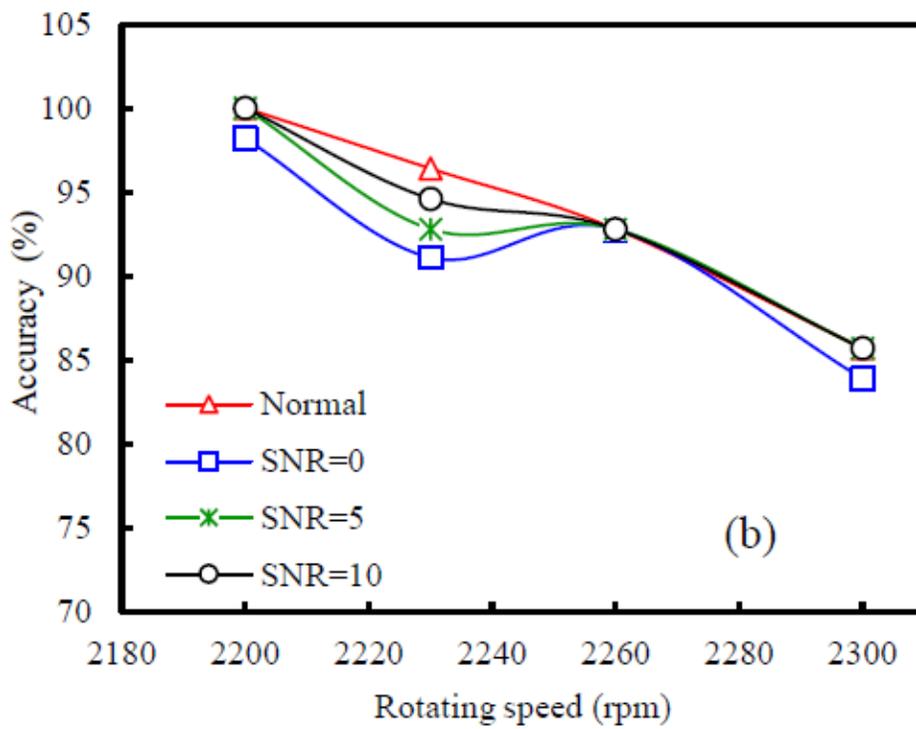
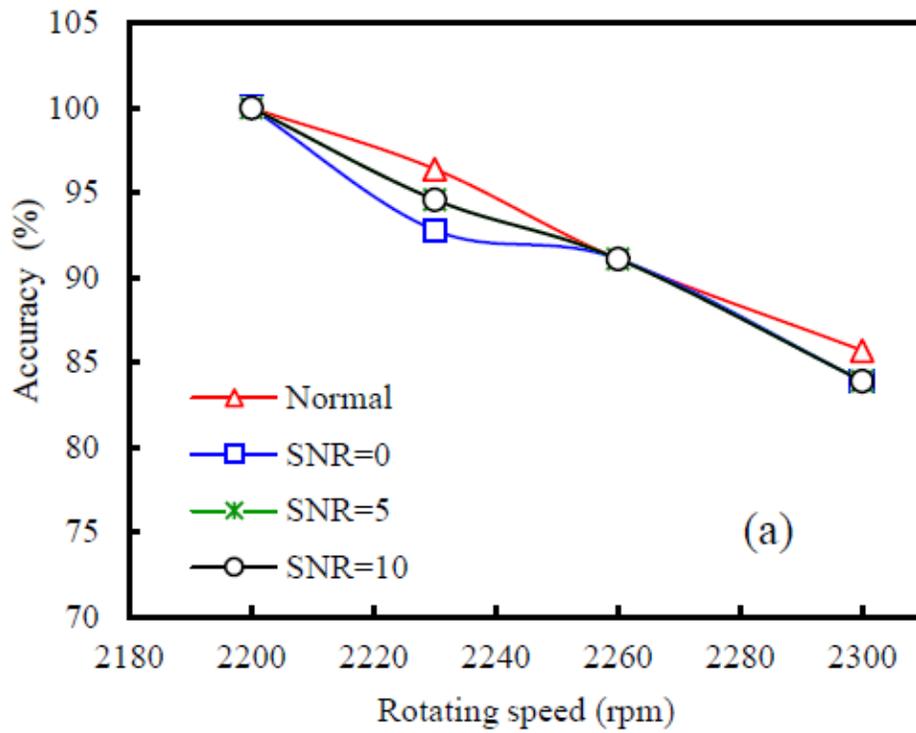


Figure 10

Synthesized signal diagnostic results using BP neural networks and BNs



**Figure 11**

Fault diagnostic results when white noise is added before signal de-noising (a) BNs, (b) BP neural networks

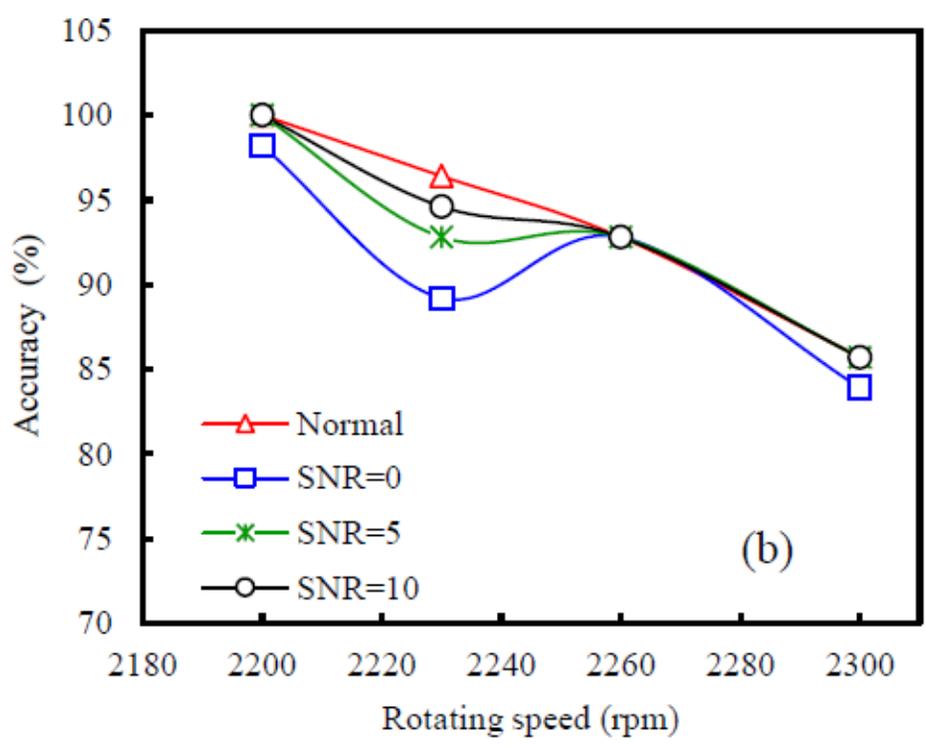
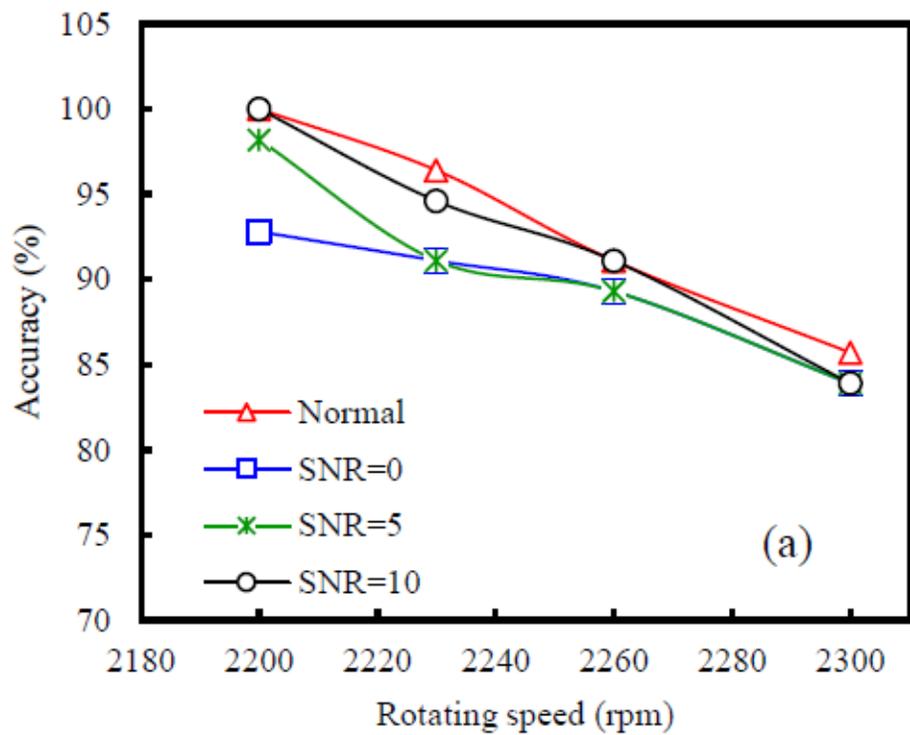
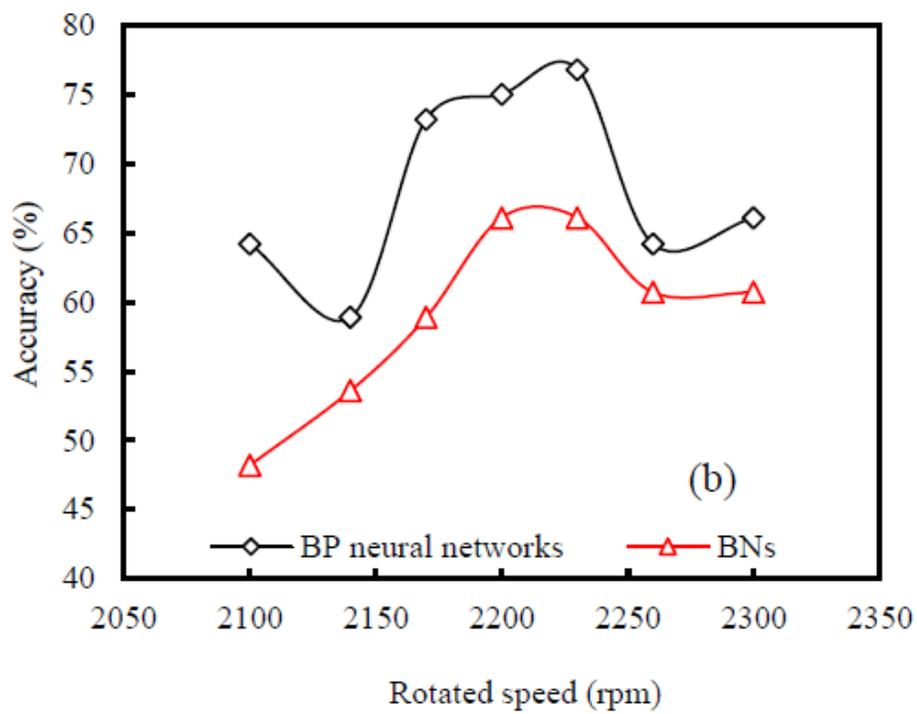
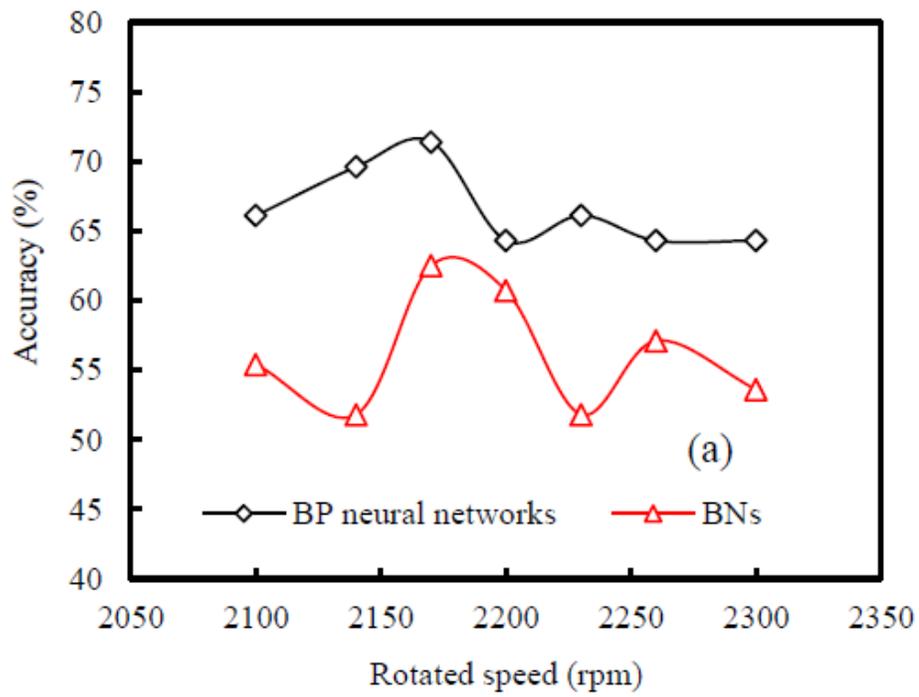


Figure 12

Fault diagnostic results when white noise is added after signal de-noising (a) BNs, (b) BP neural networks



**Figure 13**

Fault diagnostic results under different excitations (a) Continuous excitation (b) Intermittent excitation