Fault Prognosis of Hydraulic Pump Based on Bispectrum Entropy and Deep Belief Network

Hongru Li[1*], Zaike Tian[1,2], He Yu[1], Baohua Xu[1]

[1]*Army Engineering University, Heping West Road, No.97, 050003, Shijiazhuang, China, lihr168@sohu.com*
[2]*Joint Logistics College, National Defence University, Beijing, China*

Fault prognosis plays a key role in the framework of Condition-Based Maintenance (CBM). Limited by the inherent disadvantages, most traditional intelligent algorithms perform not very well in fault prognosis of hydraulic pumps. In order to improve the prediction accuracy, a novel methodology for fault prognosis of hydraulic pump based on the bispectrum entropy and the deep belief network is proposed in this paper. Firstly, the bispectrum features of vibration signals are analyzed, and a bispectrum entropy method based on energy distribution is proposed to extract the effective feature for prognostics. Then, the Deep Belief Network (DBN) model based on the Restrict Boltzmann Machine (RBM) is proposed as the prognostics model. For the purpose of accurately predicting the trends and the random fluctuations during the performance degradation of the hydraulic pump, the Quantum Particle Swarm Optimization (QPSO) is introduced to search for the optimal value of initial parameters of the network. Finally, analysis of the hydraulic pump degradation experiment demonstrates that the proposed algorithm has a satisfactory prognostics performance and is feasible to meet the requirements of CBM.

Keywords: Fault prognosis, bispectrum entropy, DBN, QPSO, hydraulic pump.

1. INTRODUCTION

Condition-based Maintenance (CBM) has been attracting more and more people's attention [1]. Technologies of fault prognostics are the core of CBM, which can estimate the residual useful life of damaged components or subsystems, and establish a targeted maintenance plan to ensure safe and reliable operation of the system [2]. Since the hydraulic pump is one of the most important components of hydraulic system, the performance of the hydraulic pump has a significant impact on the reliability of the whole system [3]. Therefore, the urgent issue is to search for a method to realize the accurate fault prognosis of hydraulic pumps.

Deep learning is a new research direction in the field of machine learning, which is developed on the basis of Artificial Neural Networks (ANN). There are several kinds of typical deep learning models, including convolutional neural network (CNN), recurrent neural network (RNN), and deep belief network (DBN). CNN models have shown their success in various computer vision applications where input data are usually 2D data [30], while our input features are mainly 1D time series data whose type is not suitable for the CNN input. RNNs are the deepest of all neural networks and are widely applied in language processing (NLP). However, there are problems such as gradient vanishing and explosion in RNNs. DBN model proposed by Hinton [4] in 2006 shows great advantages in feature learning and

processing. DBN model is composed of multiple Restrict Boltzmann Machines (RBMs). RBM consists of two layers of neurons, a hidden layer and a visible layer, which are fully and symmetrically connected between layers, but not connected within layers. Using unsupervised learning, each RBM is trained in its weight matrix to encode a probability distribution that predicts the activity of the visible layer from the activity of the hidden layer. The advantage of DBN rests on the unsupervised layer-by-layer pre-training with the Contrastive Divergence (CD) algorithm, on which supervised learning and inference can be efficiently performed [5]. Compared with the traditional shallow learning model, DBN can effectively avoid the phenomenon of over-fitting and premature convergence in the process of training a neural network model [6]. With the development of machine learning, DBN has been successfully applied in machine vision recognition [7], biometric detection [8], fault prognosis, etc. [9]. Roy et al. [10] proposed a text recognition methodology based on DBN and recurrent neural network, which further improved the recognition accuracy. Mohamed et al. [11] proposed a voice recognition method based on DBN, and used TIMIT data to verify that this method had better recognition performance. Lee et al. [12] applied convolution DBN to the field of image recognition. By compressing the input layer data according to the probability index, the model possesses translation

invariant property, and supports bottom-up and top-down probabilistic inference. In recent years, the deep learning methods based on DBN and its variants have been widely applied in fault diagnosis and prognostics. Tamilselvan et al. [26] got an accurate health state classification result by DBN. Wang et al. [27] proposed a fault diagnosis method based on the Hilbert envelope spectrum and DBN. And Shao et al. [28] put forward a novel optimization DBN based on particle swarm for rolling bearing fault diagnosis. Comparing with the applications of DBN in fault diagnosis, the applications of DBN in fault prognostics are rarely reported. Zhao et al. [29] proposed a fusion fault prognostics approach based on DBN and the Relevance Vector Machine in which DBN is only responsible for extracting features. These successful applications provide ideas for us to apply an optimized DBN to the fault prognosis of hydraulic pump.

In the process of hydraulic pump operation, vibration signals show obvious nonlinear and non-Gaussian characteristics due to the influence of fluid compressibility, fluid-structure coupling between pump source and servo system, and inherent mechanical vibration. In recent years, higher order spectra (HOS) are powerful signal processing tools that have shown significant advantages over traditional spectral analyses because of their capabilities of nonlinearity identification, phase information retention and Gaussian noise elimination [24]. Bispectrum analysis is a simple high-order spectrum analysis method, which has the above characteristics of high-order spectrum. Theoretically, bispectrum can completely suppress the Gaussian noise in the nonlinear signal, and retain the non-Gaussian components in the signal, which is suitable for processing the pump vibration signals with periodicity and modulation. Bispectrum is the two Fourier transform on the three order cumulant. It preserves the phase information of the signal and can quantitatively describe the nonlinear phase coupling in the signal which is closely related to the fault information [14]. Bispectrum can effectively restrain noises, and obtain sensitive feature information of signals simultaneously [15]. Moreover, the concept of information entropy is introduced, and the bispectrum entropy model is established. On the basis of the bispectrum analysis, the bispectrum entropy of different frequency bands of vibration signals is extracted as the prediction features.

Based on the above analysis, a novel methodology for fault prognosis of hydraulic pump based on the bispectrum entropy and QPSO-DBN is proposed in this paper. The structure of this paper is as follows. In Section 2, the basic theory and mathematical formulas of the bispectrum analysis are introduced in detail, and the extraction process of the bispectrum entropy is presented as well. In Section 3, the QPSO-DBN model is proposed for the fault prognosis of hydraulic pump. In Section 4, the effectiveness of the proposed fault prognostics methodology is verified through the experiment of the hydraulic pump. Conclusions are given in Section 5.

2. THE PROPOSED BISPECTRUM ENTROPY MODEL

2.1. Fundamentals of bispectrum analysis

Bispectrum analysis is a powerful tool for dealing with non-Gaussian signals [16]. The bispectrum can be regarded as the decomposition of the skewness of the nonlinear signal in the frequency domain, which can describe the nonlinear and asymmetric characteristics of vibration signals [17]. Assuming that $x(n)$ is zero-mean and k-order stochastic process, the third-order cumulant is defined as:

$$c_{3x}(\tau_1, \tau_2) = E(x(n)x(n+\tau_1)x(n+\tau_2)) \tag{1}$$

The bispectrum is defined as a third-order cumulant of the two-dimensional discrete Fourier transform, and it is usually represented as $B_x(\omega_1, \omega_2)$.

$$B_x(\omega_1, \omega_2) = \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} c_{3x}(\tau_1, \tau_2) e^{-j(\tau_1\omega_1+\tau_2\omega_2)} \tag{2}$$

Where ω is the circular frequency. As the bispectrum is defined as an autocorrelation discrete Fourier transform, the third-order cumulant is absolutely summable, i.e.

$$\sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} |c_{3x}(\tau_1, \tau_2)| < \infty \tag{3}$$

The bispectrum is a double period function, which can effectively preserve the phase information of the non-Gaussian signal and can be used to describe the energy information in the nonlinear signal.

2.2. The bispectrum entropy

The information entropy is a measure of the inherent complexity of the nonlinear system, which can reveal the inherent energy change of the vibration signal during the degradation of the hydraulic pump [18]. Since the measured vibration signal of the hydraulic pump is a finite length sequence, the bispectrum estimation $\hat{B}_x(\omega_1, \omega_2)$ can be obtained in practical application. Therefore, the energy entropy calculated from the bispectrum amplitude in any dual frequency domain is defined as bispectrum entropy, and the calculation steps are as follows [19]:

1) The nonparametric bispectral estimation method is applied to calculate the bispectrum estimation $\hat{B}_x(\omega_1, \omega_2)$. Firstly, the signal $x(N)$ with length N is divided into K segments, each segment has M data points, and the data of each segment is averaged.

2) Calculate the DFT coefficients of each segment.

$$Y^{(i)} = \frac{1}{M} \sum_{t=0}^{M-1} x_i(t)\exp(-j2\pi t\lambda/M) \tag{4}$$

Where $\lambda = 0, 1, \cdots, M/2$ and $x_i(t)$, $t = 0, 1, \cdots, M-1$ is the ith signal data.

3) Calculate the triple correlation of each segment of data.

$$\hat{b}_i(\lambda_1,\lambda_2) = \frac{1}{\Delta_0^2}\sum_{k_1=-L_1}^{L_1}\sum_{k_2=-L_1}^{L_1} Y^{(i)}(\lambda_1+k_1)Y^{(i)}(\lambda_2+k_2)Y^{(i)}(-\lambda_1-k_1-\lambda_2-k_2) \qquad (5)$$

Where $\Delta_0 = f_s / N_0$ is the required interval between frequency samples. $\lambda_1 + \lambda_2 \le f_s / 2$, $0 \le \lambda_2 \le \lambda_1$ and N0 and L1 satisfy $M = (2L_1 + 1)N_0$

4) The bispectral estimate is the average of k-segment data as follows.

$$\hat{B}_x(\omega_1,\omega_2) = \frac{1}{K}\sum_{i=1}^{K}\hat{b}_i(\omega_1,\omega_2) \qquad (6)$$

Where $\omega_1 = (\frac{2\pi f_1}{N_0})\lambda_1, \omega_2 = (\frac{2\pi f_2}{N_0})\lambda_2$

5) Calculate the probability of bispectrum amplitudes for all points in the definition domain.

$$P_B(\omega_1,\omega_2) = \frac{\left|\hat{B}_x^{'}(\omega_1,\omega_2)\right|}{\sum_{\omega_1=0}^{\pi}\sum_{\omega_2=0}^{\pi}\left|\hat{B}_x^{'}(\omega_1,\omega_2)\right|} \qquad (7)$$

Where $\sum_{\omega_1=0}^{\pi}\sum_{\omega_2=0}^{\pi}\left|\hat{B}_x^{'}(\omega_1,\omega_2)\right|$ is the non-Gaussian intensity eigenvalue.

6) Calculate the bispectrum entropy:

$$H_B = -\sum_{\omega_1=0}^{\pi}\sum_{\omega_2=0}^{\pi}P_B(\omega_1,\omega_2)\ln P_B(\omega_1,\omega_2) \qquad (8)$$

Similar to other information entropies, the bispectrum entropy also has non-negative, symmetry, and extreme values. When the bispectrum amplitude is more uniform, the greater is the bispectrum entropy, the smaller is the entropy.

3. Fault prognostics based on QPSO-DBN

Unlike traditional shallow learning methods, deep learning is an unsupervised learning method. Deep learning applies the bottom-up layer by layer learning to train the model, and then utilizes the top-down feedback learning to adjust the parameters of the entire model [20]. As a new intelligent prognosis algorithm, deep learning can simulate the complex structure of the human brain and achieve more accurate prognosis with a strong self-learning ability.

3.1. Fundamentals of DBN

DBN proposed by Hinton et al. is the most widely used in deep learning. DBN is composed of input layer, middle layer, and output layer, where the middle layer is composed of multiple RBMs. The network structure of DBN is shown in Fig.1. Fig.1. is a typical DBN model consisting of three layers of RBMs. RBM is a typical energy model consisting of visible and hidden layers. The first input layer v_i can be regarded as the visual layer of RBM1, and constitutes the first layer RBM together with the hidden layer h_i. RBM1

extracts the corresponding information from the visual layer and passes it to the hidden layer inside RBM1 itself. At this time, the hidden layer of RBM1 can be seen as the visual layer of the next RBM. Then, the data of the visual layer inside RBM2 can be further extracted and passed to the hidden layer of RBM2. With this operation, the output data of the DBN model can be generated from the last hidden layer of RBM. Thus, a DBN model consisting of multi-layer RBMs is implemented.
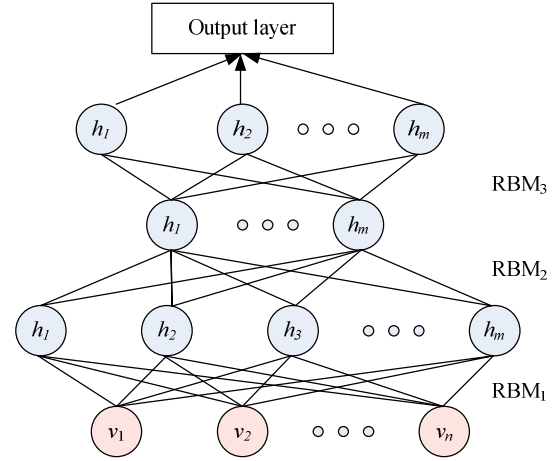

Fig.1. The structure of DBN model.

RBM is the core of the DBN model. It is a kind of energy model that obtains the dependencies between input parameters by associating the input parameters with a suitable energy function. For the energy models, the magnitude of energy is inversely proportional to the probability of combinations of parameters. This means that if a combination of parameters is considered to have a greater probability of rationality, it should also have smaller energy. Therefore, for the given set of parameter data, the configuration combinations of the parameters that minimize the corresponding energy values are obtained by training the model continuously.

Assuming that *m* and *n* are respectively the unit numbers of the hidden layer *h* and the visual layer *v*, where h_i is the *ith* unit vector of the hidden layer while v_j is the *jth* cell vector of the visual layer. The probability distribution of the hidden layer element h_i can be defined as:

$$P(h_i = 1 \mid v) = \text{sigmoid}(\sum_{j=0}^{n}W_{ij}v_j + b_i) \qquad (9)$$

Where $\text{sigmoid} = \frac{1}{1+e^{-x}}$, W_{ij} is the weight matrix between the hidden layer element and the visual layer element, b_i is the offset of hidden layer elements. In equation (6), the activation probability of the hidden layer elements is distributed according to the S type by the sigmoid function, and W_{ij} is constantly updated based on the input data. The probability distribution of the visual layer element is determined by the hidden layer data, and its probability distribution is defined as:

$$P(v_j = 1 \mid h) = \text{sigmoid}(\sum_{i=0}^{m} W_{ij} h_i + c_j) \qquad (10)$$

Where, c_j is the offset of the visual layer element. According to the above probability distribution, the joint probability distribution between the hidden layer and the visual layer can be defined as:

$$p(v, h) = \exp(-E(v, h)) / Z \qquad (11)$$

Where, $Z = \sum_{v,h} \exp(-E(v, h))$ is the normalized function, and the RBM energy function can be defined as:

$$E(v, h) = -\sum_{j=1}^{n} \sum_{i=1}^{m} W_{ij} v_j h_i - \sum_{j=1}^{n} c_j v_j - \sum_{i=1}^{m} b_i h_i \qquad (12)$$

In equation (9), $E(v,h)$ is the system energy of the RBM. According to the energy model theory, when the model energy is at minimum, the RBM is the most stable. Therefore, the optimal parameters of the RBM model can be obtained by solving the minimum value of $E(v,h)$.

When training RBM, the DBN model parameters can be optimized by calculating the minimum gradient for the log likelihood function, and then the RBM weight update model is defined as:

$$\Delta w_{ij} = E_{data}(v_j h_i) - E_{model}(v_j h_i) \qquad (13)$$

Where, $E_{data}(v_j h_i)$ is energy value expectation of training samples, $E_{model}(v_j h_i)$ is expectation defined by the model. In the process of solving $\min_w(-\lg P(v, h))$, due to the existence of the normalization factor, its computational complexity is higher, and it is difficult to accurately solve the DBN model parameters. Hinton proposed an approximate solution called the Contrastive divergence (CD) algorithm. The CD algorithm calculates the model parameters by single-step or multi-step Gibbs (Gibbs Chain) sampling and completes the updating of the above two expectations to achieve the fast learning of RBM.
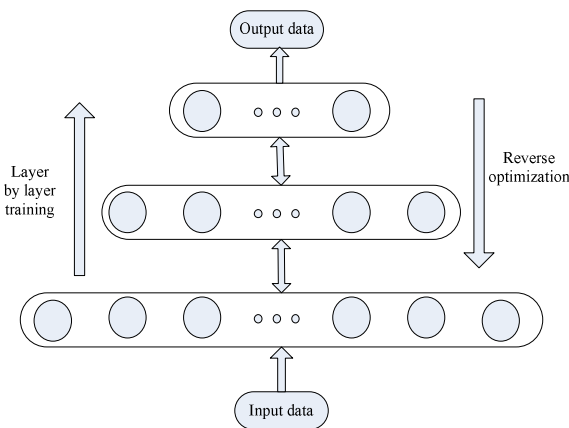

Fig.2. The training process of DBN model.

In summary, the DBN model is trained by greedy layer-wise algorithm, and the training process of the DBN model is shown in Fig.2. First of all, the first layer of RBM utilizes the CD algorithm to train the input feature data, and the model parameters of the first RBM layer are obtained. Then, the hidden layer of the first RBM layer is regarded as the visual layer of the next RBM, and this training process is continued until the top of the DBN model. Finally, the DBN model adopts the labeled input characteristic data to trim the model parameters reversely, and optimizes the model parameters by supervised training. When the model output error is less than a predetermined threshold, the model training is completed.

3.2. Parameter optimization of the prediction model based on QPSO

The DBN model includes parameters such as the number of hidden layers, the number of units per layer, and the learning rate. The choice of these parameters greatly influences the prognosis effect of the model, so the parameters of the DBN model need to be optimized.
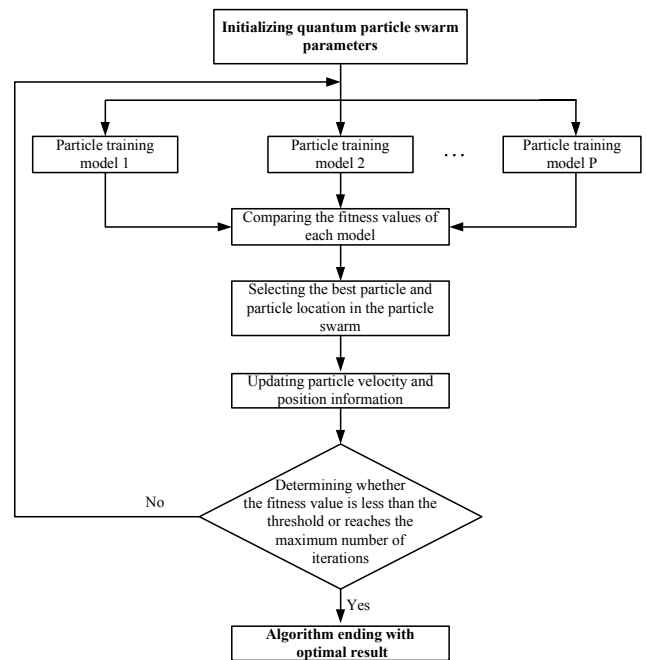

Fig.3. The flowchart of parameter optimization of the prediction model.

Quantum Particle Swarm Optimization (QPSO) is a new parameter optimization algorithm based on the quantum theory and particle swarm optimization [21]. It has been successfully applied in many fields such as artificial intelligence, image processing, automatic control, data mining, and so on. Procedures are detailed as follows [22]:

1) Set the number of particles P and the maximum number of iterations I.

2) Initialize the initial velocity of each particle $y^{k=0}$ and the initial position $s^{k=0}$.

3) Make the RMSE of the DBN model prognosis results as the fitness function, and calculate the fitness.

4) Select the global optimum particle according to the fitness, and update the velocity and the position information of each particle by the following formula.

$$s_i^{k+1} = \omega s_i^k + cr_1(y_{ibest}^k - y_i^k) + cr_2(y_{gbest}^k - y_i^k) \qquad (14)$$

$$y_i^{k+1} = y_i^k + s_i^{k+1} \qquad (15)$$

Where, ω is inertia weight, c is learning factor, and r_1, r_2 are the uniform random numbers falling into [0,1].

4. EXPERIMENTAL VALIDATION

4.1. Experimental platform

In order to validate the effectiveness and feasibility of the methodology that we proposed in this paper, the QPSO-DBN algorithm is applied in the fault prognosis of the hydraulic pump. The whole lifetime data of the hydraulic pump studied in this paper comes from the Hydraulic Pump Full Life Test Platform, which is shown in Fig.4.a). Fig.4.b) gives the schematic diagram of this platform where we can see that 603C01 Accelerometers are installed on the shell of the pump to collect vibration signals. The installation details are shown in Fig.4.c). The technical specification of the accelerometers is as follows. The sensitivity is 100 mV/g, the frequency range (±3 dB) is 0.5 Hz-10 KHz, and the range is ±50 g. The type of the tested hydraulic pump is L10VSO28DFR, which has 9 pistons. And its schematic diagram is shown in Fig.5.a). This type of pump is a constant pressure and constant flow pump of which max output pressure is 35 MPa and max oil flow is 90 L/min. On our platform, the pump is driven by a three-phase asynchronous motor whose power is 90 kW. A new hydraulic pump was taken for degradation experiment under the accelerated condition that the settled pressure was 27 MPa and the speed was 2780 rp·min⁻¹. The signals were sampled and stored by the cDAQ-9171 system of NI Corporation. The sampling frequency was10 KHz and the sampling time was 1 s. The interval time was 20 min.

The volumetric efficiency η is taken as the evaluation parameter of the hydraulic pump [23]. When the volume efficiency is less than 85 %, the hydraulic pump is judged to fail, and the test platform is automatically stopped. In this paper, a new hydraulic pump was taken as the experiment object to obtain the whole lifetime data of this pump. When the operating time was 575 h, η was less than 85 %. The hydraulic pump was judged to be totally invalid by the control system and the operating experiment platform was shut down automatically. After the experiment, the tested pump was disassembled, and it was clear that the failure mode was a severely loose slipper of one piston inside the pump, which is shown in Fig.5.b).

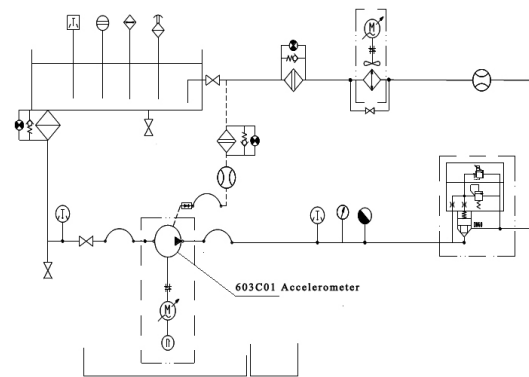Fig.4.a) Hydraulic Pump Full Life Test Platform.

Fig.4.b). The schematic diagram of the Hydraulic Pump Full Life Test Platform.
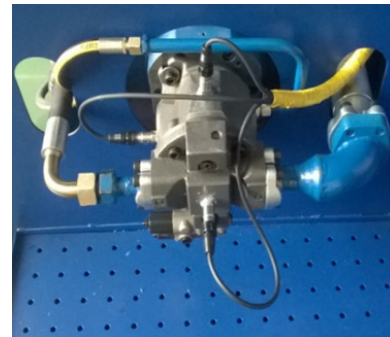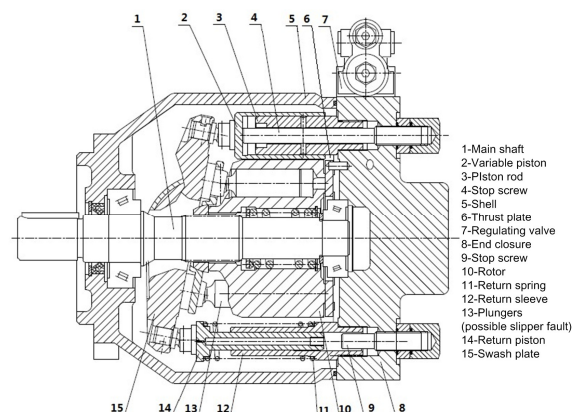

Fig.4.c) Installation of the accelerometer.

1-Main shaft
2-Variable piston
3-Plston rod
4-Stop screw
5-Shell
6-Thrust plate
7-Regulating valve
8-End closure
9-Stop screw
10-Rotor
11-Return spring
12-Return sleeve
13-Plungers
(possible slipper fault)
14-Return piston
15-Swash plate

Fig.5.a) The schematic diagram of the hydraulic pump with possible loose slipper fault.

Fig.5.b) Piston failure of loose slipper.

4.2. Results and analysis

In order to obtain the feature that can reflect the degradation state of the hydraulic pump accurately, bispectrum analysis of vibration signals is carried out in this paper, and the bispectrum entropy of different frequency bands is extracted as the prediction feature. Considering that the characteristic frequencies of the vibration signal are mainly within 3 KHz, the bispectral entropies in the frequency bands ([0,200), [200,400), … , [2800,3000)) are taken as the 15 prediction features in total. Among these 15 prediction features, corresponding to the divided frequency bands, we select four kinds of typical bispectrums along with entropy values labelled to display this process. These bispectrums in 0-200 Hz, 600-800 Hz, 1600-1800 Hz, 2400-2600 Hz are shown in Fig.6.
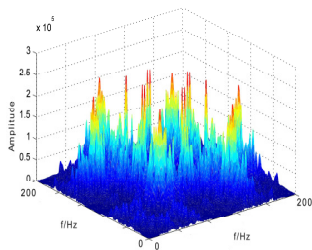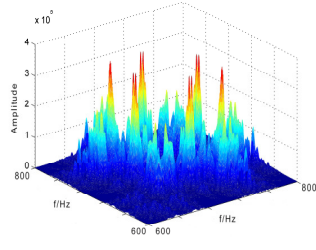

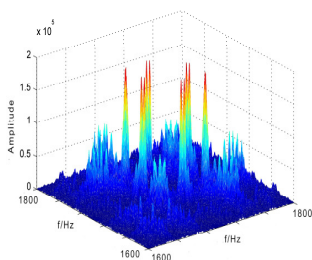Fig.6.a) H_B=0.85

Fig.6.b) H_B=0.68
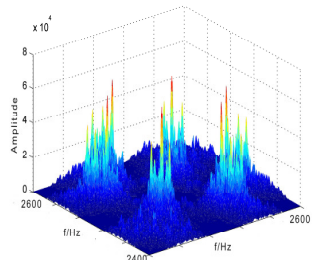

Fig.6.c) H_B=0.45

Fig.6.d) H_B=0.32

The DBN model constructed in this paper takes the 15 features as the model input layer vector, and the hydraulic pump volume efficiency is the output layer vector. First of all, in order to reduce the prognosis error of the DBN

prognosis model, the 15 features need to be normalized, and the normalized formula is shown in equation (13):

$$\tilde{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \qquad (16)$$

Before the fault prognosis of the hydraulic pump, the network structure parameters of the DBN model need to be determined. In this paper, the DBN model depth is set to 3 layers, the number of visual layer elements is 15, the learning rate for fine tuning is 0.01, and the number of cycles is 500 times. Since there is no explicit standard for setting the number of hidden layer elements and the learning rate of RBM, we optimize the above four model parameters with the QPSO optimization model constructed in Section 3.3. The RMSE of the prediction model is taken as the expected target, and the optimization result is shown in Fig.7.

The prediction error decreases gradually with the increase of the number of QPSO iterations. When the number of iterations grows to 32 times, the prediction error is minimized. The structure parameters of the DBN model can be determined as follows: the number of hidden elements is 12, the learning rate of the RBM$_1$ is 0.1, the learning rate of RBM$_2$ is 0.3, the learning rate of RBM$_3$ is 0.6.
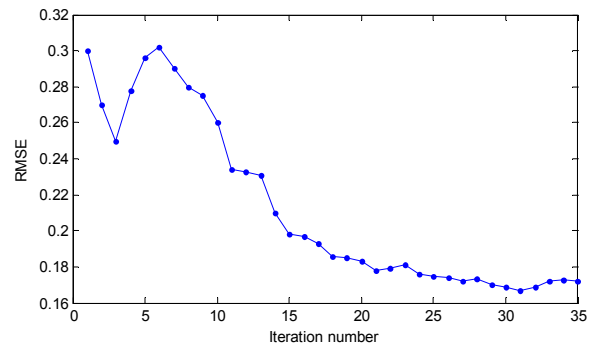

Fig.7. The prediction error of QPSO algorithm.

The volumetric efficiency η is taken as the evaluation parameter of the hydraulic pump and the threshold is settled as 85 %. Considering the failure mechanism and the fluctuation trend curve, the whole changing process of the volumetric efficiency has been divided into several statuses. During $\eta \geq 95\%$(0-145 h), the hydraulic pump is at normal status (F1). During $95\% > \eta \geq 93\%$ (145 h-312 h), the pump is at initial degradation stage (F2). During $93\% > \eta \geq 87\%$ (322 h-510 h), the pump is at slow degradation stage (F3). During $87\% > \eta \geq 85\%$ (510 h-575 h), the pump is at rapid degradation stage (F4). In the experiment, the stage F3 is selected as the training section of the DBN model, the stage F4 is selected as the predicting section. The threshold is settled according to the 1725th sample, which is considered as the occurrence time of failure. The prediction results are shown in Fig.8.
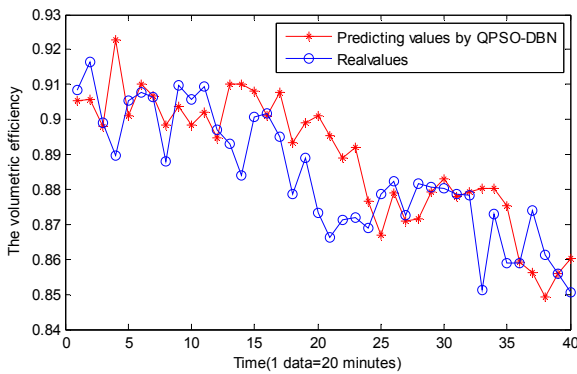
Fig.8. The predicting results by QPSO-DBN.

Fig.8. shows the results of QPSO-DBN prediction. Since the parameters of the DBN model are optimized by QPSO, the generalization ability and the predicting accuracy of the DBN model are improved. Consequently, the predicted series by the QPSO-DBN can better fit the real data and the errors are relatively small. The prediction algorithm reaches the threshold in the 38th sample and the failure is affirmed. Therefore, the error of the RUL prediction is two data points, which represents 40 min. To further verify the advantages of the proposed method, the BP neural network and the SVM model are utilized to predict the same data series with the same training data for comparison. The parameters of both models are optimized by particle swarm. For SVM, there are two critical parameters that can affect the performance of the model. These are the kernel parameter $\sigma = 4.6743$ and the punish coefficient $c = 31.7640$. For BP Neural Network, the number of input layer nodes is 15. The number of hidden layer nodes is 32. The number of output layers is 1. The number of training is set to 2000. The learning rate is set to 0.1. Results are shown in Fig.9. and Fig.10.
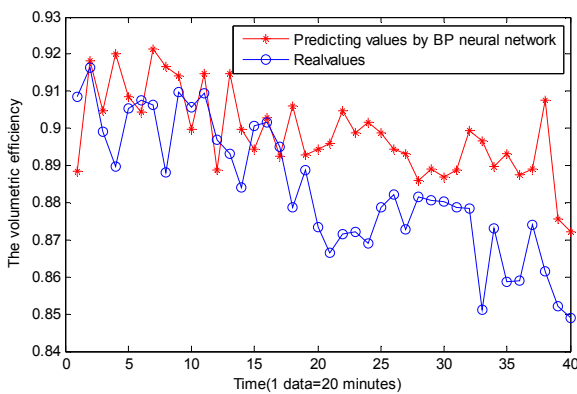

Fig.9. The predicted results by BP neural network.

Fig.9. and Fig.10. show the results predicted by the BP neural network and the SVM model, respectively. The shallow learning methods cannot accurately extract the deep correlation information of the input data, and have the problem of overfitting. Therefore, the errors between the

predicted data and the real data are obvious, and the predicted curves cannot accurately fit the degradation process of the hydraulic pump. Especially after the 30th sample, the prediction curves are almost a straight line, and the failure time of the hydraulic pump cannot be predicted accurately.
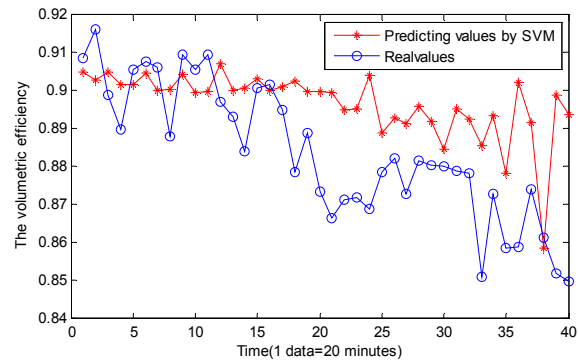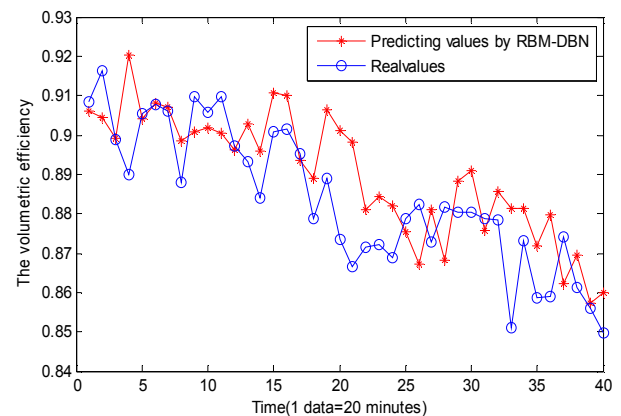

Fig.10. The predicted results by SVM.

Fig.11. The predicted results by RBM-DBN.

Fig.11. shows the prediction result by the classical RBM-DBN model which is widely applied in fault diagnosis and prognostics. The model also has three layers of RBMs and the learning rates of these layers are all set to 0.1. The other parameters of the model are given randomly. The difference between the RBM-DBN model and the QPSO-model is that the RBM-DBN model does not take parameter optimization. The prediction results demonstrate that the prediction curve is basically consistent with the actual value in the trend, which can reflect the actual degradation trend. However, the problem of insufficient gradient descent ability still exists after the 34th sample. The error of failure point is judged to be 4 sample points, that is, 80 min. Unreasonable model parameters affect the gradient descent ability of the model, and then affect the prediction performance of the model.

For further quantitative evaluation, the Root Mean Square Error (RMSE), the Mean Relative Error (MRE), and the Mean Absolute Error (MAE) are selected as the evaluation indexes. The results of the above four methodologies are shown in Table 1.

Table 1. The prediction errors of various algorithms.

Algorithms	RMSE	MAE	MRE	Error of RUL prediction
QPSO-DBN	0.032	0.062	0.107	40 min
BP neural network	0.136	0.104	0.173	120 min
SVM	0.186	0.1232	0.184	Failed
RBM DBN	0.032	0.062	0.107	80 min

Table 1. shows that the prediction accuracy of the deep learning method is obviously better than the shallow learning method. The QPSO-DBN model proposed in this paper has the highest prediction accuracy, and the final prediction error is only 40 min, which meets the needs of fault prognosis.

5. Conclusions

A novel methodology for the fault prognosis of the hydraulic pump based on the bispectrum entropy and QPSO-DBN is proposed, which is verified by the whole lifetime data sampled from the hydraulic pump degradation experiment. Conclusions can be drawn as follows:

1) Bispectrum analysis is applied to processing the vibration signal of the hydraulic pump for extracting the degenerate state information. At the same time, the bispectrum entropy of different frequency bands is extracted as the prediction features.

2) The QPSO-DBN is proposed for fault prognosis. The QPSO algorithm has a good effect on the global optimization of the structural parameters within the range of solution space for the DBN model. The problem caused by the unreasonable determination of parameters of the DBN model is solved effectively. Therefore, the generalization ability and the prediction accuracy are effectively increased.

3) Experimental results show that the proposed method is feasible and the prediction accuracy is satisfactory, which is able to meet the requirements of CBM. It is also meaningful for the improvement of system reliability.

References

[1] Ahmad, R., Kamaruddin, S. (2012). An overview of time-based and condition-based maintenance in industrial application. *Computers & Industrial Engineering*, 63 (1), 135-149.

[2] Sun, J., Li, H., Wang, W., Ye, P. (2015). Morphological undecimated wavelet decomposition fusion algorithm and its application on fault feature extraction of hydraulic pump. *Transactions of Nanjing University of Aeronautics and Astronautics*, 32 (3), 268-278.

[3] Sun, J., Li, H., Xu, B. (2016). The morphological undecimated wavelet decomposition-discrete cosine transform composite spectrum fusion algorithm and its application on hydraulic pump. *Measurement*, 94, 794-805.

[4] Hinton, G.E., Salakhutdinov, R.R. (2006). Reducing the dimensionality of data with neural networks. *Science*, 313 (5786), 504-507.

[5] Wang, Y.G., Cao, F.L., Yuan, Y.B. (2011) A study on effectiveness of extreme learning machine. *Neurocomputing*, 74 (16), 2483-2490.

[6] Kuremoto, T., Kimura, S., Kobayashi, K., Obayashi, M. (2014). Time series forecasting using a deep belief network with restricted Boltzmann machines. *Neurocomputing*, 137, 47-56.

[7] Theodorakopoulos, I., Kastaniotis, D., Economou, G., Fotopoulos, S. (2014). Pose-based human action recognition via sparse representation in dissimilarity space. *Journal of Visual Communication and Image Representation*, 25 (1), 12-23.

[8] Lefebvre, G., Berlemont, S., Mamalet, F., Garcia, C. (2013). BLSTM-RNN based 3D gesture classification. In *Artificial Neural Networks and Machine Learning – ICANN 2013*. Springer, LNCS 8131, 381-388.

[9] Tran, V.T., Althobiani, F., Ball, A. (2014). An approach to fault diagnosis of reciprocation compressor valves using Teager-Kaiser energy operator and deep belief networks. *Expert Systems with Applications*, 41 (9), 4113-4122.

[10] Roy, P.P., Chherawala, Y., Cheriet, M. (2014). Deep belief network based rescoring approach for handwritten word recognition. In *14th International Conference on Frontiers in Handwriting Recognition*. IEEE, 506-511.

[11] Mohamed, A.R., Dahl, G.E., Hinton, G. (2012). Acoustic modeling using deep belief networks. *IEEE Transactions on Audio, Speech, & Language Processing*, 20 (1), 14-22.

[12] Lee, H., Grosse, R., Ranganath, R., Ng, A.Y. (2011). Unsupervised learning of hierarchical representations with convolutional deep belief networks. *Communications of the ACM*, 54 (10), 95-103.

[13] Shen, G.J., Stephen, M., Xu, Y.C., White, P. (2014). Theoretical and experimental analysis of bispectrum of vibration signals for fault diagnosis of gears. *Mechanical Systems and Signal Processing*, 43 (1-2), 76-89.

[14] Gu, F., Shao, Y., Hu, N., Naid, A., Ball, A.D. (2011). Electrical motor current signal analysis using a modified bispectrum for fault diagnosis of downstream mechanical equipment. *Mechanical Systems and Signal Processing*, 25 (1), 360-372.

[15] Zhou, Y.B., Liu, Y.B., Li, H., Teng, W., Li, Z. (2013). Fault feature extraction for gear crack based on bispectral entropy. *China Mechanical Engineering*, 24 (2), 190-194.

[16] Hou, S., Li, Y., Xu, Y., Ma, C. (2016). Applications of dual-tree complex wavelet transform and bispectrum in roller bearing fault diagnosis. *Noise and Vibration Control*, 36 (5), 133-136.

[17] Jin, Z.Y., Lu, Y.G., Zhang, B., Yao, X.L. (2016). Applications of bispectrum analysis of intrinsic model function in the blade rack recognition. *Noise and Vibration Control*, 36 (5), 153-156.

[18] Lee, M.G., Kim, C., Pavlina, E.J., Barlat, F. (2011). Advances in sheet forming materials modeling, numerical simulation, and press technologies. *Journal of Manufacturing Science and Engineering*, 133 (6), 1001-1012.

[19] Huang, J.Y., Pan, H.X., Bi, S.H., Cui, B.Z. (2012). Fault pattern recognition based on bispectrum entropy model. *Acta Armamentarii*, 33 (6), 718-722.

[20] Qiao, J.F., Pan, G.Y., Han, H.G. (2015). Design and application of continuous deep belief network. *Acta Automatica Sinica*, 41 (12), 2138-2146.

[21] Chen, C., Liu, K., Kehtarnavaz, N. (2016). Real-time human action recognition based on depth motion maps. *Journal of Real Time Image Processing*, 12 (1), 155-163.

[22] Lu, X.H., Pan, D.F., Han, K., Li, J.S. (2016). Railway short-term wind speed prediction based on improved QPSO-WNN rolling algorithm. *Journal of Railway Science and Engineering*, 13 (5), 978-984.

[23] Sun, J., Li, H.R., Tian, Z.K. (2016). A degradation feature extraction method for hydraulic pumps based on fusion of sensitive components. *Chinese Journal of Scientific Instrument*, 37 (6), 1290-1298.

[24] Huang, B., Feng, G., Tang, X., Gu, J.X., Xu, G., Cattley, R., Gu, F., Ball, A.D. (2019). A performance evaluation of two bispectrum analysis methods applied to electrical current signals for monitoring induction motor-driven systems. *Energies*, 12 (8), 1438.

[25] Zhou, Y.B., Liu, Y.B., Li, H., Teng, W., Li, Z. (2013). Fault feature extraction for gear crack based on bispecral entropy. *China Mechanical Engineering*, 24 (2), 190-194.

[26] Tamilselvan, P., Wang, P. (2013). Failure diagnosis using deep belief learning based health state classification. *Reliability Engineering & System Safety*, 115, 124-135.

[27] Fu, Y., Zhang, Y., Qiao, H., Li, D., Zhou, H., Leopold, J. (2015). Analysis of feature extracting ability for cutting state monitoring using deep belief networks. *Procedia CIRP*, 31, 29-34.

[28] Shao, H., Jiang, H., Zhang, X., Niu, M. (2015). Rolling bearing fault diagnosis using an optimization deep belief network. *Measurement Science & Technology*, 26 (11), 1-17.

[29] Zhao, G.Q., Liu, Y.F., Zhang, B., Hu, C. (2017). Lithium-ion battery remaining useful life prediction with Deep Belief Network and Relevance Vector Machine. In *IEEE International Conference on Prognostics and Health Management (ICPHM)*. IEEE, 7-13.

[30] Zhao, R., Yan, R.Q., Chen, Z.H., Mao, K.Z., Wang, P., Gao, R.X. (2019). Deep learning and its applications to machine health monitoring. *Mechanical Systems and Signal Processing*, 115, 213-237.