# A Novel Deep Belief Network Model Constructed by Improved Conditional RBMs and its Application in RUL Prediction for Hydraulic Pumps

## He Yu

Army Engineering University, Shijiazhuang, China.

## Zaike Tian

Joint Logistics College, National Defence University, Beijing, China.

## Hongru Li and Baohua Xu

Army Engineering University, Shijiazhuang, China.

## **Guoqing An**

Hebei University of Science and Technology, Shijiazhuang, China.

#### (Received 21 October 2019; accepted 1 April 2020)

Residual Useful Life (RUL) prediction is a key step of Condition-Based Maintenance (CBM). Deep learning-based techniques have shown wonderful prospects on RUL prediction, although their performances depends on heavy structures and parameter tuning strategies of these deep-learning models. In this paper, we propose a novel Deep Belief Network (DBN) model constructed by improved conditional Restrict Boltzmann Machines (RBMs) and apply it in RUL prediction for hydraulic pumps. DBN is a deep probabilistic digraph neural network that consists of multiple layers of RBMs. Since RBM is an undirected graph model and there is no communication among the nodes of the same layer, the deep feature extraction capability of the original DBN model can hardly ensure the accuracy of modeling continuous data. To address this issue, the DBN model is improved by replacing RBM with the Improved Conditional RBM (ICRBM) that adds timing linkage factors and constraint variables among the nodes of the same layers on the basis of RBM. The proposed model is applied to RUL prediction of hydraulic pumps, and the results show that the prediction model proposed in this paper has higher prediction accuracy compared with traditional DBNs, BP networks, support vector machines and modified DBNs such as DEBN and GC-DBN.

## 1. INTRODUCTION

As the power source of the whole hydraulic system, hydraulic pump plays a pivotal role in the hydraulic system.<sup>1</sup> The performance of the hydraulic pump directly affects the safe operation of the whole system, so it is necessary to search for a method to accurately realize residual useful life prediction (RUL) prediction of the hydraulic pump.<sup>2</sup> By virtue of the remaining life prediction of the hydraulic pump, the system failure caused by hydraulic pump failure can be effectively avoided. At the same time, it can provide a favorable guarantee for the targeted performance inspection plan and the maintenance strategy.<sup>3</sup>

Deep learning is a new research direction in the field of machine learning, which is developed on the basis of Artificial Neural Networks (ANN). Deep belief network (DBN) is a classic deep learning model proposed by Hinton, the founder of deep learning theory in 2006.<sup>4</sup> 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 to encode in its weight matrix 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 has recently become a popular approach in machine learning for its promised advantages such as fast inference and the ability to encode richer and higher order network structures.<sup>6</sup> With the development of machine learning, DBN has been successfully utilized in machine vision recognition,<sup>7</sup> biometric detection,<sup>8</sup> data prediction<sup>9</sup> and so on. Roy et al.<sup>11</sup> proposed a text recognition method based on DBN and recurrent neural network, which further improved the recognition accuracy. Mohamed et al.<sup>12</sup> proposed a voice recognition method based on DBN, and used TIMIT data to verify that this method had better recognition effect. LEE et al.<sup>13</sup> applied convolution DBN to the field of image recognition. By compressing the input layer data according to the probability index, the model has translation invariant property, and supports bottom-up and top-down probabilistic inference.

DBN has achieved good application effect in the field of pattern recognition<sup>10</sup> and RUL prediction.<sup>23–25</sup> Zhao et al. proposed a fusion RUL prediction approach based on Deep Belief Network (DBN) and Relevance Vector Machine (RVM) where DBN is responsible for extracting degradation features of Lithium-ion batteries.<sup>23</sup> Zhang et al. proposed a remaining useful life estimation model based on multi-objective deep belief networks ensemble.<sup>24</sup> Sun et al. proposed a deep belief echo-state network (DBEN) model to address the issue of slow convergence and local optimum.<sup>25</sup> Compared with the traditional shallow learning method, DBN can train the model through bottom-up layer by layer learning, and then apply top-down feedback learning to adjust the parameters of the model, so as to realize fast and effective independent learning of data. For high-dimensional degradation characteristics of the hydraulic pump, the DBN model can learn and extract the deep correlation information in the input characteristics, retain key features of information at the same time effectively and reduce the interference of high-dimensional degradation characteristics contained in the component. The complicated implicit function approximation has a very good effect, thus the DBN model is very suitable for application to the hydraulic pump fault prediction field. However, using data discretization method to extract feature data leads to the limited ability of DBN gradient descent, which limits the application of DBN to the prediction of temporal data to some extent. To solve this problem, Taylor<sup>22</sup> put forward the Conditional Restrict Boltzmann Machine (CRBM) on the basis of RBM. CRBM can effectively utilize the temporal association information of the input data by adding the temporal join factor, so the DBN model can deal with the prediction of time series data stably.<sup>13</sup> Zhang et al.<sup>14</sup> applied CRBM to forecast stock data and get better prediction results. Chen et al.<sup>15</sup> proposed a continuous RBM to predict water quality parameters in the Huai River and obtained good results.

Although CRBM has increased the application of temporal association information, there is still a lack of constraints between the CRBM hidden layers, which makes it difficult for the DBN model to extract the depth association features in the input data and affects the prediction results.<sup>16</sup> To solve this problem, we added constraint variables between the hidden layer units on the basis of CRBM to adjust the activation probability between the hidden layers, and optimized the training process of CRBM. A novel RUL prediction model based on the ICRBM-DBN was established as a result. In order to verify the prediction accuracy of the ICRBM-DBN model for the time series data, performance of the ICRBM-DBN model is evaluated on the whole lifetime data of hydraulic pumps. The experimental results show that the proposed method is feasible and compared with the RBM-DBN, BP neural networks and Support Vector Machine (SVM), the predicting accuracy



Figure 1. The structure of DBN model.

is satisfactory, which is able to meet the requirements of CBM.

The structure of this article is organized as follows. In Section 2, the basic theory and mathematical formulas of DBN are introduced in detail. In Section 3, the RUL prediction model based on ICRBM-DBN is proposed to improve the accuracy of prognostics. In Section 4, the performance of the proposed approach is shown on the whole lifetime vibration data of hydraulic pump. In Section 5, conclusions are noted.

# 2. DEEP BELIEF NETWORK

DBN proposed by Hinton et al. is the most widely used in deep learning. DBN is composed of an input layer, a middle layer and an output layer, in which the middle layer is composed of multiple RBMs,17 the network structure of DBN is shown in Fig. 1. Figure 1 is a typical DBN model consisting of three layers of RBMs. RBM is a typical energy model consisting of visible and hidden layers.<sup>18</sup> The first input layer  $v_i$  can be regarded as the visual layer of RBM<sub>1</sub>, and the hidden layer  $h_i$  together constitute the first layer RBM. RBM<sub>1</sub> extracts the corresponding information from the visual layer and passes it to the hidden layer of RBM<sub>1</sub>. At this time, the hidden layer of RBM<sub>1</sub> can be seen as the visual layer of the next RBM. Then, the data of the RBM<sub>2</sub> visual layer can be further extracted and passed to the hidden layer of RBM2, and so on, the last hidden layer of RBM is the output data of the DBN model. Thus, a DBN model consisting of multi-layer RBMs is implemented.

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 a 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  $h_i$  and  $v_j$  are the *i*th unit vectors of the hidden layer and the *j*th cell vectors of the visual layer respectively. The probability distribution of the hidden layer element  $h_i$  can be defined as:

$$P(h_i = 1|\nu) = sigmoid\left(\sum_{j=0}^{vis} W_{ij}\nu_j + b_i\right); \quad (1)$$

where,  $sigmoid = \frac{1}{1+e^{-x}}$ ,  $W_i j$  is the weight matrix between the hidden layer element and the visual layer element,  $b_i$  is the offset of hidden layer elements. In Eq. (1), the activation probability of the hidden layer elements is modeled according to the S type by the sigmoid function, and  $W_i j$  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:

$$P(\nu_j = 1|h) = sigmoid\left(\sum_{i=0}^{hid} W_{ij}h_i + c_j\right); \qquad (2)$$

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:

$$p(\nu, h) = exp(-E(\nu, h))/Z; \qquad (3)$$

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

$$E(\nu,h) = -\sum_{j=1}^{\nu is} \sum_{i=1}^{hid} W_{ij}\nu_j h_i - \sum_{j=1}^{\nu is} c_j\nu_j - \sum_{i=1}^{\nu is} b_i h_i; \quad (4)$$

In Eq. (4),  $E(\nu, h)$  is the system energy of the RBM. According to the energy model theory, when the model energy is the minimum, the RBM is the most stable. Therefore, the optimal parameters of the RBM model can be obtained by solving the minimum value of  $E(\nu, h)$ .

In training RBM, the method of calculating the minimum gradient for the log likelihood function is usually used to solve the DBN model parameters, the RBM weight update model is defined as:

$$\Delta w_{ij} = E_{data}(\nu_j h_j) - E_{model}(\nu_j h_i); \tag{5}$$

where,  $E_{data}(\nu_j h_i)$  is the energy value expectation of training samples,  $E_{model}(\nu_j h_i)$  is the expectation defined by the model. In the process of solving  $min_w(-lgP(\nu,h))$ , due to the existence of normalization factor, its computational complexity is high, 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, completes the above two expectations updates, and achieves the fast learning of RBM.

In summary, the DBN model is trained by greedy algorithm layer by layer, and the training process of DBN model is shown



Figure 2. The training process of DBN model

in Fig. 2. First of all, RBM in the first layer uses CD algorithm to train the input feature data, and the model parameters of RBM in the first layer are obtained. Then, the hidden layer of RBM in the first layer is used as the visual layer of the next RBM, and continues to train until the top of the DBN model. Finally, the DBN model uses the labeled input characteristic data to reverse trimming model parameters, 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. THE THEORY OF ICRBM-DBN

#### **3.1. ICRBM**

Since RBM cannot effectively utilize the temporal association information in time-series data, the gradient descent ability of RBM is limited. This limits the application of the RBM-DBN model in the prediction of time-series data. On the basis of RBM, Montufar proposed a CRBM based on RBM.<sup>19</sup> CRBM is an extension of RBM, which inherits many good features of RBM, including simple reasoning process and effective training process, and the structure of CRBM is shown in Fig. 3. In CRBM, two kinds of connections are added to share temporal association information in time-series data. One kind of connections are the autoregressive connections between the visual layer of the previous n time RBM and the current RBM visual layer. The other kind of connection is the autoregressive connections between the visual layers of the previous n time RBM and the current RBM hidden layer.

By adding the above two kinds of connections, CRBM can make full use of the temporal association of the input data and improves the ability to deal with the time-series data. However, there is a lack of connection between each CRBM hidden layer units, which results in an unstructured weight model for the DBN model. This problem affects the prediction accuracy of the DBN model. To overcome this problem, a constraint variable between the CRBM hidden layers is added to adjust the activation probability between the hidden layers and optimize the training process of the CRBM. The new improved CRBM



Figure 3. The structure of CRBM.



Figure 4. The structure of ICRBM.

is also called ICRBM, and its structure is shown in Fig. 4.

In Fig. 4, each CRBM visual layer unit has and addition of constraint variable (represented by a star). By the constraining variables, the ability to connect and pass information between the units in the same layer is accomplished. In addition, the constraint variable of each unit is connected to the corresponding unit in the hidden layer. In the course of ICRBM training, each constraint variable can transfer information to other units of the same layer, change the weights of the variables of other units, and finally achieve the purpose of optimizing the output of the DBN model.

## 3.2. The Update Rules of ICRBM

ICRBM is composed of three layers based on a history layer, a current layer, and a hidden layer. As can be seen from Fig. 3, the visual unit of historical frame can be connected to the visual layer and the hidden layer of the target frame via the autoregressive model (AR). The input value of the previous nvisual units can be regarded as a supplementary input to the current visual unit, which directly affects the dynamic updating of the offsets in the current visual unit and the hidden unit. The update formulas are defined as:

$$b_{-now_i} = b_i + \sum_k \sum_q^{hid} \alpha_{ki}^{t-q} \nu_k^{t-q};$$
 (6)

$$c_{-now_j} = c_j + \sum_k \sum_q^{hid} \beta_{kj}^{t-q} \nu_k^{t-q};$$
 (7)

where,  $b\_now_i$  is the dynamic offset of the *ith* hidden layer unit,  $c\_now_j$  is the dynamic offset of the *jth* visual layer unit,  $\alpha_{t-q}^{ki}$  is the connection weight between the visual layer unit of t-q frame and the hidden layer unit of the target frame,  $\beta_{t-q}^{kj}$ is the connection weight between the visual layer unit of t-qframe and the visual layer unit of the target frame.

In the framework of ICRBM, an additional constraint variable is added to each visual layer units, and implicit output layer units have been connected with the visual layer unit cross variables, while the variables will cross layer unit to transmit information to the visual units of the same layer. When the output of the hidden layer element is higher than the set threshold, the constraint variable  $g_1$  of the unit is activated, and then the information is passed to the adjacent constraint variable  $g_2$ . At this point, when the output of the hidden layer unit corresponding to the variable  $g_2$  has exceeded the threshold, the constraint variable  $g_2$  is activated and the information continues until the last cell is passed. In order to simplify the computation, the direction of signal transmission is set to one-way transmission. The update rules for the hidden layer unit is defined as:

$$h_i = sigmoid(h'_i + a \times g_i); \tag{8}$$

where  $h_i$  is the updated output of hidden layer,  $g_i$  is a constraint variable, a is a tune parameter which requires manual settings to control the effect of horizontal variables on the output of the ICRBM. The larger a is the greater the impact is on the hidden layer, and vice versa. The original output  $h'_i$  of the hidden layer can be defined as:

$$h_i' = \sum_j W_{ij} \nu_j + b_n ow_i; \tag{9}$$

where  $W_{ij}$  is the weight matrix between the hidden layer unit and the visual layer unit,  $\nu_j$  is the state of the visual layer unit, and  $b_n ow_i$  is the dynamic offset of the hidden layer unit. In order to improve the training speed of ICRBM, this paper uses the activation probability of each unit as the output value directly. The constraint variable  $g_i$  is defined as:

$$g_i(t) = \begin{cases} 1, (h'_i > \theta \cup g_{i-1}(t-1) = 1) \\ \beta g_i(t-1), \text{ else}; \end{cases}$$
(10)

where  $\theta$  is the threshold of hidden layer output and  $\beta$  is the attenuation factor. Once the output of the hidden layer reaches the set threshold or the previous constraint variable is activated, the constraint variable of the unit is activated, and the signal is passed to the next constraint variable. Otherwise, the constraint variable is not activated and the weight is gradually attenuated.

### 3.3. DBN Constructed by ICRBMs

DBN prediction methods are generally divided into static multi-step prediction method and dynamic multi-step prediction method. The static multi-step prediction method, after each prediction, only updates the predicted data according to the prediction results, and does not update the input data. It is easy to cause the superposition of prediction errors and affect the prediction accuracy. The dynamic multi-step prediction method not only predicts the data according to the prediction results, but also updates the input vectors of the prediction model by using the predicted data. This ensures the dynamic adaptation of the model and improves the prediction accuracy. Therefore, this paper uses dynamic multi-step prediction method to construct DBN prediction model, and the procedures are detailed in the following.

- (1) Initialize connection weights W, learning speed  $\epsilon$  and constraint variable g;
- (2) According to the training rule of ICRBM, calculate the  $E_{data}(\nu_j h_i)$  and  $p(\nu, h)$  until the gradient of  $E_{data}(\nu_j h_i)$  is less than the set threshold, and stop calculating;
- (3) Take the hidden layer of ICRBM<sub>1</sub> as the visual layer of ICRBM<sub>2</sub> in the next layer and repeat the step (2) until the ICRBM training of the last layer is completed. The output of the last layer of ICRBM is the crude predictive value of the DBN model;
- (4) Update the ICRBM weights until the MSE is less than the set threshold, and the ICRBM-DBN is trained; and
- (5) Make a prediction by the trained ICRBM-DBN based upon the dynamic multi-steps strategy.

# 4. RUL PREDICTION BASED ON ICRBM-DBN

# 4.1. Prediction Process and Key Steps

There are mainly five steps in the process of RUL prediction for hydraulic pumps, including full life data collection, feature extraction and preprocessing, selection of training samples and testing samples, parameter optimization and ICRBM-DBN model training, ICRBM-DBN prediction and evaluation as follows.

- (1) <u>Full life data collection</u>. Several full life tests of hydraulic pumps are conducted on our full life test platform and several sets of experimental data such as vibration, oil flow, volumetric efficiency.
- (2) Feature extraction and preprocessing. The vibration data is analyzed by bispectrum and 15 bispectrum entropy features corresponding to 15 different frequency bands are extracted and normalized into [0, 1].
- (3) Sample Selection. One hundred sets of vibration data and volumetric efficiency in the slow degradation stage are taken as the training data and 40 sets of vibration data and volumetric efficiency in the rapid degradation stage are taken as the testing data.
- International Journal of Acoustics and Vibration, Vol. 25, No. 3, 2020

- (4) Parameter optimization and ICRBM-DBN model training. Three kinds of error indexes are applied to optimize the model parameters and full life data of hydraulic pumps are applied to train the ICRBM-DBN model. The stacked ICRBMs are learnt forward unsupervisedly and then the back fine tuning is conducted supervisedly.
- (5) ICRBM-DBN prediction and evaluation. Fifteen bispectrum entropy features are taken as the input data. The volumetric efficiency is taken as the output label data. Forty data samples are tested to predict the next 40 groups of volumetric efficiency. The critical moment when the volume efficiency is less than 85% is taken as the failure point of hydraulic pumps. The time difference between the predicted failure point and the real failure point is taken as the prediction error.

# 4.2. Full Life Data Collection

In order to verify the effectiveness of the method proposed in this paper, the ICRBM-DBN is applied in the RUL prediction of hydraulic pumps. The whole lifetime data of hydraulic pump used in this paper comes from the hydraulic pump full life test platform, which is shown in Fig. 5. The hydraulic pump tested is L10VSO28DFR, which has 9 pistons. A new hydraulic pump is taken for full life degradation experiment under the accelerated condition where the settled pressure is 27 MPa and the speed is 2780 rp·min-1. The signals are sampled and stored by the cDAQ-9171 system of NI Corporation. The sampling frequency is 10 KHz and the sampling time is 1 s. The interval time is 20 min. The volumetric efficiency  $\eta$ is taken as the evaluation parameter of the hydraulic pump.<sup>20</sup> When the volume efficiency is less than 85%, the hydraulic pump is judged to fall into the failure state, and the experimental platform is automatically stopped. In this paper, a new hydraulic pump is used to conduct full life test and obtain the whole lifetime data. When the operating time is 575 h,  $\eta$  is less than 85%. The hydraulic pump is judged to be totally invalid by the control system and the experimental operating is shut down automatically. After the experiment, the tested pump is disassembled, and it is clear in Fig. 6 that the failure mode is severe loose slipper.

# 4.3. Feature Extraction and Preprocessing

In order to obtain the feature that can reflect the degradation state of the hydraulic pump accurately, bispectrum analysis<sup>21</sup> 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 characteristics frequency of the vibration signal are mainly within 3 KHz, the bispectral entropy in the frequency bands  $([0, 200), [200, 400), \ldots, [2800, 3000])$  are taken as the 15 prediction features in total.<sup>26</sup> The DBN model constructed in this paper takes the 15 prediction features as the input layer vector of this model, and the hydraulic pump volume efficiency is the output layer vector. First of all, in order to reduce the prediction error of DBN prediction model, the 15 prediction features

H. Yu, et al.: A NOVEL DEEP BELIEF NETWORK MODEL CONSTRUCTED BY IMPROVED CONDITIONAL RBMS AND ITS APPLICATION IN...



Figure 5. Hydraulic Pump Full Life Test Platform.



Figure 6. Failure of loose slipper.

need to be normalized, and the normalized formula is shown in Eq. (11):

$$\tilde{x} = \frac{x - x_{min}}{x_{max} - x_{min}}.$$
(11)

# 4.4. Parameter Optimization and ICRBM-DBN Model Training

In order to better evaluate the results of the model, the root mean square error (RMSE), the mean relative error (MRE) and the mean absolute error (MAE) are used as the accuracy evaluation indexes. These three evaluation indexes can accurately and comprehensively analyze the prediction accuracy and stability of the ICRBM-DBN prediction model.

In this paper, the depth of ICRBM-DBN prediction model is



Figure 7. The prediction error of different numbers of hidden units.



Figure 8. The prediction error of different numbers of previous frames.

set to 3 layers, in which the visual layer unit is set to 15. At present, there are no explicit standards for the number of hidden layer units. We use RMSE, MRE and MAE as evaluation criteria to determine the number of hidden layer units through experiments. As can be seen from the Fig. 7, when the number of hidden layer units is 12, the model prediction error is the smallest. In Fig. 8, the effect of different previous frame number on prediction accuracy is analyzed. When the number of previous frames is 3, the model prediction error is the smallest. The experimental result shows that, when the previous frame number exceeds 3, the prediction error is not reduced but increased with the increase of previous frames. This is because the original vectors near the target frame have similar information, while the farther input vectors cannot effectively evaluate the current target vectors, resulting in error accumulation.

In the ICRBM-DBN model, the stability of the prediction model is often determined by the selection of learning rate. If the value is large, the prediction error of the model increases dramatically; if the value is small, the learning ability of the model is poor. In order to guarantee the stability of the prediction model, we calculate the prediction error of the model under different learning rates, as shown in Table 1. In the table, when the learning rate is less than 0.001, the prediction accuracy is very limited, while the training speed will decline a lot. Therefore, we will set the model learning rate as 0.001.

H. Yu, et al.: A NOVEL DEEP BELIEF NETWORK MODEL CONSTRUCTED BY IMPROVED CONDITIONAL RBMS AND ITS APPLICATION IN...

Learning rates	RMSE	MAE	MRE
0.1	0.0612	0.0536	0.1328
0.05	0.0524	0.0467	0.1047
0.01	0.0482	0.0421	0.0962
0.005	0.0421	0.0385	0.0812
0.001	0.0364	0.0316	0.0738
0.0005	0.0486	0.0412	0.1084
0.0001	0.0401	0.0306	0.0753

Table 1. The prediction errors of different learning rates.





#### 4.5. ICRBM-DBN Prediction and Evaluation

In this paper, the volumetric efficiency  $\eta$  is taken as the evaluation parameter of the hydraulic pump and the threshold is settled as 85%. Considering the failure mechanism and the fluctuation tends to curve, the whole changing process of the volumetric efficiency has been divided into several statuses. During  $\eta \ge 95\%(0-145\ h)$ , the part is the normal status (F1); During  $95\% > \eta \ge (145 \ h - 312 \ h)$ , the part is initial degradation stage (F2); During  $93\% > \eta \ge (322 h - 510 h)$ , the part is slow degradation stage (F3); During  $87\% > \eta \geq$ (510 h - 575 h), the part is rapid degradation stage (F4). In the experiment, the degradation of hydraulic pump has continuity and consistency in F3 and F4. However, there are more data samples in F3 and the model training needs enough data samples, so the training samples are selected from the stage of F3, and the testing samples are selected from the stage of F4. The threshold is settled according to the 1619th sample, which is considered as the occurrence time of failure. The prediction results are shown in Fig. 9 and Fig. 10.

Fig. 9 shows the results predicted by ICRBM-DBN. The predicting series by the ICRBM-DBN coincide with the real data and the errors are relatively small. The predicting algorithm reaches the threshold in the 39<sup>th</sup> sample and the affirmation of failure is verified. Therefore, the error of the RUL prediction is 1 data point, which is 20 min. Fig. 10 shows the results predicted by RBM-DBN. The predicted curves are basically consistent with the actual values, and can basically reflect the actual trend of degradation. And after the 38<sup>th</sup> sample, the algorithm reaches the threshold. The error of RUL is 3



Figure 10. The prediction results by RBM-DBN.



Figure 11. The prediction results by BP neural network.



Figure 12. The predicting results by SVM.

data points, which is 60 min., the prediction accuracy of RBM-DBN is lower compared with ICRBM-DBN. This is mainly due to the inability of RBM to utilize temporal association information when processing sequential data, resulting in limited gradient descent capability. ICRBM-DBN based on the CD model improves the gradient descent ability and gets higher prediction accuracy. For further indication of the advantages of the proposed method, based on the same training data and predicting data, the BP neural network and SVM method are taken for comparison. Results are shown in Fig. 11 and Fig. 12.

Fig. 11 and Fig. 12 show the results predicted by the BP neural network and SVM method. The shallow learning methods

H. Yu, et al.: A NOVEL DEEP BELIEF NETWORK MODEL CONSTRUCTED BY IMPROVED CONDITIONAL RBMS AND ITS APPLICATION IN...



Figure 13. The prediction result by DBEN.



Figure 14. The prediction result by GC-DBN.

cannot accurately extract the deep correlation information of input data, and has the problem of overfitting. Therefore, the errors between the predicted data and the real data are obvious, and the prediction curves cannot accurately fit the degradation process of the hydraulic pump. The prediction curves are almost a straight line, and the failure time of the hydraulic pump cannot be predicted accurately, especially after the 30<sup>th</sup> sample.

To further demonstrate good performance of ICRBM-DBN, two modified DBN models are applied to the same RUL prediction task. One model is the deep belief echo-state network (DBEN) proposed by Sun et al. in 2017.<sup>25</sup> Another model is a new deep belief network based on RBM with glial chains (GC-DBN) proposed by Geng et al. in 2018.<sup>27</sup> The prediction results are shown in Figs. 13–14.

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 3 methods are shown in Table 2.

Table 2 shows that the prediction accuracy of the proposed deep learning method is obviously better than the shallow learning methods. The ICRBM-DBN model proposed in this paper has the highest prediction accuracy, and the final prediction error is only 20 min, which meets the needs of CBM.

We have carried out several life tests for hydraulic pumps

Table 2. The prediction results by different algorithms.

Algorithms	RMSE	MAE	MRE	Error of RUL
				prediction
ICRBM	0.023	0.046	0.078	20 min
RBM-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
DBEN	0.027	0.042	0.084	60 min
GC-DBN	0.075	0.136	0.187	180 min



Figure 15. Pistons inside the failed pump.

of the same type while relying on the hydraulic pump life test platform. The failure mode varies from loose slipper fault to slipper wear fault. We utilized another full life dataset acquired from slipper wear fault which has been shown in Fig. 15 and Fig. 16.

The full life time is 30630 min. The sampling period is 10 min, and there are 3063 sets of data in the whole life of the hydraulic pump. The starting point is set as the 3020<sup>th</sup> set of data, and the failure point the 3063<sup>rd</sup> set of data. The prediction result of the proposed method is shown in Fig. 17. The predicted trend is consistent with the real trend and the prediction error is only one data point. That means the RUL prediction error is only 10 min. The satisfactory result demonstrates good applicability of the proposed method.

## 5. CONCLUSIONS

This paper proposes a novel prediction method called ICRBM-DBN based on deep belief network theory for the RUL prediction of hydraulic pumps. In order to improve the capability of deep belief network and increase the prediction accuracy, the conditional RBMs, serving as the core of the improved DBN model, are added with timing linkage factors and constraint variables between the same layers. And the updat-



Figure 16. The comparison between the failed piston and a normal piston.



Figure 17. The prediction result by ICRBM-DBM.

ing rule of ICRBM is also discussed in detail. The bispectrum entropy features of different frequency bands supply enough degradation information to reflect and predict the health state of the hydraulic pump which is indicated by the index of volumetric efficiency. The experiment results and the comparisons are concluded as follows:

- (1) The proposed ICRBM model solves the problems of the traditional RBM model such as inability to extract deep features and limited gradient descent.
- (2) A new prediction model, ICRBM-DBN, is constructed by multiple ICRBMs under the improved updating rule of CRBM. Experimental results and comparisons show that the proposed prediction model of ICRBM-DBN can achieve a better prediction accuracy compared with the original RBM-DBN, BP neural network, SVM and modified DBNs such as DEBN and GC-DBN when conducting RUL prediction of hydraulic pumps; and
- (3) The proposed method provides a useful tool for residual useful life prediction and the prediction accuracy is acceptable for the requirements of CBM.

## 6. ACKNOWLEDGEMENTS

This project is supported by National Natural Science Foundation of China (Grant No.51275524) and Natural Science Youth Foundation of Hebei (Grant No. E2017208086)

# REFERENCES

- <sup>1</sup> Ahmad, R. and Kamaruddin, S. An overview of time-based and condition-based maintenance in industrial application, *Computers & Industrial Engineering.*, **63**(1), 135–149, (2012). https://dx.doi.org/10.1016/j.cie.2012.02.002
- <sup>2</sup> Li, H., Wang, Y., Wang, B., Sun, J. and Li, Y. The application of a general mathematical morphological particle as a novel indicator for the performance degradation assessment of a bearing, *Mechanical Systems and Signal Processing*, **82**, 490–502, (2017). https://dx.doi.org/10.1016/j.ymssp.2016.05.038
- <sup>3</sup> Sun, J., Li, H. and Xu, B. The morphological undecimated wavelet decomposition – Discrete cosine transform composite spectrum fusion algorithm and its application on hydraulic pumps, *Measurement*, **94**, 794–805, (2016). https://dx.doi.org/10.1016/j.measurement.2016.09.024
- Hinton. G. and Salakhutdinov, R. Reducing the dimensionality of data with neural net-313(5786), 504-507, works, Science. (2006).https://dx.doi.org/10.1126/science.1127647
- <sup>5</sup> Wang, Y. , Cao, F. and Yuan, Y. A study on effectiveness of extreme learning machine, *Neurocomputing*, **74**(16), 2483–2490, (2011). https://dx.doi.org/10.1016/j.neucom.2010.11.030
- <sup>6</sup> Tamilselvan, P. and Wang, P. Failure diagnosis using deep belief learning based health state classification, *Reliability Engineering & System Safety*, **115**, 124–135, (2013). https://dx.doi.org/10.1016/j.ress.2013.02.022
- <sup>7</sup> Theodorakopoulos, I., Kastaniotis, D., Economou, G. and Fotopoulos, S. Pose-based human action recognition via sparse representation in dissimilarity space, *Journal of Visual Communication and Image Representation*, **25**(1), 12–23, (2014). https://dx.doi.org/10.1016/j.jvcir.2013.03.008
- <sup>8</sup> Jia, F., Lei, Y., Lin, J., Zhou, X. and Lu, N. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data, *Mechanical Systems and Signal Processing*, **72-73**, 303–315, (2016). https://dx.doi.org/10.1016/j.ymssp.2015.10.025
- <sup>9</sup> Huang, W., Song, G., Hong, H. and Xie, K. Deep Architecture for Traffic Flow Prediction: Deep Belief Networks With Multitask Learning, *IEEE Trans.*

#### H. Yu, et al.: A NOVEL DEEP BELIEF NETWORK MODEL CONSTRUCTED BY IMPROVED CONDITIONAL RBMS AND ITS APPLICATION IN...

*Intell. Transport. Syst.*, **15**(5), 2191–2201, (2014). https://dx.doi.org/10.1109/TITS.2014.2311123

- <sup>10</sup> Van Tran, T., AlThobiani, F. and Ball, A. An approach to fault diagnosis of reciprocating compressor valves using Teager–Kaiser energy operator and deep belief networks, *Expert Systems with Applications*, **41**(9), 4113–4122, (2014). https://dx.doi.org/10.1016/j.eswa.2013.12.026
- <sup>11</sup> Roy, P. P., Chhreawala, Y and Cheriet, M. Deep-Belief-Network Based Rescoring Approach for Handwritten Word Recognition, *Proc. International Conference on Frontiers in Handwriting Recognition*, Heraklion, Greece, (2014). https://dx.doi.org/10.1109/ICFHR.2014.91
- <sup>12</sup> Mohamed, A.-R., Dahl, G. and Hinton, G. Acoustic Modeling Using Deep Belief Networks, *IEEE Trans. Audio Speech Lang. Process.*, **20**(1), 14–22, (2012). https://dx.doi.org/10.1109/tasl.2011.2109382
- <sup>13</sup> Lee, H., Grosse, R., Ranganath, R. and Ng, A. Unsupervised learning of hierarchical representations with convolutional deep belief networks, *Commun. ACM*, **54**(10). 95-103, (2011). https://dx.doi.org/10.1145/2001269.2001295
- <sup>14</sup> Zhang R., Shen F. R. and Zhao J. X. A model with fuzzy granulation and deep belief networks for exchange rate forecasting, *Proc. International Joint Conference on Neural Networks*, Beijing, China, (2014). https://dx.doi.org/10.1109/IJCNN.2014.6889448
- <sup>15</sup> Chen, J., Jin, Q. and Chao, J. Design of Deep Belief Networks for Short-Term Prediction of Drought Index Using Data in the Huaihe River Basin, *Mathematical Problems in Engineering*, **2012**(6), 1–16, (2012). https://dx.doi.org/10.1155/2012/235929
- <sup>16</sup> Li, C., Sanchez, R. V., Zurita, G., Cerrada, M., Cabrera, D. and Vásquez, R. Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals, *Mechanical Systems and Signal Processing*, **76-77**, 283–293, (2016). https://dx.doi.org/10.1016/j.ymssp.2016.02.007
- <sup>17</sup> Kuremoto, T., Kimura, S., Kobayashi, K. and Obayashi, M. Time series forecasting using a deep belief network with restricted Boltzmann machines, *Neurocomputing*, **137**, 47–56, (2014). https://dx.doi.org/10.1016/j.neucom.2013.03.047
- <sup>18</sup> Wu, G. D. and Zhu, Z. W. An enhanced discriminability recurrent fuzzy neural network for temporal classification problems, *Fuzzy Sets and Systems*, **237**, 47–62, (2014). https://dx.doi.org/10.1016/j.fss.2013.05.007

- <sup>19</sup> Montufar G., Ay N. and Ghazi-Zahedi K. Geometry and Expressive Power of Conditional Restricted Boltzmann Machines, Retrieved from https://arxiv.org/abs/1402.3346
- <sup>20</sup> Sun J., Li H. R. and Tian Z. K. A degradation feature extraction method for hydraulic pumps based on fusion of sensitive components, *Chinese Journal* of Scientific Instrument, **37**(6), 1290-1298, (2016). https://dx.doi.org/10.3969/j.issn.0254-3087.2016.06.012
- <sup>21</sup> Guoji, S., McLaughlin, S., Yongcheng, X. and White, P. Theoretical and experimental analysis of bispectrum of vibration signals for fault diagnosis of gears, *Mechanical Systems and Signal Processing*, **43**(1-2), 76–89, (2014). https://dx.doi.org/10.1016/j.ymssp.2013.08.023
- <sup>22</sup> Taylor G. W. Hinton G. E. Factored conditional restricted Boltzmann Machines for modeling motion style, *Proc. Annual International Conference* on Machine Learning, Montreal, Canada, (2009). https://dx.doi.org/10.1145/1553374.1553505
- <sup>23</sup> Zhao G. Q., Zhang G. H., Liu Y. F. and Zhang B. Lithium-ion battery remaining useful life prediction with Deep Belief Network and Relevance Vector Machine, *IEEE International Conference on Prognostics and Health Management*, Harbin, China, (2017). https://dx.doi.org/10.1109/ICPHM.2017.7998298
- <sup>24</sup> Zhang, C., Lim, P., Qin, A. and Tan, K. Multiobjective Deep Belief Networks Ensemble for Remaining Useful Life Estimation in Prognostic, *IEEE transactions on neural networks and learning systems*, **28**(10), 2306–2318, (2017). https://dx.doi.org/10.1109/TNNLS.2016.2582798
- <sup>25</sup> Sun, X., Li, T., Li, Q., Huang, Y. and Li, Y. Deep belief echo-state network and its application to time series prediction, *Knowledge-Based Systems*, **130**, 17–29, (2017). https://dx.doi.org/10.1016/j.knosys.2017.05.022
- <sup>26</sup> Li, H., Tian, Z., Yu, H. and Xu, B. Fault Prognosis of Hydraulic Pump Based on Bispectrum Entropy and Deep Belief Network, *Measurement Science Review*, **19**(5), 195–203, (2019). https://dx.doi.org/10.2478/msr-2019-0025
- <sup>27</sup> Geng, Z., Li, Z. and Han, Y. A new deep belief network based on RBM with glial chains, *Information Sciences*, **463-464**, 294–306, (2018). https://dx.doi.org/10.1016/j.ins.2018.06.043