

Research Article

A Multiple-Kernel Relevance Vector Machine with Nonlinear Decreasing Inertia Weight PSO for State Prediction of Bearing

Sheng-wei Fei and Yong He

School of Mechanical Engineering, Donghua University, Shanghai 201620, China

Correspondence should be addressed to Sheng-wei Fei; fei1980s@163.com

Received 21 September 2014; Accepted 20 January 2015

Academic Editor: Yanxue Wang

Copyright © 2015 S.-w. Fei and Y. He. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The scientific and accurate prediction for state of bearing is the key to ensure its safe operation. A multiple-kernel relevance vector machine (MkRVM) including RBF kernel and polynomial kernel is proposed for state prediction of bearing in this study; the proportions of RBF kernel and polynomial kernel are determined by a controlled parameter. As the selection of the parameters of the kernel functions and the controlled parameter has a certain influence on the prediction results of MkRVM, nonlinear decreasing inertia weight PSO (NDIWPSO) is used to select its kernel parameters and controlled parameter. The RBF kernel RVM model with NDIWPSO (NDIWPSO-RBFRVM) and the polynomial kernel RVM model with NDIWPSO (NDIWPSO-PolyRVM) are used, respectively, to compare with the multiple-kernel RVM model with NDIWPSO (NDIWPSO-MkRVM). The experimental results indicate that NDIWPSO-MkRVM is more suitable for the state prediction of bearing than NDIWPSO-RBFRVM and NDIWPSO-PolyRVM.

1. Introduction

The scientific and accurate prediction for state of bearing is the key to ensure its safe operation [1]. Artificial neural networks (ANNs) [2, 3] and support vector machine (SVM) algorithm [4] have been used as efficient alternative tools to solve the nonlinear prediction problems. Practicability of artificial neural networks is limited due to the shortcomings of overfitting and falling into local extremum easily existed in ANNs. Support vector machine is a kind of machine learning method based on the statistical learning theory, which has better generalization performance than artificial neural networks, particularly under the condition of small training samples [5, 6].

Relevance vector machine (RVM) is an intelligent learning technique based on sparse Bayesian framework, as the number of relevance vectors in RVM is much smaller than that of support vectors in SVM, which makes RVM have a sparser representation compared with SVM [7]. Thus, RVM has better generalization ability than SVM. In order to improve the generalization ability of RVM, a multiple-kernel relevance vector machine including RBF kernel and polynomial kernel is proposed for state prediction of bearing in this

study; the proportions of RBF kernel and polynomial kernel are determined by a controlled parameter. As the selection of the parameters of the kernel functions and the controlled parameter has a certain influence on the prediction results of MkRVM, nonlinear decreasing inertia weight PSO is used to select its kernel parameters and controlled parameter. Particle swarm optimization is inspired by the social behavior of bird flocking or fish schooling [8]. Compared with traditional PSO, nonlinear decreasing inertia weight PSO has a nonlinear decreasing inertia weight instead of fixed inertia weight. The RBF kernel RVM model and the polynomial kernel RVM model are used, respectively, to compare with the multiple-kernel RVM model, and the kernel parameters of the RBF kernel RVM model and the polynomial kernel RVM model are also selected by NDIWPSO. The experimental results indicate that NDIWPSO-MkRVM is more suitable for the state prediction of bearing than NDIWPSO-RBFRVM and NDIWPSO-PolyRVM.

2. Multiple-Kernel Relevance Vector Machine

Let $T = \{\mathbf{x}_i, t_i\}_{i=1}^N$ be a set of the training data, where \mathbf{x}_i denotes the input vector and t_i denotes the corresponding

output target and the target t_i includes the additive noise [9], which can be expressed as follows:

$$t_i = y(\mathbf{x}_i, \mathbf{w}) + \varepsilon_i, \quad (1)$$

where ε_i is assumed to be mean-zero Gaussian noise with variance σ^2 .

The regression function of relevance vector machine which consists of a linear combination of the weighted kernel functions can be described as follows:

$$y(\mathbf{x}, \mathbf{w}) = \sum_{i=1}^N w_i K(\mathbf{x}, \mathbf{x}_i) + w_0, \quad (2)$$

where $K(\mathbf{x}, \mathbf{x}_i)$ is the kernel function, $\mathbf{w} = [w_1, w_2, \dots, w_N]$ is the weight vector, and w_0 is the bias.

As different kernel functions will obtain different regression functions of RVM, which can determine the operation performance of RVM, it is very important to choose a suitable kernel function of relevance vector machine. Here, two kinds of kernels including local kernel and global kernel are employed to construct the regression function of relevance vector machine. A typical local kernel is radial basis function kernel (RBF); in this study, Gaussian kernel is used as radial basis function kernel, which can be defined as follows:

$$K_{\text{RBF}}(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\gamma^2}\right), \quad (3)$$

where γ denotes the kernel parameter of the RBF kernel.

A typical global kernel is the polynomial kernel, which can be defined as follows:

$$K_{\text{Poly}}(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i \cdot \mathbf{x}_j^T + 1)^d, \quad (4)$$

where d denotes the kernel parameter of the polynomial kernel.

In order to improve the generalization ability of RVM, a multiple-kernel relevance vector machine is constructed by the local kernel function (RBF kernel K_{RBF}) and global kernel function (polynomial kernel K_{Poly}); the proportions of RBF kernel and polynomial kernel are determined by the controlled parameter u . Thus, the multiple-kernel function can be expressed as follows:

$$K_{\text{mix(RBF, Poly)}}(\mathbf{x}_i, \mathbf{x}_j) = u K_{\text{RBF}}(\mathbf{x}_i, \mathbf{x}_j) + (1 - u) K_{\text{Poly}}(\mathbf{x}_i, \mathbf{x}_j), \quad (5)$$

where u ($0 \leq u \leq 1$) denotes the controlled parameter.

When $u = 1$, relevance vector machine only employs RBF kernel function, and when $u = 0$, relevance vector machine only employs polynomial kernel function. It is obvious that the multiple-kernel function integrates all characters of independent kernel and has better data distribution performance.

3. State Prediction Method of Bearing Based on Multiple-Kernel Relevance Vector Machine with Nonlinear Decreasing Inertia Weight PSO

State prediction process of bearing based on multiple-kernel relevance vector machine with nonlinear decreasing

inertia weight PSO can be described in this section. The experimental data are normalized to the range $[0, 1]$ in order to improve the generalization ability of the prediction model. Assume that the normalized experimental data are $a_1, a_2, \dots, a_m, \dots, a_n, \dots, a_{n+k}$, among which $a_1, a_2, \dots, a_m, \dots, a_n$ are used to establish the training sample sets, and a_{n+1}, \dots, a_{n+k} are used to test the prediction model. The training sample sets can be described by the formula:

$$\mathbf{X} = \begin{bmatrix} a_1 & a_2 & \cdots & a_m \\ a_2 & a_3 & \cdots & a_{m+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-m} & a_{n-m+1} & \cdots & a_{n-1} \end{bmatrix}, \quad (6)$$

$$\mathbf{Y} = \begin{bmatrix} a_{m+1} \\ a_{m+2} \\ \vdots \\ a_n \end{bmatrix},$$

where m denotes the embedding dimension, \mathbf{X} denotes the set of input vectors, and \mathbf{Y} denotes the set of corresponding outputs.

In this paper, we use nonlinear decreasing inertia weight particle swarm optimization to select the kernel parameters γ and d and controlled parameter u . In NDIWPSO, the particle is composed of the kernel parameters γ and d and controlled parameter u ; each parameter has its value range. We perform the validation for each training sample by the MkRVM models with the different particles and obtain the corresponding validation error of each training sample. Mean validation error of all training samples can be used to evaluate the performance of the MkRVM models with the different particles. The process of the selection of the kernel parameters γ and d and controlled parameter u of MkRVM by NDIWPSO can be described as follows.

Step 1. Perform the setting of the parameters of NDIWPSO and initialization of the particles.

Step 2. Evaluate the fitness of each particle. The mean validation error of all the training samples is used to create the fitness function, which is defined as follows:

$$\tilde{e} = \frac{1}{G} \sum_{q=1}^G \left| \frac{y_q - \hat{y}_q}{y_q} \right|, \quad (7)$$

where y_q is the actual value and \hat{y}_q is the validation value; G is the number of the training samples in training sample sets.

Step 3. Update the global and personal best according to the fitness evaluation results.

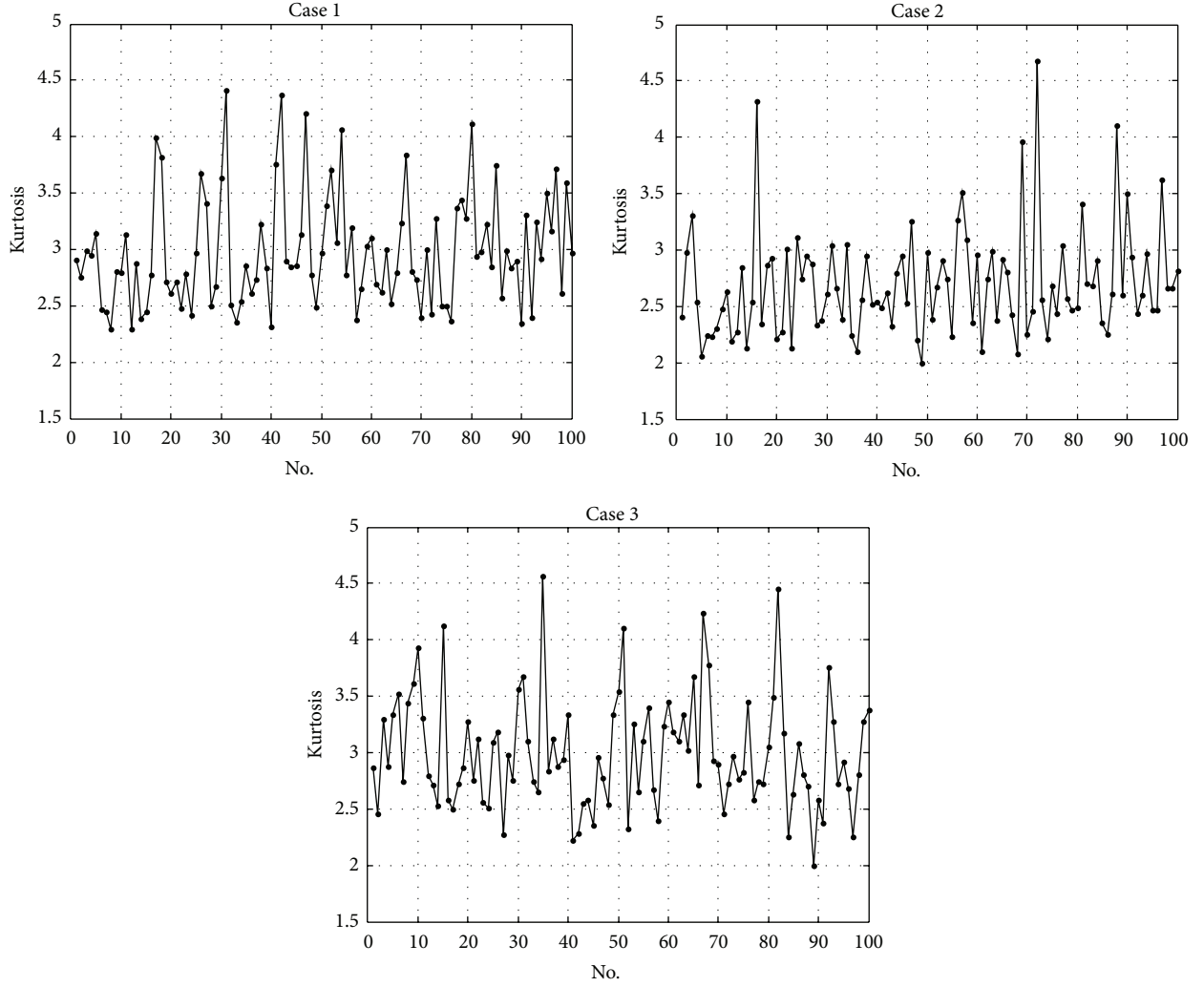


FIGURE 1: The kurtosis data of bearing vibration signal.

Step 4. The particle flies toward a new position by calculating the velocity of position change, and the velocity of each particle is calculated by the following formula:

$$\begin{aligned}
 v_{ij}(h+1) = & \omega(h) \cdot v_{ij}(h) \\
 & + c_1 \cdot \text{rand} \cdot (pbest_{ij}(h) - p_{ij}(h)) \\
 & + c_2 \cdot \text{rand} \cdot (gbest_j(h) - p_{ij}(h)),
 \end{aligned} \quad (8)$$

where h denotes the iteration counter, v_{ij} is the velocity of particle i on the j th dimension, p_{ij} is the position of particle i on the j th dimension, $pbest_{ij}$ is the personal best (pbest) position of particle i on the j th dimension, $gbest_j$ is the global best (gbest) position of the swarm, rand means the random value in the range $[0, 1]$, and the positive constants c_1 and c_2 are personal learning factor and social learning factor, respectively. ω is the inertia weight, which is used to balance the global exploration and local exploitation, $\omega(h) = \omega_{\max} - ((\omega_{\max} - \omega_{\min}) / \sqrt{H-1}) \times \sqrt{h-1}$ (ω_{\max} is the maximum

inertia weight, ω_{\min} is the minimum inertia weight, and H is the maximum iteration).

Step 5. Each particle moves to its next position according to the following formula:

$$p_{ij}(h+1) = p_{ij}(h) + \beta \cdot v_{ij}(h+1), \quad (9)$$

where β is constraint factor used to control the velocity weight.

Step 6. The same procedures from Step 2 to Step 5 are repeated until the maximum iteration is reached.

Then, the MkrVM model is trained with obtained kernel parameters γ and d and controlled parameter u . Finally, the proposed NDIWPSO-MkrVM model is tested by the testing data.

TABLE 1: The comparison of the mean absolute percentage errors among NDIWPSO-MkRVM, NDIWPSO-RBFRVM, and NDIWPSO-PolyRVM.

| | Prediction model | MAPE (%) |
|--------|------------------|----------|
| Case 1 | NDIWPSO-MkRVM | 11.32 |
| | NDIWPSO-RBFRVM | 11.99 |
| | NDIWPSO-PolyRVM | 12.08 |
| Case 2 | NDIWPSO-MkRVM | 6.44 |
| | NDIWPSO-RBFRVM | 6.63 |
| | NDIWPSO-PolyRVM | 7.41 |
| Case 3 | NDIWPSO-MkRVM | 12.79 |
| | NDIWPSO-RBFRVM | 13.93 |
| | NDIWPSO-PolyRVM | 12.96 |

4. Experimental Analysis

Kurtosis of bearing vibration signal can excellently reflect the state of bearing; that is, kurtosis prediction of bearing vibration signal can excellently reflect the future state of bearing. In this paper, the kurtosis data can be calculated and obtained from bearing vibration signal in “Bearing Data Set” [10]. As shown in Figure 1, the kurtosis data of bearing vibration signal of three cases are given, and 100 numbers of kurtosis data of bearing vibration signal are employed in each case. In each case, numbers 1~90 of kurtosis data of bearing vibration signal are used as the training data, and numbers 91~100 of kurtosis data of bearing vibration signal are used as the testing data.

The experimental data are normalized to the range $[0, 1]$ in order to improve the generalization ability of the prediction model. And we use nonlinear decreasing inertia weight particle swarm optimization to select the kernel parameters γ and d and controlled parameter u . In NDIWPSO, ω_{\max} is set to 0.9, ω_{\min} is set to 0.4, β is set to $1/1.3$, and the positive constants c_1 and c_2 are set to 2. Here, the value range of γ is $[0.5, 5]$, the value range of d is $[0, 5]$, and the value range of u is $[0, 1]$. We set the embedding dimension m to 5 to establish the training sample sets. The RBF kernel RVM model and the polynomial kernel RVM model are used, respectively, to compare with the multiple-kernel RVM model, and the kernel parameters of the RBF kernel RVM model and the polynomial kernel RVM model are also selected by NDIWPSO.

In case 1, Figure 2 gives the comparison of the prediction results for kurtosis of bearing vibration signal among the multiple-kernel RVM model with NDIWPSO (NDIWPSO-MkRVM), the RBF kernel RVM model with NDIWPSO (NDIWPSO-RBFRVM), and the polynomial kernel RVM model with NDIWPSO (NDIWPSO-PolyRVM); and Figure 3 gives the comparison of the absolute percentage prediction errors for kurtosis of bearing vibration signal among NDIWPSO-MkRVM, NDIWPSO-RBFRVM, and NDIWPSO-PolyRVM. As shown in Table 1, the mean absolute percentage error (MAPE) of NDIWPSO-MkRVM in case 1 is 11.32%, the MAPE of NDIWPSO-RBFRVM in case 1 is 11.99%, and the MAPE of NDIWPSO-PolyRVM in case 1 is 12.08%. It can

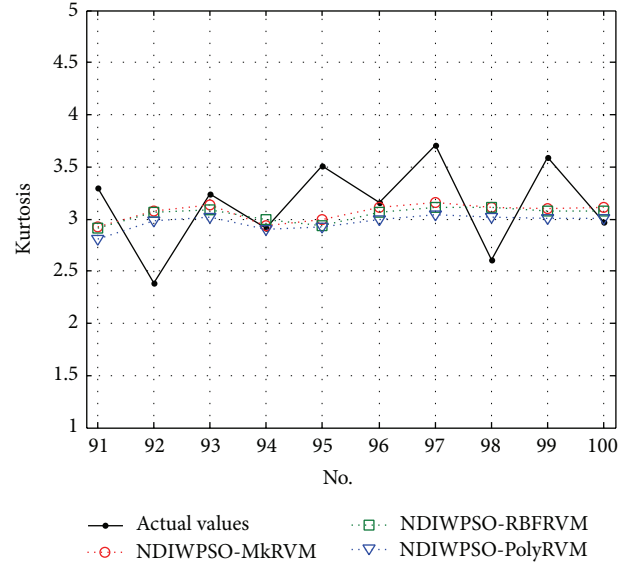


FIGURE 2: The comparison of the prediction results for kurtosis of bearing vibration signal among NDIWPSO-MkRVM, NDIWPSO-RBFRVM and NDIWPSO-PolyRVM in case 1.

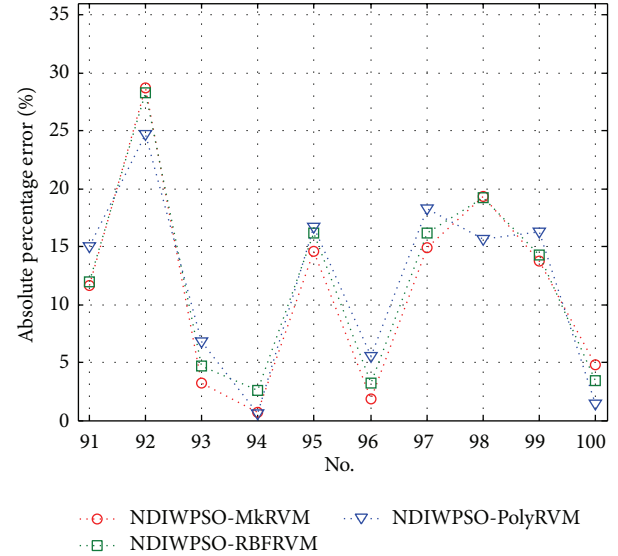


FIGURE 3: The comparison of the absolute percentage prediction errors for kurtosis of bearing vibration signal among NDIWPSO-MkRVM, NDIWPSO-RBFRVM and NDIWPSO-PolyRVM in case 1.

be seen that NDIWPSO-MkRVM has smaller MAPE than NDIWPSO-RBFRVM and NDIWPSO-PolyRVM in this case.

In case 2, Figure 4 gives the comparison of the prediction results for kurtosis of bearing vibration signal among NDIWPSO-MkRVM, NDIWPSO-RBFRVM, and NDIWPSO-PolyRVM; and Figure 5 gives the comparison of the absolute percentage prediction errors for kurtosis of bearing vibration signal among NDIWPSO-MkRVM, NDIWPSO-RBFRVM, and NDIWPSO-PolyRVM. As shown in Table 1, the MAPE of NDIWPSO-MkRVM in case 2 is

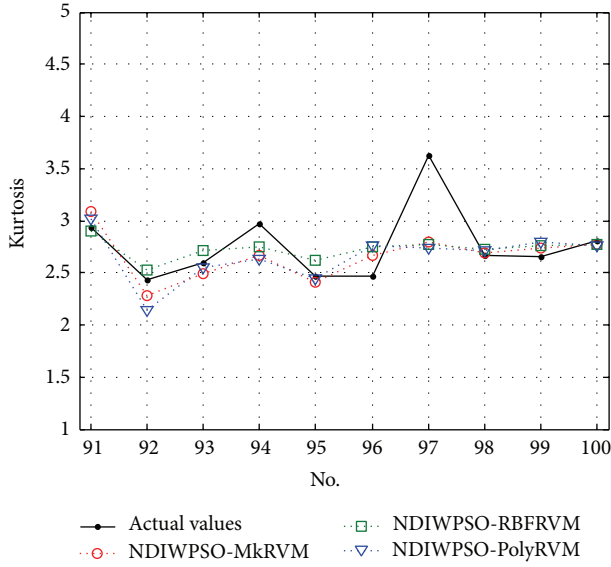


FIGURE 4: The comparison of the prediction results for kurtosis of bearing vibration signal among NDIWPSO-MkRVM, NDIWPSO-RBFRVM and NDIWPSO-PolyRVM in case 2.

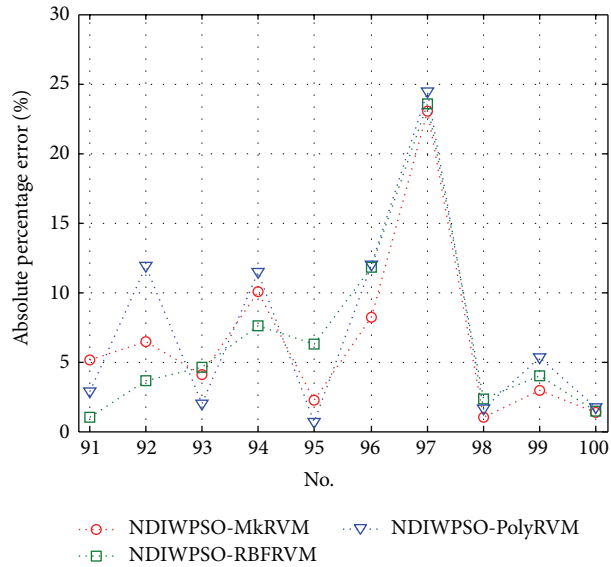


FIGURE 5: The comparison of the absolute percentage prediction errors for kurtosis of bearing vibration signal among NDIWPSO-MkRVM, NDIWPSO-RBFRVM and NDIWPSO-PolyRVM in case 2.

6.44%, the MAPE of NDIWPSO-RBFRVM in case 2 is 6.63%, and the MAPE of NDIWPSO-PolyRVM in case 2 is 7.41%. It can be seen that the MAPE of NDIWPSO-RBFRVM is near to that of NDIWPSO-MkRVM in case 2; however, the MAPE of NDIWPSO-PolyRVM is obviously higher than that of NDIWPSO-MkRVM in case 2.

In case 3, Figure 6 gives the comparison of the prediction results for kurtosis of bearing vibration signal among NDIWPSO-MkRVM, NDIWPSO-RBFRVM, and NDIWPSO-

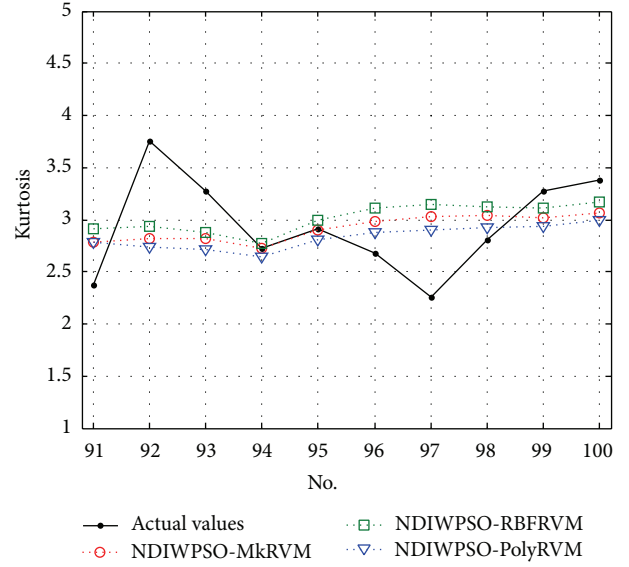


FIGURE 6: The comparison of the prediction results for kurtosis of bearing vibration signal among NDIWPSO-MkRVM, NDIWPSO-RBFRVM and NDIWPSO-PolyRVM in case 3.

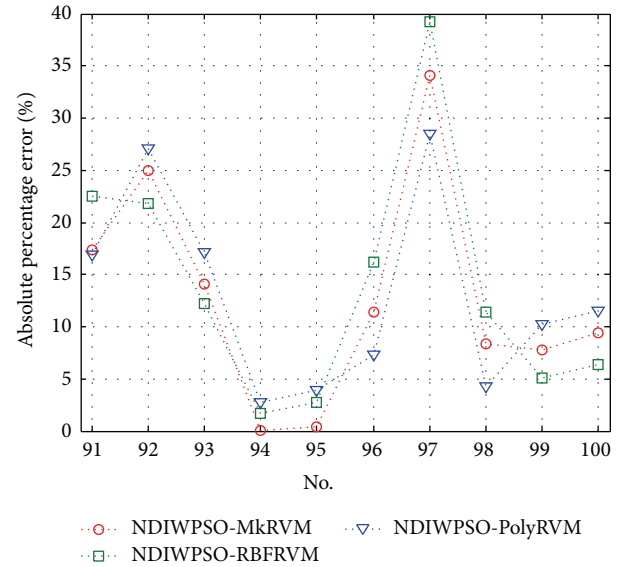


FIGURE 7: The comparison of the absolute percentage prediction errors for kurtosis of bearing vibration signal among NDIWPSO-MkRVM, NDIWPSO-RBFRVM and NDIWPSO-PolyRVM in case 3.

PolyRVM; and Figure 7 gives the comparison of the absolute percentage prediction errors for kurtosis of bearing vibration signal among NDIWPSO-MkRVM, NDIWPSO-RBFRVM, and NDIWPSO-PolyRVM. As shown in Table 1, the MAPE of NDIWPSO-MkRVM in case 3 is 12.79%, the MAPE of NDIWPSO-RBFRVM in case 3 is 13.93%, and the MAPE of NDIWPSO-PolyRVM in case 3 is 12.96%. It can be seen that the MAPE of NDIWPSO-PolyRVM is near to that of NDIWPSO-MkRVM in case 3; however, the MAPE of

NDIWPSO-RBFRVM is obviously higher than that of NDIWPSO-MkRVM in case 3.

The experimental results show that the MAPE of NDIWPSO-RBFRVM is near to that of NDIWPSO-MkRVM in case 2, but the MAPE of NDIWPSO-RBFRVM is obviously higher than that of NDIWPSO-MkRVM in case 3; in addition, the MAPE of NDIWPSO-PolyRVM is near to that of NDIWPSO-MkRVM in case 3, but the MAPE of NDIWPSO-PolyRVM is obviously higher than that of NDIWPSO-MkRVM in case 2. It is indicated that kurtosis prediction performance of bearing vibration signal of NDIWPSO-RBFRVM and NDIWPSO-PolyRVM is less stable than that of NDIWPSO-MkRVM. Thus, NDIWPSO-MkRVM is more suitable for the state prediction of bearing than NDIWPSO-RBFRVM and NDIWPSO-PolyRVM.

5. Conclusion

In this paper, multiple-kernel relevance vector machine with nonlinear decreasing inertia weight PSO is proposed for state prediction of bearing. The multiple-kernel function including two kinds of kernels, RBF kernel and polynomial kernel, is employed to improve the generalization ability of RVM, and we use nonlinear decreasing inertia weight PSO to select the kernel parameters and controlled parameter of MkRVM. The experimental results show that NDIWPSO-MkRVM has more stable kurtosis prediction performance of bearing vibration signal than NDIWPSO-RBFRVM and NDIWPSO-PolyRVM. Thus, NDIWPSO-MkRVM is more suitable for the state prediction of bearing than NDIWPSO-RBFRVM and NDIWPSO-PolyRVM.

Conflict of Interests

The authors declare that they have no conflict of interests.

Acknowledgment

This project is supported by “the Fundamental Research Funds for the Central Universities (no. 2232013D3-28).”

References

- [1] S. Hong, Z. Zhou, E. Zio, and W. Wang, “An adaptive method for health trend prediction of rotating bearings,” *Digital Signal Processing*, vol. 35, pp. 117–123, 2014.
- [2] W. Yaïci and E. Entchev, “Performance prediction of a solar thermal energy system using artificial neural networks,” *Applied Thermal Engineering*, vol. 73, no. 1, pp. 1348–1359, 2014.
- [3] L. Das, U. Maity, and J. K. Basu, “The photocatalytic degradation of carbamazepine and prediction by artificial neural networks,” *Process Safety and Environmental Protection*, vol. 92, no. 6, pp. 888–895, 2014.
- [4] I. Nejatian, M. Kanani, M. Arabloo, A. Bahadori, and S. Zendeheboudi, “Prediction of natural gas flow through chokes using support vector machine algorithm,” *Journal of Natural Gas Science and Engineering*, vol. 18, pp. 155–163, 2014.
- [5] S. Smusz, W. M. Czarnecki, D. Warszycki, and A. J. Bojarski, “Exploiting uncertainty measures in compounds activity prediction using support vector machines,” *Bioorganic & Medicinal Chemistry Letters*, vol. 25, no. 1, pp. 100–105, 2015.
- [6] M. D. C. Moura, E. Zio, I. D. Lins, and E. Droguett, “Failure and reliability prediction by support vector machines regression of time series data,” *Reliability Engineering & System Safety*, vol. 96, no. 11, pp. 1527–1534, 2011.
- [7] M. E. Tipping, “Sparse bayesian learning and the relevance vector machine,” *Journal of Machine Learning Research*, vol. 1, no. 3, pp. 211–244, 2001.
- [8] M. A. Behrang, E. Assareh, A. R. Noghrehabadi, and A. Ghanbarzadeh, “New sunshine-based models for predicting global solar radiation using PSO (particle swarm optimization) technique,” *Energy*, vol. 36, no. 5, pp. 3036–3049, 2011.
- [9] G. Valente, F. de Martino, F. Esposito, R. Goebel, and E. Formisano, “Predicting subject-driven actions and sensory experience in a virtual world with Relevance Vector Machine Regression of fMRI data,” *NeuroImage*, vol. 56, no. 2, pp. 651–661, 2011.
- [10] J. Lee, H. Qiu, G. Yu, J. Lin, and Rexnord Technical Services, “Bearing data set,” Tech. Rep., IMS, University of Cincinnati, NASA Ames Prognostics Data Repository, NASA Ames, Moffett Field, Calif, USA, 2007.

