Research Article

An Adaptive Support Vector Regression Machine for the State Prognosis of Mechanical Systems

Qing Zhang,^{1,2} Fei Liu,¹ Xiang Wan,¹ and Guanghua Xu³

¹*School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, China*

²*Key Laboratory of Education Ministry for Modern Design and Rotor-Bearing System, Xi'an Jiaotong University, Xi'an 710049, China*

³*State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, China*

Correspondence should be addressed to Qing Zhang; zhangq@mail.xjtu.edu.cn

Received 12 December 2014; Accepted 20 January 2015

Academic Editor: Yanxue Wang

Copyright © 2015 Qing Zhang et al. 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.

Due to the unsteady state evolution of mechanical systems, the time series of state indicators exhibits volatile behavior and staged characteristics. To model hidden trends and predict deterioration failure utilizing volatile state indicators, an adaptive support vector regression (ASVR) machine is proposed. In ASVR, the width of an error-insensitive tube, which is a constant in the traditional support vector regression, is set as a variable determined by the transient distribution boundary of local regions in the training time series. Thus, the localized regions are obtained using a sliding time window, and their boundaries are defined by a robust measure known as the truncated range. Utilizing an adaptive error-insensitive tube, a stabilized tolerance level for noise is achieved, whether the time series occurs in low-volatility regions or in high-volatility regions. The proposed method is evaluated by vibrational data measured on descaling pumps. The results show that ASVR is capable of capturing the local trends of the volatile time series of state indicators and is superior to the standard support vector regression for state prediction.

1. Introduction

The state prognosis of mechanical systems is of critical importance in modern industry to prevent unexpected breakdowns, to improve machine availability, and to reduce maintenance costs. Generally, the working state of mechanical systems is represented by certain indicators, which are either acquired from monitoring devices or calculated from raw monitoring signals. The primary task of state prognosis is to estimate the actual development of the state by modeling the trend of state indicators. Then, the trend model can be extrapolated to predict the upcoming failure or estimate the remaining useful life. After an accurate prognosis is achieved, timely maintenance actions can be planned to avoid catastrophic failure.

Due to the unstable operating conditions and accidental disturbances, the time series of state indicators always exhibits random fluctuations, whether the monitored system works normally or not. Therefore, many intelligent methods have been proposed to extract hidden trends from the

observed state indicators. Artificial neural network (ANN) is one of the widely used methods in the prognostics literatures [1]. Gebraeel et al. [2] developed a set of feed-forward backpropagation networks to model the degradation process of rolling element bearings and to estimate the failure time of partially degraded bearings. Tse and Atherton [3] used a recurrent neural network to determine the trend in the monitoring values and to predict the value at the next time step. Because the trend is learned and memorized by neurons and network weights, ANN provides a nontransparent solution to state prognosis, or rather, the way in which forecast results are inferred by a trained network cannot be observed. Random coefficient models are another category of prognosis method for mechanical systems. In these models, the trend in the state indicators is predefined as a linear, polynomial, exponential, or any other functional form [4, 5]. Then, the coefficients, which include the deterministic functional coefficients and the stochastic noise coefficients, are jointly estimated with historical state indicators. Due to the requirements for system-specific trend knowledge,

the applications of random coefficient models are greatly restricted. Nonparametric regression models, in which the trend needs not to take a predetermined form, overcome the barriers of prior knowledge and are also commonly used for state prognosis [6]. Among this category of models, support vector regression (SVR) [7], which has good generalization ability even if training samples are not abundant, is the most widely accepted method. By training an SVR machine, the trend of the state indicators is represented as an explicit regression function and is easily extrapolated to obtain future values. The extrapolated values are used to prognose the state evolution. Generally, while the extrapolated values reach a failure threshold predefined by theoretical or experimental analysis, a prospective failure is deduced and the failure time is estimated [8]. Therefore, SVR has been extensively studied to tackle the prognosis problem of mechanical systems or components, such as bearings, gears, and pumps [9–12].

The life of mechanical systems can be divided into a normal working stage and a deterioration stage [13]. In the first stage, the state indicators are generally shown as a stationary time series. As initial defects emerge, the system steps into the deterioration stage, including unsteady evolution and abrupt changes. These nonstationary and transient phenomena are reflected in the time series of the state indicators. To trace the evolution of states, the regression model, which is capable of adapting to staged development and volatile state indicators, is required. However, the standard SVR seeks a globally optimized regression, in which the tolerance level for noise is fixed during the entire training time series. Therefore, it lacks the flexibility to capture the local trend of a data series with time-varying variance or staged characteristics. To improve the adaptability of volatile time series, several modified SVR machines, such as localized support vector regression (LSVR) [14] and piecewise support vector regression [15], have been developed and applied in the field of financial analysis. In this paper, a novel SVR machine, called adaptive support vector regression (ASVR), is proposed to model the trend of state indicators measured from mechanical systems. We will show that, by utilizing an adaptable error-insensitive tube, ASVR can provide satisfactory performance for regression and prediction while the system is in a deterioration stage.

The rest of the paper is organized as follows. We briefly introduce related studies of standard SVR and LSVR in Section 2. The methodologies of ASVR are described in Section 3. In Section 4, standard SVR, LSVR, and ASVR are applied to address the time series of vibration acquired from actual pumps. The regression and prediction results are compared and analyzed. Finally, conclusions and future work are discussed in Section 5.

2. Related Studies

2.1. Standard Support Vector Regression. Given a time series $T = \{(t_1, y_1), (t_2, y_2), \dots, (t_n, y_n)\}$, where t_i is the time tag, y_i is the corresponding value, and n is the number of data points, the goal of SVR is to find a function $f(t)$ that has at most ε deviation from the actually obtained values y_i for all time tags while being as flat as possible [7]. By mapping the time

FIGURE 1: SVR with a soft margin.

series into a high-dimensionality feature space, the regression function $f(t)$ has the linear form as

$$f(t) = \omega \cdot \phi(t) + b, \quad (1)$$

where ω denotes a weight coefficient, $\phi(\cdot)$ is the mapping function, and b is the bias. To obtain the optimal function $f(t)$, a convex optimization problem is constructed as follows:

$$\begin{aligned} & \text{minimize} \quad \frac{1}{2} \|\omega\|^2 \\ & \text{subject to} \quad y_i - (\omega \cdot \phi(t_i) + b) \leq \varepsilon, \\ & \quad (\omega \cdot \phi(t_i) + b) - y_i \leq \varepsilon, \\ & \quad i = 1, 2, \dots, n. \end{aligned} \quad (2)$$

In the constrained minimization problem, it is assumed that a function $f(t)$ exists for which all data points in time series lie in a tube determined by $f(t) \pm \varepsilon$. ε defines the width of the error-insensitive tube, or in other words, the precision of regression is ε . However, in many applications, it is preferred to accept a number of errors, which are caused by the data points outside the error-insensitive tube, to improve the generalization ability. Therefore, the concept of a soft margin is introduced, and the original optimization problem (2) is reformed as

$$\begin{aligned} & \text{minimize} \quad \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \\ & \text{subject to} \quad y_i - (\omega \cdot \phi(t_i) + b) \leq \varepsilon + \xi_i \\ & \quad (\omega \cdot \phi(t_i) + b) - y_i \leq \varepsilon + \xi_i^* \\ & \quad \xi_i, \xi_i^* \geq 0 \\ & \quad i = 1, 2, \dots, n, \end{aligned} \quad (3)$$

where C is a positive constant, known as the regularization parameter, which determines the trade-off between the flatness of $f(t)$ and the amount up to which deviations larger than ε are tolerated. ξ_i and ξ_i^* are called slack variables and measure the deviation of $|y_i - f(t)|$ from the boundaries of the error-insensitive tube. Figure 1 provides a depiction of SVR with a soft margin.

The minimization problem denoted as (3) is called standard SVR. By the Lagrange multiplier method, this problem is transformed into its dual quadratic programming problem:

$$\begin{aligned}
 & \text{maximize} \quad -\frac{1}{2} \sum_{i,j=1}^n (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) K(t_i, t_j) \\
 & \quad + \sum_{i=1}^n y_i (\alpha_i - \alpha_i^*) - \varepsilon \sum_{i=1}^n (\alpha_i + \alpha_i^*) \\
 & \text{subject to} \quad \sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0 \\
 & \quad 0 \leq \alpha_i \leq C \\
 & \quad 0 \leq \alpha_i^* \leq C \\
 & \quad i = 1, 2, \dots, n,
 \end{aligned} \tag{4}$$

where α_i and α_i^* are the Lagrange multipliers and $K(t_i, t_j) = \phi(t_i) \cdot \phi(t_j)$ is a kernel function. After solving the dual quadratic programming problem, the regression function is formulated as

$$f(t) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(t_i, t) + b. \tag{5}$$

According to the Karush-Kuhn-Tucker conditions, $(\alpha_i - \alpha_i^*)$ equals zero when the data point (t_i, y_i) lies in the error-insensitive tube. Therefore, the regression function is simply determined by the data points that are located on the boundary or outside of the error-insensitive tube. These data points, which support the definition of the regression function, are called support vectors (SVs). Any functions that satisfy Mercer's condition can be treated as kernel functions. In this study, the Gaussian radial basis function kernel is chosen as

$$K(t_i, t_j) = \exp\left(-\frac{\|t_i - t_j\|^2}{2\sigma^2}\right), \tag{6}$$

where σ is the kernel parameter.

In standard SVR, ε is a predetermined constant. A large ε value provides a high noise-tolerance capability but may lose the local details of a trend, whereas a small ε increases the precision of regression but results in a complex learning machine. Generally, ε is chosen based on experiences. However, for a volatile time series, in particular for the state indicators of a mechanical system with staged characteristics, it is almost impossible to find a global optimal. It is more reasonable to set a narrow error-insensitive tube in the low-volatility regions and a wide error-insensitive tube in the high-volatility regions.

2.2. Localized Support Vector Regression. Many attempts have been made to adjust error-insensitive tubes based on the local characteristics of a time series. LSVR [14], which has explicit theoretical justifications, is a representative method

of these attempts. In this modified SVR, a localized region centered at the i th data point and with length $2k + 1$, that is, $(t_{i-k}, y_{i-k}), \dots, (t_{i+k}, y_{i+k})$, is considered. The standard deviation of the data points in the regions is mapped into the high-dimensionality feature space as follows:

$$s_i^\phi = \sqrt{\frac{1}{2k+1} \sum_{j=-k}^k (\phi(t_{i+j}) - \bar{\phi}_i)^2}, \tag{7}$$

where $\bar{\phi}_i = 1/(2k+1) \sum_{j=-k}^k \phi(t_{i+j})$. Then, the constrained minimization problem of LSVR is defined as

$$\begin{aligned}
 & \text{minimize} \quad \frac{1}{n} \sum_{i=1}^n u_i + C \sum_{i=1}^n (\xi_i + \xi_i^*) \\
 & \text{subject to} \quad y_i - (\omega \cdot \phi(t_i) + b) \leq \varepsilon u_i + \xi_i \\
 & \quad (\omega \cdot \phi(t_i) + b) - y_i \leq \varepsilon u_i + \xi_i^* \\
 & \quad \omega \cdot s_i^\phi \leq u_i \\
 & \quad u_i, \xi_i, \xi_i^* \geq 0 \\
 & \quad i = 1, 2, \dots, n,
 \end{aligned} \tag{8}$$

where u_i is an auxiliary variable that is determined by the upper bound of $\omega \cdot s_i^\phi$. The goal of LSVR can be interpreted as finding a regression function $f(t)$ by making the localized regions of function as low in volatility as possible while keeping the error as small as possible. By introducing the auxiliary variable u_i , LSVR can automatically adjust the error-insensitive tube. If the i th data point lies in a localized region with a larger standard deviation of noise, it will contribute to a larger $\omega \cdot s_i^\phi$ or a larger tube width εu_i . The wider error-insensitive tube reduces the impact of the noise around the data point. Conversely, if the i th data point is in a region with a smaller standard deviation of noise, it will play a greater role in the learning process of regression. In this way, the volatile noise of time series is flexibly tolerated, and the local trend of the time series is captured.

To avoid the explicit mapping operation, ω is written as a linear combination of all training data points:

$$\omega = \sum_{i=1}^n \mu_i \phi(t_i), \tag{9}$$

and it is substituted into problem (8). The computation of the kernel function $K(t_i, t_j)$ is performed to substitute for the inner products of $\phi(t_i)$ and $\phi(t_j)$ in the high-dimensionality feature space. Finally, the kernelized LSVR is transformed into a second order cone programming (SOCP) problem and solved. The regression function is obtained in the following form:

$$f(t) = \sum_{i=1}^n \mu_i K(t_i, t) + b. \tag{10}$$

Compared with standard SVR, LSVR has two disadvantages: its high computational complexity and its inadequate ability for multistep extrapolation. Because the time

FIGURE 2: Schematic diagram of ASVR.

complexity for solving the SCOP problem is an open issue, it largely restricts the computational efficiency of LSVR. Additionally, the kernel operations within each localized region also increase the consumption of computational time. On the other hand, the value of the regression function at a certain time t is largely determined by the training data points, which lie in the region neighboring t . While the regression function is extrapolated in a multistep process, the further the extrapolated time t_p is from the current time t_n , the less the training data points support the region neighboring t_p . As a result, the extrapolated value will soon approach the bias b with an increase in the extrapolated steps. Therefore, LSVR is mainly applied to the regression and one-step prediction of volatile time series.

3. Adaptive Support Vector Regression

To capture the local trend of a volatile time series while retaining the advantages of standard SVR, we made the following assumptions to build an adaptive support vector regression.

- The width of the error-insensitive tube ϵ determines the proportion of training data points excluded from the tube.
- To maintain a stabilized exclusion proportion, ϵ should keep pace with the variation of the margin of the volatile time series.
- When the margin changes, the transient distribution of a local time series is not a normal distribution but a mixed distribution or a heavily tailed distribution.

Based on these assumptions, a strategy for adaptively adjusting the error-insensitive tube is proposed. In this strategy, the constant ϵ is replaced by the variable ϵ_i determined by the distribution characteristics of the localized time series. Firstly, a sliding time window, which slides in the time axis, is adopted to continuously obtain the local regions of the training time series. For the i th sliding step, the data points within the selected local region are denoted as $L_i = \{y_{i-l+1}, y_{i-l+2}, \dots, y_i\}$, where l is the length of the time window, and $i = l, l+1, \dots, n$. Because L_i does not always follow the normal distribution, the conventional measures of scale, such as the mean and variance, are not suitable for describing the statistical distribution of L_i . Thus, a robust measure, known as the truncated range, is utilized to define

the distribution scope of L_i . It involves the calculation of the range after discarding given parts of the samples at the high and low ends and typically discarding an equal amount of both. This can be given as a percentage, but it is usually given as a fixed number of points to facilitate calculation. Suppose $L'_i = \{y'_{i-l+1}, y'_{i-l+2}, \dots, y'_i\}$ is the series of L_i in descending order, the upper bound of the truncated range is y'_{i-r} and the lower bound of the truncated range is $y'_{i-l+1+r}$, where r is the number of truncated data points. It has been verified that the truncated range is a robust estimator for mixed distributions and heavy-tailed distribution [16]. Finally, to obtain a symmetric error-insensitive tube, ϵ_i is calculated as follows:

$$\epsilon_i = \frac{y'_{i-r} - y'_{i-l+1+r}}{2}, \quad i = l, l+1, \dots, n. \quad (11)$$

It is easy to know that $\epsilon_i \geq 0$. When $i < l$ especially, ϵ_i can be set to ϵ_l . Utilizing this strategy, the error-insensitive tube is adaptive to the variation of local margins, and a stabilized tolerance level for noise is achieved, whether the time series is in a low-volatility region or a high-volatility region.

With the introduction of ϵ_i , the constrained minimization problem of ASVR is defined as

$$\begin{aligned} & \text{minimize} \quad \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \\ & \text{subject to} \quad y_i - (\omega \cdot \phi(t_i) + b) \leq \epsilon_i + \xi_i \\ & \quad (\omega \cdot \phi(t_i) + b) - y_i \leq \epsilon_i + \xi_i^* \\ & \quad \xi_i, \xi_i^* \geq 0 \\ & \quad i = 1, 2, \dots, n. \end{aligned} \quad (12)$$

This problem has the same form as that of standard SVR except for the adaptive width of the error-insensitive tube. Because ϵ_i is precomputed, the dual problem of (12) has the same solving algorithm and computational complexity as does the quadratic programming problem (4). In each local region, a fixed number of points are excluded from the error-insensitive tube. Therefore, the SVs supporting the regression function of ASVR would be more or less evenly distributed during the entire time series. This ensures that the local trend features of the volatile time series will not be omitted. The schematic diagram of ASVR is shown in Figure 2.

FIGURE 3: Regression results of the vibration RMS of a descaling pump: (a) time series of the vibration RMS; (b) regression curve of standard SVR; (c) regression curve of LSVR; (d) regression curve of ASVR.

When the regression function of ASVR is extrapolated, the current local region will exert a greater influence than the distant regions by adjusting the error-insensitive tube. This is favorable to improve the prediction accuracy for a volatile time series.

4. Experimental Verification

To demonstrate the effects of the proposed method on the state prognosis of mechanical systems, vibration signals are collected from centrifugal water-descaling pumps and processed. These descaling pumps, the function of which is to generate high-pressure and high-rate water flows to wipe away the oxide scale on a steel surface, are employed in a stainless steel plant and have an important influence on the surface quality of production. Due to working continuously under heavy loads, the bearings in water-descaling pumps are frequently damaged [17]. To monitor the working state of the bearings, high-precision velocity sensors are mounted on the input end and output end of the descaling pumps to measure the vibration velocity signals. According to the intensity of the vibrational response, the measurement ranges of the input-end sensor and the output-end sensor are set to 0–20 mm/s

and 0–50 mm/s, respectively. In this experimental research, the root mean square (RMS) is calculated from the vibrational signals and recorded at intervals of one hour to form the time series of state indicators. Because the behavior of the output-end vibration is more volatile than that of the input-end vibration, the RMS series monitored from the output end of the descaling pumps are chosen for the case studies.

4.1. Performance Analysis of the Regression. Figure 3(a) depicts a time series of vibration RMS values acquired from the output end of a descaling pump. It is composed of 265 data points, which are used to indicate the evolution process of the working state from normal to failure. Before approximately 80 h, the descaling pump runs in the later period of the normal stage, and the distribution region of the RMS is narrow and stationary. Thereafter, the working state continues to deteriorate until the descaling pump is broken down by bearing damage. In this stage, the intensity of the vibration rapidly increases to a high level, and the RMS drastically changes in a wide range. Standard SVR, LSVR, and ASVR are used to model the trend of the volatile RMS series. For comparison, similar parameters are chosen and listed in Table 1. Regression curves solved by these three learning

TABLE 1: Parameters setup for standard SVR, LSVR, and ASVR.

Parameters	C	σ	ϵ	Length of localized region	Number of truncated data points
Standard SVR	100	10	3	/	/
LSVR	100	10	3	$2k + 1 = 11$	/
ASVR	100	10	/	$l = 11$	$r = 1$

FIGURE 4: Prediction for the state of descaling pump number 1: (a) prediction results of standard SVR; (b) prediction results of ASVR.

machines, respectively, are presented in Figures 3(b)–3(d), and the margins of each error-insensitive tube are also drawn.

Figure 3(b) shows the regression results of standard SVR. To obtain global optimization, a compromise is reached between the data points in the normal stage and those in the deterioration stage. As a result, most of the regression values in the normal stage are greater than the actual values, and the drastic fluctuations in the deterioration stage are excessively smoothed. For LSVR, the regression curve, as shown in Figure 3(c), exhibits a rough trend. Although the regression precision of LSVR may be superior to that of the other regression machines, many unnecessary local details are contained in the regression function. Therefore, it is too complicated for modeling the state evolution of the mechanical system. Figure 3(d) depicts the adaptive margins of the error-insensitive tube, that is, the truncated range, and the regression curve obtained by ASVR. In the normal stage, high precision of the regression is achieved with the help of the narrow error-insensitive tube. When the state enters the deterioration stage, the error-insensitive tube is expanded to adapt the high-volatility RMS. ASVR captures the local trend features of the volatile RMS series well and provides a more practical solution than LSVR for trending the state indicators of mechanical systems.

In our research, the algorithm of standard SVR is performed by a Matlab toolbox, SVM-KMToolbox [18]. The ASVR algorithm is written based on the toolbox as well. According to [14], the SOCP problem in LSVR is solved using the software package CVX [19]. The regression results shown in Figures 3(b)–3(d) are obtained by running these algorithms on a PC with a 3 GHz Intel core processor and 2 GB of RAM. The average computational times are 0.188,

9.360, and 0.189 s for standard SVR, LSVR, and ASVR, respectively. The algorithm of ASVR has approximately the same computational speed as standard SVR and is suitable for dealing with the monitored state indicators.

4.2. Performance Analysis of the Prediction. The RMS series indicating two other state evolution processes of the descaling pumps are applied to evaluate the prediction performance of our method. Due to the disadvantage of LSVR for multistep extrapolation, only standard SVR is used for comparison. The parameters listed in Table 1 are still chosen for the algorithms of standard SVR and ASVR. In general, the longer the prediction step is, the greater the prediction error is. In this case study, the prediction step is set as 5. This means that the time series of RMS are separated into two parts by the time point of five hours before breakdown. The previous part is used to train the regression model, and the later part is used to examine the prediction results. To quantify the prediction accuracy, two criteria, including the root mean square error (RMSE) and the mean absolute percentage error (MAPE), are introduced:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=n-4}^n [y_i - f(t_i)]^2}{5}}, \quad (13)$$

$$\text{MAPE} = \frac{1}{5} \sum_{i=n-4}^n \left| \frac{y_i - f(t_i)}{y_i} \right| \times 100\%.$$

The prediction results for these two descaling pumps are shown in Figures 4 and 5, respectively. It can be observed that the predicted values of ASVR conform better to the actual values than those of standard SVR, even though

TABLE 2: Prediction errors of standard SVR and ASVR.

Error criteria	Descaling pump number 1		Descaling pump number 2	
	RMSE	MAPE	RMSE	MAPE
Standard SVR	3.23	14.76%	4.12	10.44%
ASVR	2.87	11.31%	3.57	8.75%

FIGURE 5: Prediction for the state of descaling pump number 2: (a) prediction results of standard SVR; (b) prediction results of ASVR.

the RMS series in the deterioration stage have dissimilar volatile behaviors. The error criteria are calculated and listed in Table 2. ASVR obtains better prediction performance than standard SVR. For example, the prediction MAPE on descaling pump number 2 utilizing standard SVR is 10.44%, whereas the prediction MAPE utilizing ASVR is only 8.75%. In conclusion, ASVR provides a capability for predicting the state trend of mechanical systems with volatile time series of state indicators. The results of the proposed regression machine are superior to those of standard SVR.

5. Conclusion and Discussion

It is common for a deteriorating mechanical system to generate volatile time series of state indicators. Due to the fixed error-insensitive tube, the traditional support vector regression is ill-suited to modelling a nonstationary state trend. In this paper, an adaptive support vector regression machine is proposed to capture the local trend of volatile state indicators and to predict the deterioration behavior of mechanical systems. Compared with traditional SVR, ASVR has the significant characteristic that the error-insensitive tube is adaptively adjusted according to the transient distribution boundary of local regions in the training time series. Considering the nonnormality of the transient distribution, a truncated range is used to define the distribution scope of localized regions and calculate the width of the error-insensitive tube. The experimental results demonstrate that ASVR has the same computational efficiency as standard SVR and provides a more practical solution than LSVR for trending state indicators. Moreover, the prediction accuracy of ASVR for volatile state indicators is higher than those for standard SVR.

However, the algorithm of ASVR on the entire training time series will be implemented again while new state indicators are obtained. For the data flow acquired from mechanical systems in long-term operation, the calculation is too huge to satisfy the trend analysis online. Thus, an incremental algorithm for ASVR is required to meet the further demand of state prognosis. Therefore, our research work will focus on this topic next.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was supported by 973 Program of China under Grant no. 613237.

References

- [1] A. Heng, S. Zhang, A. C. C. Tan, and J. Mathew, "Rotating machinery prognostics: state of the art, challenges and opportunities," *Mechanical Systems and Signal Processing*, vol. 23, no. 3, pp. 724–739, 2009.
- [2] N. Gebraeel, M. Lawley, R. Liu, and V. Parmeshwaran, "Residual life predictions from vibration-based degradation signals: a neural network approach," *IEEE Transactions on Industrial Electronics*, vol. 51, no. 3, pp. 694–700, 2004.
- [3] P. W. Tse and D. P. Atherton, "Prediction of machine deterioration using vibration based fault trends and recurrent neural networks," *Transactions of the ASME—Journal of Vibration and Acoustics*, vol. 121, no. 3, pp. 355–362, 1999.

- [4] N. Gebraeel, "Sensory-updated residual life distributions for components with exponential degradation patterns," *IEEE Transactions on Automation Science and Engineering*, vol. 3, no. 4, pp. 382–393, 2006.
- [5] M. Guida and G. Pulcini, "The inverse Gamma process: a family of continuous stochastic models for describing state-dependent deterioration phenomena," *Reliability Engineering and System Safety*, vol. 120, pp. 72–79, 2013.
- [6] X.-S. Si, W. Wang, C.-H. Hu, and D.-H. Zhou, "Remaining useful life estimation—a review on the statistical data driven approaches," *European Journal of Operational Research*, vol. 213, no. 1, pp. 1–14, 2011.
- [7] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, no. 3, pp. 199–222, 2004.
- [8] C. Hua, Q. Zhang, G. Xu, Y. Zhang, and T. Xu, "Performance reliability estimation method based on adaptive failure threshold," *Mechanical Systems and Signal Processing*, vol. 36, no. 2, pp. 505–519, 2013.
- [9] T. Benkedjouh, K. Medjaher, N. Zerhouni, and S. Rechak, "Remaining useful life estimation based on nonlinear feature reduction and support vector regression," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 7, pp. 1751–1760, 2013.
- [10] J. Qu and M. J. Zuo, "An LSSVR-based algorithm for online system condition prognostics," *Expert Systems with Applications*, vol. 39, no. 5, pp. 6089–6102, 2012.
- [11] V. T. Tran, H. Thom Pham, B.-S. Yang, and T. Tien Nguyen, "Machine performance degradation assessment and remaining useful life prediction using proportional hazard model and support vector machine," *Mechanical Systems and Signal Processing*, vol. 32, pp. 320–330, 2012.
- [12] C. Shen, D. Wang, Y. Liu, F. Kong, and P. W. Tse, "Recognition of rolling bearing fault patterns and sizes based on two-layer support vector regression machines," *Smart Structures and Systems*, vol. 13, no. 3, pp. 453–471, 2014.
- [13] W. Wang, "A two-stage prognosis model in condition based maintenance," *European Journal of Operational Research*, vol. 182, no. 3, pp. 1177–1187, 2007.
- [14] H. Yang, K. Huang, I. King, and M. R. Lyu, "Localized support vector regression for time series prediction," *Neurocomputing*, vol. 72, no. 10–12, pp. 2659–2669, 2009.
- [15] J.-L. Wu and P.-C. Chang, "A trend-based segmentation method and the support vector regression for financial time series forecasting," *Mathematical Problems in Engineering*, vol. 2012, Article ID 615152, 20 pages, 2012.
- [16] R. A. Maronna, R. D. Martin, and V. J. Yohai, *Robust Statistics: Theory and Methods*, Wiley Series in Probability and Statistics, John Wiley & Sons, Chichester, UK, 2006.
- [17] Q. Zhang, C. Hua, and G. Xu, "A mixture Weibull proportional hazard model for mechanical system failure prediction utilising lifetime and monitoring data," *Mechanical Systems and Signal Processing*, vol. 43, no. 1–2, pp. 103–112, 2014.
- [18] S. Canu, Y. Grandvalet, V. Guigue, and A. Rakotomamonjy, "SVMand kernel methods Matlab toolbox," in *Perception Systems et Information*, INSA de Rouen, Rouen, France, 2005.
- [19] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.0 beta," 2013, <http://cvxr.com/cvx/>.