A Model-Based Method for Remaining Useful Life Prediction of Machinery

Yaguo Lei, *Member, IEEE*, Naipeng Li, Szymon Gontarz, Jing Lin, Stanislaw Radkowski, and Jacek Dybala

***Abstract*—Remaining useful life (RUL) prediction allows for predictive maintenance of machinery, thus reducing costly unscheduled maintenance. Therefore, RUL prediction of machinery appears to be a hot issue attracting more and more attention as well as being of great challenge. This paper proposes a model-based method for predicting RUL of machinery. The method includes two modules, i.e., indicator construction and RUL prediction. In the first module, a new health indicator named weighted minimum quantization error is constructed, which fuses mutual information from multiple features and properly correlates to the degradation processes of machinery. In the second module, model parameters are initialized using the maximum-likelihood estimation algorithm and RUL is predicted using a particle filtering-based algorithm. The proposed method is demonstrated using vibration signals from accelerated degradation tests of rolling element bearings. The prediction result identifies the effectiveness of the proposed method in predicting RUL of machinery.**

***Index Terms*—Health indicator, parameter initialization, particle filtering, remaining useful life (RUL) prediction.**

NOMENCLATURE

BMU	Best matching unit.
MLE	Maximum-likelihood estimation.
MQE	Minimum quantization error.
PDF	Probability density function.
PF	Particle filtering.
RUL	Remaining useful life.
SOM	Self-organizing map.
SIS	Sequential importance sampling.
SD of IHC	Standard deviation of inverse hyperbolic cosine.
SD of IHS	Standard deviation of inverse hyperbolic sine.
WMQE	Weighted minimum quantization error.
$f(\cdot)$, $h(\cdot)$	Transition function and measurement function.
x_k, z_k	System state and measured data at time t_k.
ω_k, v_k	Process noise and measurement noise at time t_k.
$p(x_k \| z_{1:k-1})$, $p(x_k \| z_{1:k})$	Prior state PDF and posterior state PDF at time t_k.
$q(x_k \| x_{k-1}, z_{1:k})$	Importance PDF.
Ns	Number of particles.
x_k^i, w_k^i	ith particle and its weight at time t_k.
N_{eff}, N_T	Effective sample size and its threshold for resampling.
$\{y_k^i\}_{k=1:K}$, $\{Y_k^i\}_{1:K}$	ith feature sequence with a length of K and the rank of the feature sequence.
$\{t_k\}_{k=1:K}$, $\{T_k\}_{1:K}$	Time sequence with a length of K and the rank of the time sequence.
ρ^i	Trendability of the ith feature.
H	Number of clusters.
$u_{hc,k}$	Membership degree of the kth feature to the hth cluster.
$R_{i,j}$	Correlation coefficient of the ith feature and the jth feature.
PBM(H)	PBM index of the clustering result when the cluster number equals H.
s_k	WMQE value at time t_k.
$\Theta = (\mu_\alpha, \sigma_\alpha^2, \beta, \sigma_\nu^2)'$	Unknown model parameters.
$\boldsymbol{S}_{0:M}$	Set of the WMQE values up to t_M with $\mathbf{S}_{0:M} = (s_0, \ldots, s_M)'$.
$\ell(\cdot)$	Log-likelihood function.
l_k	RUL at time t_k.
λ	Prespecified failure threshold for RUL prediction.
Er_i	Percent error of the RUL prediction result for the ith testing dataset.

Manuscript received July 12, 2015; revised November 3, 2015, January 22, 2016, and March 21, 2016; accepted April 10, 2016. Date of publication June 29, 2016; date of current version August 30, 2016. This work was supported by the National Natural Science Foundation of China under Grants 51475355 and 51421004 Young Talent Support Plan of Central Organization Department, and Fundamental Research Funds for the Central Universities under Grant 2012jdgz01 and CXTD2014001. Associate Editor: M. J. Zuo.

Y. G. Lei, N. P. Li, and J. Lin are with the State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: yaguolei@mail.xjtu.edu.cn; li3112001096@stu.xjtu.edu.cn; jinglin@mail.xjtu.edu.cn).

S. Gontarz, S. Radkowski, and J. Dybala are with the Institute of Automotive Engineering, Warsaw University of Technology, 02-524 Warsaw, Poland (e-mail: sgontarz@simr.pw.edu.pl; ras@simr.pw.edu.pl; jdybala@simr.pw.edu.pl).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TR.2016.2570568

I. INTRODUCTION

THE last few decades have witnessed an increasing interest in health monitoring of machinery, such as rolling element bearings, gears and rotors [1]–[4]. Machinery health monitoring includes three stages, i.e., fault detection, fault diagnosis, and remaining useful life (RUL) prediction [5]. Fault detection and diagnosis are processes of determining fault presence in machinery as early as possible and identifying kinds, locations, and degrees of faults [2]. RUL prediction is a process using prediction methods to forecast the future performance of machinery and obtain the time left before machinery loses its operation ability [5]. Most work reported in literature focuses on the first two stages. However, the last stage, i.e., RUL prediction, has not been well developed to date.

As a fault is detected or diagnosed, it is common to shut down the machinery as soon as possible to avoid catastrophic consequences. Performing such an action, which usually occurs at inconvenient times, typically causes substantial time and economic losses. It is, therefore, essential to schedule the maintenance strategy in a predictive rather than diagnostic manner. Accurate RUL prediction of machinery allows for effective predictive maintenance, thus reducing costly unscheduled maintenance on machinery [6]–[8]. Consequently, RUL prediction appears to be a hot issue attracting more and more attention in recent years.

The RUL prediction methods could be roughly classified into data-driven methods and model-based methods [9]. Data-driven methods attempt to derive the degradation process of a machine from measured data using machine learning techniques [10]. These methods rely on the assumption that the statistical characteristics of data are relatively consistent unless a fault occurs. They produce RUL prediction results based on history data measured from machinery. Therefore, the prediction accuracy of data-driven methods depends on not only the quantity but also the quality of the history data [9]. However, it is generally difficult to collect enough qualified data in real cases.

On the contrary, model-based methods attempt to set up mathematical or physical models to describe degradation processes of machinery, and update model parameters using measured data [11], [12]. The commonly used models include the Markov process model [13], [14], the Winner process model [15], [16], the Gaussian mixture model [17], [18], etc. Model-based methods could incorporate both expert knowledge and real-time information from machinery. Consequently, they may work well in RUL prediction of machinery. However, there still exist some problems in model-based methods.

One of the problems is how to construct a health indicator, which is suitable for RUL prediction. A suitable health indicator is supposed to reflect the degradation of machinery and have an obvious monotonic trend [19], which may simplify modeling and produce accurate prediction results. Many studies have been conducted on constructing such health indicators. Camci et al. extracted time-domain indicators from vibration signals of rolling element bearings and evaluated the suitability of the indicators using a monotonicity metric [19]. Li and He extracted acoustic emission indicators based on empirical mode decomposition and used them to monitor health condition of gearboxes [20]. Qian et al. applied recurrence quantification analysis to extracting recurrence plot entropy indicators and used them to predict bearing RUL [21]. The aforesaid studies try to use original features that are directly extracted from measured data as the health indicators. Actually, each original feature is only sensitive to a certain fault in a certain degradation stage [22]. A suitable health indicator should take advantage of mutual information from multiple features for degradation assessment and RUL prediction. Therefore, researchers have tried to fuse multiple features and construct a fusion indicator. Qiu et al. used a self-organizing map (SOM) to fuse time-domain features into a health indicator named minimum quantization error (MQE) [22]. After that, Huang et al. applied MQE into RUL prediction of rolling element bearings [23]. Yu used dynamic principal component analysis to extract principal components from multiple features and constructed a hidden Markov model-based Mahalanobis distance as a health indicator [24]. These fusion indicators may incorporate mutual information from multiple features. However, some original features may be not informative for the degradation processes or even have a negative influence on the fusion result. Therefore, it is necessary to evaluate the original features and select typical ones before fusion.

Another problem in model-based methods is to estimate model parameters according to the real-time information from measured data. Particle filtering (PF) is suitable to deal with this problem because of its capability in combining information from real-time data and expert knowledge from empirical models. PF is derived from traditional filtering algorithms, such as Kalman filtering and Extended Kalman filtering [25]. It is based on the Bayesian theory that attempts to evaluate the state of a system based on measurements. A sequential importance sampling (SIS) algorithm is developed and applied in PF [26]. Based on the Bayesian theory and SIS, PF is particularly effective in estimating states and model parameters of degradation models according to real-time information. As a result, it has been introduced into the field of RUL prediction [27]. An et al. presented a Matlab-based tutorial for PF-based prognostic algorithm [28]. Chen et al. combined the adaptive neuro-fuzzy inference system with the high-order PF and used this method to predict the RUL of planetary gear plates [29]. Corbetta et al. described the fatigue crack propagation using a stochastic dynamic state space model and updated the model parameters using PF [30]. Orchard et al. presented a PF-based real-time prognosis algorithm for the estimation of Li-ion battery discharge time, as well as a risk measure that includes information about the risk of battery failure and a measure of confidence for the prognostic algorithm itself [31]. These methods have successfully utilized the PF algorithm to estimate health states and update model parameters according to real-time information. However, the model parameters should be initialized before updating. The above research either sets the initial parameters based on empirical knowledge or predetermines the initial parameters arbitrarily, which may influence the performance of the PF-based prediction algorithm.

In order to deal with the aforementioned problems, this paper proposes a model-based method for predicting RUL of machinery. The method is composed of two modules, i.e., indicator construction and RUL prediction. In the first module, a new health indicator named weighted minimum quantization error (WMQE) is constructed, which fuses mutual information from multiple features and properly correlates to the degradation processes of machinery. This indicator is developed from the MQE proposed by Qiu et al. in [22]. The major improvement of WMQE compared with MQE is that original features are selected and weighted before fusion, which can further enhance the superiority of the fusion health indicator for RUL prediction. In the second module, a commonly used degradation model, i.e., the Paris–Erdogan model [32], is employed to describe the degradation process of machinery, and the RUL is predicted using a PF-based algorithm [28]. Our major contribution in the RUL prediction module is that, model parameters are initialized using the maximum likelihood estimation (MLE) algorithm instead of prespecified artificially. The initialization process is conducted based on the available measurements of specified machinery. Therefore, it can adjust the model parameters adaptively according to the degradation trend of the specified machinery.

The rest of the paper is organized as follows. Section II introduces the basic theory of PF. Section III shows the proposed prediction method. In Section IV, the method is demonstrated using data of accelerated degradation tests on rolling element bearings. Conclusion is drawn in Section V.

II. Basic Theory of PF

In PF, the state of a system is described using the following transition function [25]:

$$x_k = f\left(x_{k-1}, \omega_k\right) \tag{1}$$

where x_{k-1} is the state of the system at t_{k-1}, and ω_k is the independent identically distributed (i.i.d.) process noise.

In most of the time, the state of the system cannot be observed directly. But we may estimate it based on the relationship between the state and measured data. The relationship between them at t_k is as follows:

$$z_k = h\left(x_k, \nu_k\right) \tag{2}$$

where z_k is the measured data of the system at t_k, $h(\cdot)$ is the measurement function, and ν_k is the i.i.d. measurement noise.

Based on (1) and (2), the transition probability density function (PDF) $p(x_k|x_{k-1})$ and the state PDF $p(x_{k-1}|z_{1:k-1})$ at t_{k-1} is obtained, respectively. Then, the prior state PDF at t_k is predicted using the Chapman–Kolmogorov equation:

$$p\left(x_k|z_{1:k-1}\right) = \int p\left(x_k|x_{k-1}\right) p\left(x_{k-1}|z_{1:k-1}\right) \mathrm{d}x_{k-1}. \tag{3}$$

When new measured data are available, the prior state PDF is updated based on the Bayesian rule to generate the posterior state PDF following:

$$p\left(x_k|z_{1:k}\right) = \frac{p\left(z_k|x_k\right) p\left(x_k|z_{1:k-1}\right)}{p\left(z_k|z_{1:k-1}\right)}. \tag{4}$$

The recursive propagation of the posterior state PDF is only a conceptual solution, which cannot be obtained analytically, because it requires the evaluation of complex high-dimensional integrals. Therefore, a SIS algorithm is employed to approximate the posterior state PDF numerically. SIS approximates the state PDF using a set of particles $\{x_k^i\}_{i=1:Ns}$ sampled from an importance PDF $q(x_k|x_{k-1}, z_{1:k})$, where Ns is the number of particles. Using SIS, the state of the system is approximated using the following discrete density:

$$p\left(x_k|z_{1:k}\right) \approx \sum_{i=1}^{Ns} w_k^i \delta\left(x_k - x_k^i\right) \tag{5}$$

where w_k^i is the weight of x_k^i, which is updated and normalized following:

$$w_k^i = w_{k-1}^i \frac{p\left(z_k|x_k^i\right) p\left(x_k^i|x_{k-1}^i\right)}{q\left(x_k^i|x_{k-1}^i, z_{1:k}\right)}, w_k^i = w_k^i \Big/ \sum_{i=1}^{Ns} w_k^i. \tag{6}$$

When the importance PDF $q(x_k|x_{k-1}, z_{1:k})$ is chosen as $p(x_k|x_{k-1})$, (6) is simplified as follows:

$$w_k^i = w_{k-1}^i p\left(z_k|x_k^i\right), w_k^i = w_k^i \Big/ \sum_{i=1}^{Ns} w_k^i. \tag{7}$$

One problem in the SIS algorithm is the particle degeneracy, i.e., after a few iterations, all but one of the particles have negligible weights. To deal with this problem, a resampling step is utilized. The basic idea of resampling is to eliminate particles with small weights and to concentrate upon particles with large weights. A suitable measure of the degeneracy of particles is the effective sample size:

$$N_{\text{eff}} = 1 \Big/ \sum_{i=1}^{Ns} \left(w_k^i\right)^2. \tag{8}$$

When N_{eff} falls below a threshold N_T, a new particle set $\{x_k^{j*}\}_{j=1:N_s}$ will be generated by resampling Ns times following the principle of $p(x_k^{j*} = x_k^i) = w_k^i$, and the weights are reset to $w_k^{j*} = 1/Ns$.

III. Proposed Prediction Method

The model-based method proposed in this paper is shown in Fig. 1. This method is composed of two modules, i.e., indicator construction and RUL prediction.

In the indicator construction module, original features are extracted from vibration signals of machinery. Then, the degradation trends of these features are evaluated and some features are selected from them. The selected features are further clustered using a correlation clustering algorithm and typical features of the clusters are utilized to construct the health indicator WMQE. In the RUL prediction module, the WMQE is input into the state space model and model parameters are initialized using the MLE algorithm. Then health states and model parameters are updated adaptively using the PF algorithm. Finally, the degradation trend in the future and the RUL of the machinery are predicted based on the estimated states and parameters. More details about these two modules are presented as follows.

A. Indicator Construction

In this module, original features are extracted from vibration signals of machinery. Since machinery always suffers from an irreversible degradation process, an original feature sensitive to the degradation process is supposed to have a monotonic degradation trend. To evaluate the monotonicity of the original features, a trendability measure of each feature is calculated, which is defined as the Spearman coefficient of the feature with

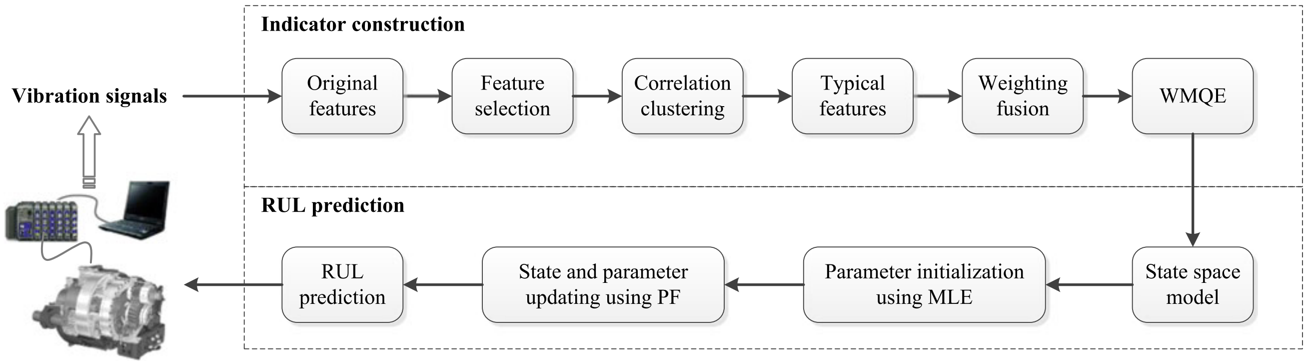

Fig. 1. Flowchart of the proposed method.

TABLE I
PROCESS OF THE CORRELATION CLUSTERING ALGORITHM

Step 1: Calculate the absolute correlation matrix of the selected features and let $H = 1$.
Step 2: Find the first center and the second center which have the lowest correlation coefficient among these features.
Step 3: Find the next center which has the lowest mean correlation coefficient with the determined centers.
Step 4: Repeat Step 3 until all of the H centers are determined.
Step 5: Classify each of the remaining features into one cluster whose center has the highest correlation coefficient with this feature.
Step 6: Set the membership degree $u_{hc,k}$ according to the clustering result, i.e., $u_{hc,k} = 1$ if the kth feature belongs to the hth cluster, otherwise $u_{hc,k} = 0$. Calculate the value of $PBM(H)$ using (10).
Step 7: Let $H = H + 1$ and repeat Step 3 to Step 6 until H equals the number of the selected features.
Step 8: Report the cluster number H which has the maximum PHM value and output the clustering result.

time [33]

$$\rho^i = \frac{\left|\sum_{k=1}^{K}\left(T_k - \bar{T}_k\right)\left(Y_k^i - \bar{Y}_k^i\right)\right|}{\sqrt{\sum_{k=1}^{K}\left(T_k - \bar{T}_k\right)^2 \sum_{k=1}^{K}\left(Y_k^i - \bar{Y}_k^i\right)^2}} \tag{9}$$

where $\{T_k\}_{k=1:K}$ and $\{Y_k^i\}_{k=1:K}$ are the ranks of the time $\{t_k\}_{k=1:K}$ and the ith feature $\{y_k^i\}_{k=1:K}$, respectively. K is the length of the time sequence. $\bar{T}_k$ and $\bar{Y}_k^i$ are the means of $\{T_k\}_{1:K}$ and $\{Y_k^i\}_{1:K}$, respectively. When a feature monotonously increases or decreases with time, the trendability value is 1. In contrast, if a feature is constant or varies randomly with time, the value equals zero. So a higher trendability means a better monotonicity.

The Spearman coefficient is generally used to assess how well the relationship between two variables can be described using a monotonic function. It is derived from the Pearson coefficient, which is also commonly used in the relationship assessment of two variables. The Pearson coefficient has already been used to measure the trendability of features by Javed, Gouriveau, Zerhouni, and Nectoux. But the Pearson coefficient is only suitable to measure the linear degradation trend of a feature, whereas machinery generally performs nonlinear degradation trends in reality. Different from the Pearson coefficient, the Spearman coefficient evaluates the trendability of a feature using the rank of the feature instead of the feature itself. Therefore, it is able to evaluate the nonlinear relationship and is more suitable to measure the monotonicity of the features.

The trendability value calculated by (9) belongs to the interval $[0, 1]$. In order to select the original features which have obvious monotonic trends, the features whose trendability values are higher than $1/2$ are selected. Some selected original features possibly have redundant information that represents the similar degradation trends of machinery. To identify the redundant information, a correlation clustering algorithm is established to classify these features. This algorithm measures the similarity of these selected features and classifies them into different clusters according to their correlation matrix. The process of the correlation clustering algorithm is shown in Table I. In this algorithm, a PBM index, which is commonly used for verifying the cluster validity [35], is employed to determine an appropriate number of clusters. The PBM index is defined as follows:

$$\mathrm{PBM}(H) = \left(\frac{1}{H} \cdot \frac{E_H}{E_1} \cdot D_H\right)^2 \tag{10}$$

where H is the number of clusters,

$$E_H = \sum_{h=1}^{H}\left(\sum_{k=1}^{K} u_{hc,k} \cdot R_{hc,k}\right),\ D_H = \max_{i,j=1}^{H} R_{ic,jc},\ u_{hc,k}$$

is the membership degree of the kth feature belonging to the hth cluster, $R_{hc,k}$ is the correlation coefficient of the hth center and the kth feature, and $R_{ic,jc}$ is correlation coefficient of the ith center and the jth center.

After the correlation clustering process, all of the selected features are classified into H different clusters. Features in the same cluster are assumed to have the similar information for the degradation process. In order to reduce the redundancy of the features, one typical feature which has the highest trendability is selected from each cluster. As a result, H typical features are selected.

The last step is to fuse the H typical features using a SOM-based method [22]. At first, The SOM is trained by the feature vectors in the normal operation stage. Then, the feature vector under an unidentified condition is compared with the weight vector of its best matching unit (BMU) in the SOM. From the point of view of the degradation monitoring, the distance between the BMU and input feature vector indicates how far the unidentified condition deviates from the normal operation stage. The indicator MQE [22] is calculated depending on the distance between the BMU and input feature vector. One major concern during the fusion process is that different features generally present distinct sensitivities to the degradation process of machinery. The features which are more sensitive to the degradation process are supposed to make more contribution to the fusion process. Therefore, a WMQE indicator is proposed and denoted as follows:

$$s_k = \sum_{h=1}^{H} \left(\tilde{\rho}^h \left| y_k^h - m_{\mathrm{BMU}}^h \right| \right) \tag{11}$$

where s_k is the WMQE value at t_k, $\tilde{\rho}^h$ is the normalized trendability of the hth typical feature with $\tilde{\rho}^h = \rho^h / \sum_{h=1}^{H} \rho^h$, which measures the sensitivity of the feature to the degradation process, y_k^h is the value of the hth typical feature at t_k, and m_{BMU}^h is the hth weight of the BMU.

B. RUL Prediction

In this module, a PF-based prediction algorithm is utilized to predict the RUL of machinery with degradation processes described using a variant of Paris-Erdogan model. Paris–Erdogan model was first proposed by Paris and Erdogen in [32] to explain the propagation of micro fatigue cracks in the mechanical material. Since most degradation processes of machinery are directly or indirectly caused by the micro fatigue cracks, Paris–Erdogan model is effective in describing the degradation process of machinery. So far, many variants have been developed from the Paris–Erdogan model and utilized to the RUL prediction of machinery [28]. The original formula of the Paris–Erdogan model is

$$\frac{dx}{dn} = c(\Delta K)^{\gamma},\ \Delta K = \varepsilon \sqrt{x} \tag{12}$$

where x represents the semicrack length, n is the number of stress cycles (i.e., the fatigue life), c, γ, and ε are material constants that are determined by tests, and ΔK is the amplitude of stress intensity factor roughly proportional to the square root of x.

It is seen from (12) that several model parameters should be predetermined by tests, i.e., c, γ, and m. What is more, the semicrack length x is difficult to be measured during the operation process of machinery. For convenient application, the Paris–Erdogan model is transformed into the following formula with $\alpha = c\varepsilon^{\gamma}$ and $\beta = \gamma/2$:

$$\frac{\mathrm{d}x}{\mathrm{d}n} = \alpha x^{\beta}. \tag{13}$$

With the aforesaid function rewritten in the form of (1) and (2), the state space model of machinery is obtained as follows:

$$\begin{cases} x_k = x_{k-1} + \alpha_{k-1} x_{k-1}^{\beta} \Delta t_k \\ \alpha_k = \alpha_{k-1} \\ s_k = x_k + \nu_k \end{cases} \tag{14}$$

where α_{k-1} is a random variable following a normal distribution of $N(\mu_\alpha, \sigma_\alpha^2)$, β is a constant parameter, and $\Delta t_k = t_k - t_{k-1}$.$x_k$ is the health state at t_k, which is hard to be acquired in real applications and is generally represented by a health indicator [28], [29], [34]. Here, it is represented by the WMQE value. Considering the influence of the measurement noise, the measured WMQE value s_k is denoted as x_k plus a measurement noise $\nu_k \sim N(0, \sigma_\nu^2)$.

After acquiring the state space model, the WMQE values are input into the model, and model parameters are initialized using MLE. The unknown model parameters are denoted as $\Theta = (\mu_\alpha, \sigma_\alpha^2, \beta, \sigma_\nu^2)'$, where $(\cdot)'$ denotes the vector transposition. It is assumed that there are a series of measurements $\boldsymbol{S}_{0:M} = (s_0, \ldots, s_M)'$ at ordered times $t_0, \ldots, t_M$. According to (14), s_k can be expressed as follows:

$$s_k = x_{k-1} + \alpha_k x_{k-1}^{\beta} \Delta t_k + \nu_k. \tag{15}$$

The health sate x_{k-1} has the following relationship with the measurement s_{k-1}:

$$x_{k-1} = s_{k-1} - \nu_{k-1}. \tag{16}$$

The measurement noise ν_{k-1} is generally small enough to be ignored compared with the measurement itself [28]. Therefore, x_{k-1} is able to be approximated by s_{k-1}. Let $\Phi = (s_0^{\beta} \Delta t_1, \ldots, s_{M-1}^{\beta} \Delta t_M)'$. It is assumed that $\boldsymbol{S}_{1:M} = (s_1, \ldots, s_M)'$ is multivariate normally distributed, which is denoted as follows:

$$\boldsymbol{S}_{1:M} \sim N(\boldsymbol{S}_{0:M-1} + \mu_\alpha \Phi, \sigma_\alpha^2 \Phi\Phi' + \sigma_\nu^2 I_M) \tag{17}$$

where I_M is an identity matrix of order M.

Let $\Delta \boldsymbol{S}_{1:M} = (s_1 - s_0, \ldots, s_M - s_{M-1})'$, then the log-likelihood function of the unknown parameters based on the measurements is expressed as

$$\begin{aligned} \ell(\Theta | \boldsymbol{S}_{0:M}) = & -\frac{M}{2} \ln(2\pi) - \frac{1}{2} \ln \left| \sigma_\alpha^2 \Phi\Phi' + \sigma_\nu^2 I_M \right| \\ & -\frac{1}{2} (\Delta \boldsymbol{S}_{1:M} - \mu_\alpha \Phi)' (\sigma_\alpha^2 \Phi\Phi' + \sigma_\nu^2 I_M)^{-1} \\ & (\Delta \boldsymbol{S}_{1:M} - \mu_\alpha \Phi) \\ = & -\frac{M}{2} \ln(2\pi) - \frac{M}{2} \ln \sigma_\alpha^2 - \frac{1}{2} \ln \left| \Phi\Phi' + \tilde{\sigma}_\nu^2 I_M \right| \\ & -\frac{1}{2\sigma_\alpha^2} (\Delta \boldsymbol{S}_{1:M} - \mu_\alpha \Phi)' (\Phi\Phi' + \tilde{\sigma}_\nu^2 I_M)^{-1} \\ & (\Delta \boldsymbol{S}_{1:M} - \mu_\alpha \Phi) \end{aligned} \tag{18}$$

with $\tilde{\sigma}_\nu^2 = \sigma_\nu^2/\sigma_\alpha^2$. Then, first partial derivatives of $\ell(\Theta|\boldsymbol{S}_{0:M})$ with respect to μ_α and σ_α^2 are expressed as follows:

$$\frac{\partial \ell(\Theta|\boldsymbol{S}_{0:M})}{\partial \mu_\alpha} = \frac{1}{\sigma_\alpha^2}\Phi'\left(\Phi\Phi' + \tilde{\sigma}_\nu^2 I_M\right)^{-1}(\Delta\boldsymbol{S}_{1:M} - \mu_\alpha\Phi), \tag{19}$$

$$\frac{\partial \ell(\Theta|\boldsymbol{S}_{0:M})}{\partial \sigma_\alpha^2} = -\frac{M}{2\sigma_\alpha^2} + \frac{1}{2\sigma_\alpha^4}(\Delta\boldsymbol{S}_{1:M} - \mu_\alpha\Phi)' \left(\Phi\Phi' + \tilde{\sigma}_\nu^2 I_M\right)^{-1}(\Delta\boldsymbol{S}_{1:M} - \mu_\alpha\Phi). \tag{20}$$

Let $\partial\ell(\Theta|\boldsymbol{S}_{0:M})/\partial\mu_\alpha = 0$ and $\partial\ell(\Theta|\boldsymbol{S}_{0:M})/\partial\sigma_\alpha^2 = 0$. The results of the MLE for μ_α and σ_α^2 are

$$\mu_\alpha = \frac{\Phi'\left(\Phi\Phi' + \tilde{\sigma}_\nu^2 I_M\right)^{-1}\Delta\boldsymbol{S}_{1:M}}{\Phi'(\Phi\Phi' + \tilde{\sigma}_\nu^2 I_M)^{-1}\Phi} \tag{21}$$

$$\sigma_\alpha^2 = \frac{(\Delta\boldsymbol{S}_{1:M} - \mu_\alpha\Phi)'\left(\Phi\Phi' + \tilde{\sigma}_\nu^2 I_M\right)^{-1}(\Delta\boldsymbol{S}_{1:M} - \mu_\alpha\Phi)}{M}. \tag{22}$$

With (21) and (22) substituted into (18), the log-likelihood function is reduced to a two-variable function about β and $\tilde{\sigma}_\nu^2$

$$\ell(\Theta|\boldsymbol{S}_{0:M}) = -\frac{M}{2}\ln(2\pi) - \frac{M}{2}\ln\sigma_\alpha^2 - \frac{1}{2}\ln\left|\Phi\Phi' + \tilde{\sigma}_\nu^2 I_M\right| - \frac{M}{2}. \tag{23}$$

The MLE values of β and $\tilde{\sigma}_\nu^2$ are obtained by maximizing the log-likelihood function (23) through two-dimensional optimizing. Then, the MLE values of β and $\tilde{\sigma}_\nu^2$ are substituted into (21) and (22), and the MLE values of μ_α and σ_α^2 are acquired. The value of σ_ν^2 is calculated with $\tilde{\sigma}_\nu^2$ multiplied by σ_α^2. By now, all of the unknown parameters $\Theta = (\mu_\alpha, \sigma_\alpha^2, \beta, \sigma_\nu^2)'$ are initialized.

After parameter initialization, the model parameters are further updated and the RUL is predicted using a PF-based prediction algorithm. Based on the initialized parameters, a series of initial particles $\{\mathbf{z}_0^i\}_{i=1:Ns}$ are sampled from $p(\mathbf{z}_0^i|\Theta_0)\sim N(\mathbf{z}_0, Q_0)$ with

$$\mathbf{z}_0 = \begin{bmatrix} x_0 \\ \mu_\alpha \end{bmatrix} \text{ and } Q_0 = \begin{bmatrix} 0 & 0 \\ 0 & \sigma_\alpha^2 \end{bmatrix}. \tag{24}$$

The number of particles is Ns and the weight of each particle is $1/Ns$. Then, new particles $\{\mathbf{z}_k^i\}_{i=1:Ns}$ are obtained as follows:

$$\mathbf{z}_k^i = \begin{bmatrix} x_k^i \\ \mu_\alpha^i \end{bmatrix} = \begin{bmatrix} x_{k-1}^i + \mu_\alpha^i\left(x_{k-1}^i\right)^\beta \Delta t_k \\ \mu_\alpha^i \end{bmatrix}. \tag{25}$$

When the new measurement s_k at t_k is available, each particle weight is updated and normalized by

$$w_k^i = w_{k-1}^i p\left(s_k|\mathbf{z}_k^i\right),\ w_k^i = w_k^i \Big/ \sum_{i=1}^{N_s} w_k^i \tag{26}$$

where

$$p\left(s_k|\mathbf{z}_k^i\right) = \frac{1}{\sqrt{2\pi}\sigma_\nu}\exp\left[-\frac{1}{2}\left(\frac{s_k - x_k^i}{\sigma_\nu}\right)^2\right]. \tag{27}$$

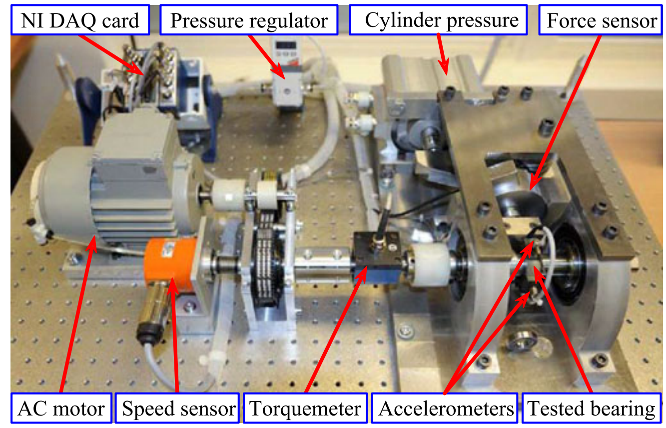

Fig. 2. Overview of the experimental system.

According to the particle weights, the particles are resampled and their weights are reset to be $1/Ns$.

After that, the RUL is predicted based on the resampled particles. The RUL l_k at t_k is defined as

$$l_k = \inf\{l_k : x(l_k + t_k) \geq \lambda|\boldsymbol{x}_{0:k}\} \tag{28}$$

where λ is a prespecified failure threshold, $\boldsymbol{x}_{0:k}$ is the estimated state values at $t_0, \ldots, t_k$, and $x(l_k + t_k)$ is the predicted state value at $l_k + t_k$. Since the estimated state value and the model parameter of the ith particle at t_k are x_k^i and μ_α^i, respectively, the future state value of the ith particle can be predicted using the following transition function:

$$x^i((j+1)\Delta l + t_k) = x^i(j\Delta l + t_k) + \mu_\alpha^i x^i(j\Delta l + t_k)^\beta \Delta l \tag{29}$$

where Δl is the time interval and $j \in \mathrm{N} = \{0, 1, 2, \ldots\}$. The transition process is stopped when the predicted state value exceeds the threshold. Then, the RUL l_k^i of the ith particle is acquired from (28). Finally, the PDF of the RUL is approximated by

$$p(l_k|\boldsymbol{s}_{0:k}) = \sum_{i=1}^{Ns} w_k^i \delta\left(l_k - l_k^i\right). \tag{30}$$

IV. Experimental Demonstration

In this section, the proposed method is demonstrated by predicting RUL of rolling element bearings. Accelerated degradation tests of rolling element bearings are conducted and vibration signals are acquired during the tests to verify the effectiveness of the method.

A. Introduction to the Experimental System and Vibration Data

The experimental system named PRONOSTIA is shown in Fig. 2 [37]. It is designed to test methods proposed for fault detection, diagnosis, and RUL prediction of rolling element bearings. This experimental system is able to conduct accelerated degradation tests on bearing in a few hours. Seventeen bearings are tested under three different operation conditions. The

TABLE II
OPERATION CONDITIONS OF TESTED BEARINGS

Operation conditions	Load (N)	Speed (rpm)	Training dataset	Testing dataset
Condition 1	4000	1800	Bearing 1_1 Bearing 1_2	Bearing 1_3 Bearing 1_4 Bearing 1_5 Bearing 1_6 Bearing 1_7
Condition 2	4200	1650	Bearing 2_1 Bearing 2_2	Bearing 2_3 Bearing 2_4 Bearing 2_5 Bearing 2_6 Bearing 2_7
Condition 3	5000	1500	Bearing 3_1 Bearing 3_2	Bearing 3_3

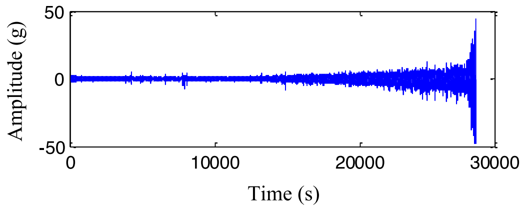

Fig. 3. Vibration signals of a tested bearing.

TABLE III
ORIGINAL FEATURES EXTRACTED FROM VIBRATION SIGNALS

Type	Feature	
Time-domain features	F1: Standard deviation F2: Peak-Peak F3: Mean-absolute F4: Root mean square F5: Shape factor	F6: Crest factor F7: Impulse factor F8: Skewness F9: Kurtosis F10: Entropy
Time-frequency-domain features	F11–F18: Energies of eight bands F19–F26: Energy ratios of eight bands	
Features based on trigonometric functions	F27: SD of IHC F28: SD of IHS	

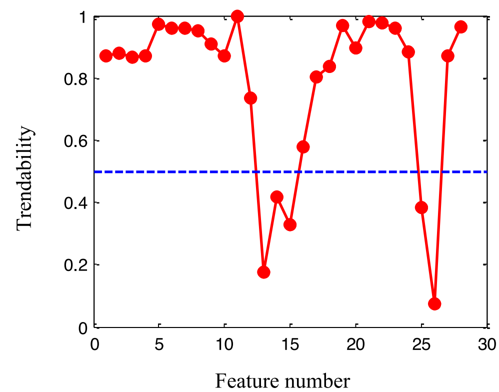

Fig. 4. Feature selection result of Bearing 1_1.

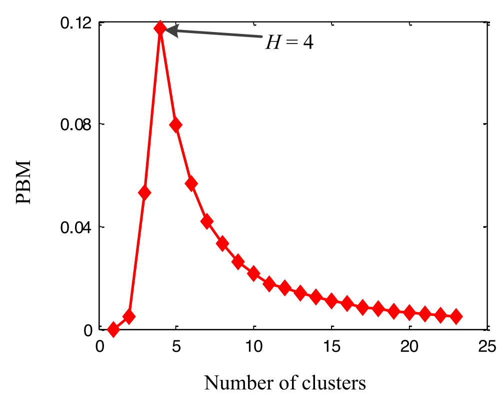

Fig. 5. Variation of the PBM index with the number of clusters for Bearing 1_1.

operation conditions are shown in Table II. Two bearings under each operation condition are used for training, and the others are used for testing. Accelerometers are fixed on the outer race of the bearings and vibration signals are captured. The sampling frequency is 25.6 kHz. Each sample contains 2560 points, i.e., 0.1 s, and sampling is repeated every 10 s. Vibration signals of a tested bearing are shown in Fig. 3.

B. Indicator Construction for the Bearings

Original features listed in Table III are extracted from vibration signals of the bearings, i.e., ten time-domain features [38], 16 time-frequency-domain features, and two features based on trigonometric functions [34]. The 16 time-frequency-domain features are energies and energy ratios of eight frequency bands generated by performing three-level wavelet packet decomposition on vibration signals. The two features based on trigonometric functions are defined by (31), and in particularly the standard deviation of inverse hyperbolic cosine (SD of IHC) and the standard deviation of inverse hyperbolic sine (SD of IHS), respectively,

$$\begin{cases} \text{SD of IHC} = \sigma\left[\log\left(x_i + \sqrt{x_i^2 - 1}\right)\right] \\ \text{SD of IHS} = \sigma\left[\log\left(x_i + \sqrt{x_i^2 + 1}\right)\right] \end{cases}. \tag{31}$$

These original features are utilized to construct the fusion health indicator WMQE. Taking Bearing 1_1, for example, the detailed information of the WMQE construction process is displayed as follows. First, the trendability values of the 28 original features are calculated and the features whose trendability values are higher than $1/2$ are selected from the original features. As shown in Fig. 4, 23 features are selected from the 28 original features. Then, the 23 selected features are input into the correlation clustering algorithm presented in Table I. According to the variation of the PBM index with the number of clusters shown in Fig. 5, the 23 selected features are classified into four different clusters. One typical feature is selected from each cluster and the selected typical features are shown in Fig. 6. It is seen that these four typical features show different types of trends during the whole lifetime and represent distinct information about the degradation process. F11 does not show any degradation information until the bearing approaches the end

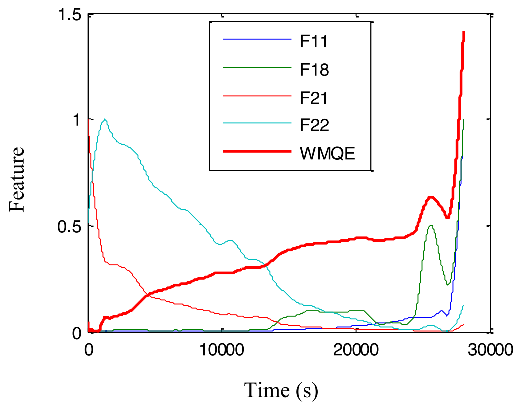

Fig. 6. Typical features and the WMQE of Bearing 1_1.

of the lifetime. F18 shows abrupt fluctuation from the middle of the lifetime. F21 decreases abruptly in the beginning but becomes stable gradually. F22 has a more obvious degradation trend than other typical features, but it is not informative for the final failure. In order to take advantage of the mutual information from these typical features, they are fused into WMQE. With the fusion of these typical features, WMQE shows more obvious degradation trends than the typical features and comprehensively describes the degradation process of Bearing 1_1 during the whole lifetime as shown in Fig. 6.

The WMQE values of the 17 bearings are constructed using the same method. Fig. 7(a) shows the WMQE of the six training datasets. In order to identify the benefits of the WMQE for RUL prediction of the bearings, the MQE [22] values of the bearings are also constructed using the 28 original features displayed in Table III, and the MQE values of the six training datasets are shown in Fig. 7(b). It is seen that the WMQE has more obvious degradation trends than the MQE. To compare the degradation trends of the WMQE and MQE quantitatively, the trendability values of them are calculated and displayed in Table IV. It is seen that, for most bearings, the trendability values of the WMQE are higher than those of the MQE. In addition, the WMQE values have similar amplitude ranges, which is helpful for determining a uniform threshold value for all of these bearings. In contrast, the MQE values have distinct amplitude ranges among different bearings. In Fig. 7(a), the degradation process can be roughly classified into two distinct stages. From 0 to 0.6, the WMQE values increase gradually, which corresponds to the long-time fault development process. From 0.6 to 1.2, the WMQE values show rapid increasing trends, which means that the bearings run to failure abruptly once the faults get severe. To prevent the happening of accidents, the bearings should be shut down before the faults get severe. Therefore, the threshold value of the bearings is set to be 0.6 for RUL prediction.

C. RUL Prediction for the Tested Bearings

The WMQE values are input into the state space model and the model parameters are initialized using the MLE algorithm. Then, the model parameters are updated and the RUL values are predicted using the PF-based prediction algorithm. The particle number is set to be 1000. The parameter updating process of the 11 testing datasets are shown in Fig. 8. It is seen that the parameter α is justified according to measurements in real time. As time goes on, the distribution of parameter α gradually converges to a narrow range and finally approaches a constant value. Taking Bearing 1_3, for example, the parameter initialization results are $\mu_\alpha = 8.2325e{-}4$, $\sigma_\alpha^2 = 5.2877e{-}7\,\beta = 0.7840$, and $\sigma_\nu^2 = 3.6655e{-}6$. As time goes on, the distribution of parameter α gradually converges to a narrow range. After about $t_k = 7000$s, α converges to a constant value 1.0116e-3.

The degradation trend of Bearing 1_3 is predicted as shown in Fig. 9(a). It is seen that the prediction result roughly represents the gradual degradation trend of the WMQE curve. The distribution of the predicted RUL of 1000 particles is shown in Fig. 9(b). Each particle represents a possible degradation process in the future. It is seen that the predicted RUL values of the particles are approximately normally distributed. To evaluate the uncertainty associated with the estimates of the RUL, the median and the 95% confidence interval (CI) of the predicted RUL are calculated and given in Fig. 9(b). The median and 95% CI are 5750 s and [5150, 6590 s], respectively.

The RUL prediction results of the other ten testing datasets are given in Fig. 10. It is seen that most of the bearings have distinctly different degradation trends before and after the current time. Taking Bearing 3_3, for example, its WMQE presents accelerated degradation trends before the current time. While after the current time, its WMQE begins to decrease gradually. The time-varying degradation trends are actually caused by the randomness of the degradation processes and increase the difficulty of RUL prediction. Our method predicts the future degradation trends of the bearings depending on their historical degradation trends. Since the degradation trends are different before and after the current time, the prediction results deviate from the actual values. Thus, the degradation trends of most test cases are outside of the predicted 95% CI. Actually, it is impractical to predict the degradation trend of each case accurately using one degradation model. Our major concern is to improve the average accuracy of the RUL prediction results for the whole cases. Therefore, a score in the IEEE PHM 2012 challenge [37] is used to comprehensively evaluate the performance of the prediction method, which is defined as follows:

$$\text{Score} = \frac{1}{11}\sum_{i=1}^{11} A_i \tag{32}$$

where

$$A_i = \begin{cases} \exp(-\ln(0.5)\cdot(Er_i/5)) & \text{if } Er_i \le 0 \\ \exp(+\ln(0.5)\cdot(Er_i/20)) & \text{if } Er_i > 0 \end{cases} \tag{33}$$

and Er_i represents the percent error of prediction result for the ith case

$$Er_i = \frac{\text{ActRUL}_i - \text{RUL}_i}{\text{ActRUL}_i} \times 100\% \tag{34}$$

where ActRUL_i and RUL_i are the actual RUL and the predicted RUL of the ith testing dataset, respectively. For convenient comparison in this paper, the median of the distribution is denoted as

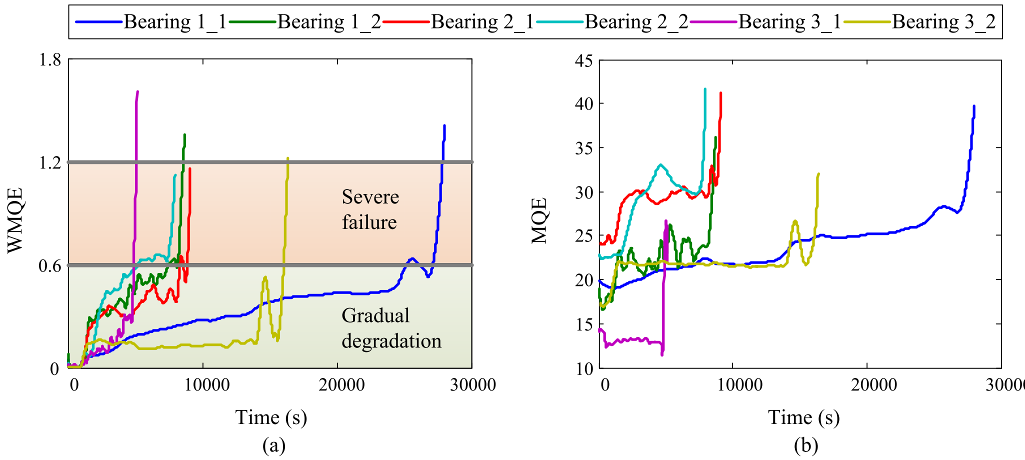

Fig. 7. WMQE and MQE of the six training datasets: (a) WMQE and (b) MQE.

TABLE IV
TRENDABILITY OF WMQE AND MQE FOR THE TESTED BEARINGS

Case	WMQE	MQE	Case	WMQE	MQE	Case	WMQE	MQE
Bearing 1_1	0.9960	0.9863	Bearing 2_1	0.9314	0.7146	Bearing 3_1	0.9764	0.0577
Bearing 1_2	0.9797	0.7770	Bearing 2_2	0.9852	0.7466	Bearing 3_2	0.5000	0.1613
Bearing 1_3	0.9818	0.9976	Bearing 2_3	0.1368	0.1359	Bearing 3_3	0.9706	0.8250
Bearing 1_4	0.9868	0.9790	Bearing 2_4	0.9854	0.8651			
Bearing 1_5	0.9905	0.9403	Bearing 2_5	0.6764	0.8932			
Bearing 1_6	0.9533	0.9190	Bearing 2_6	0.9909	0.9774			
Bearing 1_7	0.9895	0.8921	Bearing 2_7	0.8434	0.3363			

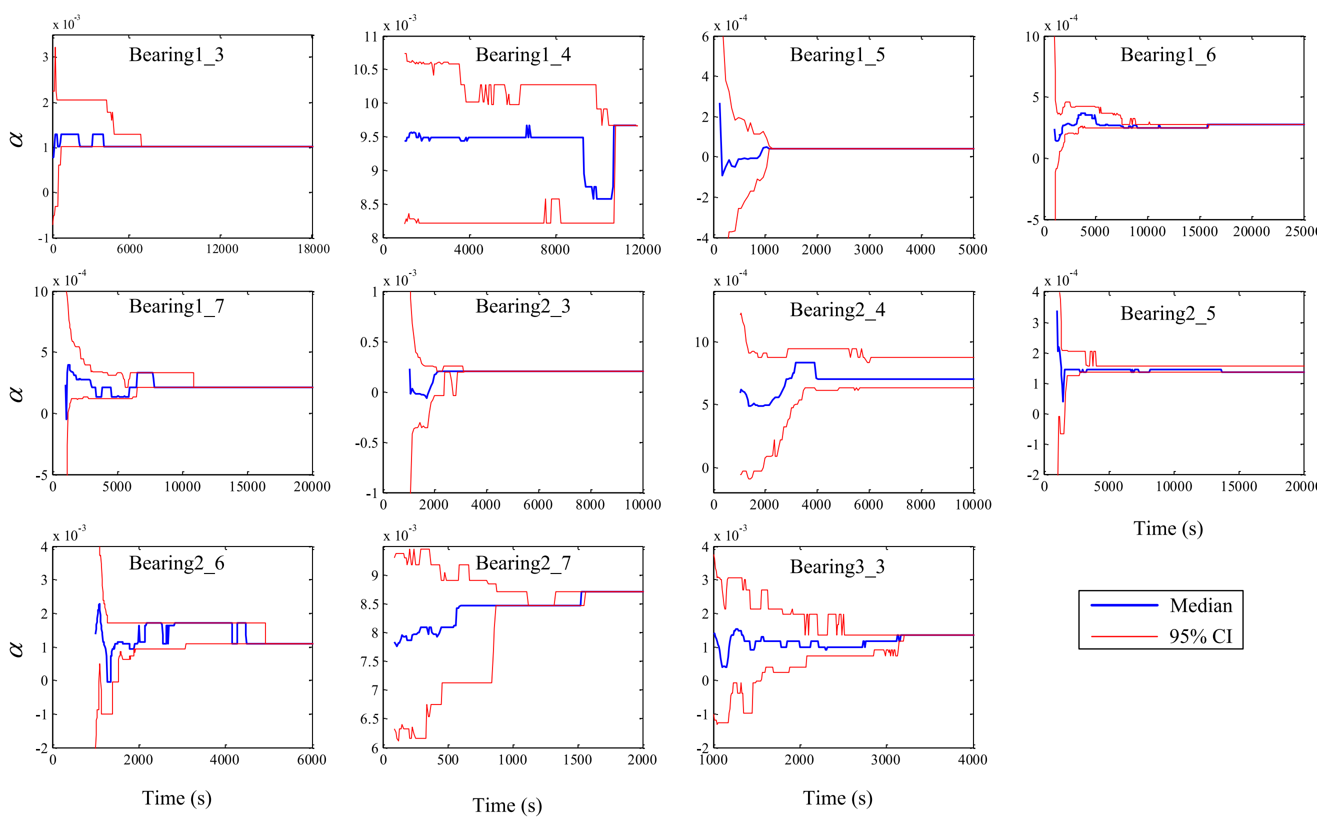

Fig. 8. Updating processes of parameter α for all of the testing datasets.

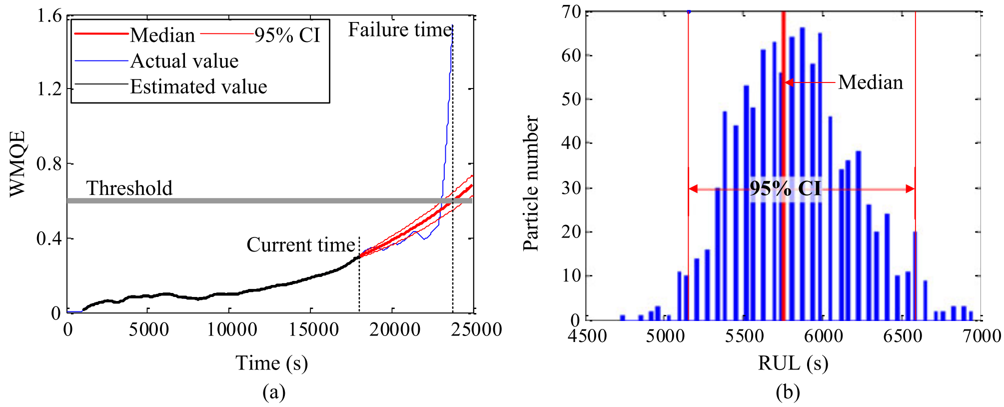

Fig. 9. RUL prediction result of Bearing 1_3: (a) prediction of the degradation trend and (b) distribution of the predicted RUL.

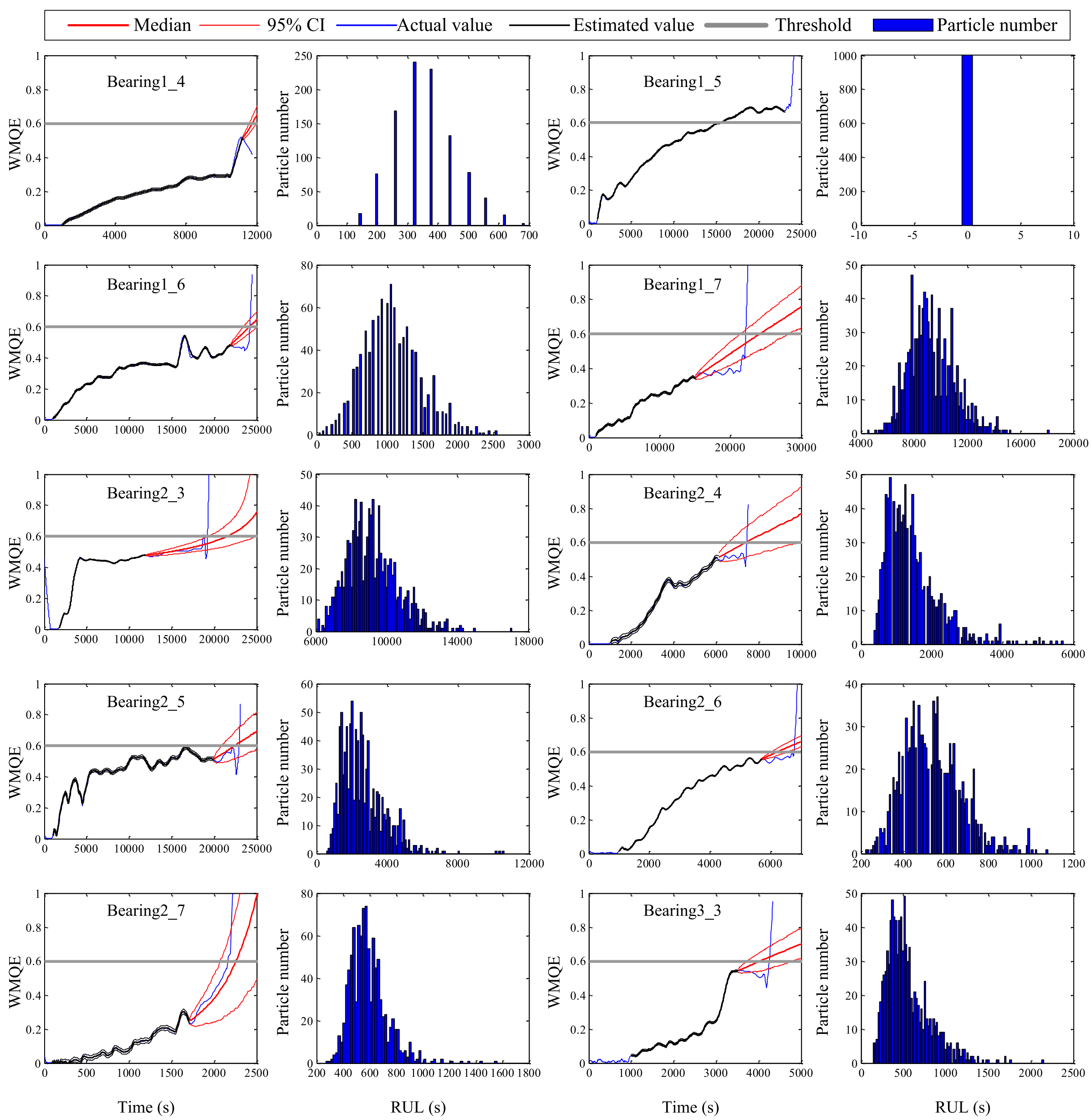

Fig. 10. RUL prediction results of the other ten testing datasets.

TABLE V
RUL PREDICTION RESULTS OF THE TESTING DATASETS

Testing dataset	Current time (s)	Actual RUL (s)	Predicted RUL (s)	Er 1 (%)	Er 2 (%)	Er 3 (%)
Bearing 1_3	18 010	5 730	5 750	−0.35	−1.04	37
Bearing 1_4	11 380	339	320	5.60	−20.94	80
Bearing 1_5	23 010	1 610	0	100	−278.26	9
Bearing 1_6	23 010	1 460	1 050	28.08	19.18	−5
Bearing 1_7	15 010	7 570	9 050	−19.55	−7.13	−2
Bearing 2_3	12 010	7 530	9 050	−20.19	10.49	64
Bearing 2_4	6 110	1 390	1 270	8.63	51.80	10
Bearing 2_5	20 010	3 090	2 370	23.30	28.80	−440
Bearing 2_6	5 710	1 290	530	58.91	−20.93	49
Bearing 2_7	1 710	580	550	5.17	44.83	−317
Bearing 3_3	3 510	820	490	40.24	−3.66	90
Score				0.4285	0.3550	0.3066
SD of Er				35.4135	90.2924	173.2757

the predicted RUL of the rolling element bearing. And all RUL values of the 11 testing datasets are predicted and displayed in Table V.

The percent errors of the proposed method are displayed in the column of Er 1 in Table V. In order to identify the effectiveness of the proposed method in the RUL prediction of rolling element bearings, the prediction results are compared with those of two similar studies where the same dataset is used [37]. The first study was published by Hong, Zhou, Zio, and Hong[39]. This study has the similar strategy with ours, both attempting to construct a fusion indicator, which is developed from the MQE. And its prediction results are outstanding among the published papers using the same dataset. Therefore, it is selected to be compared with our method. However, the limitation of the fusion indicator in this method is that original features are directly fused without selecting and weighting processes. The second one is selected since it is the winner of the IEEE PHM 2012 prognostic challenge [40]. This method predicts the RUL of most bearings based on the survival time ratios, except for Bearing 1_3, where a step of feature extrapolation is added. However, the survival time ratios need to be artificially determined by searching the anomaly time points from the frequency spectrum. Therefore, the prediction accuracy of this method is influenced by subjective factors. The percent errors of the two studies are shown in the column of Er 2 and Er 3, respectively.

The scores of the three methods are calculated and displayed in Table V. To compare the stability of the RUL prediction results, the SD of the percent errors for each method is calculated and displayed in the last row of Table V as well. Comparatively speaking, our method obtains the highest score among them and its SD is also the smallest. This implies that our proposed method presents more accurate and stable RUL prediction results than the others. The potential superiorities of our method compared with the other two methods should be concluded as follows. Compared with the first method, our method constructs the fusion health indicator by selecting and weighting the typical features, which are more sensitive to the degradation process, instead of fusing all of the original features without destination. Therefore, our method is supposed to enhance the superiority of the fusion health indicator for the RUL prediction. Compared with the second method, our method is less dependent on subjective factors, thus, acquiring more robust and more stable RUL prediction results among the testing datasets.

V. CONCLUSION

To deal with two problems in model-based RUL prediction methods, i.e., health indicator construction and model parameter estimation, this paper develops a model-based method for RUL prediction of machinery. In this method, a new health indicator called WMQE is constructed, which fuses mutual information from multiple features and properly correlates with the degradation processes of machinery, and RUL is predicted using a PF-based algorithm with model parameters initialized using the MLE algorithm.

Vibration signals collected from accelerated degradation tests of rolling element bearings are used to demonstrate the effectiveness of the method. Two similar studies that use the same dataset are compared with the proposed method. The results show that the proposed model-based method performs best among the three methods in RUL prediction of machinery.

However, there are still some limitations in the proposed method. For example, our intention in the indicator construction module is to construct a health indicator with monotonic trend. In some real cases, however, the degradation trends are actually nonmonotonic or even change abruptly. In these cases, the proposed method is unable to predict the future degradation trend of the machinery. Therefore, much research work is still needed to establish more robust degradation models, which are able to describe nonmonotonic or abruptly varying degradation trends.

REFERENCES

[1] Y. G. Lei, J. Lin, Z. J. He, and M. J. Zuo, "A review on empirical mode decomposition in fault diagnosis of rotating machinery," *Mech. Syst. Signal Process.*, vol. 35, nos. 1−2, pp. 108–126, Feb. 2013.

[2] Y. G. Lei, Z. J. He, and Y. Y. Zi, "Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAs," *Mech. Syst. Signal Process.*, vol. 21, no. 5, pp. 2280–2294, Jul. 2007.

[3] F. Q. Zhao, Z. G. Tian, E. Bechhoefer, and Y. Zeng, "An integrated prognostics method under time-varying operating conditions," *IEEE Trans. Rel.*, vol. 64, no. 2, pp. 673–686, Jun. 2015.

[4] Y. G. Lei, J. Lin, M. J. Zuo, and Z. J. He, "Condition monitoring and fault diagnosis of planetary gearboxes: A review," *Meas.*, vol. 48, pp. 292–305, 2014.

[5] J. Lee, F. J. Wu, W. Y. Zhao, M. Ghaffari, L. X. Liao, and D. Siegel, "Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications," *Mech. Syst. Signal Process.*, vol. 42, nos. 1−2, pp. 314−334, Jan. 2014.

[6] L. R. Cui, S. Du, and A. G. Hawkes, "A study on a single-unit repairable system with state aggregations," *IIE Trans.*, vol. 44, pp. 1022−1032, 2012.

[7] X. F. Zhao, T. Nakagawa, and M. J. Zuo, "Optimal replacement last with continuous and discrete policies," *IEEE Trans. Rel.*, vol. 63, no. 4, pp. 868−880, Dec. 2014.

[8] Z. Q. Wang, W. B. Wang, C. H. Hu, X. S. Si, and W. Zhang, "A prognostic-information-based order-replacement policy for a non-repairable critical system in service," *IEEE Trans. Rel.*, vol. 64, no. 2, pp. 721−735, Jun. 2015.

[9] J. Liu, W. Wang, F. Ma, Y. B. Yang, and C. S. Yang, "A data-model-fusion prognostic framework for dynamic system state forecasting," *Eng. Appl. Artif. Intell.*, vol. 25, no. 4, pp. 814–823, Jun. 2012.

[10] J. Liu, W. Wang, and F. Golnaraghi, "A multi-step predictor with a variable input pattern for system state forecasting," *Mech. Syst. Signal Process.*, vol. 23, pp. 1586–1599, 2009.

[11] Z. Y. Huang, Z. G. Xu, W. H. Wang, and Y. X. Sun, "Remaining useful life prediction for a nonlinear heterogeneous Wiener process model with an adaptive drift," *IEEE Trans. Rel.*, vol. 64, no. 2 pp. 687–700, Jun. 2015.

[12] H. Hanachi, J. Liu, A. Banerjee, Y. Chen, and A. Koul, "A physics-based modeling approach for performance monitoring in gas turbine engines," *IEEE Trans. Rel.*, vol. 64, no. 1, pp. 197–205, Mar. 2015.

[13] H. Y. Dui, S. B. Si, M. J. Zuo, and S. D. Sun, "Semi-Markov process-based integrated importance measure for multi-state systems," *IEEE Trans. Rel.*, vol. 64, no. 2, pp. 754−765, Jun. 2015.

[14] L. R. Cui, Y. Xu, and X. Zhao, "Developments and applications of the finite Markov chain imbedding approach in reliability," *IEEE Trans. Rel.*, vol. 59, no. 4, pp. 685−690, Dec. 2010.

[15] X. S. Si, W. B. Wang, M. Y. Chen, C. H. Hu, and D. H. Zhou, "A degradation path-dependent approach for remaining useful life estimation with an exact and closed-form solution," *Eur. J. Oper. Res.*, vol. 226, no. 1, pp. 53−66, Apr. 2013.

[16] X. S. Si, W. B. Wang, C. H. Hu, D. H. Zhou, and M. G. Pecht, "Remaining useful life estimation based on a nonlinear diffusion degradation process," *IEEE Trans. Rel.*, vol. 61, no. 1, pp. 50−67, Mar. 2012.

[17] J. B. Yu, "Health degradation detection and monitoring of lithium-ion battery based on adaptive learning method," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 7, pp. 1709–1721, Jul. 2014.

[18] J. B. Yu, "A nonlinear probabilistic method and contribution analysis for machine condition monitoring," *Mech. Syst. Signal Process.*, vol. 37, no. 1, pp. 293–314, 2013.

[19] F. Camci, K. Medjaher, N. Zerhouni, and P. Nectoux, "Feature evaluation for effective bearing prognostics," *Quality Rel. Eng. Int.*, vol. 29, no. 4, pp. 477–486, 2013.

[20] R. Li and D. He, "Rotational machine health monitoring and fault detection using EMD-based acoustic emission feature quantification," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 4, pp. 990−1001, Apr. 2012.

[21] Y. N. Qian, S. J. Hu, and R. Q. Yan, "Bearing performance degradation evaluation using recurrence quantification analysis and auto-regression model," in *Proc. IEEE Int. Conf. Instrum. Meas. Tech.*, Minneapolis, MN, USA, May 6−9, 2013, pp. 1713−1716.

[22] H. Qiu, J. Lee, J. Lin, and G. Yu, "Robust performance degradation assessment methods for enhanced rolling element bearing prognostics," *Adv. Eng. Inform.*, vol. 17, no. 3−4, pp. 127−140, Jul. 2003.

[23] R. Q. Huang, L. F. Xi, X. L. Li, C. R. Liu, H. Qiu, and J. Lee, "Residual life predictions for ball bearings based on self-organizing map and back propagation neural network methods," *Mech. Syst. Signal Process.*, vol. 21, no. 1, pp. 193−207, Jan. 2007.

[24] J. B. Yu, "Health condition monitoring of machines based on hidden Markov model and contribution analysis," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 8, pp. 2200−2211, Aug. 2012.

[25] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[26] A. Doucet, N. D. Freitas, and N. Gordon, "An introduction to sequential Monte Carlo methods," in *Sequential Monte Carlo Methods in Practice*, New York, NY, USA: Springer, 2001, ch. 1, pp. 3–14.

[27] M. E. Orchard, B. Wu, and G. Vachtsevanos, "A particle filtering framework for failure prognosis," in *Proc. World Tribology Congr. III*, 2005, pp. 883–884.

[28] D. An, J. H. Choi, and N. H. Kim, "Prognostics 101: A tutorial for particle filter-based prognostics algorithm using Matlab," *Rel. Eng. Syst. Safety*, vol. 115, pp. 161−169, 2013.

[29] C. C. Chen, G. Vachtsevanos, and M. E. Orchard, "Machine remaining useful life prediction: An integrated adaptive neuro-fuzzy and high-order particle filtering approach," *Mech. Syst. Signal Process.*, vol. 28, pp. 597−607, Apr. 2012.

[30] M. Corbetta, C. Sbarufatti, A. Manes, and M. Giglio, "Real-time prognosis of crack growth evolution using sequential Monte Carlo methods and statistical model parameters," *IEEE Trans. Rel.*, vol. 64, no. 2, pp. 736–753, Jun. 2015.

[31] M. E. Orchard, P. Hevia-Koch, B. Zhang, and L. Tang, "Risk measures for particle-filtering-based state-of-charge prognosis in lithium-ion batteries," *IEEE Trans. Ind. Electron.*, vol. 60, no. 11, pp. 5260–5269, Nov. 2013.

[32] P. C. Paris and F. Erdogan, "A critical analysis of crack propagation laws," *J. Fluids Eng.*, vol. 85, no. 4, pp. 528–533, 1963.

[33] C. Spearman, "The proof and measurement of association between two things," *Amer. J. Psychol.*, vol. 15, no. 1, pp. 72−101, 1904.

[34] K. Javed, R. Gouriveau, N. Zerhouni, and P. Nectoux, "Enabling health monitoring approach based on vibration data for accurate prognostics," *IEEE Trans. Ind. Electron.*, vol. 62, no. 1, pp. 647−656, Jan. 2015.

[35] M. K. Pakhira, S. Bandyopadhyay, and U. Maulik, "Validity index for crisp and fuzzy clusters," *Pattern Recog.*, vol. 37, pp. 487−501, 2004.

[36] Y. G. Lei, Z. J. He, Y. Y. Zi, and Q. Hu, "Fault diagnosis of rotating machinery based on a new hybrid clustering algorithm," *Int. J. Adv. Manuf. Technol.*, vol. 35, no. 9-10, pp. 968–977, 2006.

[37] P. Nectoux, R. Gouriveau, K. Medjaher, E. Ramasso, B. Chebel-Morello *et al.*, "PRONOSTIA: An experimental platform for bearings accelerated degradation tests," in *Proc. Int. Conf. Prognost. Health Manage.*, Jun. 18–21, 2012, pp. 1–8.

[38] Y. G. Lei and M. J. Zuo, "Gear crack level identification based on weighted K nearest neighbor classification algorithm," *Mech. Syst. Signal Process.*, vol. 23, no. 5, pp. 1535−1547, 2009.

[39] S. Hong, Z. Zhou, E. Zio, and K. Hong, "Condition assessment for the performance degradation of bearing based on a combinatorial feature extraction method," *Digit. Signal Process.*, vol. 27, pp. 159−166, 2014.

[40] E. Sutrisno, H. Oh, A. S. S. Vasan, and M. Pecht, "Estimation of remaining useful life of ball bearings using data driven methodologies," in *Proc. Int. Conf. Prognost. Health Manage.*, June 18–21, 2012, pp. 1–7.

Yaguo Lei (M'15) received the B.S. and Ph.D. degrees in mechanical engineering from Xi'an Jiaotong University, Xi'an, China, in 2002 and 2007, respectively.

He is a Full Professor of mechanical engineering with Xi'an Jiaotong University. Prior to joining Xi'an Jiaotong University in 2010, he was a Post-Doctoral Research Fellow with the University of Alberta, Edmonton, AB, Canada. He was also an Alexander von Humboldt Fellow with the University of Duisburg-Essen, Duisburg, Germany. His research interests include machinery condition monitoring and fault diagnosis, mechanical signal processing, intelligent fault diagnostics, and remaining useful life prediction.

Dr. Lei is a Member of the American Society of Mechanical Engineers, a Senior Member of the Chinese Mechanical Engineering Society, and an Editorial Board Member of *Neural Computing and Applications, Advances in Mechanical Engineering, The Scientific World Journal, International Journal of Mechanical Systems Engineering, Chinese Journal of Engineering,* and the *International Journal of Applied Science and Engineering Research.*

Naipeng Li received the B.S. degree in mechanical engineering from Shandong Agricultural University, Tai'an, Shandong, China, in 2012. He is currently working toward the Ph.D. degree in mechanical engineering at the State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an, China.

His research interests include machinery condition monitoring, intelligent fault diagnostics, and remaining useful life prediction of rotating machinery.

Szymon Gontarz received the B.S. and Ph.D. degrees in mechanical engineering from the Faculty of Automotive and Construction Machinery Engineering, Warsaw University of Technology, Warsaw, Poland, in 2004 and 2011, respectively. Parallel accomplished Interdisciplinary Study of Insurance Technical Systems during individual program of study.

He is currently an Assistant Professor with the Faculty of Automotive and Construction Machinery Engineering, Warsaw University of Technology since 2012. He was a Senior Research Assistant with the Department of Manufacturing Engineering and Engineering Management, City University of Hong Kong, in 2005. He held a scientific internship at the Department of Industrial Chemistry and Materials Engineering, University of Messina, Messina, Italy, in 2006 and with Faculty of Geotechnics, Materialprüfungsanstalt Universität Stuttgart, Stuttgart, Germany,in 2007. Since 2006,he has been a member of the team of the Integrated Laboratory of the Mechatronics System of Vehicles and Construction Machinery, Warsaw University of Technology. His research interests include vibration based diagnostic as well as passive magnetic technique for stress assessment.

Prof. Gontarz is a Member of the Polish Society of Technical Diagnostics.

Stanislaw Radkowski received the degree in construction machinery from Warsaw University of Technology, Warsaw, Poland, in 1973.

He has been the Dean of the Faculty of Automotive and Construction Machinery Engineering at Warsaw University of Technology, Poland, since 2012. He was with the Team of Machine Design Fundamentals from 1977 to 1998, and with the Team of Fundamentals of Machine Design and Maintenance since 1998. He has been a Full Professor since 2009 and the Director of the Institute of Vehicle since 2010. He is a coauthor of 12 deployments, eightpatents, and more than 100 scientific papers in the range of vibroacoustics, in particular to minimize vibration and noise in machines, vibroacoustic diagnostics of constructions, and components of machines and technical risk assessment. His research interests include machine design and maintenance, technical risk analysis, technical diagnostics, and diagnostics prestressed structures.

Dr. Radkowski is a Member of the Committee of Mechanical Engineering of the Polish Academy of Sciences, Polish Society of Technical Diagnostics, Polish Society of Ecological Vehicles, International Society of Condition Monitoring, and the Chairman of the committee section of maintenance.

Jing Lin received the B.S., M.S., and Ph.D. degrees from Xi'an Jiaotong University, Xi'an, China, in 1993, 1996, and 1999, respectively, all in mechanical engineering.

He is currently a Professor with the State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University. From July 2001 to August 2003, he was a Post-Doctoral fellow with the University of Alberta, Edmonton, AB, Canada, and a Research Associate with the University of Wisconsin, Milwaukee, Milwaukee, WI, USA. From September 2003 to December 2008, he was a Research Scientist with the Institute of Acoustics, Chinese Academy of Sciences, Beijing, China, under the sponsorship of the Hundred Talents Program. His current research directions are in mechanical system reliability, fault diagnosis, and wavelet analysis.

Dr. Lin received the National Science Fund for Distinguished Young Scholars in 2011.

Jacek Dybala received the Ph.D. and Habilitation Doctorate (D.Sc.) degrees in mechanical engineering from the Faculty of Automotive and Construction Machinery Engineering and the Institute of Organization of Production Systems, Warsaw University of Technology, Warsaw, Poland, in 2000 and 2010, respectively.

He is an Associate Professor with the Laboratory of Mechatronics, Diagnostics and Technical Safety, Institute of Vehicles, Warsaw University of Technology. In 2013, he was a Research Fellow with the University of Michigan, Ann Arbor, MI, USA. His major research interests include machine condition monitoring, structural health monitoring, vibroacoustic diagnostics, intelligent diagnostics, signal processing and analysis, and data mining.

Prof. Dybalais a Member of the Polish Society of Technical Diagnostics.