Development of Vibration-Based Health Indexes for Bearing Remaining Useful Life Prediction

Xiaohang Jin[1,2], Zijun Que[2], Yi Sun[1,2]

[1] Ninghai ZJUT Academy of Science and Technology
Zhejiang University of Technology
Ninghai, Zhejiang, China
xhjin@zjut.edu.cn

[2] College of Mechanical Engineering
Zhejiang University of Technology
Hangzhou, China

***Abstract*—Bearing failure can cause their host system shutdown, and even some catastrophic accidents. These will lead to a high maintenance cost and a huge economic loss. Thus, health monitoring and fault prognosis for bearings becomes increasingly important. Developing an effective health index (HI) will do help in these works. Hence, three different HIs are developed by using root mean square, Kolmogorov-Smirnov test, and Mahalanobis distance to reflect bearings' online health conditions. Four degradation models are constructed to estimate bearings remaining useful life (RUL) by using particle filter algorithm. Bearing life data are used to test the performance of fault prognostic approaches. Results show that all HIs reflect the degradation process of bearing effectively, and the proposed degradation model has the best performance in bearing RUL prediction.**

Keywords- Bearing; fault prognostics; health index; particle filter; remaining useful life; vibration signals

I. INTRODUCTION

Bearing plays a critical role in wind turbines and induction motors [1, 2]. However, bearings are easy to be damaged due to their harsh operating conditions [3]. Bearing failures will lead to wind turbine cannot operate properly and cause high maintenance costs. Condition based maintenance (CBM) techniques have helped to save up to £1.3 billion per year in wind energy industry [4]. Thus, estimating bearings' online health status and predicting their future health conditions becomes increasingly important [5]. These works could provide the optimized maintenance actions to avoid bearing failure [6,7].

Many approaches, such as data-driven approach and physics-of-failure (PoF) approach, can be used for fault prognostics [7]. Since different components interact each other and there are also many failure modes in bearings, it is difficult to build a PoF model to estimate bearing's future health status. The data-driven approach analyzes monitored data with aid of artificial intelligence models, statistical methods, and deep learning algorithms to fault prognostics for bearings [7]. This approach can learn system behaviors from the data, infer health status, predict future health conditions, and estimate the RUL [7]. Many signals, such as vibration, current, and temperature, can be used in data-driven approach. Among them, vibration signals that contain rich health information on bearings are widely used for fault prognosis [8]-[12].

Bearing vibration signals are composed of random noise, periodic fluctuations and fault-related features, etc. They have the non-Gaussian and non-stationary characteristics. Thus, it is challenging to find the fault-related features and diagnose typical faults. One of typical solutions is to calculate many statistics in time- and frequency-domain to reflect bearing's health condition, and then machine learning algorithms are used to analyze these data for fault diagnostics and prognostics, such as Jin calculated RMS, skewness, kurtosis, and other statistics based on the vibration signal and then built a Mahalanobis distances (MDs)-based health index (HI) to indicate cooling fan's degradation process [13]. It should be noted that different statistics reflect different characteristics of the vibration signal. Some statistics have good performance on indicating incipient faults while others indicate bearing's degradation trend effectively. Therefore, appropriate statistics should be selected out and suitable health indexes should be constructed to predict bearing's RUL [1].

The contribution of this paper is that different HIs and degradation models of bearings are constructed based on statistical characteristics of vibration signals to predict the RUL. First, three different HIs (RMS, K-S, MD) are built to reflect the health conditions of bearings. Second, according to different HIs, different degradation models are proposed to quantify the process of bearing degradation. Finally, the particle filter (PF) algorithm is used to update model's parameters and then predict the future health status of bearing. The rest of this paper is organized as follows. Section II presents the development of HIs. Section III reports different degradation models and the particle filter algorithm. the performance of bearing fault prognostics is tested and validated based on bearing data in Section IV. Finally, conclusions are drawn in Section V.

II. DEVELOPMENT OF HEALTH INDEX

When bearing goes into degradation stage, some faults have happened in bearing components. The rolling elements of

This work was supported in part by Ningbo Natural Science Foundation under Grant 2018A610045, and in part by the National Natural Science Foundation of China under Grants 51505424 and 51675484, and in part by Zhejiang Provincial Natural Science Foundation of China under Grants LY15E050019.

978-1-7281-0860-5/19/$31.00 ©2019 IEEE

bearing pass these faults, impulses are generated. These will lead to the acquired vibration signals vary in amplitude and energy distribution, compared to those bearings under fault-free conditions [5]. Time-domain features of vibration signals are sensitive to bearing health conditions and have clear physical meanings Thus, three HIs are constructed to indicate bearing health conditions and analyze the degradation process of bearings in this section based on the time-domain features of vibration signal.

A. Root Mean Square-based HI

The root mean square (RMS) of vibration signals are expressed in (1). Since RMS reflects the strength of vibration signal, it is widely used to indicate bearings health conditions.

$$\mathrm{RMS} = \sqrt{\frac{1}{M}\sum_{i=1}^{M} x^2(i)} \tag{1}$$

where $x(i)$ is ith data of the vibration signal; M is the total number of sample and length of vibration signal.

B. Kolmogorov-Smirnov test-based HI

When anomalies and faults happen in bearing components, they will cause severe bearing vibration, such as shock and the increasing amplitude can be observed in the vibration signals. These lead the faulty bearing vibration signals to deviate from the healthy ones. Kolmogorov-Smirnov (K-S) test, which is a non-parametric test, is commonly used to determine whether two distributions are statistically similar or not. With the K-S test result, it is possible to compare the testing conditions with the known healthy conditions. As a result, the testing conditions could be determined.

The empirical cumulative distribution function (CDF) of a snapshot vibration signal, $F(x)$, could be defined as

$$F(x) = \frac{\text{number of samples} \le h}{M} \tag{2}$$

where h is the amplitude of the vibration signals; M is the total number of samples in a snapshot vibration signal.

The K-S statistic, D_t, which quantifies the maximum difference between vibration signals at time k and healthy bearing vibrations signals, could be used to indicate bearing condition. Small K-S statistic indicates healthy bearings while big K-S statistic indicates bearings may have some faults. The bigger this value of K-S statistic is, the worse condition bearing is.

TABLE I STEPS FOR DEVELOPING MD-BASED HI

1.	Collect vibration signal from healthy products.
2.	Calculate time-domain feature data from vibration signal.
3.	Calculate the mean and standard deviation of different features, respectively; normalize the training data set, and then calculate their correlation matrix.
4.	Calculate MDs of the training data to construct a Mahalanobis space as a reference.
5.	Collect data from test products, and normalize these test data by using healthy products' means and standard deviation.
6.	Calculate MDs of the test data and use MDs to indicate the health condition of test products.

C. Mahalanobis Distance-Based HI

MD is used to determine the similarities between test samples and training samples [14]. It is a generalized distance that combines multi-feature information to a scalar value by considering the correlations among features.

$$\mathrm{MD} = \frac{1}{n} y_i C^{-1} y_i^T \tag{3}$$

where $y_i = [y_{i1}, y_{i2}, \ldots, y_{in}]$ is the normalized input feature vector, y_i^T is the transposed vector of y_i; n is the number of features; C is the correlation matrix between features. In this paper, bearings healthy data are defined as the training samples. The step for developing MD-based HI is shown in table I.

III. BEARING FAULT PROGNOSTICS

Generally, bearing's life can be divided into run-in, normal and degradation stages. These three stages are well-described in [3, 6]. Different components adjust each other to enter into a good operating condition gradually in run-in stage. This stage could be described by a decreasing HI. In the normal stage of bearings, they undergo a comparable long and stable condition as indicated by small values of HIs. When bearings begin to degrade, they will have a dynamic process. Bearings degrade slowly when incipient faults happen, as indicated by a slowly increasing HIs. After that, bearing will degrade fast as indicated by a fast-growing HIs.

The RUL can be estimated by analyzing the bearing data in the degradation stage. In this section, four degradation models are used to study bearings' degradation process. And then the particle filter (PF) algorithm is used to update the parameters of bearings' degradation models, predict bearings' future health condition, and get the RUL finally.

A. Degradation model

When bearings begin to degrade, their vibration will be strengthened, as reflected by the increased amplitude of vibration signals. Bearing's degradation process is closely related to its internal health conditions, such as faults in its components and the lubrication condition. Time-dependent models can be constructed to study bearing's degradation process with the aid of regression analysis.

Since different HIs reflect difference characteristics of bearings, several functions are suggested to analyze bearings degradation data. These functions include exponential function, the sum of two exponential functions [3, 6], polynomial function and the sum of one exponential function and polynomial function, which are described as follows:

Model A: Exponential function:

$$HI(k) = a_A \cdot \exp(b_A \cdot k) \tag{4}$$

where HI is bearing's HI; k is the time index; a_A and b_A are the parameters associated with this function.

Model B: The sum of two exponential functions:

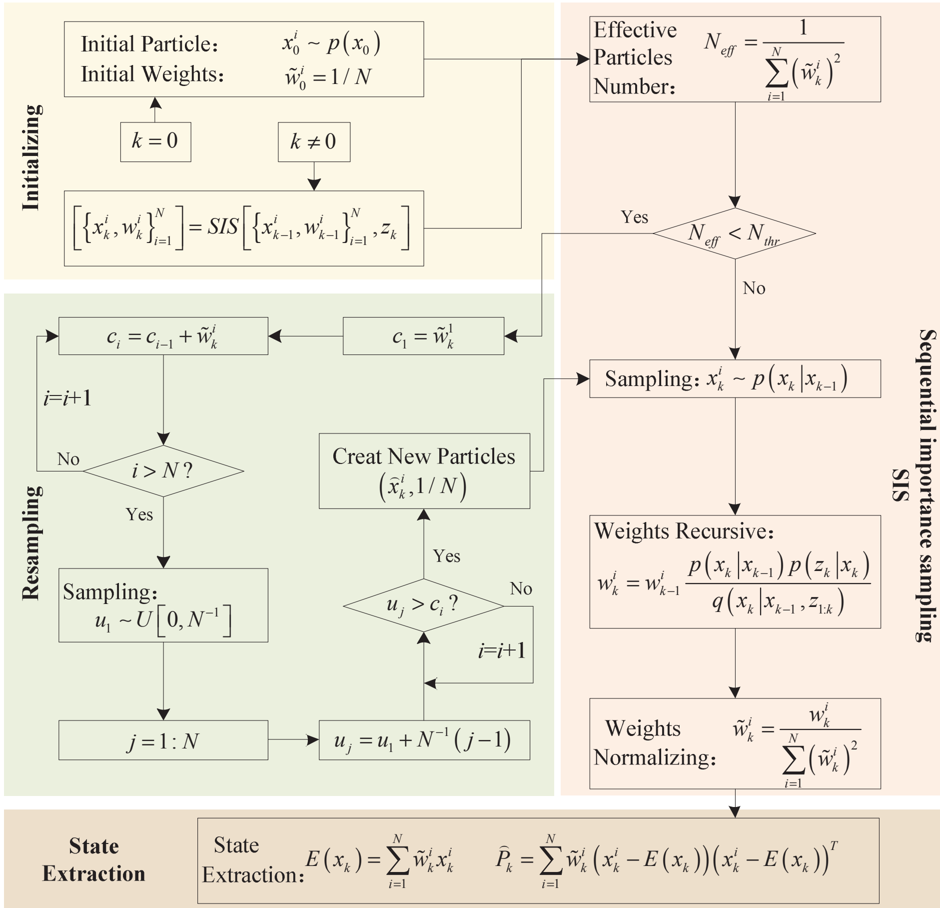

Figure 1. Algorithm of PF [20].

$$HI(k) = a_B \cdot \exp(b_B \cdot k) + c_B \cdot \exp(d_B \cdot k) \tag{5}$$

where a_B, b_B, c_B and d_B are the parameters, respectively.

Model C: The polynomial function:

$$HI(k) = a_C \cdot k^3 + b_C \cdot k^2 + c_C \cdot k + d_C \tag{6}$$

where a_C, b_C, $c_C$and d_C are the parameters, respectively.

Model D: The sum of one exponential and polynomial function:

$$HI(k) = a_D \cdot \exp(b_D \cdot k) + c_D \cdot k^2 + d_D \cdot k + e_D \tag{7}$$

where a_D, b_D, c_D, d_D and e_D are the parameters, respectively.

Different HIs have different degradation trends. The performance of different degradation models is evaluated by the curve fitting results. The fitting error $R^2$is used to evaluate the performance of regression analysis. The closer the value of R^2 to 1, the better the fitting effect is.

B. RUL estimation

RUL is defined as the time between the current inspection time and the end life of a system [6, 15]. RUL estimation is triggered when bearings begin to degrade. Considering the nonlinear degradation process, particle filter (PF) algorithm is used to update the degradation model parameters with the new observing data. PF is one of Bayesian filter algorithms. It is based on Monte Carlo approximation, which is able to deal with nonlinear and non-Gaussian problems effectively [16].

Based on the Bayesian' theorem, the poster distribution of the states x_k can be defined as:

$$p(x_{0:k} | z_{1:k}) = p(x_{0:k-1} | z_{1:k-1}) \frac{p(z_k | x_k) p(x_k | x_{k-1})}{p(z_k | z_{1:k-1})} \tag{8}$$

where $x_{0:k} = \{x_0, x_1, \ldots, x_k\}$ is the sequences of states; $z_{1:k} = \{z_1, \ldots, z_k\}$ is the sequences of observations; k is the time index; $p(\cdot)$ is the probability density function (PDF).

PF approximates the posterior distribution with a number of weighted samples (or particles) $x_{0:k}^i \left(i = 1, \ldots, N\right)$ [17], in which N is the number of particles. The sequential importance sampling (SIS) is used to acquire samples from the posterior distribution. Choose the importance distribution as:

$$q(x_{0:k} | z_{1:k}) = q(x_{0:k-1} | z_{1:k-1}) q(x_k | x_{0:k-1}, z_{1:k}) \tag{9}$$

The importance weight for each particle can be formulated in a recursive form:

$$w_k^i \propto w_{k-1}^i \frac{p(z_k | x_k^i) p(x_k^i | x_{k-1}^i)}{q(x_k^i | x_{0:k-1}^i, z_{1:k})} \tag{10}$$

where w_k^i is the weight. Thus the poster distribution can be

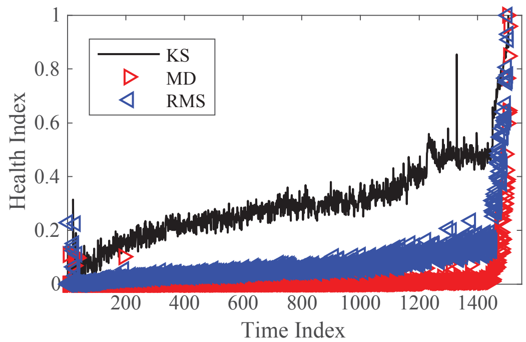

Figure 2 Degradation trend described by different HIs.

TABLE II GOODNESSOF FIT STATISTICS

HIs	R^2			
	Model A	Model B	Model C	Model D
RMS-	0.6168	0.9242	0.6278	0.5219
KS-	0.8493	0.9133	0.9001	0.8372
MD-	0.9233	0.9233	0.2865	0.1873

approximated by particles as:

$$p\left(x_{0:k} \left| z_{1:k}\right.\right) \approx \sum_{i=1}^{N} \tilde{w}_k^i \delta\left(x_{0:k} - x_{0:k}^i\right) \tag{11}$$

where $\tilde{w}_k^i$ is the normalized w_k^i , $\delta(\cdot)$ is the Dirac delta measure. Fig.1 illustrates the algorithm of PF, in which SIS technique is used to solve the particle degeneracy problem [18, 19].

IV. RESULTS AND DISCUSSIONS

A PRONOSTIA platform-based bearing life test data is used to test the performance of different HIs and degradation models in this paper [21]. Bearing operates continuously at its rated load (4000N) until it fails. Two vibration signals in horizontal and vertical directions are collected every 10 seconds. These signals are acquired by a National Instrument (NI) DAQ system at a sampling rate of 25.6 kHz.

As talked in Section II, three HIs are constructed in this paper. They are RMS-, KS- and MD-based HIs. Different HIs have different degradation tendency as shown in Fig. 2. The RMS-based HI is relatively smooth. At the early period of the degradation stage, as the fault features are not obvious, the change of RMS-based HI is relatively small. As the fault became worse, the vibration of the bearing increased gradually. These could be indicated by an increasing HI. At the end of the degradation stage, the HI increases sharply because of the violent vibration. Before constructing the KS-based HI, a sample of vibration signals corresponding to normal working conditions is chosen randomly. This sample is used as a benchmark to build the KS-based HI. As it can be observed on Fig. 2, KS-based HI increases but the fluctuations are large compared to other two HIs. Such fluctuation means that it is sensitive to small changes in vibration signals. The MD-based HI is similar with the RMS-based HI. However, the MD-based HI grows very fast at the end of bearing life. This indicates that there are significant changes in bearing status. In other

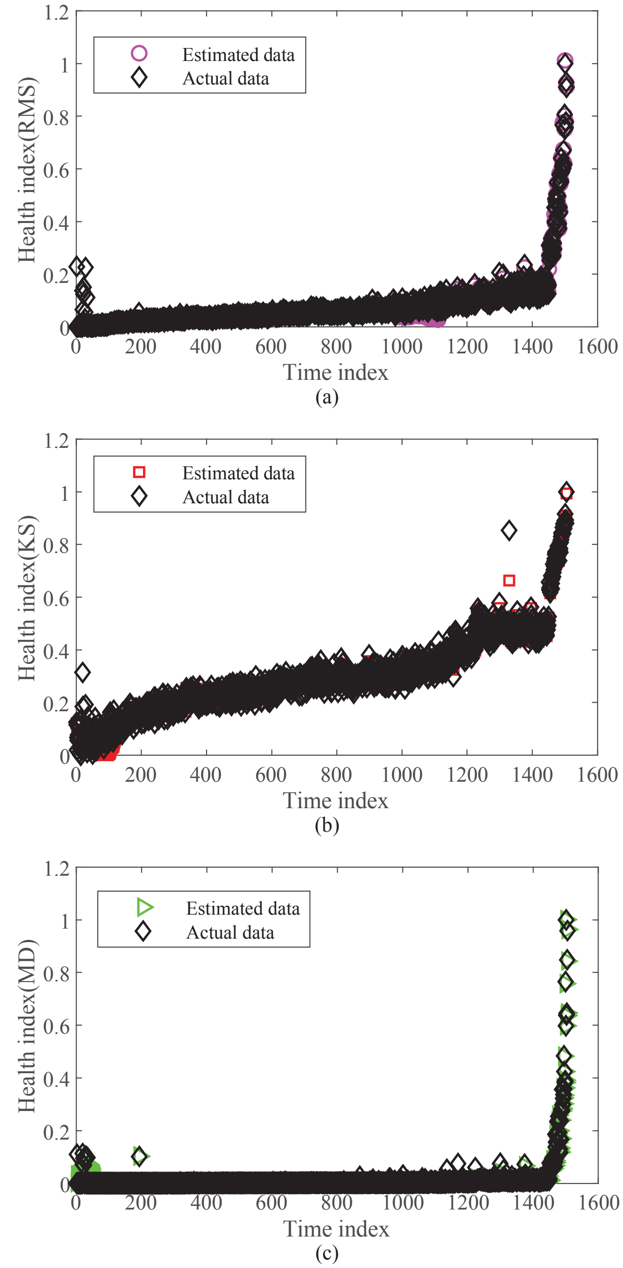

Figure 3 Degradation tracking using different HIs (a) RMS-based HI (b) KS-based HI (c) MD-based HI.

words, the MD-based HI has the capability to reflect the dynamic degradation process of bearing clearly. In addition, this HI is good at detect bearing incipient faults.

As talked in section III, four different models are used to describe different HIs trend. Parameters in these models are got by curve fitting techniques with aid of MATLAB. The fitting performance is evaluated by the R^2 as reported in Table II. It can be found that model B fits all HIs very well since their R^2 are close to 1, compared with the other three models. The R^2 of model A and B are same for MD-based HI. Both of them are 9.233. This is because the parameter a_B in model B is 0. Thus, model A and B are same. Then their performances are same. Based on the results presented in Table II, it can be concluded that it is better to use model B to construct

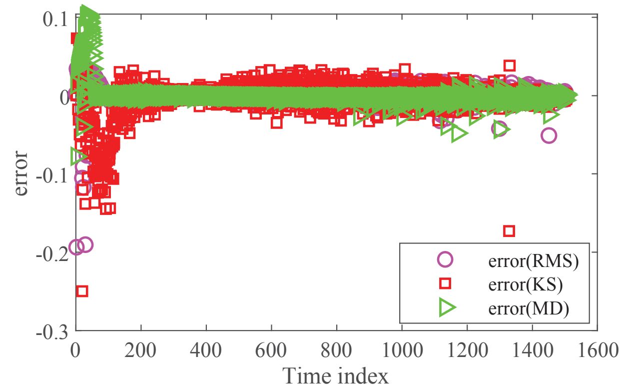

Figure 4. Tracking error using different HIs

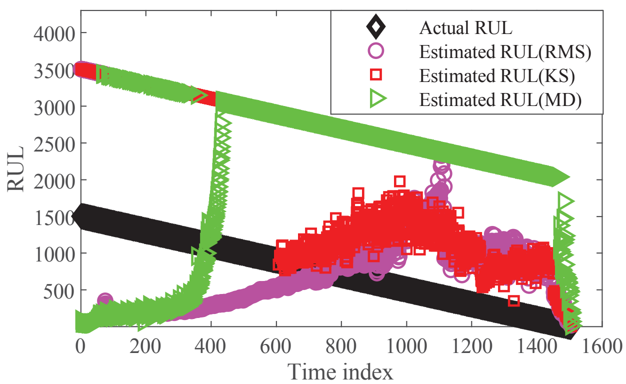

Figure 5. RUL estimation using different HIs by PF

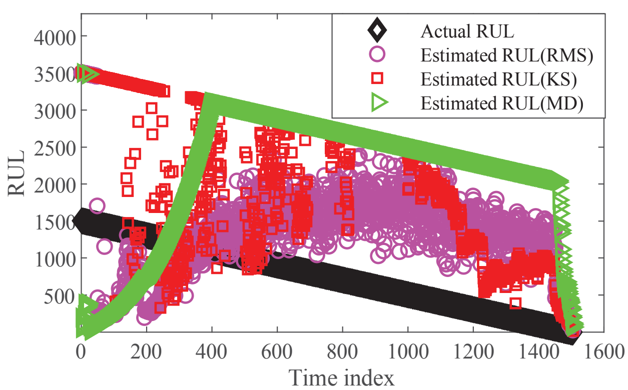

Figure 6. RUL estimation using different HIs by UKF

bearing's degradation model by referring to these three HIs.

After constructing the degradation model corresponding to different HIs, the PF is used to update model's parameters. The actual HIs and the estimated HIs are almost overlapping as shown in Fig. 3. Fig. 4 shows the tracking error corresponding to different HIs. Except for the initial time, the errors are small. This means that the estimated HI could converge the actual HIs fast with new observing data.

At last, the three HIs are used to predict bearing's RUL by using PF algorithm. The result of RUL estimation is shown in Fig.5. At the very beginning, RUL estimation of the three HIs has a large deviation due to the lack of observation data. When more and more bearing's online vibration signals are acquired, there have more data for updating the parameters of the nonlinear degradation model, the estimated RUL converge to the actual RUL. It should be noted that if the estimated HI does not reach the failure threshold after 2000 time indexes, bearing's RUL is defined as 2000. As it can be seen in Fig. 5, some estimated RULs in KS- and MD-HI cases deviate a lot from the actual RUL. The estimated RULs in RMS-HI case have the smallest deviations. Thus, it can be concluded that the RMS-based HI has the best performance in the RUL estimation. The unscented Kalman filter (UKF) algorithm is also used to predict bearing's RUL. The result is shown in Fig. 6. By comparing with Fig. 5, it can be found that PF has better performance than UKF in bearing RUL estimation.

V. CONCLUSION

RUL prediction for bearings plays a critical role in reliable operation of many rotary machines. Data-driven approach is a feasible way to predict bearing RUL. Different HIs and degradation models can be used to predict bearings' RUL. Which one is best? To compare the performances of different methods, three HIs (RMS-, KS- and MD-based HIs) are built and four degradation models are proposed to analyze bearing nonlinear degradation process in this paper. Bearing's RUL is predicted with the aid of PF finally. Results show that different HIs indicate different degradation characteristics of bearing, and the sum of two exponential functions has the best performance in the estimation of RUL of bearing.

REFERENCES

[1] X. Jin, M Zhao, T. W. S. Chow, and M. Pecht, "Motor bearing fault diagnosis using trace ratio linear discriminant analysis,"IEEE Trans. Ind. Electron., vol. 61, pp. 2441-2451, May 2014.

[2] X. Jin, W. Qiao, Y. Peng, F. Cheng, and L. Qu, "Quantitative evaluation of wind turbine faults under variable operational conditions," IEEE Trans. on Ind. Appl., vol. 52, no. 3, pp.2061-2069, May/June. 2016.

[3] Y. Wang, Y. Peng, Y. Zi, X. Jin, and K-L Tsui, "A two-stage data-driven based prognostic approach for bearing degradation problem," IEEE Trans. Ind. Inf., vol. 12, pp. 924-932, June 2016.

[4] A. Purarjomandiangrudi, G. Nourbakhsh, H. Ghaemmaghami, and A. Tan," Application of anomaly technique in wind turbine bearing fault detection," Conference of the IEEE industrial Electronics Society, 2014:1984-1988.

[5] X. Jin, Y. Sun, Z. J. Que, Y. Wang, and T. W. S. Chow, "Anomaly detection and fault prognosis for bearings," IEEE Trans. Instrum. Meas., vol. 65, pp. 2046-2054, Sep. 2016.

[6] X. Jin, Z. Que, Y. Sun, Y. Guo, and W. Qiao, "A data-driven approach for bearing fault prognostics," IEEE Trans. on Ind. Appl., vol.55, no. 4, pp.3394-3401, July/August 2019.

[7] M. Pecht and M. Kang, Prognostics and Health Management of Electronics: Fundamentals, Machine Learning, and the Internet of Things. New Jersey: John Wiley & Sons, 2018.

[8] X. S. Si, W. Wang, C. H. Hu, D. H. Zhou, and M. G. Pecht, "Remaining useful life estimation based on a nonlinear diffusion degradation process," IEEE Trans. Reliab., vol. 61, pp. 50-61, Mar. 2012.

[9] X.S. Si, W. Wang, C.H. Hu, and D.H. Zhou, "Estimating remaining useful life with three-source variability in degradation modeling,"IEEE Trans. Rel., vol. 63, pp. 167-190, Mar. 2014.

[10] X. Jin, F. Cheng, Y. Peng, W. Qiao, and L. Qu, "Drivetrain gearbox fault diagnosis: vibration- and current-based approaches," Ind. Appl. Mag., vol. 24, no.6, pp. 56-66, Dec. 2018.

[11] 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, pp. 647-656, Jan. 2015.

[12] N. Gebraeel, M. Lawley, R. Liu, and V. Parmeshwaran, "Residual life predictions from vibration-based degradation signals: A neural network approach," IEEE Trans. Ind. Electron., vol. 51, pp. 694-700, June 2004.

[13] X. Jin, E. W. M. Ma, L. L. Cheng, and M. Pecht, "Health monitoringof cooling fans based on Mahalanobis distance with mRMR feature selection," IEEE Trans. Instrum. Meas., vol. 61, pp. 2222–2229, Aug. 2012.

[14] J. Mei, M. Liu, Y. Wang, and H. Gao, "Learning a Mahalanobis distance-based dynamic time warping measure for Multivariate time series classification," IEEE Trans.Cybern., vol. 46, pp. 1363-1374, Jun. 2016.

[15] Z. Li and Q. He, "Prediction of railcar remaining useful life by multiple data source fusion," IEEE Trans. Intell. Transp. Syst., vol. 16, pp. 2225-2236, Aug. 2015.

[16] J. Wu, S. Song, W. An, D. Zhou, and H. Zhang, "Defect detection and localization of nonlinear system based on particle filter with an adaptive parametric model," Math. Probl. Eng, vol. 2015, pp. 1-12, Nov. 2015.

[17] M. Jouin, R. Gouriveau, D. Hissel, M. Pera, and N. Zerhouni, "Partical filter-based prognostics: Review, discussion and perspectives," Mech. Syst. Signal. Pr., vol. 72-73, pp. 2-31, 2016.

[18] M. Qiu, W. Li, F. Jiang, and Z. Zhu, "Remaining useful life estimation for rolling bearing with SIS-based indicator and particle filtering," IEEE Access, vol. 6, pp. 24531-24532, May. 2018.

[19] S. Yin and X. Zhu, "Intelligent particle filter and its application to fault detection of nonlinear system," IEEE Trans. Ind. Electron., vol. 62, June 2015.

[20] Z. Que, "Remaining useful life estimation for offshore wind turbine bearings," M.S. thesis, Department of Mechanical Engineering, Zhejiang University of Technology, Hangzhou, China, 2017.

[21] P. Nectoux, R. Gouriveau, K. Medjaher, E. Ramasso, B. Morello, N. Zerhoui, and C. Varnier, "PRONOSTIA: An experimental platform for bearings accelerated degradation tests," in Proc. IEEE Int. Conf. Prognostics Health Manag. (PHM), Denver, CO, USA, pp. 1-8, June 2012.