2021 CAA Symposium on Fault Detection, Supervision, and Safety for Technical Processes (SAFEPROCESS) | 978-1-6654-0115-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/SAFEPROCESS52771.2021.9693575

A Data Driven RUL Prediction and Predictive Maintenance for Stochastic Degrading Systems

Yuhan Zhang, Xianchao Xiu and Ying Yang

***Abstract*—This paper proposes a novel remaining useful life (RUL) prediction approach with application to predictive maintenance for stochastic degrading systems. By modeling the degradation state with a set of weighted basis functions, it can achieve a more flexible representation for modeling the nonlinear behaviors within degradation process. Then, the parameters including the weights is computed by a sparse Bayesian algorithm. After parameter identification, we obtain an explicit form of the RUL distribution based on the proposed degradation process model. Further, the optimal replacement time for predictive maintenance is obtained by minimizing the expected cost rate which is calculated with the aid of RUL distribution. Finally, the effectiveness of the proposed method in RUL prediction and predictive maintenance is verified by a numerical simulation.**

***Index Terms*—Wiener process (WP), Remaining useful life (RUL), predictive maintenance.**

I. Introduction

The prediction of residual effective life (RUL) has attracted extensive attention in the field of prognosis and health management (PHM), due to that it can provide essential instructions for maintenance decisions in PHM [1]. Because of the considerable ability of stochastic models in describing stochastic dynamics in degradation processes, stochastic-model-based degradation analysis and RUL prediction have gained much popularity among many researchers [2]. Stochastic models including the inverse Gaussian process [3], [4], Wiener process (WP) [5], [6], [7], [8], [9], [10] and Gamma process [11], [12] are used typically to model degradation processes. However, inverse Gaussian process and gamma process are only suitable for modeling degradation processes with monotonic path. Wiener process is suitable for modeling non-monotonic degradation process compared with them. Therefore, WP gains more popularity adapting to degradation modeling for degraded systems, such as batteries [8], bearings [7] and blast furnaces [13].

Si et al. [10] uses $dX(t) = \sigma(t;\boldsymbol{\theta})dB(t) + \mu(t;\boldsymbol{\theta})dt$ to model a nonlinear diffusion process for representing the stochasticity and nonlinearity within the degradation. Based on the model, [10] provides closed-formed approximation formulas for calculating the distributions of the first hitting time (FHT) when the degradation status crosses a preset threshold and its RUL. In [10], the exponential-law or power-law drift coefficient function $\mu(t;\boldsymbol{\theta}) = ab\cdot\exp(bt)$ or $\mu(t;\boldsymbol{\theta}) = ab\cdot t^{b-1}$ is used for modeling the nonlinear behaviors of degradation processes. However, in more actual case, some degradation processes possess more complex nonlinear behaviors that the two models given by [10] cannot describe appropriately. To solve this problem, [7] presents a more general modeling method for nonlinear degradation processes. It depicts the increment of the Wiener process with a sum of weighted kernel functions. Moreover, [7] utilizes a sparse Bayesian learning algorithm for calculating model's parameters. However, the method in [7] performs not well at the early stage of some degradation processes. To improve it in [8], a Wiener process (WP) whose drift increment is a weighted combination of basis functions is used to model the degradation process and depicts the nonlinearity within degradation trend. In addition, long short term memory (LSTM) network is introduced to capture the long-term dependence between offline and online degradation degradation samples, and the future degradation increment prediction is obtained. [7] and [8] both gave the numerical approximation computation formula for the probability density function (PDF) of RUL, but an explicit form of RUL distribution is essential for an timely replacement plan [14]. In this paper, a new degradation modeling method based on generalized nonlinear diffusion process model is proposed and an explicit form of RUL distribution calculation is obtained. The nonlinear behaviors of the model in this paper are modeled in degradation state itself directly, instead of being modeled in the increments of degradation in [7] and [8]. They are modeled with a series of basis functions which are weighted combined with each other. A sparse Bayesian algorithm is used to calculate the weights with the corresponding specific kernel functions. Then, a new RUL calculation theorem is deduced from the lemma in [10]. Finally, by adopting the optimal replacement time definition in [14], an optimal replacement plan can be developed for degrading systems.

The contributions of this proposed approach can be given as follows:

- In this paper, we model the degradation state with a sum of weighted basis functions. This provides flexible representation for nonlinear behaviors of degradation processes.
- This paper gives an explicit form of RUL distribution for RUL estimation.
- This paper extends the RUL distribution to calculate the optimal replacement time for predictive maintenance.

The paper has the following structure. In Section II, a novel model is presented for nonlinear degradation processes. Moreover, based on the model, the parameter calculation method

This work is supported by National Natural Science Foundation (NNSF) of China under Grant 12001019 and Grant 61633001. (*corresponding author: Ying Yang*)

Yuhan Zhang, Xianchao and Ying Yang are with College of Engineering, Peking University, Beijing 100871, P. R. China.(email: yuhanzhang@pku.edu.cn; yy@pku.edu.cn)

under a Bayesian framework is proposed. Section III gives gives the main theoretical results of RUL distribution calculation, which is followed by the corresponding predictive maintenance strategy. Finally, Section VII concludes this paper.

II. Degradation Process Modeling

In order to flexibly represent the nonlinear behaviors of degradation process, we present a general nonlinear degradation model based the diffusion process as follows. $X(t)$ denotes the degradation state at time t.

$$X(t) = X(0) + \sum_{i=1}^{m} \omega_i K(t, t_i) + \omega_0 t + \sigma B(t), \tag{1}$$

where $X(0)$ denotes the initial value of degradation state that can be converted to 0 without loss of generality. The diffusion coefficient σ is constant. $B(t)$ is a standard Brownian motion (BM). $\sum_{i=1}^{m} \omega_i K(t,t_i) + \omega_0 t$ represents the time-dependent drift coefficient function, where Gaussian basis function is usually selected as $K(t, t_i)$.

$$K(t, t_i) = exp\left(-\frac{(t-t_i)^2}{q^2}\right). \tag{2}$$

Relying on (1) and (2), the parameter identification algorithm is derived under a sparse Bayesian framework as follows.

$$X(t_k) = X(t_{k-1}) + \sum_{i=1}^{m} \omega_i (K(t_k, t_i) - K(t_{k-1}, t_i)) + \omega_0 \tau + \eta(t_k), \tag{3}$$

where $t_k - t_{k-1} = \tau$, $\eta(t_k) = \sigma B(t_k) - \sigma B(t_{k-1}) \sim N(0, \sigma^2 \tau)$ due to the property of Brownian motion $B(t)$. Because $\eta(t_k)$ is Gaussian distributed and independent, the degradation increments $\Delta X_\tau(t_k) = X(t_k) - X(t_{k-1})$ are independent and we can compute the parameters in (1) with the sparse Bayesian algorithm using the increment data set of paired input-targets $\{t_k, \Delta X_\tau(t_k)\}_{k=1}^{m}$.

We follow the assumption that $\Delta X_\tau(t_k)$ is independent. Then, the likelihood of the complete data set can be obtained as follows.

$$p(\boldsymbol{\Delta X}|\boldsymbol{\omega}, \sigma^2\tau) = (2\pi\sigma^2\tau)^{-m/2} exp\left(-\frac{\|\boldsymbol{\Delta X} - \boldsymbol{\Phi\omega}\|^2}{2\sigma^2\tau}\right), \tag{4}$$

where $\boldsymbol{\Delta X} = [\Delta X_\tau(t_1), \ldots, \Delta X_\tau(t_m)]^T$, $\boldsymbol{\omega} = [\omega_0, \omega_1, \ldots, \omega_m]^T$, and $\boldsymbol{\Phi}$ is an $m \times (m+1)$ matrix with $\boldsymbol{\Phi} = [\phi(t_1), \phi(t_2), \ldots, \phi(t_m)]^T$ and $\phi(t_i) = [\tau, K(t_i, t_1) - K(t_i - \tau, t_1), \ldots, K(t_i, t_m) - K(t_i - \tau, t_m)]$.

The direct maximum likelihood estimation (MLE) of $\boldsymbol{\omega}$ and σ from (4) may result in over-fitting [15]. For avoiding this, we set additional constraints on the parameters by explicitly defining a Gaussian prior distribution with the mean of zero over $\boldsymbol{\omega}$.

$$p(\boldsymbol{\omega}|\boldsymbol{\alpha}) = \prod_{i=0}^{m} \frac{\alpha_i}{\sqrt{2\pi}} exp\left(-\frac{\omega_i^2 \alpha_i^2}{2}\right), \tag{5}$$

where $\boldsymbol{\alpha}$ is the $m+1$-dimension hyper-parameter vector, namely $\boldsymbol{\alpha} = [\alpha_0, \alpha_1, \ldots, \alpha_m]^T$.

Therefore, we can derive the posterior distribution over ω by

$$p(\boldsymbol{\omega}|\boldsymbol{\Delta X}, \boldsymbol{\alpha}, \sigma^2\tau) = (2\pi)^{-\frac{m+1}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} exp\{-\frac{(\boldsymbol{\omega} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\omega} - \boldsymbol{\mu})}{2}\}, \tag{6}$$

where the posterior mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ are given as follows.

$$\boldsymbol{\mu} = \sigma^{-2}\tau^{-1} \boldsymbol{\Sigma\Phi}^T \boldsymbol{\Delta X}, \tag{7}$$

$$\boldsymbol{\Sigma} = (\sigma^{-2}\tau^{-1} \boldsymbol{\Phi}^T \boldsymbol{\Phi} + \boldsymbol{A})^{-1}, \tag{8}$$

where $\boldsymbol{A} = diag(\alpha_0, \alpha_1, \ldots, \alpha_m)$.

We maximize the following marginal likelihood and derive the optimal values of α and σ. The marginal likelihood can be determined by

$$p(\boldsymbol{\Delta X}|\boldsymbol{\alpha}, \sigma^2\tau) = \int p(\boldsymbol{\Delta X}|\boldsymbol{\omega}, \sigma^2\tau) \cdot p(\boldsymbol{\omega}|\boldsymbol{\alpha}) d\boldsymbol{\omega}, \tag{9}$$

then the marginal likelihood can be computed by:

$$p(\boldsymbol{\Delta X}|\boldsymbol{\alpha}, \sigma^2\tau) = (2\pi)^{-\frac{m}{2}} \cdot |\sigma^2\tau \boldsymbol{I} + \boldsymbol{\Phi A}^{-1}\boldsymbol{\Phi}^T|^{-\frac{1}{2}} \cdot exp\left(-\frac{1}{2} \boldsymbol{\Delta X}^T (\sigma^2\tau\boldsymbol{I} + \boldsymbol{\Phi A}^{-1} \boldsymbol{\Phi}^T)^{-1} \boldsymbol{\Delta X}\right). \tag{10}$$

Due to the fact that acquiring the closed-form optimal $\boldsymbol{\alpha}$ and σ by maximizing (10) is difficult, we re-estimate the solution of this optimization problem iteratively. $\boldsymbol{\alpha}$ equals 0 with the differential of (10). Based on the method in [15], we can get the following results.

$$\alpha_i^{new} = \frac{\gamma_i}{\mu_i^2}, \tag{11}$$

where μ_i is the ith posterior mean weight, and γ_i is defined as

$$\gamma_i = 1 - \alpha_i \Sigma_{ii}, \tag{12}$$

Where σ_{ii} denotes the the ith diagonal element of the posterior weight covariance calculated by (8) with the current $\boldsymbol{\alpha}$ and σ^2. Next, we update σ^2, and the differentiation makes the following re-estimation

$$(\sigma^2\tau)^{new} = \frac{\|\boldsymbol{\Delta X} - \boldsymbol{\Phi\mu}\|^2}{m - \Sigma_{i=1}^{m} \gamma_i}, \tag{13}$$

where m represents the number of data samples.

Therefore, when (11) and (12) are applied repeatedly, the algorithm updates the posterior statistics μ and $\Sigma$of (8) and (7) simultaneously until some appropriate convergence criteria are met. Then, $\boldsymbol{\omega}$ by (7) and σ by (13) can be calculated. Therefore, $\boldsymbol{\theta_{RVM}}$ and σ are calculated.

When the parameter estimation process converges, we can use another new test sample t_*, the maximum value μ_{MP} and σ_{MP}^2p of the corresponding degradation increment target ΔX_* can be found according to the prediction of the posterior distribution on ω. Further, the distribution of ΔX_* is a normal distribution like

$$p(\Delta X_* | \Delta X, \boldsymbol{\alpha_{MP}}, \sigma_{MP}^2 \tau) = N(\Delta X_* | f_*, \sigma_*^2 \tau) \tag{14}$$

where the mean f_* and the variance $\sigma_*^2\tau$ are respectively

$$f_* = \boldsymbol{\mu}^T \boldsymbol{\phi}(\boldsymbol{t_*}) \tag{15}$$

$$\sigma_*^2\tau = \sigma_{MP}^2\tau + \boldsymbol{\phi}(t_*)^T \boldsymbol{\Sigma}\boldsymbol{\phi}(t_*) \tag{16}$$

The predicted mean is the sum of the basis functions, weighted by a posterior mean weight, most of which are usually zero to achieve sparsity. The training vectors corresponding to the remaining non-zero weights are called relevance vectors.

III. RUL ESTIMATION AND PREDICTIVE MAINTENANCE

In this section, we investigate the derivation of the RUL distribution. First, we introduce the definition of lifetime and RUL of degradation processes. The lifetime is usually defined by the first hitting time (FHT) of the degradation state exceeding a preset threshold J

$$T = inf\{t : X(t) \geq J | X(0) \leq J\}. \tag{17}$$

Based on FHT, the RUL at t_k time point can be defined as

$$L_k = inf\{l : X(t_k + l) \geq J | x(t_k) \leq J\}. \tag{18}$$

Both the life time T and the RUL L_k defined by (17) and (18) are considered as random variables because of the stochastic property in the degradation process. For deriving the distribution of RUL, we first introduce the following lemma.

Lemma 1. *The PDF of the FHT for $X(t)$ exceeding a constant threshold can be approximated as*

$$p_{X(t)}(\xi, t) \cong \frac{1}{\sqrt{2\pi t}} \cdot \left(\frac{S(t)}{t} + \frac{1}{\sigma} \cdot \mu(t; \boldsymbol{\theta})\right) \cdot exp\left(-\frac{S^2(t)}{2t}\right), \tag{19}$$

where

$$S(t) = \frac{1}{\sigma} \cdot \left(\xi - \int_0^t \mu(\tau; \boldsymbol{\theta}) d\tau\right). \tag{20}$$

Then, we obtain the following Theorem to compute the RUL distribution.

Theorem 1. *Given the degradation measurement $X(t_k)$ that satisfies $X(t) = X(0) + \sum_{i=1}^m \omega_i K(t, t_i) + \omega_0 t + \sigma B(t)$, the PDF of the RUL of $X(t)$ can be formulated at time t_k as*

$$\begin{aligned} f_L(l) =& \frac{1}{\sqrt{2\pi l}} \cdot exp\left(-\frac{S^2(l)}{2l}\right) \\ & \cdot \left(\frac{S(l)}{l} + \sum_{i=1}^m \frac{\omega_i}{\sigma} \cdot \frac{dK(l + t_k, t_i)}{dl}\right), \end{aligned} \tag{21}$$

where

$$\begin{aligned} S(l) =& \frac{1}{\sigma} \cdot (\xi - X(t_k)) \\ & - \frac{1}{\sigma}\left(\omega_0 l + \sum_{i=1}^m \omega_i \left(K(l + t_k, t_i) - K(t_k, t_i)\right)\right). \end{aligned} \tag{22}$$

Based on Theorem 1, in the typical degradation case $X(t) = X(0) + \sum_{i=1}^m \omega_i K(t, t_i) + \omega_0 t + \sigma B(t)$ where $K(t, t_i) = exp\left(-\frac{(t-t_i)^2}{q^2}\right)$, the following corollary is derived.

Corollary 1. *For the degradation process $X(t) = X(0) + \sum_{i=1}^m \omega_i K(t, t_i) + \omega_0 t + \sigma B(t)$ where $K(t, t_i) = exp\left(-\frac{(t-t_i)^2}{q^2}\right)$, the PDF of the RUL is formulated at time t_k with an explicit form as*

$$\begin{aligned} f_L(l) =& \frac{1}{\sqrt{2\pi l}} \cdot exp\left(-\frac{S^2(l)}{2l}\right) \\ & \cdot \left(\frac{S(l)}{l} + \sum_{i=1}^m \frac{2\omega_i(l + t_k - t_i)}{\sigma * b^2} \cdot exp\left(-(\frac{l + t_k - t_i}{b})^2\right)\right), \end{aligned} \tag{23}$$

where

$$\begin{aligned} S(l) =& \frac{1}{\sigma} \cdot (\xi - X(t_k)) \\ & - \frac{1}{\sigma}\left(\omega_0 l + \sum_{i=1}^m \omega_i \left(K(l + t_k, t_i) - K(t_k, t_i)\right)\right), \\ & and \quad K(t, t_i) = exp\left(-\frac{(t - t_i)^2}{q^2}\right). \end{aligned} \tag{24}$$

Algorithm 1 RUL Estimation and Predictive Maintenance

Input:

$X_{t_1:t_k} = \{X(t_1), ..., X(t_k)\}$: Degradation history measurements to t_k;

Output:

$f_L(t_k)$: RUL's PDF at t_k;

$C_R(t_k, T_R)$: the expected cost per unit time at t_k;

1: Design the step length τ, then obtain the degradation increment data set $\{t_k, \Delta X_\tau(t_k)\}_{k=1}^m$;

2: Initialize $\{\alpha_i\}$ and σ^2;

3: **repeat**

4: Calculate $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ by (7) and (8);

5: Calculate all $\{\gamma_i\}$ by (12), then re-estimate $\{\alpha_i\}$ and σ^2 by (11) and (13);

6: **until** convergence

7: Predict new drift increment by (14);

8: Delete basis functions for an optimal $\alpha_i = \infty$;

9: Calculate RUL's PDF at t_k by (23);

10: Calculate CDF value F_L by (27);

11: Compute the expected cost per unit time $C_R(t_k, T_R)$ at t_k by (25);

12: **return** $f_L(t_k)$ and $C_R(t_k, T_R)$

Based on Corollary 1, we adopt the replacement decision for the degrading system at the t_k CM point in [14] via the minimisation of the following expected cost per unit time, i.e. the expected cost rate.

$$C_R(t_k, T_R) = \frac{C_f F_L(T_R) + C_p[1 - F_L(T_R)]}{t_k + \int_0^{T_R}[1 - F_L(z)]dz}, \tag{25}$$

where F_L is the cumulative density function (CDF) of RUL. C_p refers to the cost of a preventive replacement, and C_f represents the replacement cost associated with the failure of an individual component of the degrading system. Then, the optimal replacement time T_R^* is

$$T_R^* = minC_R(t_k, T_R) \tag{26}$$

Due to the complexity of f_L, the value of F_L with respect to T_R is obtained by calculating the integral of f_L with the numerical technique as follows.

$$F_L(T_R) = \sum_{i=1}^{\frac{T_R}{\tau}} f_L(i * \tau). \tag{27}$$

Then, T_R^* is the minimum point of $C_R(t_k, T_R)$, where $\int_0^{T_R}[1-F_L(z)]dz$ also needs a numerical approximation similar to (27).

Moreover, the RUL estimation and predictive maintenance algorithm is summarized as Algorithm 1.

IV. Numerical Example

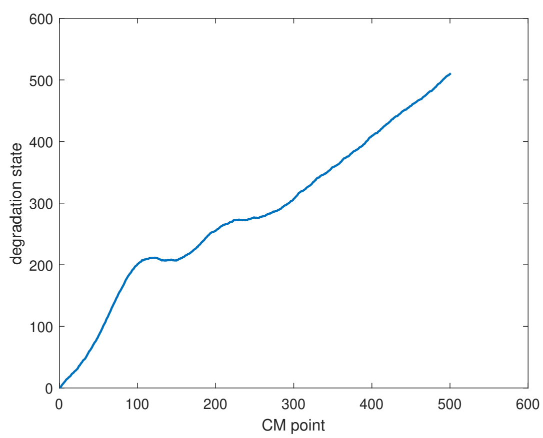

Fig. 1. Degradation process.

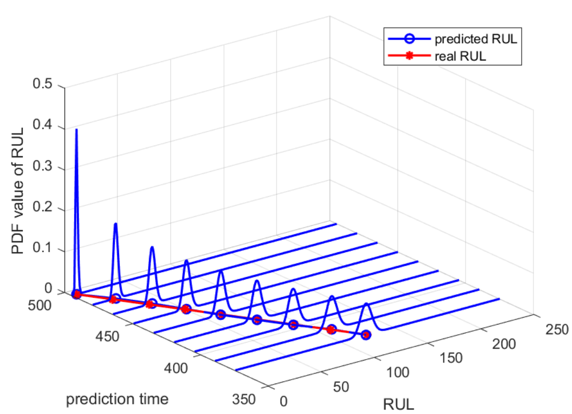

Fig. 2. Predicted RUL distribution.

One set of numerically simulated data is used to verify the proposed method. The degradation process is driven by (1) where the basis function is Gussian basis function (2) where $q = 50$ and the weights are most zeros and a small amount of nonzero constants that $\omega_0 = 1, \omega_{100} = 100$ and $\omega_{200} = 50$. Besides, the diffusion coefficient is $\sigma = 0.5$. Fig.1 shows that the degradation state $X(t)$ crosses the failure threshold at the $500th$ CM point along such a nonlinear stochastic degradation trend. The non-zero weights of individual basis functions lead to the complex nonlinearity of the degradation process. The results of modeling this degradation process using a common nonlinear modeling approach like [10] will lead to bias of RUL estimation. this case requires an effective method as the proposed method.

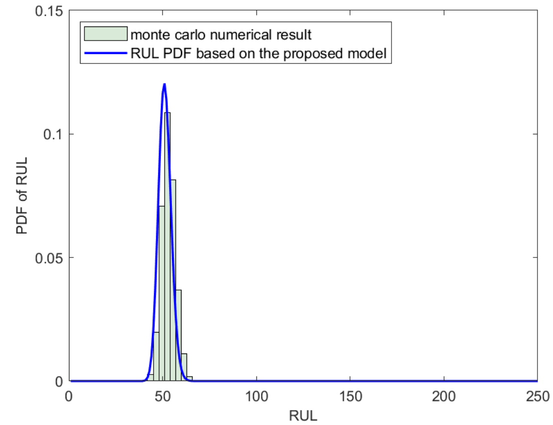

Fig. 3. Comparison result of the RUL distribution of monte carlo numerical result and the proposed method at the 450th CM point.

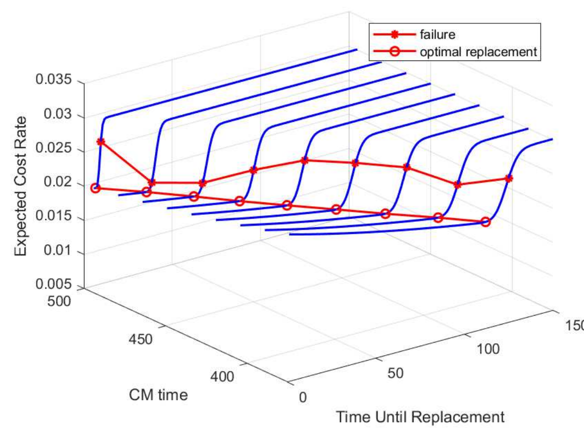

Fig. 4. The evaluation of the replacement decision at each CM point.

Preparing for identifying the parameters with this set of data, we set $\tau = 1$ and apply some basic data preprocessing work to obtain the degradation increment data. For testing the proposed method for RUL estimation, Fig.2 shows the estimated RUL distributions at each CM point. The blue circles are the mean value of the RUL at each CM point, i.e the predicted RUL results. The red asterisks represent the real RUL at these CM points. The proximity of the blue circles and the red asterisks can prove the validity of the proposed method in RUL estimation.

Further, we compare the proposed method and the corresponding Monte Carlo simulation result to test its effectiveness. The RUL estimation result using our proposed method is compared to the RUL results using Monte Carlo method with 10000 simulated sample paths. It can be seen in Fig.3

that our RUL estimation result is basically consistent with the numerical RUL distribution. Moreover, our proposed method can approximate the distribution of the RUL well.

Finally, we evaluate the proposed method's replacement decision at the ith CM point with the lifetime of the component, the average cost of failure and preventive replacement, and RUL's distribution. The ideal characteristic of the replacement decision is to provide the maximum possible operational availability, while component failures depend on the fitting of the actual underlying RUL distribution. In this numerical example, the cost of a preventive replacement is set as $C_p = 10$, and the replacement cost of the component is $C_f = 15$. After utilizing Algorithm 1, Figs. 4 shows all optimal replacement time, i.e. the minimum points of expected cost rate is earlier than the failure time. Therefore, it demonstrates the proposed method gives reasonable replacement decisions at each CM points.

V. Conclusion

This paper presents a new RUL prediction method for stochastic degrading systems, which is extended to the predictive maintenance. It adopts a weighted combination of basis function for modeling the degradation process. The parameters of the model are identified by a sparse Bayesian algorithm. Then, the derived explicit form of RUL distribution facilitates the calculation of the optimal replacement time for predictive maintenance. A numerical case demonstrates its effectiveness.

References

[1] M. Pecht, "Prognostics and health management of electronics," *Encyclopedia of Structural Health Monitoring*, 2009.

[2] Z. Zhang, X. Si, C. Hu, and Y. Lei, "Degradation data analysis and remaining useful life estimation: A review on wiener-process-based methods," *European Journal of Operational Research*, vol. 271, no. 3, pp. 775–796, 2018.

[3] Z.-S. Ye and N. Chen, "The inverse gaussian process as a degradation model," *Technometrics*, vol. 56, no. 3, pp. 302–311, 2014.

[4] W. Peng, Y.-F. Li, Y.-J. Yang, H.-Z. Huang, and M. J. Zuo, "Inverse gaussian process models for degradation analysis: A bayesian perspective," *Reliability Engineering & System Safety*, vol. 130, pp. 175–189, 2014.

[5] C.-H. Hu, H. Pei, X.-S. Si, D.-B. Du, Z.-N. Pang, and X. Wang, "A prognostic model based on dbn and diffusion process for degrading bearing," *IEEE Transactions on Industrial Electronics*, 2019.

[6] N. Li, Y. Lei, T. Yan, N. Li, and T. Han, "A wiener-process-model-based method for remaining useful life prediction considering unit-to-unit variability," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 3, pp. 2092–2101, 2018.

[7] Y. Zhang, Y. Yang, H. Li, X. Xiu, and W. Liu, "A data-driven modeling method for stochastic nonlinear degradation process with application to rul estimation," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–12, 2021.

[8] Y. Zhang, Y. Yang, X. Xiu, H. Li, and R. Liu, "A remaining useful life prediction method in the early stage of stochastic degradation process," *IEEE Transactions on Circuits and Systems II: Express Briefs*, pp. 1–1, 2020.

[9] Q. Guan, Y. Tang, and A. Xu, "Objective bayesian analysis accelerated degradation test based on wiener process models," *Applied Mathematical Modelling*, vol. 40, no. 4, pp. 2743–2755, 2016.

[10] 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 Transactions on Reliability*, vol. 61, no. 1, pp. 50–67, 2012.

[11] J. H. Cha and G. Pulcini, "Optimal burn-in procedure for mixed populations based on the device degradation process history," *European Journal of Operational Research*, vol. 251, no. 3, pp. 988–998, 2016.

[12] M. H. Ling, H. Ng, and K. L. Tsui, "Bayesian and likelihood inferences on remaining useful life in two-phase degradation models under gamma process," *Reliability Engineering & System Safety*, vol. 184, pp. 77–85, 2019.

[13] Z. Hanwen, C. Maoyin, and Z. Donghua, "Remaining useful life prediction for a nonlinear multi-degradation system with public noise," *Journal of Systems Engineering and Electronics*, vol. 29, no. 2, pp. 429–435, 2018.

[14] M. J. Carr and W. Wang, "An approximate algorithm for prognostic modelling using condition monitoring information," *European journal of operational research*, vol. 211, no. 1, pp. 90–96, 2011.

[15] M. E. Tipping, "Sparse bayesian learning and the relevance vector machine," *Journal of machine learning research*, vol. 1, no. Jun, pp. 211–244, 2001.