Article

A Prognostic Framework for Wheel Treads Integrating Parameter Correlation and Multiple Uncertainties

Guifa Huang ^{1,2}, Yu Zhao ¹, Han Wang ^{1,*}, Xiaobing Ma ¹ and Deyao Tang ²

¹ School of Reliability and Systems Engineering, Beihang University, Beijing 100191, China; hgf@tangzhi.com (G.H.); zhaoyu@buaa.edu.cn (Y.Z.); maxiaobing@buaa.edu.cn (X.M.)

² Beijing Tangzhi Science and Technology Development Co., Ltd., Beijing 100043, China; tdy@tangzhi.com

* Correspondence: wh216@buaa.edu.cn

Received: 28 November 2019; Accepted: 6 January 2020; Published: 8 January 2020

Abstract: As crucial rotary components of high-speed trains, wheel treads in realistic operation environment usually suffer severe cyclic shocks, which damage the health status and ultimately cause safety risks. Timely and precise health prognosis based on vibration signals is an effective technology to mitigate such risks. In this work, a new parameter-related Wiener process model is proposed to capture multiple uncertainties existed in on-site prognosis of wheel treads. The proposed model establishes a quantitative relationship between degradation rate and variations, and integrates uncertainties via heterogeneity analysis of both criterions. A maximum-likelihood-based method is presented to initialize the unknown model parameters, followed by a recursive update algorithm with fully utilization of historical lifetime information. An investigation of real-world wheel tread signals demonstrates the superiority of the proposed model in accuracy improvement.

Keywords: condition monitoring; remaining useful life; Wiener process model; recursive algorithm; multiple uncertainty; wheel tread

1. Introduction

As a fast and convenient transportation asset, high-speed trains have drawn world-wide attention in recent years. Under high-speed operating conditions, health monitoring for rotary machinery systems is a crucial technology to guarantee the safety and reliability of high speed trains [1]. With the rapid development of sensor technologies, multiple sources of operational/health data can be collected, supporting online diagnosis and prognosis [2–4]. For high-speed trains, engineers collect vibration signals of rotary machinery systems such as bearings, gears, axles and wheel treads via on-board monitoring device and then conduct fault recognition, location and isolation based on signal processing technologies. However, diagnosis only provides failure information and provides no guidance for preventive maintenance, which is quite crucial to mitigate malfunction risks. Therefore, it is of urgent necessary to develop advanced prognostic methodologies, which predicts real-time health information to support condition-based maintenance plan.

Remaining useful lifetime (RUL) is a main objective of health prognostic, which measures the distance between the latest monitoring time and system malfunction. Currently, RUL prediction methodologies can be substantially partitioned into two types, data-driven approaches and model-based approaches [5–7]. Data-driven approaches immediately utilize historical sensor data as input and outputs lifetime information. Representative data-driven technology is machine learning, including neural network [8–11], support vector machine [12,13] and artificial intelligence approaches [14–16]. Such technologies can adopt diverse application scenarios, but heavily rely on the quality and size of data. Model-based approaches, on the other hand, utilize statistical or physical modes to characterize the degradation paths. Combined with the measured data, model parameters can be identified and

the future degradation behavior can be predicted. So far, lots of RUL prediction models have been developed for rotary machinery systems such as degradation path models [17,18], Wiener-process-based models [19,20], Gamma-process-based models [21,22], as well as some physical models. Particularly, Wiener-process-based models have received more attention as they provide a flexible description of non-monotonous process, which are frequently encountered in practice. Zhang et al. conducted a comprehensive review of the development of Wiener-process-based models for degradation analysis and RUL estimation [23]. Li et al. proposed a novel Wiener-process-based exponential model to describe the degradation signals of bearings [24]. In their research, the particle filtering algorithm is used to estimate the degradation of the bearings and therefore reduce random errors of the stochastic process. Wang et al. proposed a novel two-stage Wiener-process-based linear model to interpret two-phase degradation behaviors of bearings [25]. Different from other studies, they incorporated stage correlation into the prior distributions of model parameters to improve the prediction accuracy. More detailed discussions on Wiener-process-based models can be found in [26–28]. Except for data-driven approaches and model-based approaches, some other methods based on signal processing can also be used to predict remaining useful life of structures such as vibro-acoustic modulation [29,30] and so on.

Despite the advancement of RUL technology, its application in high-speed wheel train is rather limited without consideration of actual operational environment and failure mechanism. As a susceptible part of a high-speed train, wheel treads are exposed to severe cyclic shocks causing several failure modes including stripping, crack propagation, out-of-round, internal defect, etc. Any failure mode may cause an immediate breakdown of the entire system and leads to catastrophic consequences. Furthermore, condition monitoring of wheel treads has two notable features: (a) there exists multiple uncertainty sources including failure dependence, inherent temporal uncertainty, unit-to-unit heterogeneity due to materials and measurement errors and (b) the variation degree of vibration sensors is closely related to degradation velocity. Currently, there lacks a systematic prognostic framework that can effectively capture both features of wheel treads.

To address the above-mentioned challenges, this paper proposes a parameter-related Wiener-process-based prognostic framework. Unlike existing RUL prediction models, the proposed model relates degradation variation to degradation rate quantitatively. To the best of our knowledge, this is the first attempt to formulate such correlation, which better reflects evolution pattern of degradation. Furthermore, multiple uncertainties including unit-to-unit uncertainty, failure mode uncertainty, measurement uncertainty as well as temporal uncertainty can be integrated into our framework, which better characterizes the health status of treads under non-stationary operation environment. Based on this framework, a recursive updating algorithm based on Kalman filter (KF) and expectation maximization (EM) technologies is proposed to realize dynamical prediction of RUL.

The main contributions of this paper are threefold.

- (1) A novel parameter-related Wiener process model with consideration of multiple uncertainties is proposed for accurately characterizing the real degradation process of wheel treads.
- (2) A recursive algorithm, which integrates KF and EM algorithm, is established to update model parameters and RUL distribution with the updating of online monitoring data.
- (3) An investigation of real-world wheel tread signals is used to demonstrate the superiority of the proposed prognostic framework in accuracy improvement.
- (4) The rest of the paper is organized as follows. Section 2 introduces the proposed RUL prediction methodology. Section 3 conducts experiment using real monitoring data of wheel treads of high-speed train. Section 4 draws the main conclusions.

2. Methodology

We collected vibration signals from vibration sensors mounted on the axial boxes and on-board monitoring system of a train. Root mean square (RMS) of original vibration signals is selected as

the health index since it effectively describes the signal strength and reveals the impact of failure development on vibration energy. A careful review of RMS data states that there are two degradation phases, i.e., normal operation phase I and rapid degradation phase II, as shown in Figure 1. In phase I, the RMS fluctuates trivially, indicating a relatively healthy state. In phase II, the degradation speed grows abruptly until the occurrence of failure. The demarcation point between the two phases is defined as phase changing time (PCT).

Figure 1. A descriptive illustration of the degradation process of wheel tread.

Precise detection of PCT is important since it influences the lifetime prediction accuracy. We introduce a 3σ -interval criterion for screening of PCT. Assume that RMS in phase I follows a normal distribution with mean μ and variance σ^2 . With both μ and σ obtained from historical data, the 3σ -interval of the extracted indicator can be constructed as $[\mu - 3\sigma, \mu + 3\sigma]$. We used this interval to monitor the state of the wheel tread by a comparison with latest RMS data. The wheel tread was treated normal if the RMS value fell within the interval, and abnormal otherwise. The abnormal state might be caused by either a fault or random noises. However, when multiple consecutive data were detected as abnormal, there was a high probability that a fault had occurred. In light of this, the following procedure was proposed for PCT detection (Algorithm 1).

Algorithm 1. PCT detection procedure

- (1) Let $w = 0$, find the first time τ_0 when the extracted RMS value exceeds 3σ -interval;
 - (2) Let $w = w + 1$, find the time τ_w when the consecutive $w + 1$ RMS values $\{y_{f+k}\}_{k=0:w}$ exceed 3σ -interval, i.e., $\{|y_{f+k} - \mu| > 3\sigma\}_{k=0:w}$;
 - (3) Compare τ_w with τ_{w-1} . If $\tau_w \neq \tau_{w-1}$, return to Step (2); else, output τ_w as the PCT value.
-

According to the detected PCT, we could divide the whole degradation process of the wheel tread into two phases and the RMS values in phase II were further used for RUL prediction.

2.1. RUL Prediction Model

For a specific unit, let $v(t)$ denote the degradation rate of the extracted RMS. We constructed $v(t)$ as

$$v(t) = \lambda \cdot \varphi(t; \theta) + \sqrt{m\lambda} \cdot dB(t), \quad (1)$$

where λ represents the basic degradation rate, $\varphi(t; \theta)$ is a function of time with parameter θ , m is a constant and $B(t)$ is the standard Brownian motion. Below are some interpretations of Equation (1).

- (1) The basic degradation rate λ is determined by material properties. Due to the heterogeneity among different units, λ is a random variable instead of deterministic. Here Gaussian distribution is chosen to describe such uncertainty, i.e., $\lambda \sim N(\mu_\lambda, \sigma_\lambda^2)$ [31];
- (2) The function of time $\varphi(t; \theta)$ is determined by specific failure mode;
- (3) Standard Brownian motion $B(t)$ represents temporal uncertainty of the degradation process.

Accordingly, the constructed degradation rate $v(t)$ follows a normal distribution, i.e., $v(t) \sim N(\lambda \cdot \varphi(t; \theta), m\lambda)$. Note that $\lambda \cdot \varphi(t; \theta)$ denotes the deterministic term related to material performance and failure mode, and $\sqrt{m\lambda} \cdot dB(t) \sim N(0, m\lambda)$ denotes the stochastic term which describes random errors caused by inhomogeneity of material. The fluctuation caused by inhomogeneity of material is proportional to the material itself.

Taking integration on both sides of (1), we could obtain the degradation path as

$$x(t) = \int_0^t v(\tau) d\tau = x_0 + \lambda \int_0^t \varphi(\tau; \theta) d\tau + \sqrt{m\lambda} B(t), \quad (2)$$

where x_0 describes initial value. Without loss of generality, we assumed $x_0 = 0$. Since measurement errors are inevitable in condition monitoring, the observation function is formulated as

$$y(t) = x(t) + \varepsilon, \quad (3)$$

where ε describes measurement uncertainty, which follows a normal distribution with zero mean and variance σ_m^2 .

A notable feature of the above-mentioned degradation model is that it integrates multiple uncertainties extensively existed in wheel tread, which can better reflect the actual health condition. In addition, the model relates degradation variation (described by the diffusion parameter in Equation (2), i.e., $\sqrt{m\lambda}$) to degradation rate (described by the drift parameter in Equation (2), i.e., λ) quantitatively. This is consistent with actual observation that a severer degradation provides an acute vibration variation [28].

2.2. Parameter Initialization and Updating Algorithm

RUL prediction starts with offline parameter estimation. A total of c monitoring time points starting from the PCT, i.e., $t_1, t_2, \dots, t_c$, were selected as the training set, as shown in Figure 2. Maximum-likelihood was chosen as the estimation algorithm.

Figure 2. Offline parameter initialization and online remaining useful lifetime (RUL) prediction in phase II.

Denote RMS values at these time points as $\mathbf{y}_c = [y_1, \dots, y_c]'$, where $[\cdot]'$ represents the vector transposition, and $y_i = y(t_i)$, $i = 1, 2, \dots, c$. Based on (3), y_i can be expressed as

$$y_i = x_0 + \lambda \int_0^{t_i} \varphi(\tau; \theta) d\tau + \sqrt{m\lambda} + \varepsilon. \quad (4)$$

Proposition 1. Let $\mathbf{t}_c = [\int_0^{t_1} \varphi(\tau; \theta) d\tau, \dots, \int_0^{t_c} \varphi(\tau; \theta) d\tau]'$, then $\mathbf{y}_c = [y_1, \dots, y_c]'$ is multivariate normally distributed, characterized by

$$\mathbf{y}_c \sim \text{MN}(\mu_\lambda \mathbf{t}_c, \sigma_\lambda^2 \mathbf{t}_c \mathbf{t}_c' + \mathbf{\Omega}), \quad (5)$$

with $\mathbf{\Omega} = m\mu_\lambda \mathbf{Q}_c + \sigma_m^2 \mathbf{E}_c$, where $\mathbf{Q}_c = [\min\{t_i, t_j\}]_{1 \leq i, j \leq c}$, $\mathbf{E}_c = \text{diag}\{1, \dots, 1\}_{c \times c}$ and MN represents the multivariate normal distribution.

Proof of Proposition 1 can be found in Appendix A. Denote unknown model parameters as $\Theta = [\mu_\lambda, \sigma_\lambda^2, \theta, m, \sigma_m^2]'$, then the log-likelihood function based on the extracted RMS up to t_c is

$$\ln l(\Theta | \mathbf{y}_c) = -\frac{c}{2} \ln(2\pi) - \frac{1}{2} \ln |\sigma_\lambda^2 \mathbf{t}_c \mathbf{t}_c' + \mathbf{\Omega}| - \frac{1}{2} (\mathbf{y}_c - \mu_\lambda \mathbf{t}_c)' (\sigma_\lambda^2 \mathbf{t}_c \mathbf{t}_c' + \mathbf{\Omega})^{-1} (\mathbf{y}_c - \mu_\lambda \mathbf{t}_c). \quad (6)$$

For simplification, a parameter transformation is carried out for $\mathbf{\Omega}$, i.e.,

$$\tilde{\mathbf{\Omega}} = \mathbf{\Omega} / \sigma_\lambda^2 = \frac{m\mu_\lambda}{\sigma_\lambda^2} \mathbf{Q}_c + \frac{\sigma_m^2}{\sigma_\lambda^2} \mathbf{E}_c, \quad (7)$$

and the log-likelihood function is transformed into

$$\ln l(\Theta | \mathbf{y}_c) = -\frac{c}{2} \ln(2\pi) - \frac{c}{2} \ln \sigma_\lambda^2 - \frac{1}{2} \ln |\mathbf{t}_c \mathbf{t}_c' + \tilde{\mathbf{\Omega}}| - \frac{1}{2\sigma_\lambda^2} (\mathbf{y}_c - \mu_\lambda \mathbf{t}_c)' (\mathbf{t}_c \mathbf{t}_c' + \tilde{\mathbf{\Omega}})^{-1} (\mathbf{y}_c - \mu_\lambda \mathbf{t}_c). \quad (8)$$

The first partial derivatives of $\ln l(\Theta | \mathbf{y}_c)$ with respect to μ_λ and σ_λ^2 are given as

$$\begin{cases} \frac{\partial \ln l(\Theta | \mathbf{y}_c)}{\partial \mu_\lambda} = \frac{\mathbf{t}_c' (\mathbf{t}_c \mathbf{t}_c' + \tilde{\mathbf{\Omega}})^{-1} (\mathbf{y}_c - \mu_\lambda \mathbf{t}_c)}{\sigma_\lambda^2} \\ \frac{\partial \ln l(\Theta | \mathbf{y}_c)}{\partial \sigma_\lambda^2} = -\frac{c}{2\sigma_\lambda^2} + \frac{1}{2\sigma_\lambda^4} (\mathbf{y}_c - \mu_\lambda \mathbf{t}_c)' (\mathbf{t}_c \mathbf{t}_c' + \tilde{\mathbf{\Omega}})^{-1} (\mathbf{y}_c - \mu_\lambda \mathbf{t}_c) \end{cases}. \quad (9)$$

Let $\partial \ln l(\Theta | \mathbf{y}_c) / \partial \mu_\lambda = \partial \ln l(\Theta | \mathbf{y}_c) / \partial \sigma_\lambda^2 = 0$, then the maximum likelihood estimations (MLEs) of μ_λ and σ_λ^2 based on the monitoring data up to t_c , denoted as $\mu_{\lambda,0}$ and $\sigma_{\lambda,0}^2$, are derived as

$$\begin{cases} \mu_{\lambda,0} = \frac{\mathbf{t}_c' (\mathbf{t}_c \mathbf{t}_c' + \tilde{\mathbf{\Omega}})^{-1} \mathbf{y}_c}{\mathbf{t}_c' (\mathbf{t}_c \mathbf{t}_c' + \tilde{\mathbf{\Omega}})^{-1} \mathbf{t}_c} \\ \sigma_{\lambda,0}^2 = \frac{(\mathbf{y}_c - \mu_{\lambda,0} \mathbf{t}_c)' (\mathbf{t}_c \mathbf{t}_c' + \tilde{\mathbf{\Omega}})^{-1} (\mathbf{y}_c - \mu_{\lambda,0} \mathbf{t}_c)}{c} \end{cases}. \quad (10)$$

Substitute Equation (10) into Equation (8), we obtained a simplified version of the log-likelihood function with model parameters θ , m and σ_m^2 . In this way, the MLEs of θ , m and σ_m^2 , denoted as θ_0 , m_0 and $\sigma_{m,0}^2$, can be obtained using some numerical optimization algorithms. Finally, we could obtain $\mu_{\lambda,0}$ and $\sigma_{\lambda,0}^2$ by substituting θ_0 , m_0 and $\sigma_{m,0}^2$ into Equation (10).

After offline training of model parameters based on the extracted RMS values up to t_c , we could further carry out online RUL prediction when a new monitoring data is available. Based on the proposed model, the RUL of the unit at t_i , denoted as r_i , is defined as

$$R_i = \inf\{r_i : x(t_i + r_i) \geq d | \mathbf{x}_i\} = \inf\{r_i : y(t_i + r_i) \geq d + \varepsilon | \mathbf{y}_i\}, \quad (11)$$

where d represents the failure threshold, $\mathbf{x}_i = \{x_1, \dots, x_i\}$, $\mathbf{y}_i = \{y_1, \dots, y_i\}$ and $i = c + 1, c + 2, \dots, n$.

According to the definition, the conditional probability density function (PDF) of r_i can be derived as

$$f_{R_i | \lambda, \mathbf{y}_i}(r_i | \lambda, \mathbf{y}_i) = \frac{d_i + \varepsilon - \lambda(p_i - r_i q_i)}{\sqrt{2\pi m \lambda r_i^3}} \cdot \exp\left[-\frac{(d_i + \varepsilon - \lambda p_i)^2}{2m \lambda r_i}\right], \quad (12)$$

where $d_i = d - y_i$, $p_i = \int_{t_i}^{t_i+r_i} \varphi(\tau; \theta) d\tau$ and $q_i = \varphi(t_i + r_i; \theta)$.

Equation (12) states that the conditional PDF of r_i uses the latest monitoring data instead of the entire historical data. This is conflict with the degradation evolution pattern introduced in Equation (11). To incorporate the history of observations while maintain the nice property of Wiener process model, we considered an updating procedure for parameter λ by using a linear random walk model [32], i.e.,

$$\lambda_i = \lambda_{i-1} + \eta, \quad (13)$$

where $\eta \sim N(0, \kappa)$. This model is chosen because of its simplicity and practical rationality. According to the property of random walk, the hidden drift at time point t_i , λ_i is a time-dependent variable dependent on λ_{i-1} and adjusted by a noise term. Accordingly, the state space model for observation y_i can be constructed as

$$y_i = y_{i-1} + \lambda_{i-1} \int_{t_{i-1}}^{t_i} \varphi(\tau; \theta) d\tau + \sqrt{m \lambda_{i-1}} \Delta B_i + \Delta \varepsilon_i, \quad (14)$$

where $t_0 = 0$, $y_0 = 0$, $\Delta B_i = B(t_i) - B(t_{i-1}) \sim N(0, t_i - t_{i-1})$ and $\Delta \varepsilon_i = \varepsilon_i - \varepsilon_{i-1} \sim N(0, 2\sigma_m^2)$.

Since the existence of hidden states in both state transition function Equation (13) and observation function Equation (14), we could not utilize maximum likelihood estimation method to update the model parameters and system states directly. To this end, we developed a recursive algorithm by integrating KF and EM for parameter update and estimation. The proposed algorithm consists of two main steps. The KF-step focuses on the estimate and update of the system state, and the EM-step aims to estimate other unknown and non-updating parameters in the observation equation.

Recall that the initial mean and variance of drift is $\mu_{\lambda,0}$ and $\sigma_{\lambda,0}^2$, i.e., $\lambda_0 \sim N(\mu_{\lambda,0}, \sigma_{\lambda,0}^2)$. In addition, the hidden state λ_i can be estimated once the observed monitoring data up to t_i , i.e., $\mathbf{y}_i = \{y_1, \dots, y_i\}$, are available. Denote the mean and variance of λ_i by $E_{\lambda_i} = E(\lambda_i | \mathbf{y}_i)$ and $\text{Var}_{\lambda_i} = \text{Var}(\lambda_i | \mathbf{y}_i)$ respectively. To obtain E_{λ_i} and Var_{λ_i} , we need to calculate the conditional PDF of λ_i given $\mathbf{y}_i$, denoted by $p(\lambda_i | \mathbf{y}_i)$. Recursion solution of $p(\lambda_i | \mathbf{y}_i)$ can be estimated from $p(\lambda_{i-1} | \mathbf{y}_{i-1})$ using the well-adopted Bayesian rule, i.e.,

$$\begin{aligned} p(\lambda_i | \mathbf{y}_i) &= \int_{-\infty}^{+\infty} p(\lambda_i | \lambda_{i-1}) p(\lambda_{i-1} | \mathbf{y}_i) d\lambda_{i-1} \\ &= \frac{\int_{-\infty}^{+\infty} p(\lambda_i | \lambda_{i-1}) p(y_i | \lambda_{i-1}, \mathbf{y}_{i-1}) p(\lambda_{i-1} | \mathbf{y}_{i-1}) d\lambda_{i-1}}{p(y_i | \mathbf{y}_{i-1})}. \end{aligned} \quad (15)$$

Combing Equation (13) to Equation (15), we could obtain that $p(\lambda_i | \mathbf{y}_i)$ follows Gaussian distribution with mean E_{λ_i} and variance Var_{λ_i} , i.e.,

$$p(\lambda_i | \mathbf{y}_i) = \frac{1}{\sqrt{2\pi \text{Var}_{\lambda_i}}} \exp\left[-\frac{(\lambda_i - E_{\lambda_i})^2}{2\text{Var}_{\lambda_i}}\right], \quad (16)$$

where both E_{λ_i} and Var_{λ_i} can be estimated by using the following KF-step (Algorithm 2).

Algorithm 2. KF-step

- (1) Initialize the parameters at t_c : $E_{0|0} = \mu_{\lambda,0}$, $\text{Var}_{0|0} = \sigma_{\lambda,0}^2$, $\theta = \theta_0$, $m = m_0$, $\sigma_m^2 = \sigma_{m,0}^2$, $\kappa = \kappa_0$;
- (2) Estimate the mean of λ_i at t_{i-1} : $E_{i|i-1} = E_{i-1|i-1}$;
- (3) Estimate the variance of λ_i at t_{i-1} : $\text{Var}_{i|i-1} = \text{Var}_{i-1|i-1} + \kappa$;
- (4) Estimate the filter gain: $K_i = \left(\int_{t_{i-1}}^{t_i} \varphi(\tau; \theta) d\tau \right)^2 \text{Var}_{i|i-1} + m E_{i|i-1} (t_i - t_{i-1}) + 2\sigma_m^2$;
- (5) Update the mean of λ_i at t_i : $E_{i|i} = E_{i|i-1} + \text{Var}_{i|i-1} \int_{t_{i-1}}^{t_i} \varphi(\tau; \theta) d\tau K_i^{-1} (y_i - y_{i-1} - E_{i|i-1} \int_{t_{i-1}}^{t_i} \varphi(\tau; \theta) d\tau)$;
- (6) Update the variance of λ_i at t_i : $\text{Var}_{i|i} = \text{Var}_{i|i-1} - \text{Var}_{i|i-1} \left(\int_{t_{i-1}}^{t_i} \varphi(\tau; \theta) d\tau \right)^2 K_i^{-1} \text{Var}_{i|i-1}$.

Following the KF-step, the historical monitoring data can be captured via recursively updating $E_{i|i}$ and $\text{Var}_{i|i}$, based on which the conditional PDF of λ_i , i.e., $p(\lambda_i | \mathbf{y}_i)$, is derived.

On the basis of the estimated $p(\lambda_i | \mathbf{y}_i)$, we need to update the unknown parameters including μ_λ , σ_λ^2 , m , θ , σ_m^2 and κ . Define parameter vector $\tilde{\Theta} = [\Theta, \kappa]'$, then the log-likelihood function of $\mathbf{y}_i$ can be given as

$$\begin{aligned} \ln L_i(\tilde{\Theta}) &= \ln p(\mathbf{y}_i, \lambda_i | \tilde{\Theta}) \\ &= \ln p(\mathbf{y}_i | \lambda_i, \tilde{\Theta}) + \ln p(\lambda_i | \tilde{\Theta}) \\ &= \ln p(\lambda_0 | \tilde{\Theta}) + \sum_{j=1}^i \ln p(\lambda_j | \lambda_{j-1}, \tilde{\Theta}) + \sum_{j=1}^i \ln p(y_j | \lambda_{j-1}, \tilde{\Theta}), \end{aligned} \quad (17)$$

where $\lambda_i = \{\lambda_0, \lambda_1, \dots, \lambda_{i-1}\}$ denotes the history of the updated drift parameters, and $p(\mathbf{y}_i, \lambda_i | \tilde{\Theta})$ denotes the joint PDF of the observed monitoring data $\mathbf{y}_i$. According to Equations (13) and (14), we have $\lambda_j | \lambda_{j-1} \sim N(\lambda_{j-1}, \kappa)$ and $y_j | \lambda_{j-1} \sim N(y_{i-1} + \lambda_{i-1} \int_{t_{i-1}}^{t_i} \varphi(\tau; \theta) d\tau, m \lambda_{i-1} (t_j - t_{j-1}) + 2\sigma_m^2)$. When ignoring the constant item, the log-likelihood function can be formulated by

$$\begin{aligned} \ln L_i(\tilde{\Theta}) &= -\left\{ \ln \sigma_{\lambda,0}^2 + (\lambda_0 - \mu_{\lambda,0})^2 / \sigma_{\lambda,0}^2 + \sum_{j=1}^i \left[\ln \kappa + (\lambda_j - \lambda_{j-1})^2 / \kappa \right] \right. \\ &\quad \left. + \sum_{j=1}^i \left[\ln (m \lambda_{j-1} (t_j - t_{j-1}) + 2\sigma_m^2) + (y_j - y_{j-1} - \lambda_{j-1} \int_{t_{j-1}}^{t_j} \varphi(\tau; \theta) d\tau)^2 / (m \lambda_{j-1} (t_j - t_{j-1}) + 2\sigma_m^2) \right] \right\} / 2. \end{aligned} \quad (18)$$

Note that Equation (18) involves hidden variables $\lambda_i = \{\lambda_0, \lambda_1, \dots, \lambda_{i-1}\}$, so estimating unknown model parameters by maximizing this function is intractable. To solve the problem, we considered EM algorithm that provides a possible framework for estimating the unknown parameters involving hidden variables. A basic assumption of EM algorithm is that the hidden variables can be estimated by the monitoring data. Our problem meets the assumption since $p(\lambda_i | \mathbf{y}_i)$ was obtained by Equation (16). A detail of the EM-step for our problem is given as follows (Algorithm 3).

To calculate the conditional expectation $E_{\lambda_i | \mathbf{y}_i, \hat{\Theta}_i^{(k)}} \{\ln L_i(\tilde{\Theta})\}$, we need to obtain the values of $E_{\lambda_i | \mathbf{y}_i, \hat{\Theta}_i^{(k)}}(\lambda_j)$, $E_{\lambda_i | \mathbf{y}_i, \hat{\Theta}_i^{(k)}}(\lambda_j^2)$, $E_{\lambda_i | \mathbf{y}_i, \hat{\Theta}_i^{(k)}}(\lambda_j \lambda_{j-1})$, $E_{\lambda_i | \mathbf{y}_i, \hat{\Theta}_i^{(k)}}(\lambda_j^{-1})$ and $E_{\lambda_i | \mathbf{y}_i, \hat{\Theta}_i^{(k)}}(\ln \lambda_j)$. In this paper, these conditional expectations are estimated based on the properties of variance-covariance and Rauch–Tung–Striebel (RTS) smoothing algorithm [32].

Finally, the PDF of the updated RUL at the current monitoring time point t_i can be obtained by using the law of total probability, i.e.,

$$f_{R_i | \mathbf{y}_i}(r_i | \mathbf{y}_i) = \int_{-\infty}^{+\infty} f_{R_i | \lambda, \mathbf{y}_i}(r_i | \lambda, \mathbf{y}_i) p(\lambda | \mathbf{y}_i) d\lambda = E_{\lambda | \mathbf{y}_i} [f_{R_i | \lambda, \mathbf{y}_i}(r_i | \lambda, \mathbf{y}_i)], \quad (19)$$

where the model parameters have been derived above.

Algorithm 3. EM-step

- (1) E-step: Taking the expectation operator on the side of $\ln L_i(\tilde{\Theta})$ with respect to $\lambda_i|y_i, \hat{\Theta}_i^{(k)}$, we have

$$\ln L_i\left(\tilde{\Theta}|\hat{\Theta}_i^{(k)}\right) = E_{\lambda_i|y_i, \hat{\Theta}_i^{(k)}}\left\{\ln L_i(\tilde{\Theta})\right\},$$

where k is the number of iterations and $\hat{\Theta}_i^{(k)} = [\mu_{\lambda,i}^{(k)}, \sigma_{\lambda,i}^{2(k)}, \theta_i^{(k)}, m_i^{(k)}, \sigma_{m,i}^{2(k)}, \kappa_i^{(k)}]'$ denotes the estimated parameters in the k -th step conditional on y_i .

- (2) M-step: Calculate unknown parameters $\hat{\Theta}_i^{(k+1)}$ by maximizing $\ln L_i\left(\tilde{\Theta}|\hat{\Theta}_i^{(k)}\right)$ with $\tilde{\Theta}$, i.e.,

$$\hat{\Theta}_i^{(k+1)} = \underset{\tilde{\Theta}}{\operatorname{argmax}} \left\{ E_{\lambda_i|y_i, \hat{\Theta}_i^{(k)}}\left\{\ln L_i(\tilde{\Theta})\right\} \right\},$$

which can be solved by directly taking $\partial \ln L_i\left(\tilde{\Theta}|\hat{\Theta}_i^{(k)}\right) / \partial \tilde{\Theta} = 0$.

- (3) Given a precision requirement ρ , iterate E-step and M-step until a criterion of convergence is satisfied, i.e.,

$$\left| \left(\hat{\Theta}_i^{(k+1)} - \hat{\Theta}_i^{(k)} \right) / \hat{\Theta}_i^{(k)} \right| \leq \rho.$$

3. Case Study

So far, some types of train have installed on-board real-time monitoring system, based on which online vibration signals of wheel treads can be monitored. A complete set of on-board real-time monitoring system is shown in Figure 3. We could find that it includes shock vibration sensor, data pre-processor, diagnostic instrument, display terminal, etc. When the train is running, the system can output original vibration data to data analysis center via the wireless local area network (WLAN).

Figure 3. Illustration of the structure of the on-board real-time monitoring system.

To demonstrate the applicability of the proposed RUL framework, vibration signals of wheel treads of a high-speed train collected from the system were adopted as the data set. Figure 4 shows a schematic for the wheel tread with sensor. The marked point 11 and point 16 are installation positions of the sensors used for monitoring the wheel treads. Figure 5 presents the original vibration signals of two wheel treads. It is easy to find that the monitoring system recorded the running mileage of wheel treads and corresponding amplitude (dB) at each monitoring point. The reference value for measuring dB values was 10^{-5} mm/s. Note that the recorded amplitude reflected the normalized strength of vibration shocks by eliminating the effects of running speed and axle diameter of the train. The recorded running mileage started from 402,590 km and ended at 447,870 km and 450,000 km for the

first and the second wheel tread, respectively. In addition, because of human factors, the amplitudes at some monitoring points were missing and this could be reflected by the blank intervals in Figure 5.

Figure 4. A schematic for the wheel tread with a sensor in our monitoring program.

Figure 5. Original vibration signals of two wheel treads collected via on-board real-time monitoring system. (a) Original vibration signals of the 1-st wheel tread and (b) original vibration signals of the 2-nd wheel tread.

To understand the failure processes of two wheel treads more clearly, failure mode analysis were carried out based on physical technologies. Results show that the degradation of the first wheel tread was caused by out-of-round and the second one was caused by stripping. Engineering experience shows that out-of-round and stripping are two most common failure modes of the wheel tread of train. For an illustrative purpose, an illustration of the two failure modes was provided, as shown in Figure 6a,b respectively. In particular, out-of-round is always caused by uneven wear of wheel treads. During the running state, because of various factors such as possible scratches in traction and braking, foreign body bruising and so on, the wheel rolling circle loses its roundness and presents a polygon phenomenon. Vibration and noise caused by wheel polygon not only affect the comfort of the ride but also increase the vertical contact force of the wheel and rail, resulting in accelerated damage of the track and the components of train and shortening the service life. Different from this, stripping is always caused by fatigue stress, which is inevitable in all of rotary machinery systems. Tiny defects in the material will be continuously amplified under the effect of fatigue stress until part of the material is stripped from the entire wheel tread. Around the occurrence of stripping, degradation velocity of tread varies significantly because the balance of rotary system has been broken. This phenomenon will be verified in the following. Note that when the degradation caused by either out-of-round or stripping reaches a certain extent, appropriate and seasonable maintenance measures must be taken. Otherwise, the train may happen catastrophic failure during its high-speed running state. For the wheel treads provided in our case, turning maintenance was carried out at the ending

points of their monitoring processes, which means that the wheel treads have been considered to be failed in practical applications.

Figure 6. Illustration of the two failure modes of wheel tread. (a) Out-of-round and (b) stripping.

According to the characteristics of the vibration signals, the mileage, i.e., the X-axis of Figure 5, was selected as a generalized time, which would be used in our proposed model. Meanwhile, the RMS of original vibration signals was extracted as the degradation indicator. To obtain the RMS, the recorded mileages were separated into equally distributed sliding intervals, each of which was 200 km in length. In this way, the first RMS value was calculated based on the amplitudes between 0 and 200 km, and the second RMS value was calculated based on the amplitudes between 100 and 300 km, and so on.

Figure 7 shows the extracted RMS values of the two wheel treads. There were in total 389 extracted data points for the first wheel tread and 456 data points for the second wheel tread. We found that the extracted RMS values were relatively stable with some fluctuations at the beginning, but present obvious increasing trends within the following monitoring process. Therefore, we detected the PCT of each wheel tread and divided the degradation path into normal operation phase and rapid degradation phase. Results are also presented in Figure 7. It is observed that the RMS was sensitive to the incipient fault of wheel treads. Except this, we found that the second wheel tread performed a two-phase degradation behavior within its rapid degradation phase. Combined with its failure mode, we thought that this was reasonable as the health states of the wheel tread could be different from each other before and after the occurrence of stripping. In this view, the changing point of the two-state degradation should be the time of stripping, and when the balance of rotary wheel tread was broken, more severe vibrations accelerated its degradation. For convenience, the changing point was denoted as $t_{stripping}$ in the following.

To carry out RUL prediction, we need to determine the failure threshold of wheel treads. As mentioned above, the two wheel treads were considered to be failed in practical applications. Thus, we selected the mean of the final RMS values of the two wheel treads as the threshold, i.e., $d = 35$. In addition, $\varphi(t; \theta) = \theta t^{\theta-1}$ was selected for the 1-st wheel tread and $\varphi(t; \theta) = \theta$ was selected for the 2-nd wheel tread because the extracted RMS values always increase with the running mileage in a power-law form (or a linear form) for the wheel treads whose failure mode is out-of-round (or stripping) according to our previous engineering experience. Under these assumptions, we illustrated our proposed prediction model and corresponding parameter updating algorithm based on RMS values of the rapid degradation phase.

Figure 7. Extracted degradation features and phase changing times (PCTs) of the two wheel treads. (a) Extracted root mean square (RMS) values and PCT of the 1-st wheel tread and (b) extracted RMS values and PCT of the 2-nd wheel tread.

To verify the rationality of the proposed model, we separated the degradation paths of the two wheel treads into 10 equal intervals respectively and calculated degradation rate and degradation diffusion of each interval. According to the results, a correlation test between degradation rate and degradation diffusion was carried out. Figure 8 shows the test results, which mean that there really existed a linear relationship between the drift parameters and the square of diffusion parameters, as mentioned in our proposed model.

Figure 8. Correlation test between degradation rate and degradation diffusion. (a) Correlation test result of the 1-st wheel tread and (b) correlation test result of the 2-nd wheel tread.

Then, based on the parameter initialization method, we initialized the unknown parameters by using the first twenty RMS values from each changing point including two PCTs and a stripping point. Results are shown in Table 1. Next, the model parameters were updated continuously based on KF and EM algorithm with the increasing of RMS values. The parameter updating procedure is presented in Figure 9. It is observed that all parameters were not estimated accurately at the beginning of the predicting process. However, when enough RMS values were put into parameter estimation, these parameters gradually converged to fixed values. Note that, for the second wheel tread whose failure mode was stripping, the parameter updating procedure included two phases. This was consistent with the above-mentioned explanations.

Table 1. Parameter initialization results of the two wheel treads.

Parameter	$\mu_{\lambda,0}$	$\sigma_{\lambda,0}^2$	θ_0	m_0	$\sigma_{m,0}^2$
PCT of the 1-st wheel tread	1.13×10^{-8}	6.40×10^{-10}	15.62	176.11	1.44×10^{-2}
PCT of the 2-nd wheel tread	1.07×10^{-4}	3.40×10^{-3}	2.80	44.12	9.60×10^{-3}
$t_{stripping}$ of the 2-nd wheel tread	1.80×10^{-3}	2.70×10^{-3}	0.61	6.56	1.07×10^{-2}

Figure 9. Parameter updating procedure based on Kalman filter (KF) and expectation maximization (EM) algorithm. (a) The updating procedure of the mean and variance of drift parameter for the 1-st wheel tread and (b) the updating procedure of the mean and variance of drift parameter for the 2-nd wheel tread.

To identify the effectiveness of the proposed model, we predicted the degradation state and the RUL distribution at different monitoring points of the two wheel treads. Figure 10 compares the predicted degradation path with the original degradation path. We can observe that the prediction results matched the original degradation path well except for the first twenty points, which means that our proposed model had a better goodness-of-fit with vibration monitoring data of the wheel treads. Figure 11 presents the PDFs of the RUL of the wheel treads at different monitoring points calculated by our proposed model. We found that our proposed model could cover actual RULs with an acceptable precision.

Figure 10. The predicted degradation paths of the two wheel treads. (a) The predicted RMS values of the 1-st wheel tread and (b) the predicted RMS values of the 2-nd wheel tread.

Figure 11. The estimated probability density functions (PDFs) of the RUL at various monitoring points. (a) The estimated PDFs of the RUL of the 1-st wheel tread and (b) the estimated PDFs of the RUL of the 2-nd wheel tread.

For comparison, we further compared the expected RULs derived by our model (defined as Model 1) and the following model (defined as Model 2), i.e.,

$$x(t) = x_0 + \lambda \int_0^t \varphi(\tau; \theta) d\tau + \sigma B(t) + \varepsilon, \varepsilon \sim N(0, \sigma_m^2), \quad (20)$$

which does not consider the quantitative relationship between degradation rate and degradation variation, with the actual RULs at various monitoring points.

Table 2 shows the relative errors (REs) of the predicted results. Noticeably, our proposed model incurred smaller REs compared with the model in Equation (20). Therefore, the results could demonstrate that our proposed model could help to improve the accuracy of degradation path tracking and RUL estimation. In the meanwhile, it should be mentioned that the REs at the first two monitoring points of the 2-nd wheel tread were a little big, which means that the predicted results were not quite satisfactory. This is reasonable because the two points were particularly close to the stripping point of the wheel tread and there was not much degradation information available for RUL prediction. With the increasing of degradation data, we could see that the REs of the predicted results became smaller and smaller, especially for our proposed model.

Table 2. Comparison of the predicted RULs and the actual RULs.

	The 1-st Wheel Tread				The 2-nd Wheel Tread			
Monitoring point/km $\times 10^5$	4.30	4.35	4.40	4.45	4.30	4.35	4.40	4.45
Actual RUL/km $\times 10^4$	1.79	1.29	0.79	0.29	2.00	1.50	1.00	0.50
REs of Model 1	11.08%	7.32%	4.91%	2.44%	93.90%	130.77%	13.63%	3.54%
REs of Model 2	16.95%	11.01%	6.54%	3.19%	91.68%	142.53%	14.13%	5.27%

4. Conclusions

This paper constructed a parameter-related Wiener process model with a recursive algorithm for RUL prediction of wheel treads. A distinguished superiority of our model was the consideration of parameter correlation and multiple uncertainties, which better fits the evolution pattern of degradation path. A recursive algorithm based on KF and EM technologies was developed to realize joint estimation and update of lifetime parameters. Experiments on real-world vibration signals of wheel treads demonstrated the superiority of the proposed methodology in prediction accuracy.

A possible extension is to develop an early warning system based on RUL prediction results. According to the safety requirements, we could design different warning thresholds so that precise time points for making maintenance decisions could be given. In other words, when the degradation value

of the wheel tread surpasses a given percentage of the failure threshold, we expected that early warning would occur. This technology could contribute to safety-critical predictive maintenance plans [33].

Author Contributions: Methodology, H.W. and X.M.; validation, G.H.; writing—original draft preparation, H.W.; writing—review and editing, G.H.; supervision, Y.Z. and D.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by NATIONAL NATURAL SCIENCE FOUNDATION OF CHINA, grant number 61473014.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Proof of Proposition 1. According to the statistical properties of Brownian motion, we have $E(B(t_i)) = 0$ and $D(B(t_i)) = t_i$, where $E(\cdot)$ and $D(\cdot)$ denote the mean and variance, respectively. In this way, the mean of the health indicators at t_i is clearly derived as

$$E(y_i) = E\left(\lambda \int_0^{t_i} \varphi(\tau; \theta) d\tau\right) + E\left(\sqrt{m\lambda}B(t_i)\right) + E(\varepsilon) = \mu_\lambda \int_0^{t_i} \varphi(\tau; \theta) d\tau, \quad (\text{A1})$$

and then

$$E(\mathbf{y}_c) = [E(y_1), \dots, E(y_c)]' = \mu_\lambda \mathbf{t}_c. \quad (\text{A2})$$

For the covariance between y_i and y_j , we calculate it by using

$$\begin{aligned} \text{Cov}(y_i, y_j) &= \text{Cov}\left(\lambda \int_0^{t_i} \varphi(\tau; \theta) d\tau, \lambda \int_0^{t_j} \varphi(\tau; \theta) d\tau\right) \\ &\quad + \text{Cov}\left(\sqrt{m\lambda}B(t_i), \sqrt{m\lambda}B(t_j)\right) \\ &\quad + \text{Cov}(\varepsilon, \varepsilon) \end{aligned} \quad (\text{A3})$$

where each item can be given as

$$\begin{aligned} &\text{Cov}\left(\lambda \int_0^{t_i} \varphi(\tau; \theta) d\tau, \lambda \int_0^{t_j} \varphi(\tau; \theta) d\tau\right) \\ &= E\left(\lambda \int_0^{t_i} \varphi(\tau; \theta) d\tau \cdot \lambda \int_0^{t_j} \varphi(\tau; \theta) d\tau\right) - E\left(\lambda \int_0^{t_i} \varphi(\tau; \theta) d\tau\right) \cdot E\left(\lambda \int_0^{t_j} \varphi(\tau; \theta) d\tau\right) \\ &= \int_0^{t_i} \varphi(\tau; \theta) d\tau \cdot \int_0^{t_j} \varphi(\tau; \theta) d\tau \cdot (E(\lambda^2) - E^2(\lambda)) \\ &= \sigma_\lambda^2 \cdot \int_0^{t_i} \varphi(\tau; \theta) d\tau \cdot \int_0^{t_j} \varphi(\tau; \theta) d\tau \end{aligned} \quad (\text{A4})$$

$$\begin{aligned} &\text{Cov}\left(\sqrt{m\lambda}B(t_i), \sqrt{m\lambda}B(t_j)\right) \\ &= E\left(\sqrt{m\lambda}B(t_i) \cdot \sqrt{m\lambda}B(t_j)\right) - E\left(\sqrt{m\lambda}B(t_i)\right) \cdot E\left(\sqrt{m\lambda}B(t_j)\right) \\ &= \begin{cases} E\left(\sqrt{m\lambda}B(t_i)\left(\sqrt{m\lambda}B(t_j) - \sqrt{m\lambda}B(t_i)\right) + m\lambda B(t_i)^2\right) = E\left(m\lambda B(t_i)^2\right) = m\mu_\lambda t_i, & i \leq j \\ E\left(\sqrt{m\lambda}B(t_j)\left(\sqrt{m\lambda}B(t_i) - \sqrt{m\lambda}B(t_j)\right) + m\lambda B(t_j)^2\right) = E\left(m\lambda B(t_j)^2\right) = m\mu_\lambda t_j, & i > j \end{cases} \end{aligned} \quad (\text{A5})$$

$$\text{Cov}(\varepsilon, \varepsilon) = \begin{cases} \sigma_m^2, & i = j \\ 0, & i \neq j \end{cases}. \quad (\text{A6})$$

Denote the results in matrix and we derive

$$D(\mathbf{y}_c) = \sigma_2^\lambda \mathbf{t}_c \mathbf{t}_c' + m\mu_\lambda \mathbf{Q}_c \sigma_2^m \mathbf{E}_c. \quad (\text{A7})$$

□

References

1. Jin, X.; Xiao, X.; Ling, L.; Zhou, L.; Xiong, J. Study on safety boundary for high-speed train running in severe environments. *Int. J. Rail Transit.* **2013**, *1*, 87–108. [\[CrossRef\]](#)
2. Hu, H.; Tang, B.; Gong, X.; Wei, W.; Wang, H. Intelligent Fault Diagnosis of the High-Speed Train with Big Data Based on Deep Neural Networks. *IEEE Trans. Ind. Inform.* **2017**, *13*, 2106–2116. [\[CrossRef\]](#)
3. Zhang, B.; Tan, A.C.C.; Lin, J.H. Gearbox fault diagnosis of high-speed railway train. *Eng. Fail. Anal.* **2016**, *66*, 407–420. [\[CrossRef\]](#)
4. Qin, N.; Jin, W.D.; Huang, J. Fault Diagnosis of High Speed Train Bogie Based on Ensemble Empirical Mode Decomposition. *Comput. Eng.* **2013**, *39*, 1–10.
5. An, D.; Kim, N.H.; Choi, J.H. Practical options for selecting data-driven or physics-based prognostics algorithms with reviews. *Reliab. Eng. Syst. Saf.* **2015**, *133*, 223–236. [\[CrossRef\]](#)
6. Si, X.S.; Wang, W.; Hu, C.H.; Zhou, D.H. Remaining useful life estimation-A review on the statistical data driven approaches. *Eur. J. Oper. Res.* **2011**, *213*, 1–14. [\[CrossRef\]](#)
7. Tsui, K.L.; Chen, N.; Zhou, Q.; Hai, Y.; Wang, W. Prognostics and Health Management: A review on data driven approaches. *Math. Probl. Eng.* **2015**, *2015*, 1–17. [\[CrossRef\]](#)
8. Gebraeel, N.; Lawley, M.; Liu, R. Residual life predictions from vibration-based degradation signals: A neural network approach. *IEEE Trans. Ind. Inform.* **2004**, *51*, 694–700. [\[CrossRef\]](#)
9. Guo, L.; Li, N.; Jia, F.; Lei, Y.; Lin, J. A recurrent neural network based health indicator for remaining useful life prediction of bearings. *Neurocomputing* **2017**, *240*, 98–109. [\[CrossRef\]](#)
10. Rigamonti, M.; Baraldi, P.; Zio, E.; Roychoudhury, I.; Goebel, K.; Poll, S. Ensemble of optimized echo state networks for remaining useful life prediction. *Neurocomputing* **2018**, *281*, 121–138. [\[CrossRef\]](#)
11. Elsheikh, A.; Yacout, S.; Ouali, M.S. Bidirectional Handshaking LSTM for Remaining Useful Life Prediction. *Neurocomputing* **2019**, *323*, 148–156. [\[CrossRef\]](#)
12. Maio, F.D.; Tsui, K.L.; Zio, E. Combining Relevance Vector Machines and exponential regression for bearing residual life estimation. *Mech. Syst. Signal Proc.* **2012**, *31*, 405–427. [\[CrossRef\]](#)
13. Tran, V.T.; Hong, T.P.; Yang, B.S.; Nguyen, T.T. Machine Performance Degradation Assessment and Remaining Useful Life Prediction Using Proportional Hazard Model and SVM. *Mech. Syst. Signal Proc.* **2012**, *32*, 320–330. [\[CrossRef\]](#)
14. Soualhi, A.; Razik, H.; Clerc, G.; Doan, D.D. Prognosis of Bearing Failures using Hidden Markov Models and the Adaptive Neuro-Fuzzy Inference System. *IEEE Trans. Ind. Electron.* **2014**, *61*, 2864–2874. [\[CrossRef\]](#)
15. Liu, J.; Wang, W.; Golnaraghi, F. A multi-step predictor with a variable input pattern for system state forecasting. *Mech. Syst. Signal Proc.* **2009**, *23*, 1586–1599. [\[CrossRef\]](#)
16. Liu, Z.; Cheng, Y.; Wang, P.; Yu, Y.; Long, Y. A method for remaining useful life prediction of crystal oscillators using the Bayesian approach and extreme learning machine under uncertainty. *Neurocomputing* **2018**, *305*, 27–38. [\[CrossRef\]](#)
17. Ali, J.B.; Chebel-Morello, B.; Saidi, L.; Malinowski, S.; Fnaiech, F. Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network. *Mech. Syst. Signal Proc.* **2015**, *56*, 150–172.
18. Caesarendra, W.; Widodo, A.; Yang, B.S. Application of relevance vector machine and logistic regression for machine degradation assessment. *Mech. Syst. Signal Proc.* **2010**, *24*, 1161–1171. [\[CrossRef\]](#)
19. Liao, H.T.; Tian, Z.G. A framework for predicting the remaining useful life of a single unit under time varying operating conditions. *Lie Trans.* **2013**, *45*, 964–980. [\[CrossRef\]](#)
20. Zhai, Q.; Ye, Z.S. RUL Prediction of Deteriorating Products Using an Adaptive Wiener Process Model. *IEEE Trans. Ind. Inform.* **2017**, *13*, 2911–2921. [\[CrossRef\]](#)
21. Bian, L.; Gebraeel, N. Computing and updating the first-passage time distribution for randomly evolving degradation signals. *Lie Trans.* **2012**, *44*, 974–987. [\[CrossRef\]](#)
22. Son, K.L.; Fouladirad, M.; Barros, A. Remaining useful lifetime estimation and noisy gamma deterioration process. *Reliab. Eng. Syst. Saf.* **2016**, *149*, 76–87. [\[CrossRef\]](#)
23. Zhang, Z.; Si, X.; Hu, C.; Lei, Y. Degradation Data Analysis and Remaining Useful Life Estimation, A Review on Wiener-Process-Based Methods. *Eur. J. Oper. Res.* **2018**, *271*, 775–796. [\[CrossRef\]](#)
24. Li, N.; Lei, Y.; Lin, J.; Ding, S.X. An improved exponential model for predicting remaining useful life of rolling element bearings. *IEEE Trans. Ind. Electron.* **2015**, *62*, 7762–7773. [\[CrossRef\]](#)

25. Wang, H.Y.; Zhao, Y.; Ma, X.B. Remaining Useful Life Prediction Using a Novel Two-stage Wiener Process with Stage correlation. *IEEE Access* **2018**, *6*, 65227–65238. [[CrossRef](#)]
26. Ye, Z.S.; Wang, Y.; Tsui, K.L.; Pecht, M. Degradation data analysis using wiener processes with measurement errors. *IEEE Trans. Reliab.* **2013**, *62*, 772–780. [[CrossRef](#)]
27. Bian, L.; Gebraeel, N.; Kharoufeh, J.P. Degradation modeling for real-time estimation of residual lifetimes in dynamic environments. *IIE Trans.* **2015**, *47*, 471–476. [[CrossRef](#)]
28. Zhai, Q.; Chen, P.; Hong, L.; Shen, L. A random-effects Wiener degradation model based on accelerated failure time. *Reliab. Eng. Syst. Saf.* **2018**, *180*, 94–103. [[CrossRef](#)]
29. Ramezani, M.G.; Golchinfar, B.; Donskoy, D.; Hassiotis, S.; Venkiteela, G. Steel Material Degradation Assessment Via Vibro-Acoustic Modulation Technique. *Transp. Res. Record.* **2019**, *2673*, 579–585. [[CrossRef](#)]
30. Donskoy, D.; Golchinfar, B.; Ramezani, M.; Rutner, M.; Hassiotis, S. Vibro-acoustic amplitude and frequency modulations during fatigue damage evolution. *AIP Conf. Proc.* **2019**, *2102*, 040004.
31. Li, N.P.; Lei, Y.G.; Yan, T.; Li, N.; Han, T. A Wiener Process Model-based Method for Remaining Useful Life Prediction Considering Unit-to-Unit Variability. *IEEE Trans. Ind. Electron.* **2018**, *66*, 2092–2101. [[CrossRef](#)]
32. Si, X.S.; Wang, W.; Hu, C.H.; Chen, M.Y.; Zhou, D.H. A Wiener-process-based degradation model with a recursive filter algorithm for remaining useful life estimation. *Mech. Syst. Signal Proc.* **2013**, *35*, 219–237. [[CrossRef](#)]
33. Yang, L.; Zhao, Y.; Ma, X.B. Group maintenance scheduling for two-component systems with failure interaction. *Appl. Math. Model.* **2019**, *71*, 118–137. [[CrossRef](#)]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).