Stochastic Prognostics under Multiple Time-varying Environmental Factors

Salman Jahani JAHANI@WISC.EDU

Shiyu Zhou SHIYUZHOU@WISC.EDU

Dharmaraj Veeramani RAJ.VEERAMANI@WISC.EDU

Department of Industrial and Systems Engineering

University of Wisconsin-Madison

Madison, WI 53706, USA

Abstract

Prediction of the remaining useful life of in-field components, traditionally, relies on condition monitoring signals which are correlated with the physical degradation of the system. Many models assume that condition monitoring signals behave under similar environmental conditions (e.g. pressure, temperature, workload and relative humidity) or these conditions have no effect on degradation process. In this paper, we propose a Brownian motion process with a stress-dependent drift to model multiple time-varying environmental covariates. A semiparametric regression approach utilizing penalized splines is, further, proposed to model the environmental covariates-drift relationship. The unique feature of our approach is that it does not assume a functional form for the degradation process drift and models multiple environmental covariates' effect on the degradation process. Moreover, the model is combined with in situ degradation measurements of the in-field unit and its environmental conditions to predict the unit's remaining useful life through a Bayesian updating scheme. The performance of the proposed framework is investigated and benchmarked through analysis based on numerical studies and a case study using real-world data of frying oil degradation collected from connected fryers.

Keywords: Environmental covariates, Remaining useful life estimation, Degradation modeling, B-spline regression, Bayesian updating

1. Introduction

Recent advances in sensor technology and wireless communication systems are playing a significant role in what is referred to as the Internet of Things (IoT). Remote condition monitoring of physical assets using sensor data provides unprecedented opportunity for assessing the health and performance of engineering systems. Many statistical models have used sensor data, typically referred to as condition monitoring (CM) signals e.g. internal resistance of automotive battery or vibration signal of a rotating machinery, to predict the remaining useful life (RUL) of in-service units (19; 10). In such models, a historical database of CM signals is utilized to predict the RUL of an in-service unit through a linkage between the historical data and real-time CM information collected from the in-service units. A unit is typically deemed failed once its CM signal reaches a pre-specified threshold value (30).

In most of the existing RUL prediction models, it is assumed that the units perform under a fixed operating condition (e.g. pressure, temperature, workload and relative humidity), i.e. the environmental condition has no effect on the failure and degradation process (19; 18; 36; 24; 39). However, in many cases, the environmental condition varies with time and units may be exposed to different operating conditions. Under such circumstances, the degradation of an operating unit can greatly decelerate or accelerate, and the CM signal may exhibit different evolution rate over time. For products with degradation driven by environmental conditions, information about these variables can be critical for modeling the degradation process. For example, the degradation of the frying oil is primarily driven by a series of physical and chemical reactions as a result of change in temperature, humidity, and other factors affecting oil quality during frying. There are many other examples where degradation is driven by the usage and environmental variables, such as degradation of ball bearings, the loss of color and gloss of an automobile coating, and corrosion in oil transportation pipeline.

Development in technology makes collection and transmission of massive amount of data possible in many systems. Nowadays, it is common to dynamically record product/system usage as well as other environmental condition information like temperature, humidity, and etc. Environmental condition data contains rich information and can be utilized for modeling and predicting product reliability. Most of the literature that incorporates environmental condition in reliability modeling is based on proportional hazards models. Such models handle the environmental conditions by incorporating them as covariates in the parametric hazard functions. For instance, Jardine and Makis (15) used the hazard model for condition-based maintenance by modeling different attributes of CM signals as covariates. Along this line, Banjevic and Jardine (3) modeled the operating condition as Markov processes and developed approximations for the failure time distribution. Moreover, Rizopoulos (34) proposed a joint framework that incorporates a polynomial mixed effects model and the proportional hazards model. This framework was later extended in Zhou et al. (45) through updating the parameters of the polynomial mixed effects models. Other extensions and applications of the proportional hazards model can be found in (12) and (44).

For degradation data analysis, the operating condition and environmental effects are also available in several settings such as accelerated repeated measure degradation tests (28) and accelerated destructive degradation tests (9). Most of these approaches proposed in the literature for modeling the environmental conditions estimate the lifetime of the population rather than that of a single unit operating in the field. However the few approaches that utilize CM signals for RUL estimation are usually based on mixed effects model, gamma processes, Brownian motion process and other Lévy processes. Gebraeel and Pan (11), for instance, extended the mixed effects model with an exponential function to provide individualized prognosis for an in-service unit with different attributes using a Bayesian approach. Doksum and Hoyland (7) proposed a stress-dependent drift to account for environmental condition in accelerated life testing experiments. Later, Whitmore and Schenkelberg (40) proposed a time-scale transformation for accelerated experiments to account for the level of stress under which the CM signals are operated. Lawless and Crowder (20) modeled CM signals as a gamma process with random effects and covariates. Li et al. (22) modeled the effect of a single time-varying environmental covariate based on a two-factor state-space model considering signal jumps at condition change points. Pang et al. (31) proposed a non-linear diffusion process model to characterize the degradation process of a product and then a parametric model was developed to establish the relationship between a single environmental covariate and the model parameters. Similar ideas can be found in (23), (25) and (2).

Most of the existing statistical models, including the above-mentioned studies, consider only the effect of a single environmental covariate on the degradation signal. Few studies extended this assumption to incorporate the effects of multiple environmental factors on CM signal modeling. Park and Padgett (32), and Ye and Chen (43), for instance, modeled the effect of multiple environmental covariates by assuming a known physics-based model. A limitation of these approaches is the need to specify the correct physical model for the covariate-CM signal relationship. However, there exist numerous practical situations where this parametric form is unknown or difficult to specify. For instance, consider the degradation of ball bearings operated under different loads and speeds where the parametric form of degradation signal as a function of environmental covariates is hard to specify (4). As another case in point, consider the degradation of frying oil in fast food restaurants. During the frying process a complex series of physical and chemical reactions takes place, resulting in degradation of frying oil. In the field of food science, the physical and chemical changes of oil and its degradation during frying operations have been extensively studied through lab testing and experiments (6; 16; 5). Although very useful, most of these studies are qualitative or use a simple quantitative model under a very specific well-controlled lab environment to describe oil degradation. Thus, the results in above studies cannot be directly used to model the oil degradation in a real-life setting, such as in a restaurant environment. An exception is the work by Hong et al. (13) where they modeled the effect of multiple covariates on degradation process using a non-parametric general path model and considered a parametric model to describe the covariate process. While this approach works well in case all the units are exposed to the same environmental condition, it does not apply to the scenario that the environmental condition covariates are evolving uniquely for each unit. For instance, the case study of this paper models the degradation of frying oil in different frying pots where the environmental conditions (e.g. temperature, cooking time, humidity, and etc.) are changing uniquely for each pot. An ideal degradation model in this case should be able to take into account this unique evolution of environmental conditions. With the growing usage of IoT-enabled fryers in fast food restaurants, the comprehensive dataset collected from the connected fryers enables us to establish a quantitative data-driven model to consider multiple dynamic factors in the oil degradation model. Little work has been done in degradation data modeling that considers the uniquely evolving environmental covariates. Table 1 summarizes the result of literature survey on prognostics models considering time-varying environmental condition.

Table 1: Summary of existing literature on prognostics models considering time-varying environmental condition.

Number of covariates	**Covariate effect model**	**Online updating**	**Parameter estimation procedure**	**Unique evolution of environmental covariates**	**Sources**
Single	Parametric	No	Monte Carlo	No	(15; 9; 3); etc.
Single	Parametric	Yes	MLE	No	(34; 45; 11; 25; 31; 22; 2); etc.
Single	Parametric	No	MLE	No	(20); etc
Multiple	Parametric	No	MLE	No	(32; 43); etc
Multiple	Non-parametric	Yes	MLE	No	(13; 41)
Multiple	Non-parametric	Yes	EM algorithm with penalized splines	Yes	Current work

The purpose of this paper is to develop a degradation modeling framework that does not depend on a specific parametric form and considers the effect of multiple time-varying environmental covariates as well as their unit-to-unit variability. Specifically, we propose a stochastic Brownian motion process with a drift dependent on the environmental conditions to incorporate the effects of multiple operational covariates. A semi-parametric regression model using penalized splines is considered to model the drift of proposed Brownian motion process. Estimation is then performed through casting the problem into a mixed effects modeling framework. A Expectation-Maximization (EM) procedure is further developed to estimate the parameters of the proposed framework. This framework offers a number of advantages. First, it is semiparametric and does not assume any rigid parametric form on the Brownian motion process drift. The penalized splines that combine a set of spline basis functions with a quadratic penalty on the corresponding coefficients, also, strike a balance between ovefitting and accuracy of covariate-drift relationship. Moreover, the mixed modeling framework proposed offers a natural way for choosing the smoothing parameters. The proposed framework allows for the unique evolution of environmental covariates for each unit. Also, the stochastic model can further be updated through a Bayesian procedure combining the degradation signal observations and environmental conditions to predict the RUL of the in-field unit in real-time.

The remainder of this paper is organized as follows. Section 2 provides a detailed framework of our modeling approach and estimation procedure and presents the RUL prediction using online updating for the in-field units. To illustrate the effectiveness of our proposed approach, extensive numerical study and a case study based on the degradation data of frying oil are conducted in Section 3 and Section 4. Finally, Section 5 concludes the paper with a short discussion.

2. Model Development

In this section, we first introduce the notation and then develop an offline modeling framework using a semi-parametric based Brownian motion process modeling the effect of multiple time varying environmental covariates on the degradation process. Then, an online updating procedure is proposed to combine the degradation data and environmental condition observations of the in-service unit with the developed offline model.

2.1 Notation for the data

Here, we introduce some notations for the degradation data model and the environmental covariate information. Suppose the database contains a historical dataset of N units. For the ith unit, denote the degradation measurements at time $t_{i,j}$ by $X_i(t_{i,j})$, $i = 1, ..., N$, $j = 0, ..., n_i$, and n_i marks the last timepoint where the degradation measurement was taken. The history of degradation signal observations for unit i is denoted as $\boldsymbol{X}_i(\boldsymbol{t}_{i,0:n_i}) = [X_i(t_{i,0}), ..., X_i(t_{i,n_i})]^t$. Let $\boldsymbol{\omega}_i(\tau) = [\omega_{i,1}(\tau), \omega_{i,2}(\tau), ..., \omega_{i,q}(\tau)]^t$ be a vector denoting the values of q environmental covariates at time τ for unit i. The value of covariate l for unit i at the time τ is denoted by $\omega_{i,l}(\tau)$. The history of covariate process for unit i is denoted by $\boldsymbol{\Omega}_i(\boldsymbol{t}_{i,n_i}) = \{\boldsymbol{\omega}_i(\tau) : 0 \leq \tau \leq t_{i,n_i}\}$, which records the dynamic covariates information from time 0 to time t_{i,n_i} for unit i. We consider all the observed degradation signals and the corresponding operating condition for each unit i in our historical database as $\mathcal{D}_i$ and all observed data from N units as $\mathcal{D} = \{\mathcal{D}_i\}_{i=1}^N$. It should be noted that in any

Figure 1 shows the degradation process measurements and two corresponding environmental covariates. The CM signal measurements are done at equal time intervals ΔT while the two environmental covariates continuously change over time and are measured within each time intervals ΔT.

2.2 Offline degradation modeling based on the Brownian motion process

2.2.1 Model Structure

The Borwnian motion process has several good properties and is often used to describe the degradation of products (21; 43; 1; 33). In this paper, we model the degradation based on a Brownian motion process as follows:

$$X(t) = x_0 + \int_0^t \Gamma(\boldsymbol{\omega}(\tau))d\tau + \sigma W(t), \tag{1}$$

where x_0 is the initial degradation, $\Gamma : \mathbb{R}^q \to \mathbb{R}$ is the drift dependent on the level of environmental covariates $\boldsymbol{\omega}(\tau)$ at time τ, and σ is the diffusion parameter. $W(t)$ denotes the standard Brownian motion with the following three properties:

1. $W(0) = 0$, where $W(t) \in (-\infty, +\infty)$.
2. $W(t + \Delta t) - W(t) \sim \mathcal{N}(0, \Delta t)$.
3. $W(t) \sim \mathcal{N}(0, t)$.

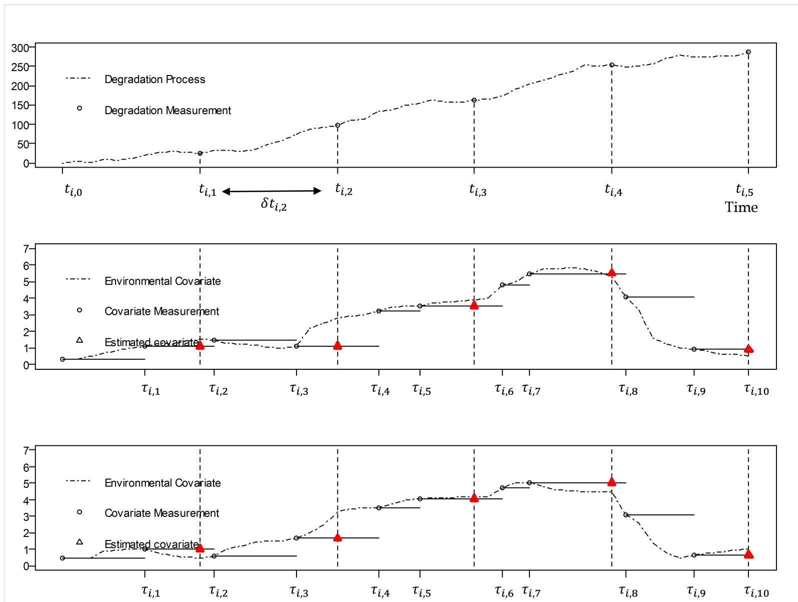


Figure 1: Unit degradation in response to time-varying environmental covariates.

We introduce an additive regression model to incorporate operating condition information into the degradation model. In particular, the drift of process for unit i at time t can be modeled as follows:

$$\Gamma(\boldsymbol{\omega}_i(t)) = \beta_0 + \sum_{l=1}^{q} f_l\left(\omega_{i,l}(t); \boldsymbol{\beta}_l\right), \tag{2}$$

where each function $f_l(.)$ models the effect of an environmental covariate. Here β_0 is the intercept and $\boldsymbol{\beta}_l$ denotes the parameter(s) in each function $f_l(.), l \in \{1, ..., q\}$. The coefficient vector of the functions parameters can be denoted as $\boldsymbol{\beta'} = \left[\boldsymbol{\beta}_1^t, ..., \boldsymbol{\beta}_q^t\right]^t$. For each covariate l, the function $f_l(\omega_{i,l}(\tau); \boldsymbol{\beta}_l)$ represents the effect of $\omega_{i,l}(\tau)$ at time τ on the drift of degradation process.

The cumulative damage model considered by equations (1) and (2) is motivated by the cumulative damage model for the accelerated failure-time model in Nelson (29). For certain degradation mechanisms (e.g. wear out, crack growth, and decomposition of chemical structures), the assumption of cumulative effects is appropriate. In the motivating example of this article, the environmental variables cause certain chemical reactions which leads to degradation of oil over the time. Thus, the assumption of cumulative effects is appropriate for the application.

Two approaches are typically available to choose the functional form for the effect functions $f_l(.)$. The first approach is based on the models motivated by physical, chemical and engineering knowledge. For example, the Arrhenius relationship (27) is typically used to model the temperature effect on the degradation rate. When there is not enough information on the correct form of $f_l(.)$ from physical/engineering knowledge or when such models do not fit the data well, nonparametric models can be used. For this approach, the functions $f_l(.)$ can be estimated as the linear combination of the spline bases. Considering the B-spline bases to non-parametrically model each covariate l, we have that:

$$f_l(\omega_{i,l}(\tau);\boldsymbol{\beta}_l) = \sum_{d=1}^{d_l} B^t_{d,l}(\omega_{i,l}(\tau))\beta_{d,l}, \tag{3}$$

where $B_{d,l}(\omega_{i,l}(\tau))$ is the value of the corresponding B-spline basis evaluated at $\omega_{i,l}(\tau)$, $\beta_{d,l}$'s are B-spline coefficients and d_l is the number of B-spline bases for each covariate l.

Based on the independent increment property of the Brownian motion process $\Delta X_i(t_{i,j}) = X_i(t_{i,j}) - X_i(t_{i,j-1})$ is independent and identically distributed with a normal distribution $\mathcal{N}\left(\int_{t_{i,j-1}}^{t_{i,j}} \Gamma(\boldsymbol{\omega}(\tau))d\tau, (\sigma\sqrt{\delta t_{i,j}})^2\right)$ where σ is the diffusion of the process and $\delta t_{i,j} = t_{i,j} - t_{i,j-1}$. Considering the additive regression model in (2), the mean of $\Delta X_i(t_{i,j})$ in each interval $[t_{i,j-1}, t_{i,j}]$ can be written as follows:

$$\mu_{ij} = \int_{t_{i,j-1}}^{t_{i,j}} \Gamma(\boldsymbol{\omega}_i(\tau))d\tau = \beta_0 \delta t_{i,j} + \int_{t_{i,j-1}}^{t_{i,j}} \sum_{l=1}^{q} f_l\left(\omega_{i,l}(\tau);\boldsymbol{\beta}_l\right) d\tau. \tag{4}$$

It should be noted that most of the IoT systems record the environmental covariates information in discrete time points. Therefore, a discrete-data version of (4) should be estimated for any real life application. Such an estimate of equation (4) can be obtained as follows:

$$\mu_{ij} = \beta_0 \delta t_{i,j} + \sum_{l=1}^{q} \sum_{t_{i,j-1} \leq \tau_{i,k} < t_{i,j}} f_l(\omega_{i,l}(\tau_{i,k});\beta_l)(\tau_{i,k+1} - \tau_{i,k}). \tag{5}$$

Here, $\tau_{i,k}$ are the time points where the environmental covariates were recorded for unit i. Thus, considering B-splines to model $f_l(.)$ as in (3), equation (4) can be rewritten as follows:

$$\mu_{ij} = \beta_0 \delta t_{i,j} + \sum_{l=1}^{q} \sum_{d=1}^{d_l} G^t_{d,l}(t_{i,j})\beta_{d,l}, \tag{6}$$

where $G_{d,l}(t_{ij}) = \sum_{t_{i,j-1} \leq \tau_{i,k} < t_{i,j}} B_{d,l}(\omega_{i,l}(\tau_{i,k}))(\tau_{i,k+1} - \tau_{i,k})$.

One popular way of modeling the additive regression is through penalized splines, and here we focus on the penalized B-splines or P-splines (8). For covariate $l = 1, ..., q$ suppose we choose B-spline of degree m_l and we consider K_l interior knots equally spaced on the input interval. Then, the cumulative effect of each covariate in each time interval can be modeled as $s_l(t_{i,j}) = \boldsymbol{G}^t_l(t_{i,j})\boldsymbol{\beta}_l$ where $\boldsymbol{G}_l(t_{i,j})$ is constructed based on a set of $d_l = K_l + m_l - 1$ basis functions evaluated at $t_{i,j}$ and $\boldsymbol{\beta}_l = [\beta_{1,l}, ..., \beta_{d_l,l}]^t$ is the corresponding d_l-vector of coefficients.

To avoid overfitting, we augment the bases with a quadratic penalty on the coefficients, typically based on the squared difference of the adjacent coefficients. Such penalty can be written in the form of $P_l = \lambda_l \sum_{d=1}^{d_l-1} (\beta_{d+1,l} - \beta_{d,l})^2$ which has a matrix form as follows:

$$P_l = \lambda_l \boldsymbol{\beta}_l^t \begin{bmatrix} 1 & -1 & 0 & \dots \\ -1 & 2 & -1 & \dots \\ 0 & -1 & 2 & \dots \\ \vdots & \vdots & \vdots & \dots \end{bmatrix} \boldsymbol{\beta}_l = \lambda_l \boldsymbol{\beta}_l^t \boldsymbol{S}_l \boldsymbol{\beta}_l, \tag{7}$$

where $\lambda_l \geq 0$ is a smoothing parameter. Increase in the λ_l forces a smoother curve, while $\lambda_l = 0$ implies no penalization. In this paper, we focus on the cubic B-splines and set the degree as 3 for all covariates, i.e. $m_1 = ... = m_q = 3$. Regarding the number of knots, for simplicity, we use the same number of interior knots for all q covariates.

2.2.2 Model Estimation

Let $\Psi = [\beta_0, \sigma, \boldsymbol{\lambda}^t]^t$ denote the parameters corresponding to the parametric component of the model where $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, ..., \lambda_q]^t$. Also, let $\boldsymbol{\beta}' = [\boldsymbol{\beta}_1^t, ..., \boldsymbol{\beta}_q^t]^t$ denote the coefficients in the non-parametric component corresponding to the B-splines. The penalized log-likelihood estimate of model parameters considering the independent increment property of Brownian motion process can be written as follows:

$$\ell_P(\Psi, \boldsymbol{\beta}') = \sum_{i=1}^{N} \sum_{j=1}^{n_i} \ln\left(p(\Delta X_i(t_{i,j}) | \Psi, \boldsymbol{\beta}')\right) - 2^{-1} \sum_{l=1}^{q} \lambda_l \boldsymbol{\beta}_l^t \boldsymbol{S}_l \boldsymbol{\beta}_l, \tag{8}$$

where $p(\Delta X_i(t_{i,j})|\Psi, \boldsymbol{\beta}') = \mathcal{N}(\beta_0 \delta t_{i,j} + \boldsymbol{G}^t(t_{i,j})\boldsymbol{\beta}', (\sigma\sqrt{\delta t_{i,j}})^2)$ and $\boldsymbol{G}(t_{i,j}) = [G_1^t(t_{i,j}), ..., \ G_q^t(t_{i,j})]^t$. Since the works of (17), (38) and (35), it has been recognized that this penalized likelihood estimation can be reformulated as a mixed effects model by considering that the quadratic penalty amounts to assuming a normal random effects distribution. Specifically for $l = 1, ..., q$, let $p(\boldsymbol{\beta}_l|\lambda_l) = \mathcal{N}_{d_l}\left(0, (\lambda_l \boldsymbol{S}_l)^{-1}\right) = c_0 \lambda^{2^{-1} d_l} \exp(-2^{-1}\lambda_l \boldsymbol{\beta}_l^T \boldsymbol{S}_l \boldsymbol{\beta}_l)$ where c_0 is the normalizing constant independent of $\boldsymbol{\beta}_l$ and λ_l. Noting that $\boldsymbol{S}_l$ is of full rank, we then have the marginal log-likelihood as follows:

$$\ell_P(\Psi) = \ln\left(\int \prod_{i=1}^{N} \prod_{j=1}^{n_i} p(\Delta X_i(t_{i,j})|\Psi, \boldsymbol{\beta}') \prod_{l=1}^{q} p(\beta_l|\lambda_l) d\beta_l\right). \tag{9}$$

Formulating the additive regression drift as a mixed effects model allows us to exploit the wealth of mixed model methodology for inference. We can utilize an EM algorithm to estimate the model parameters.

The EM algorithm is an iterative method for computing maximum likelihood estimators by alternating between the expectation step, the E-step, and its maximization, the M-step, at every iteration until convergence. We notice here that the coefficients $\boldsymbol{\beta}'$ are actually hidden random variables in our mixed effects framework. The EM algorithm proceeds by computing the expectation of the log-likelihood of the complete data with respect to the posterior $p(\boldsymbol{\beta}'|\mathcal{D};\Psi)$ in the E-step. It is straightforward to show that the posterior distribution of $\boldsymbol{\beta}'$ is multivariate normal $\mathcal{N}_D(\boldsymbol{\beta}'|\boldsymbol{\mu}_{\beta'},\boldsymbol{\Sigma}_{\beta'})$ with :

$$\begin{aligned}
\boldsymbol{\Sigma}_{\beta'} =&\Big(\frac{1}{\sigma^2}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\frac{\boldsymbol{G}(t_{i,j})\boldsymbol{G}(t_{i,j})^t}{\delta t_{i,j}}+\boldsymbol{S}_\lambda\Big)^{-1},\\
\boldsymbol{\mu}_{\beta'} =&\Big(\frac{1}{\sigma^2}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\frac{\boldsymbol{G}(t_{i,j})\boldsymbol{G}(t_{i,j})^t}{\delta t_{i,j}}+\boldsymbol{S}_\lambda\Big)^{-1}\times\sum_{i=1}^{N}\sum_{j=1}^{n_i}(\Delta X_i(t_{i,j})-\beta_0)\boldsymbol{G}(t_{i,j}),
\end{aligned} \tag{10}$$

where $\boldsymbol{S}_\lambda$ is a $D\times D$ block diagonal matrix formed by taking the blocks of $\lambda_l\boldsymbol{S}_l$ for $l=1,...,q$ and $D=\sum_{l=1}^{q}d_l$. At each iteration ν of the EM algorithm, the expected value of the logarithm of complete likelihood with respect to the posterior $p(\boldsymbol{\beta}'|\mathcal{D};\Psi)$ is as follows:

$$Q^{(\nu)}(\boldsymbol{\Delta X},\boldsymbol{\beta}';\Psi)=E_{p(\boldsymbol{\beta}'|\boldsymbol{\Delta X};\Psi^{(\nu)})}\{\ell_P(\Psi,\boldsymbol{\beta}')\}, \tag{11}$$

where $\boldsymbol{\Delta X}=\Big\{\{\Delta X_i(t_{i,j})\}_{j=1}^{n_i}\Big\}_{i=1}^{N}$ is the collection of degradation process shift observations from historical data. The E-step corresponds to analytical calculation of the function $Q^{(\nu)}$, considering the model parameters $\Psi^{(\nu)}$ calculated in the previous iteration. The M-step entails maximizing the $Q^{(\nu)}(\boldsymbol{\Delta X},\boldsymbol{\beta}';\Psi)$ with respect to the parameters β_0,σ and λ. In this regard, we get the following formulas to update the parameters β_0, σ and λ:

$$\begin{aligned}
\beta_0^{(\nu+1)} =&\frac{1}{\sum_{i=1}^{N}\sum_{j=1}^{n_i}\delta t_{i,j}}\sum_{i=1}^{N}\sum_{j=1}^{n_i}(\Delta X_{ij}-\boldsymbol{G}^t(t_{ij})\boldsymbol{\mu}_{\beta'}^{(\nu)}),\\
\sigma^{(\nu+1)} =&\Big(\frac{1}{M}\sum_{i=1}^{N}\sum_{j=1}^{n_i}(\frac{\big(R_i(t_{ij})-\boldsymbol{G}^t(t_{ij})\boldsymbol{\mu}_{\beta'}^{(\nu)}\big)^2}{\delta t_{i,j}}+\frac{\boldsymbol{G}^t(t_{ij})\boldsymbol{\Sigma}_{\beta_l}^{(\nu)}\boldsymbol{G}(t_{ij})}{\delta t_{i,j}})\Big)^{1/2},\\
\lambda_p^{(\nu+1)} =&d_l\left({\boldsymbol{\mu}_{\beta_l}^{(\nu)}}^t\boldsymbol{S}_l\boldsymbol{\mu}_{\beta_l}^{(\nu)}-\mathrm{tr}(\boldsymbol{S}_l\boldsymbol{\mu}_{\beta_l}^{(\nu)})\right)^{-1},
\end{aligned} \tag{12}$$

where $M=\sum_{i=1}^{N}n_i$ and $R_i(t_{i,j})=\Delta X_i(t_{i,j})-\beta_0\delta t_{i,j}$ denotes the residual observations from parametric component. Please refer to appendix for a detailed derivation of these updating formulas. Notice that after initializing the parameters to some values $(\beta_0^{(0)},\sigma^{(0)},\boldsymbol{\lambda}^{(0)})$, the algorithm proceeds by iteratively performing the E-step and M-step until some convergence criterion is satisfied. Here, we note that inference for $\boldsymbol{\beta}'$ can be directly obtained since the sufficient statistic for its posterior is computed in the E-step. Using this property, we propose a Bayesian updating scheme for online updating based on the new data available from the in-service unit in the next section.

2.3 RUL prediction via online Bayesian updating

The proposed RUL prediction approach requires the simultaneous monitoring of both degradation signal and operating conditions. For in-service unit r at the current time point t_h when the prediction is to be made, assume s_r is the recent value of degradation signal observation, i.e. $\boldsymbol{X}_r(\boldsymbol{t}_{r,0:h}) = [X_r(t_{r,0}), ..., X_r(t_{r,s_r})]^t$ and $\boldsymbol{t}_{r,0:h} = [t_{r,0}, t_{r,1}, ..., t_{r,s_r}]^t$ where $t_{r,s_r} \leq t_h$. Moreover, we consider $\Delta\boldsymbol{X}_r(\boldsymbol{t}_{r,1:h}) = [\Delta X_r(t_{r,1}), ..., \Delta X_r(t_{r,s_r})]^t$ and $\hat{\boldsymbol{R}}_r(\boldsymbol{t}_{r,1:h}) = [\hat{R}_r(t_{r,1}), ..., \hat{R}_r(t_{r,s_r})]^t$ where $\hat{R}_r(t_{r,j}) = \Delta X_r(t_{r,j}) - \hat{\beta}_0 \delta t_{r,j}$ as the vector of estimated residual observations for in-service unit r. In addition, let $\boldsymbol{\Omega}_r(\boldsymbol{t}_{r,h}) = \{\boldsymbol{\omega}_r(\tau) : 0 \leq \tau \leq t_{r,s_r}\}$ denote the environmental covariates observations from unit r up to s_r. The posterior distribution of the environmental effect coefficients $\boldsymbol{\beta}'^{(r)}$ based on the newly observed data from the in-service unit up to time s_r, can be computed as $p(\boldsymbol{\beta}'^{(r)}|\hat{\boldsymbol{R}}_r(\boldsymbol{t}_{r,1:h}), \boldsymbol{\Omega}_r(\boldsymbol{t}_{r,h}), \hat{\boldsymbol{\Psi}}) \propto p(\hat{\boldsymbol{R}}_r(\boldsymbol{t}_{r,1:h}), \boldsymbol{\Omega}_r(\boldsymbol{t}_{r,h})|\boldsymbol{\beta}'^{(r)}, \hat{\boldsymbol{\Psi}})\pi(\boldsymbol{\beta}'^{(r)}, \hat{\boldsymbol{\Psi}})$ where $p(\hat{\boldsymbol{R}}_r(\boldsymbol{t}_{r,1:h}), \boldsymbol{\Omega}_r(\boldsymbol{t}_{r,h})|\boldsymbol{\beta}'^{(r)}, \hat{\boldsymbol{\Psi}})$ denotes the likelihood for in-service unit r, which by the property of Brownian motion process is normally distributed, and $\pi(\boldsymbol{\beta}'^{(r)}, \hat{\boldsymbol{\Psi}})$ refers to the prior distribution estimated in the offline stage and in the E-step of EM algorithm. Assuming normally distributed $\boldsymbol{\beta}'^{(r)}$, and following the conjugate property of multivariate normal distribution, the posterior is also normally distributed $p(\boldsymbol{\beta}'^{(r)}|\hat{\boldsymbol{R}}_r(\boldsymbol{t}_{r,1:h}), \boldsymbol{\Omega}_r(\boldsymbol{t}_{r,h}); \hat{\boldsymbol{\Psi}}) \sim \mathcal{N}(\hat{\boldsymbol{\mu}}_{\beta'_{r,h}}, \hat{\boldsymbol{\Sigma}}_{\beta'_{r,h}})$ where $\hat{\boldsymbol{\mu}}_{\beta'_{r,h}}$ and $\hat{\boldsymbol{\Sigma}}_{\beta'_{r,h}}$ represent the mean and covariance matrix of the posterior. Considering the prior parameter estimates $\boldsymbol{\mu}_{\beta'_0}$ and $\boldsymbol{\Sigma}_{\beta'_0}$, it is straightforward to derive the closed form expression for the posterior mean and covariance matrix as follows:

$$\begin{aligned}
\hat{\boldsymbol{\mu}}_{\beta'_{r,h}} &= \hat{\boldsymbol{\Sigma}}_{\beta'_{r,h}} \left[\frac{\boldsymbol{G}^t(\boldsymbol{t}_{r,1:h})\Delta T^{-1}\hat{R}_r(t_{r,j})}{\sigma^2} + \boldsymbol{\Sigma}_{\beta'_0}^{-1}\boldsymbol{\mu}_{\beta'_0} \right], \\
\hat{\boldsymbol{\Sigma}}_{\beta'_{r,h}} &= \left[\boldsymbol{\Sigma}_{\beta'_0}^{-1} + \frac{\boldsymbol{G}^t(\boldsymbol{t}_{r,1:h})\Delta T^{-1}\boldsymbol{G}(\boldsymbol{t}_{r,1:h})'}{\sigma^2} \right]^{-1},
\end{aligned} \tag{13}$$

where $\boldsymbol{G}^t(\boldsymbol{t}_{r,1:h}) = [\boldsymbol{G}(t_{r,1}), ..., \boldsymbol{G}(t_{r,s_r})]$, $\Delta T = \text{diag}(\delta t_{r,1}, ..., \delta t_{r,s_r})$. Also $\boldsymbol{\mu}_{\beta'_0}$, $\boldsymbol{\Sigma}_{\beta'_0}$ and σ can be replaced by their estimates $\hat{\boldsymbol{\mu}}_{\beta'_0}, \hat{\boldsymbol{\Sigma}}_{\beta'_0}$ and $\hat{\sigma}$ obtained in the previous section from the offline stage. It should be noted that the model updating performed in equation (13) involves the inversion of $\boldsymbol{\Sigma}_{\beta'_0}$ which has a computational complexity of $O((q \times d_l)^3)$. We note that $q \times d_l$ is always reasonably small and therefore updating of the parameters does not take much of time.

The RUL of a unit is defined as the remaining time until its degradation level reaches a pre-defined failure threshold D_F. Two assumptions for RUL prediction under time-varying environmental covariates can be considered. The first one assumes that the future environmental covariates are unknown and can be estimated from their most recent observation. Let T_r denote the RUL of unit r, i.e. $T_r = \inf\{t \geq 0 : X_r(t_{r,s_r}) + X_r(t) \geq D_F\}$ where $D_F \geq X_r(t_{r,s_r})$. Then assuming constant stress level $\boldsymbol{\omega}_r(t_{r,s_r})$ after t_{r,s_r} i.e. $\{\boldsymbol{\omega}_r(\tau) = \boldsymbol{\omega}_r(t_{r,s_r}) : t_{r,s_r} \leq \tau\}$, the RUL of unit r conditional on $\hat{\Psi}$ and the updated coefficients $\boldsymbol{\beta}'^{(r)}$ follows an inverse Gaussian (IG) distribution as:

$$P(T_r = t|X_r(t_{r,s_r}), \boldsymbol{\Omega}_r(t_{r,h}), \boldsymbol{\beta}'^{(r)}, \hat{\Psi}) \sim \mathcal{IG}(t; \mu_{\mathcal{IG}}(t_h), \lambda_{\mathcal{IG}}(t_h)), \tag{14}$$

where $\mathcal{IG}(t; ., .)$ represents the pdf of an IG distribution, $\mu_{\mathcal{IG}}(t_h) = \frac{D_F - X_r(t_{r,s_r})}{\hat{\beta}_0 + \boldsymbol{B}(\boldsymbol{\omega}_r(t_{r,h}))^t \hat{\boldsymbol{\mu}}_{\beta'_{r,h}}}$ is the mean parameter of IG distribution, and $\lambda_{\mathcal{IG}}(t_h) = \frac{(D_F - X_r(t_{r,s_r}))^2}{\hat{\sigma^2}}$ is the shape parameters of the IG distribution. Here, recall that $\boldsymbol{\omega}_r(t_{r,h})$ is the most recent observation of the environmental conditions for in-service unit r and $\boldsymbol{B}(\boldsymbol{\omega}_r(t_{r,h}))^t = [\boldsymbol{B}_1(\boldsymbol{\omega}_r(t_{r,h})), ..., \boldsymbol{B}_q(\boldsymbol{\omega}_r(t_{r,h}))]^t$ are the B-spline bases.

Typically we have control over the environmental covariates and they can be set according to the specific application. The second assumption, which is used in this study, assumes that the future operating conditions can be set in advance; therefore, the RUL of the component can be predicted based on the profiled environmental covariates. Assume that the stresses in the future after t_{r,s_r} is a time-varying series and denoted as $\boldsymbol{\Omega}'(t_{r,s_r}) = \{\boldsymbol{\omega}_r(\tau) : s_r \leq \tau\}$. It is quite difficult to obtain an analytical formula of the RUL distribution for a Brownian motion process under varying drift parameters. For this reason, we use a Monte Carlo simulation method to estimate the RUL distribution approximately. The basic idea is to simulate the evolution of the degradation process in the future with the knowledge at time t_{r,s_r}. For simplicity, the time interval for predictions are assumed to be uniform and equal to unit interval 1, namely $t_{r,s_r+j} - t_{r,s_r+j-1} = 1, j = 1, 2, ..., m$. For a simulated degradation path, the m-step prediction of the future state, that is $\boldsymbol{X}_r(t_{r,s_r})$, under varying stress series $\boldsymbol{\Omega}'(t_{r,s_r})$ is given by the following procedure:

Step 1: Generate a sample $\boldsymbol{\beta}'^{(r)^*}$ from $\mathcal{N}(\hat{\boldsymbol{\mu}}_{\beta'_{r,h}}, \hat{\boldsymbol{\Sigma}}_{\beta'_{r,h}})$;

for $j = 1, 2, ..., m$;

Step 2: Generate a sample $\Delta X_r(t_{r,s_r+j})$ from $\mathcal{N}(\hat{\beta}_0 + \sum_{l=1}^{q} \sum_{t_{r,s_r+j-1} \leq \tau_{r,k} < t_{r,s_r+j}} f_l(\omega_{r,l}(\tau_{r,k}); \beta_l'^{(r)^*})(\tau_{r,k+1} - \tau_{r,k}), \sigma^{*2})$

Step 3: Let $X_r(t_{r,s_r+j}) = X_r(t_{r,s_r+j-1}) + \Delta X_r(t_{r,s_r+j})$.

Repeat Step 1-3 U times (say $U = 1000$), we can obtain U simulated paths to predict the states of $X_r(t)$ after time t_{r,s_r}. To obtain the distribution of the RUL, we can determine when the U samples fail, respectively. As mentioned above, the RUL for a sample can be obtained by the first simulated steps when $\boldsymbol{X}_r(\boldsymbol{t}_{r,s_r+1:s_r+m})$ exceeds D_F. Then, we have U simulated RULs, which are denoted as $\{T_r^1, T_r^2, ..., T_r^U\}$. The mean and confidence interval can be estimated approximately using these simulated RULs. The mean RUL at time t_{r,s_r} is estimated by $\frac{1}{U}\sum_{u=1}^{U} T_r^u$. The approximated $100(1-\alpha)\%$ CI for RUL at time t_{r,s_r} is estimated by $\left[T_r^{(LB)}, T_r^{(UB)}\right]$, where $LB = \lfloor(\alpha/2) \cdot U\rfloor$, $UB = \lfloor(1 - (\alpha/2)) \cdot U\rfloor$ and $\lfloor\cdot\rfloor$ means round to the nearest integer and $T_r^{(u)}$ is the u-th ordered statistic of the set of RULs for unit r.

3. Numerical Study

In this section we conduct simulations to validate the performance of the proposed modeling approach. Specifically, we first discuss the general procedure to generate degradation signals considering the operating conditions and to evaluate the performance of different methods. Then using the simulated signals, we demonstrate the advantage of our proposed framework.

3.1 Simulation Setup

We simulate degradation signals from three model settings. In each model setting, we report the RUL prediction accuracy of a partially observed in-service unit at varying time points t^*. Specifically, we report the prediction accuracy for different percentiles of the observed in-service signal (i.e. 35%, 50% and 85%). Our procedure is based on simulating 20 degradation signals from a specific model setting and then randomly selecting one of them as the in-service unit, and treating the rest as our historical dataset. This procedure is repeated 1000 times, and each time we report the absolute error (AE) between the true RUL ($T_r(t)$) and the estimated RUL ($\hat{T}_r(t^*)$ as follows:

$$AE(t^*) = |\hat{T}_r(t^*) - T_r(t)|, \tag{15}$$

where, as mentioned before, a unit is deemed failed once its degradation signal passes its failure threshold. Specifically, we compare the performance of our approach with the degradation modeling approach proposed in (25) where a Brownian motion process with a parametric stress-dependent drift is defined to model the degradation process.

Let us first assume that we have signals generated from a Brownian motion process with a drift dependent on the effect of one environmental covariate where $\mu_{ij} = \beta_0 \delta t_{i,j} + \alpha_i \sin(\pi\omega_i(t_j))(t_{j+1} - t_j)$, $i \in \{1, ..., 20\}$, $j \in \{1, ..., 20\}$. Specifically, we assume that $\alpha_i \sim U[3, 3.5]$ and we generate ω_i from a uniform distribution $U[0, 1]$. The initial degradation measure is normally distributed as $x_0 \sim \mathcal{N}(20, 1)$, the diffusion is set to $\sigma = 1$ and degradation signal observations are made at equal time intervals $\delta t_{i,j} = 0.5, \forall i, j$. Figure 2 demonstrates the signals generated in such a setting.

As shown in Figure 2, the generated degradation signals are heterogeneous and do not follow a common functional form. The reason is that unlike the traditional RUL prediction models (14; 19; 18) where the degradation signals are generated from a common function and have similar temporal evolution, here in each degradation signal observation stage, the Brownian motion drift is generated according to a random covariate effect function. Considering these historical signal observations, we can use the framework proposed in this study which utilizes environmental covariate observations to model the degradation process. For the sake of comparison, we can fit the mixed effects model proposed in (14) to predict the RUL of a new in-service unit. The model in (14) utilizes B-splines to model the temporal evolution of degradation signals and does not consider the effect of environmental covariates. Figure 3 demonstrates the prognosis results for the new

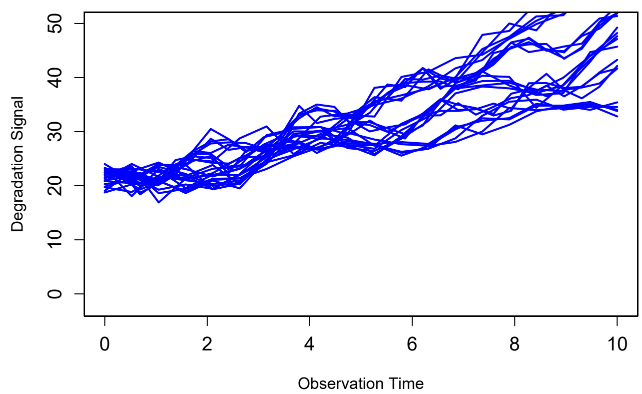


Figure 2: Historical signal observations

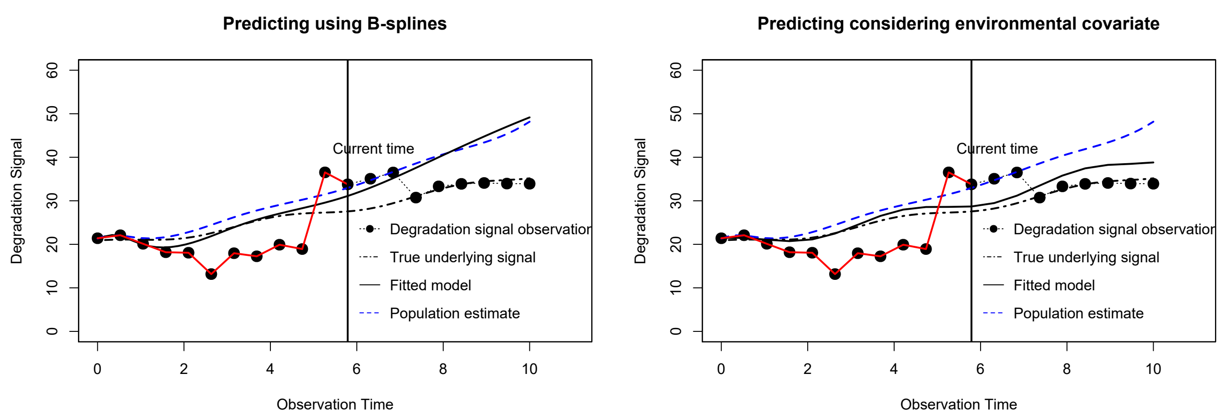


Figure 3: Prognostics using B-spline vs. using proposed method considering the environmental covariate.

in-service unit considering the mean of the parameters' posterior and the future environmental conditions as discussed in section 2.3.

From Figure 3, it can be observed that the prediction results based on B-splines as proposed in (14) are misleading. The main reason is that the historical signal data is heterogeneous hence making it difficult, even for a non-parametric approach like B-splines, to converge to the true underlying signal of the in-service unit. However, the model based on the environmental covariate observations can quickly adapt to the trend of the new in-service unit. The reason is that this model is not limited to the temporal evolution of the degradation signal and performs regression on the environmental covariate which controls the degradation process evolution over the time in future.

In order to get insight on the performance of the proposed framework, we perform further numerical study. We assume that the initial degradation measure is $x_0 = 0.5$ and the diffusion parameter is set to $\sigma = 1$, which are identical across all units. Moreover, it is assumed that the degradation signal observations are made at equal time intervals $\delta t_{i,j} = 1, \forall i, j$, the environmental covariates are fixed during the time intervals, and the failure threshold is set to $D_F = 350$. We simulate degradation signals dependent on the effect of three environmental covariates from the Brownian motion with the following drift in each interval j of unit i:

$$\mu_{ij} = \beta_0 \delta t_{i,j} + \sum_{l=1}^{3} f_l(\omega_{i,l}(t_j); \beta_l)(t_{j+1} - t_j), \tag{16}$$

Specifically, for $j \in \{1, ..., 30\}$ and $i \in \{1, ..., 20\}$ a vector of three covariates $(\omega_{i,l}(t), l = 1, 2, 3)$ is generated by simulating $\omega_{i,1}$ from a uniform distribution $U[0, 2]$, $\omega_{i,2} = 0.9\omega_{i,1} + e_1$ where $e_1 \sim U[0, 1]$ and $\omega_{i,3}$ from a uniform distribution $U[0, 1]$. We adopt the linear function of stress as proposed in (25) with some modifications to model the stress-drift relationship of $\omega_{i,1}$. To be more specific, we consider $f_1(\omega_{i,1}) = \alpha_{i,1}\omega_{i,1}$ where $\alpha_{i,1} \sim \mathcal{N}(0.1, 0.25^2)$. The second covariate effect function is designed as $f_2(\omega_{i,2}) = \exp(-\alpha_{i,2}\omega_{i,2})$ where $\alpha_{i,2} \sim \mathcal{N}(2.5, 0.25^2)$. The effect of the third covariate is modeled as $f_3(\omega_{i,3}) = \alpha_{i,3} \sin(\pi\omega_{i,3})$ where $\alpha_{i,3} \sim U[1.2, 2.5]$. It is also considered that $\beta_0 = 20$.

3.2 Performance Comparison

We first generate the degradation signals as discussed in section 3.1. In model setting I, we initially only consider the effect of the first covariate for degradation modeling using the procedure proposed in section 2. Then, we add the effect of the second covariate to the proposed framework and finally we consider the effect of all three covariates in RUL prediction. These three models are denoted as "Additive Regression I", "Additive Regression II", "Additive Regression III", respectively. In each of these three models, we compare the performance of our proposed framework with (25)'s approach where the drift is modeled to be linearly dependent on $\omega_{i,1}$ for each unit i denoted as "Linear Drift Model". Figure 4 shows the results of numerical study in this model setting.

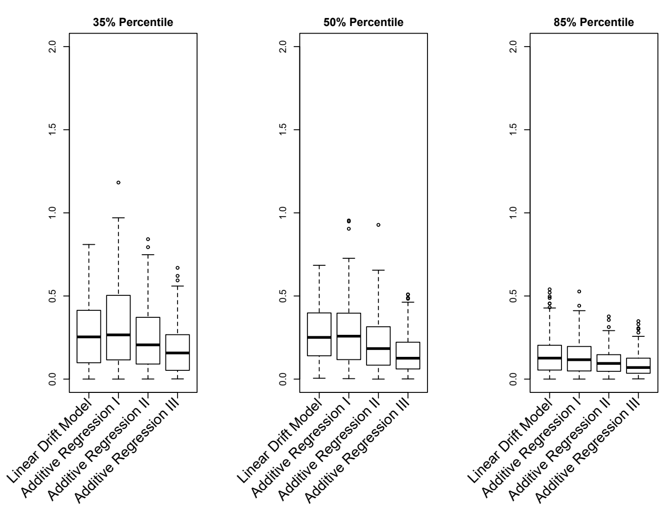


Figure 4: Results of model setting I

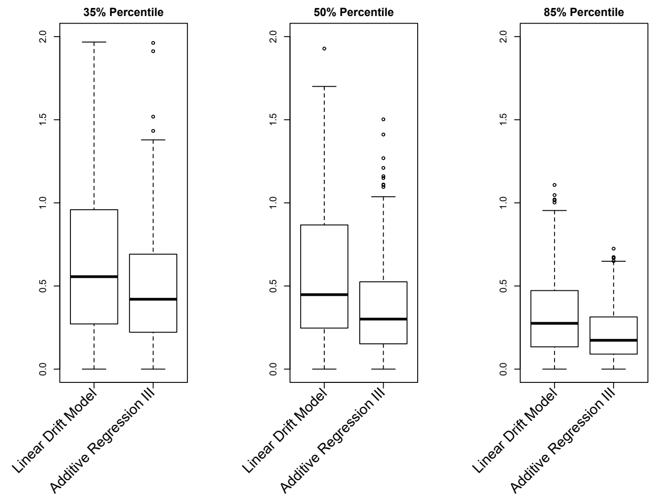


Figure 5: Results of model setting II

Figure 4 shows the performance of the proposed framework in comparison to the Brownian motion process with parametric linear drift model proposed in (25). As can be seen from this figure, with increase in the number of observations from the in-service unit the prediction accuracy generally improves. Moreover, figure 4 shows that as we include more covariates in the proposed degradation modeling framework, the prediction accuarcy increases. We would, in fact, expect these results as the originally generated degradation signals are affected by all three environmental covariates. Removing covariates from the degradation modeling framework introduces extra bias in estimating the diffusion term of the Brownian motion process and consequently causes inaccurate RUL prediction.

Figure 4 also shows that in the case where we only consider the effect of the first covariate, the non-parametric approach proposed in this paper performs relatively similar to the linear drift model based on this covariate. In the earlier stage of RUL prediction when only 35% of in-service unit degradation signal is observed, however, the parametric approach performs relatively better than our non-parametric approach. This is intuitively understandable as the linear drift model is based on the true parametric function and can inherently capture the correct unit specific parameters even with small number of observations from the in-service unit.

In order to further investigate the performance of our approach, we conduct two more numerical studies. Specifically in model setting II, we investigate the effect of proposed approach when the noise level is high. In this model setting, we set the diffusion parameter of the Brownian motion process to $\sigma = 5$. Figure 5 demonstrates the performance of the parametric model which only considers the effect of one covariate with the proposed procedure with all three covariates. It can be seen that increasing the diffusion generally increases the prediction error of different approaches; however, the proposed framework of this paper which considers all three covariates effect remains superior.

In model setting III, we add extra environmental covariates' observations in each time interval i, j. TO be more specific, we assume that in each time interval i, j of degradation shift, we observe environmental covariates in four equally-spaced time intervals. The drift-effect relationship for each covariate l would then be modeled as $\sum_{t_{i,j-1} \leq \tau_{i,k} < t_{i,j}} f_l(\omega_{i,l}(\tau_{i,k}); \beta_l)(\tau_{i,k+1} - \tau_{i,k})$ where $\tau_{i,k+1} - \tau_{i,k} = \frac{1}{4}\ \forall i, j$. Specifically, we compare the performance of our proposed framework with the case where we ignore the extra environmental covariate observations and only consider the observations made at the beginning of each degradation measurement interval. We denote the former model with extra environmental observations modeling as "Additive Regression V" and the latter as "Additive Regression IV" model. Figure 6 demonstrates the performance of two models. As can be seen from Figure 6, ignoring extra environmental covariates information affecting the degradation process in our modeling, indeed, hurts the RUL prediction accuracy.

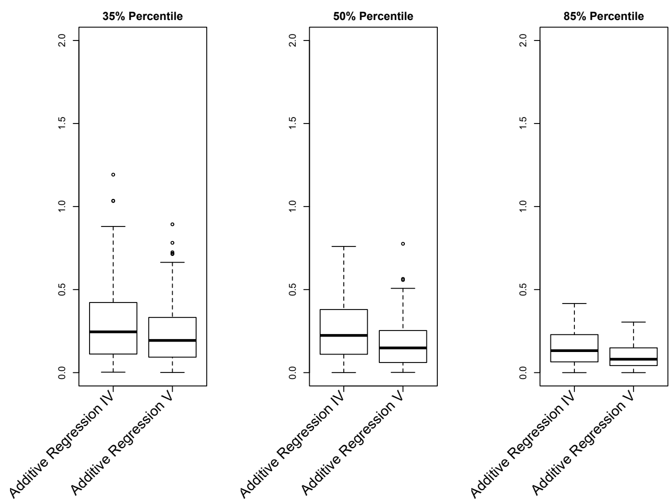


Figure 6: Results model setting III

4. Case Study

This section applies the proposed framework on a real-world case study of frying oil degradation modeling. During the frying process, typically, a complex series of physical and chemical reactions occur leading to the degradation of oil. The total polar component (TPC) which is a measure of proportion of polar materials in the oil, is the most predominant indicator for oil degradation. TPC can be obtained by measuring the dielectric of the oil. Fortunately, modern fryers are equipped with a sensor that can provide TPC measurements and daily TPC readings from the connected fryers can be collected. Per oil disposal, the frying vat is refilled with fresh oil and the TPC values reset to a lower value. The oil is typically deemed bad once its TPC value hits a pre-specified threshold value. We note that this threshold value is domain specific and depends on the degradation process under study. Also, this threshold value does not necessarily indicates hard failure and merely suggests that the oil is not healthy any more with TPC above this value.

This pre-specified threshold value is considered to be 25 in this study. Figure 7 demonstrates the TPC measurements for 12 such oil disposal cycles.

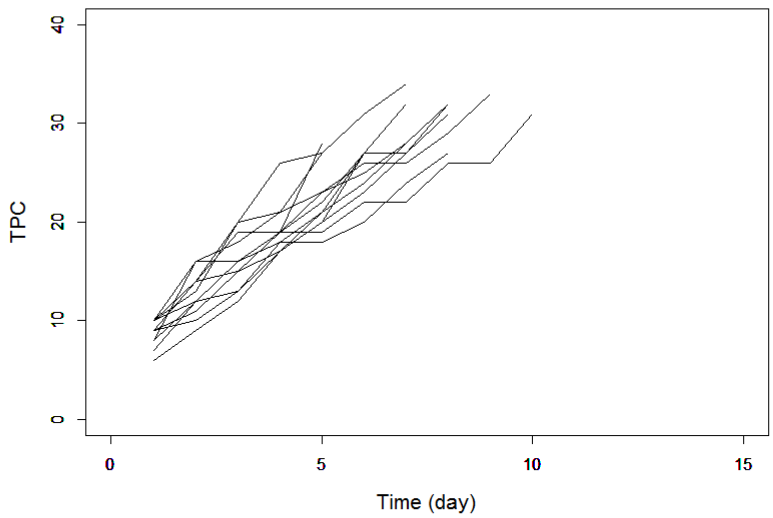
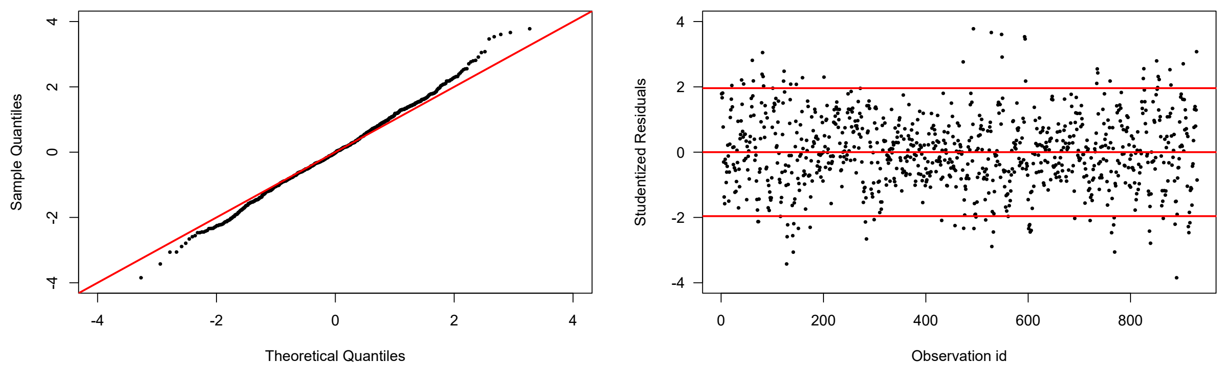


Figure 7: Degradation paths of frying oil

Figure 7 shows that the degradation rates of frying oil are considerably heterogeneous in different oil disposal cycles. Moreover, the degradation rate of each oil disposal cycle is different in different time intervals. This is, in fact, directly related to the condition that the frying oil was exposed to. Oil temperature is the most significant time-varying covariate contributing to the oil degradation. Beside temperature, other covariates like number of cooks and number of filterings in each time interval influence the corresponding degradation rate. We also checked the degradation data model fit based on the Brownian motion process for the oil degradation data. The results are in figure 8. The plot of studentized residuals in figure 8 shows that the constant variance assumption holds reasonably well. We can also check the normality assumption using the QQ-plot. The plot shows that the normality assumptions holds well. These graphical check shows that the overall the model assumptions holds reasonably well.

Figure 9 shows the plot of estimated functions for the drift-environmental covariates relationship estimated by the procedure introduced in Section 2.2. This figure, further, demonstrates that there is a strong non-linear relationship between the temperature and the drift of Brownian motion process defining the degradation.

In order to investigate the performance of our proposed method, we pursue the leave-one-out cross validation scheme for the available 117 oil disposal cycles. To be more specific, we leave one of the available cycles as the testing data and fit the model using all other available data. Then, the prediction for the in-service unit is performed for different percentiles of its lifespan. The whole procedure is repeated 117 times and the AE is calculated each time. Each time, we fit three models similar to numerical study section.



7(a) Q-Q plot

7(b) Residuals

Figure 8: QQ-plot and the studentized residual plots for the residuals of the degradation data.

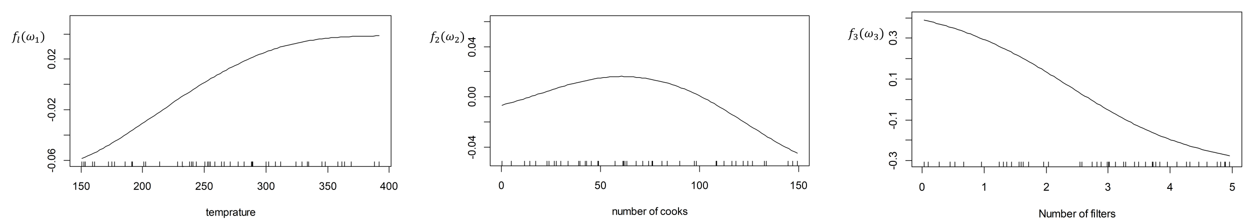


Figure 9: Estimated covariate effect functions and the corresponding approximate 95% confidence interval.

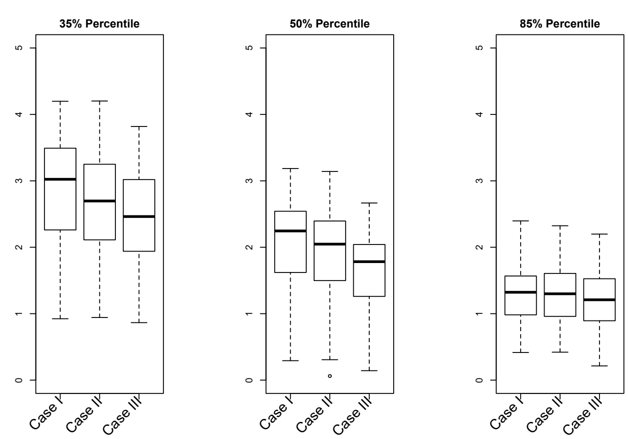


Figure 10: Results of case study

The first model considers only the effect of temperature, the second model adds the effect of number of cooks and the third model considers all three covariates. We denote these three models as “Case I”, “Case II” and “Case III” respectively. Figure 10 summarizes the results for the case study.

As can be seen from Figure 10, increasing the number of covariates using the proposed framework increases the RUL prediction accuracy. We would, indeed, expect this result as increasing the number of covariates increases the available information affecting the change in the oil degradation process.

5. Conclusion

In this paper, we propose a statistical framework for modeling degradation signals using a Brownian motion process with a drift dependent on the time-varying environmental covariates. Specifically, our proposed framework considers a semi-parametric regression based on penalized splines to model the environmental covariates-drift relationship. An EM algorithm is further developed to estimate the model parameters. The advantage of the proposed approach is that while it models the effect of multiple time-varying environmental covariates on the degradation process, it does not assume any rigid parametric assumption on the environmental covariate-drift relationship of the Brownian motion process. The penalized splines that combine a set of splines basis functions with a quadratic penalty on the corresponding coefficients, also, strike a balance between ovefitting and accuracy of covariate-drift relationship while not being limited to specifying the exact number and location of knot values. Moreover leveraging the mixed effect modeling framework and EM algorithm, it offers a natural way of choosing the smoothing parameters of penalty term. Combining the offline model estimation with degradation data and environmental covariates observations from the in-service unit, we proposed a Bayesian updating scheme to conduct individualized RUL prediction. We evaluated the performance of our approach using both simulated and real data of frying oil degradation collected from connected fryers. The numerical study and case study results confirmed that ignoring the extra information available through multiple time varying covariates affecting the degradation process hurts the RUL prediction accuracy.

This study considers a penalty based on the squared difference of the adjacent coefficients which results in smoother functional forms. One can also use the adaptive LASSO penalty or any other appropriate penalty form which provides a natural framework for variable selection for stress-drift relationship. Moreover, the proposed model in this study only considers the additive effect of multiple covariates. However, one can consider interactions among some of the covariates in the degradation modeling framework using a tensor based spline approach. Also, the degradation process developed in this study mainly concerns modeling the evolution of a one-dimensional degradation process. However, one can extend this framework by considering techniques like data fusions to develop a composite health index for degradation modeling and prognostics analysis based on multidimensional degradation process (26; 42; 37). We will investigate along this lines and report the results in a future study.

Appendices

The equation in (11) can be further expanded as follows:

$$Q^{(\nu)}(\boldsymbol{\Delta X}, \boldsymbol{\beta'}; \Psi) = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\frac{\left(R_i(t_{ij}) - \boldsymbol{G}^t(t_{ij})\boldsymbol{\mu}_{\beta'}^{(\nu)}\right)^2}{\delta t_{i,j}} - \frac{1}{2\sigma^2}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\frac{\boldsymbol{G}^t(t_{ij})\boldsymbol{\Sigma}_{\beta'}^{(\nu)}\boldsymbol{G}(t_{ij})}{\delta t_{i,j}}$$
$$+\frac{1}{2}\sum_{l=1}^{q}\left\{d_l\ln(\lambda_l) - \lambda_l\boldsymbol{\mu}_{\beta_l}^{(\nu)^t}\boldsymbol{S}_l\boldsymbol{\mu}_{\beta_l}^{(\nu)} - \lambda_l\,\mathrm{tr}(\boldsymbol{S}_l\boldsymbol{\Sigma}_{\beta_l}^{(\nu)})\right\} + \mathrm{const}, \tag{17}$$

where $\boldsymbol{\mu}_{\beta_l}^{(\nu)}$ and $\boldsymbol{\Sigma}_{\beta_l}^{(\nu)}$ correspond to coefficients posterior of l^{th} covariate and $\boldsymbol{\Sigma}_{\beta_l}^{(\nu)}$ is of dimension $d_l \times d_l$ and denotes the l^{th} diagonal block in $\boldsymbol{\Sigma}_{\beta'}^{(\nu)}$ estimated in iteration ν. Here $\boldsymbol{\mu}_{\beta'}^{(\nu)}$ and $\boldsymbol{\Sigma}_{\beta'}^{(\nu)}$ are computed using the current estimates of $\Psi^{(\nu)}$. The M-step entails maximizing the $Q^{(\nu)}(\boldsymbol{\Delta X}, \boldsymbol{\beta'}; \Psi)$ with respect to the parameters β_0, σ and λ. The derivatives of $Q^{(\nu)}(\boldsymbol{\Delta X}, \boldsymbol{\beta'}; \Psi)$ with respect to these parameters are as follows:

$$\begin{aligned}
\frac{dQ^{(\nu)}(\boldsymbol{\Delta X}, \boldsymbol{\beta'}; \Psi)}{d\beta_0} &= \frac{1}{\sigma^2}\sum_{i=1}^{N}\sum_{j=1}^{n_i}(\Delta X_i(t_{i,j}) - \beta_0\delta t_{i,j} - \boldsymbol{G}^t(t_{i,j})\boldsymbol{\mu}_{\beta'}^{(\nu)}),\\
\frac{dQ^{(\nu)}(\boldsymbol{\Delta X}, \boldsymbol{\beta'}; \Psi)}{d\sigma} &= -\sum_{i=1}^{N}\sum_{j=1}^{n_i}\frac{1}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\left(\frac{\left(R_i(t_{i,j}) - \boldsymbol{G}^t(t_{i,j})\boldsymbol{\mu}_{\beta'}^{(\nu)}\right)^2}{\delta t_{i,j}} + \frac{\boldsymbol{G}^t(t_{i,j})\boldsymbol{\Sigma}_{\beta_l}^{(\nu)}\boldsymbol{G}(t_{i,j})}{\delta t_{i,j}}\right),\\
\frac{dQ^{(\nu)}(\boldsymbol{\Delta X}, \boldsymbol{\beta'}; \Psi)}{d\lambda_l} &= \frac{d_l}{\lambda_l} - \boldsymbol{\mu}_{\beta_l}^{(\nu)^t}\boldsymbol{S}_l\boldsymbol{\mu}_{\beta_l}^{(\nu)} - \mathrm{tr}(\boldsymbol{S}_l\boldsymbol{\mu}_{\beta_l}^{(\nu)}).
\end{aligned} \tag{18}$$

Setting these to zero, we get the formulas to update the parameters β_0, σ and λ in equation (12).

References

[1] Abdel-Hameed, M. *Lévy processes and their applications in reliability and storage.* Springer, 2014.

[2] Al-Dahidi, S., Di Maio, F., Baraldi, P., and Zio, E. Remaining useful life estimation in heterogeneous fleets working under variable operating conditions. *Reliability Engineering & System Safety 156* (2016), 109–124.

[3] Banjevic, D., and Jardine, A. Calculation of reliability function and remaining useful life for a markov failure time process. *IMA journal of management mathematics 17*, 2 (2006), 115–130.

[4] Bian, L., and Gebraeel, N. Stochastic methodology for prognostics under continuously varying environmental profiles. *Statistical Analysis and Data Mining: The ASA Data Science Journal 6*, 3 (2013), 260–270.

[5] Brühl, L. Fatty acid alterations in oils and fats during heating and frying. *European Journal of Lipid Science and Technology 116*, 6 (2014), 707–715.

[6] Choe, E., and Min, D. Chemistry of deep-fat frying oils. *Journal of food science 72*, 5 (2007), R77–R86.

[7] Doksum, K., and Hoyland, A. Models for variable-stress accelerated life testing experiments based on wiener processes and the inverse gaussian distribution. *Theory of Probability & Its Applications 37*, 1 (1993), 137–139.

[8] Eilers, P. H., and Marx, B. D. Flexible smoothing with b-splines and penalties. *Statistical science* (1996), 89–102.

[9] Escobar, L. A., Meeker, W. Q., Kugler, D. L., and Kramer, L. L. Accelerated destructive degradation tests: data, models, and analysis. In *Mathematical and statistical methods in reliability*. World Scientific, 2003, pp. 319–337.

[10] Gebraeel, N., Elwany, A., and Pan, J. Residual life predictions in the absence of prior degradation knowledge. *IEEE Transactions on Reliability 58*, 1 (2009), 106–117.

[11] Gebraeel, N., and Pan, J. Prognostic degradation models for computing and updating residual life distributions in a time-varying environment. *IEEE Transactions on Reliability 57*, 4 (2008), 539–550.

[12] Gorjian, N., Ma, L., Mittinty, M., Yarlagadda, P., and Sun, Y. A review on reliability models with covariates. In *Engineering Asset Lifecycle Management*. Springer, 2010, pp. 385–397.

[13] Hong, Y., Duan, Y., Meeker, W. Q., Stanley, D. L., and Gu, X. Statistical methods for degradation data with dynamic covariates information and an application to outdoor weathering data. *Technometrics 57*, 2 (2015), 180–193.

[14] Jahani, S., Kontar, R., Zhou, S., and Veeramani, D. Remaining useful life prediction based on degradation signals using monotonic b-splines with infinite support. *IISE Transactions 52*, 5 (2020), 537–554.

[15] Jardine, A., Banjevic, D., and Makis, V. Optimal replacement policy and the structure of software for condition-based maintenance. *Journal of Quality in Maintenance Engineering 3*, 2 (1997), 109–119.

[16] Juárez, M. D., Osawa, C. C., Acuña, M. E., Sammán, N., and Gonçalves, L. A. G. Degradation in soybean oil, sunflower oil and partially hydrogenated fats after food frying, monitored by conventional and unconventional methods. *Food Control 22*, 12 (2011), 1920–1927.

[17] Kimeldorf, G. S., and Wahba, G. A correspondence between bayesian estimation on stochastic processes and smoothing by splines. *The Annals of Mathematical Statistics 41*, 2 (1970), 495–502.

[18] Kontar, R., Son, J., Zhou, S., Sankavaram, C., Zhang, Y., and Du, X. Remaining useful life prediction based on the mixed effects model with mixture prior distribution. *IISE Transactions 49*, 7 (2017), 682–697.

[19] Kontar, R., Zhou, S., Sankavaram, C., Du, X., and Zhang, Y. Nonparametric modeling and prognosis of condition monitoring signals using multivariate gaussian convolution processes. *Technometrics 60*, 4 (2018), 484–496.

[20] Lawless, J., and Crowder, M. Covariates and random effects in a gamma process model with application to degradation and failure. *Lifetime Data Analysis 10*, 3 (2004), 213–227.

[21] Le Son, K., Fouladirad, M., Barros, A., Levrat, E., and Iung, B. Remaining useful life estimation based on stochastic deterioration models: A comparative study. *Reliability Engineering & System Safety 112* (2013), 165–175.

[22] Li, N., Gebraeel, N., Lei, Y., Bian, L., and Si, X. Remaining useful life prediction of machinery under time-varying operating conditions based on a two-factor state-space model. *Reliability Engineering & System Safety 186* (2019), 88–100.

[23] Liao, C.-M., and Tseng, S.-T. Optimal design for step-stress accelerated degradation tests. *IEEE Transactions on Reliability 55*, 1 (2006), 59–66.

[24] Liao, G., Yin, H., Chen, M., and Lin, Z. Remaining useful life prediction for multi-phase deteriorating process based on wiener process. *Reliability Engineering & System Safety 207* (2021), 107361.

[25] Liao, H., and Tian, Z. A framework for predicting the remaining useful life of a single unit under time-varying operating conditions. *Iie Transactions 45*, 9 (2013), 964–980.

[26] Liu, K., Gebraeel, N. Z., and Shi, J. A data-level fusion model for developing composite health indices for degradation modeling and prognostic analysis. *IEEE Transactions on Automation Science and Engineering 10*, 3 (2013), 652–664.

[27] Meeker, W. Q., and Escobar, L. A. *Statistical methods for reliability data.* John Wiley & Sons, 2014.

[28] Meeker, W. Q., Escobar, L. A., and Lu, C. J. Accelerated degradation tests: modeling and analysis. *Technometrics 40*, 2 (1998), 89–99.

[29] Nelson, W. Prediction of field reliability of units, each under differing dynamic stresses, from accelerated test data. *Handbook of Statistics 20* (2001), 611–621.

[30] Nelson, W. B. *Accelerated testing: statistical models, test plans, and data analysis*, vol. 344. John Wiley & Sons, 2009.

[31] Pang, Z., Si, X., Hu, C., Du, D., and Pei, H. A bayesian inference for remaining useful life estimation by fusing accelerated degradation data and condition monitoring data. *Reliability Engineering & System Safety 208* (2021), 107341.

[32] Park, C., and Padgett, W. J. Stochastic degradation models with several accelerating variables. *IEEE Transactions on Reliability 55*, 2 (2006), 379–390.

[33] Riascos-Ochoa, J., Sánchez-Silva, M., and Klutke, G.-A. Modeling and reliability analysis of systems subject to multiple sources of degradation based on lévy processes. *Probabilistic Engineering Mechanics 45* (2016), 164–176.

[34] Rizopoulos, D. Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics 67*, 3 (2011), 819–829.

[35] Silverman, B. W. Some aspects of the spline smoothing approach to non-parametric regression curve fitting. *Journal of the Royal Statistical Society: Series B (Methodological) 47*, 1 (1985), 1–21.

[36] Son, J., Zhou, S., Sankavaram, C., Du, X., and Zhang, Y. Remaining useful life prediction based on noisy condition monitoring signals using constrained kalman filter. *Reliability Engineering & System Safety 152* (2016), 38–50.

[37] Song, C., and Liu, K. Statistical degradation modeling and prognostics of multiple sensor signals via data fusion: A composite health index approach. *IISE Transactions 50*, 10 (2018), 853–867.

[38] Wahba, G. Bayesian "confidence intervals" for the cross-validated smoothing spline. *Journal of the Royal Statistical Society: Series B (Methodological) 45*, 1 (1983), 133–150.

[39] Wen, Y., Wu, J., Das, D., and Tseng, T.-L. B. Degradation modeling and rul prediction using wiener process subject to multiple change points and unit heterogeneity. *Reliability Engineering & System Safety 176* (2018), 113–124.

[40] Whitmore, G., and Schenkelberg, F. Modelling accelerated degradation data using wiener diffusion with a time scale transformation. *Lifetime data analysis 3*, 1 (1997), 27–45.

[41] Xu, Z., Hong, Y., and Jin, R. Nonlinear general path models for degradation data with dynamic covariates. *Applied Stochastic Models in Business and Industry 32*, 2 (2016), 153–167.

[42] Yan, H., Liu, K., Zhang, X., and Shi, J. Multiple sensor data fusion for degradation modeling and prognostics under multiple operational conditions. *IEEE Transactions on Reliability 65*, 3 (2016), 1416–1426.

[43] Ye, Z.-S., and Chen, N. The inverse gaussian process as a degradation model. *Technometrics 56*, 3 (2014), 302–311.

[44] Zhao, X., Fouladirad, M., Bérenguer, C., and Bordes, L. Condition-based inspection/replacement policies for non-monotone deteriorating systems with environmental covariates. *Reliability Engineering & System Safety 95*, 8 (2010), 921–934.

[45] Zhou, Q., Son, J., Zhou, S., Mao, X., and Salman, M. Remaining useful life prediction of individual units subject to hard failure. *IIE Transactions 46*, 10 (2014), 1017–1030.