Reliability Engineering and System Safety 180 (2018) 94–103

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

A random-effects Wiener degradation model based on accelerated failure time

Qingqing Zhai[a], Piao Chen[b], Lanqing Hong[c], Lijuan Shen[*,c]

[a] School of Management, Shanghai University, Shanghai 200444, China
[b] Institute of High Performance Computing, 138632, Singapore
[c] Department of Industrial Systems Engineering & Management, National University of Singapore, 117576, Singapore

ARTICLE INFO

Keywords:
Unit-to-unit heterogeneity
Inverse Gaussian distribution
Maximum likelihood estimation
EM algorithm
Accelerated degradation test

ABSTRACT

Due to the variability of raw materials and the fluctuation in the manufacturing process, degradation of products may exhibit unit-to-unit variability in a population. The heterogeneous degradation rates can be viewed as random effects, which are often modeled by a normal distribution. Despite of its mathematical convenience, the normal distribution has certain limitations in modeling the random effects. In this study, we propose a novel random-effects Wiener process model based on ideas from accelerated failure time principle. An inverse Gaussian (IG) distribution can be used to characterize the unit-specific heterogeneity in degradation paths, which overcomes the disadvantages of the traditional models and provides more flexibility in the degradation modeling using Wiener processes. Properties of the model are investigated, and statistical inference based on the maximum likelihood estimation and the EM algorithm is established. An extension of the model to the constant-stress accelerated degradation test (ADT) is developed. The effectiveness and applicability of the proposed model are validated using a laser degradation dataset and an LED ADT dataset.

1. Introduction

The Wiener process is one of the most popular degradation models in recent decades [5,8,20,26,30]. Denote the degradation of a quality characteristic by $X(t)$, e.g., the capacity of a battery cell, the following Wiener process is often used to model its evolution with time:

$$X(t) = vt + \sigma\mathcal{B}(t). \quad (1)$$

In this basic Wiener process model, $\mathcal{B}(t)$ is the standard Brownian motion, the drift rate $v > 0$ corresponds to the mean rate of the degradation process, and the diffusion coefficient $\sigma > 0$ quantifies the magnitude of the process fluctuation.

The basic Wiener process model assumes all the units in a population have the same drift rate v and diffusion coefficient σ. However, this assumption has certain limitations in many practical applications. Particularly, units from a same batch, due to the variability of raw materials and the fluctuation in the manufacturing process, probably exhibit heterogeneities in their degradation paths [2,17]. When using the basic Wiener process to analyze such degradation data, the heterogeneities among individuals are neglected, which may lead to inaccurate statistical inference. To capture the heterogeneities among units, random-effects Wiener process models, which allow some parameters of the degradation process to be unit-specific and follow certain distributions across the population, have been proved to be useful [3,21,24,28,29]. In this context, most studies assume that the diffusion coefficient σ is invariant while the drift rate v is normally distributed for the units in a population [3,21,29]. Some studies also take the heterogeneity of the diffusion coefficient into account. Wang [24] allowed both drift and diffusion to be random, and assumed the diffusion σ^2 follows an inverse Gamma distribution and v follows a Gaussian distribution conditional on σ. Ye et al. [28] assumed that σ is a linear function of v and assigned a Gaussian distribution for the reciprocal of the drift rate.

The Gaussian distribution for the heterogeneous drift, although widely adopted, is restrictive in the following aspects. First, the Gaussian distribution has a support over the whole real line, which is often conflicted with the practical applications where the drift rate is believed to be one-sided, i.e., positive or negative. To address this incompatibility, many studies assume that the negative part of the Gaussian distribution is negligible, e.g., [28]. However, when the mean of v is not significantly larger than the standard deviation of v, the probability of a negative realization of v may not be neglected. Some studies truncated the Gaussian distribution [22], but it makes the model much more complicated. Second, the Gaussian distribution has a

* Corresponding author.
E-mail addresses: zhaiqing59@126.com (Q. Zhai), isesl@nus.edu.sg (L. Shen).

https://doi.org/10.1016/j.ress.2018.07.003
Received 4 January 2018; Received in revised form 27 June 2018; Accepted 4 July 2018
Available online 11 July 2018

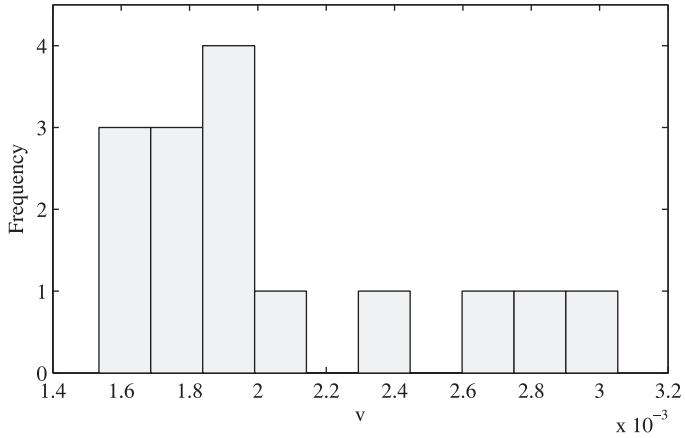


Fig. 1. Histogram of the drift rates of 15 laser devices when fitting each path individually by a basic Wiener process.

symmetric, bell-shaped probability density function (PDF), which is restrictive in modeling unit-to-unit heterogeneity. For example, we examine the degradation data of 15 laser devices from Meeker and Escobar [11], and fit each degradation path individually by a basic Wiener process in Eq. (1). As shown in Fig. 1, a skewed distribution appears to be more suitable to the unit-specific drift rates of the laser devices. Therefore, the performance of the Wiener process model with Gaussian drift may not be satisfactory due to these restrictions.

To overcome the deficiency, this paper proposes a novel random-effects Wiener process model exploiting the ideas in accelerated test modeling. In accelerated life/degradation tests (ALT/ADT), two or more groups of units are tested in harsh operation stresses to accelerate the failure/degradation process. Thus, sufficient reliability information can be observed which otherwise may be difficult to obtain under normal use conditions within practical test duration [4,15,23]. Since the reliability in normal stress is concerned, it is of vital importance to extrapolate the model in higher stress levels to the use stress level. The accelerated failure time (AFT) is a widely accepted concept, where the degradation path/lifetime under accelerated conditions is equivalent to the degradation path/lifetime that the same unit would have had under the normal use condition after properly scaling the time axis by the acceleration factor [12,13]. Therefore, the degradation paths/lifetimes under different stress levels are comparable after transformed to a same baseline stress level.

Inspired by this idea, the heterogeneity in the degradation rate can be viewed as a random acceleration phenomenon caused by unobserved random effects, which can be modeled by random scaling of the time axis. In this context, an inverse Gaussian (IG) distribution can be used to characterize the random effects, and a novel random-effects Wiener process model is developed. As a model for random effects, the IG assumption naturally accommodates the practical requirement that the drift rate should be positive. Moreover, it promises a wide applicability of the Wiener process model with IG random effects as the shape of the IG distribution is quite flexible [19]. Because the proposed model inherits the ideas in accelerated testing, it is also convenient to extend the proposed model to analyze the ADT data.

The remainder of the paper is organized as follows. Section 2 presents the detailed model formulation and the relevant properties. Section 3 develops an EM algorithm for the maximum likelihood (ML) estimation of the proposed model. In Section 4, we extend the proposed model to constant-stress ADT modeling. The application of the proposed model to real degradation data is illustrated in Section 5. Conclusions are given in Section 6.

2. Wiener process model with IG drift

2.1. Model formulation

In this study, we consider the following random-effects Wiener process model

$$X(t) = v\Lambda(t) + \kappa\mathcal{B}(v\Lambda(t)), \tag{2}$$

where $v > 0$ is the unit-specific drift rate, $\kappa > 0$ is a diffusion parameter, and $\mathcal{B}(\cdot)$ is the standard Brownian motion. $\Lambda(t) = \Lambda(t; \boldsymbol{\theta})$ is a determined function with parameter $\boldsymbol{\theta}$ to capture possible non-linear degradation patterns [18,26]. For example, $\Lambda(t)$ can be linear $\Lambda(t) = t$ or of the power law form $\Lambda(t) = t^{\theta}$. Generally, $\Lambda(t)$ is specified according to degradation physics or empirical observations [29]. To ease the notation, we will omit $\boldsymbol{\theta}$ whenever confusion is impossible. By convention, it is assumed that $\Lambda(0) = 0$ and $\Lambda(t)$ is monotonically increasing.

This model is proposed based on ideas from accelerated testing. In ALT, the AFT model is a widely applied model in quantifying the acceleration effects [7,13,14]. In its simplest form, the linear AFT model assumes the lifetime distribution under different stresses can be specified as

$$F_{\text{elevated stress}}(t|r) = F_0(rt),$$

where r denotes the joint acceleration effects and $F_0(\cdot)$ denotes the cumulative distribution function (CDF) of the lifetime at a baseline stress level. This relationship implies that the lifetime of a product at an elevated stress level is equivalent to the lifetime that the same unit would have had at a baseline stress by properly scaling the time axis. Thus, the lifetimes under different stress levels are unified and become comparable when transformed to the baseline stress. Inspired by this idea, the unit-specific random effects in the degradation paths can be modeled as random acceleration effects that result in random scaling of the time axis. Specifically, the time scale of each unit is assumed to be subject to a random scaling factor r, where r follows a continuous distribution with mean 1. Applying this formulation to the basic Wiener process in Eq. (1), the degradation process of the unit can be modeled as

$$X(t) = \mu rt + \sigma\mathcal{B}(rt), \tag{3}$$

where μ is the nominal drift rate for the units in a population. Since the random scaling factor r is hidden, we reformulate the model in (3) by resorting to the observable drift rate $v = \mu r$ and obtain the model in (2).

We note the following two features of the proposed model in (2). First, the observed degradation rate and diffusion coefficient are v and $\kappa^2 v$, respectively, which inherently interprets the heterogeneity in the degradation rate and process noise. Second, the random scaling factor introduces the dependence between the observed drift rate and diffusion coefficient, which accommodates the common phenomenon in practice that the unit with a larger degradation rate often possesses a larger variation [28].

To model the unit-specific drift rate v, a mathematically tractable choice is to adopt an IG distribution $\mathcal{IG}(\mu, \zeta)$:

$$f(v) = \sqrt{\frac{\zeta}{2\pi v^3}}\exp\left(-\frac{\zeta(v-\mu)^2}{2\mu^2 v}\right), \tag{4}$$

where $\mathbb{E}[v] = \mu$ and $\mathrm{Var}[v] = \mu^3/\zeta$. In addition to the mathematical convenience, using an IG distribution to model the unit-to-unit variability has considerable benefits. First, the IG distribution has a support on $(0, +\infty)$, which is compatible to the assumption that $v > 0$. On the contrary, the Gaussian distribution assumption in existing models is only an approximation that assumes the probability of $v < 0$ is negligible. Second, depending on the distribution parameters (μ, ζ), the IG distribution has a quite flexible shape and can accommodate a wide range of situations. Note that if we let $\zeta \to +\infty$, the proposed model degenerates to the commonly-used Wiener process model without random-effects.

Conditional on the drift rate v, the degradation of a single unit $X(t)$ is Gaussian distributed:

$$X(t)|v \sim \mathcal{N}(v\Lambda(t), v\kappa^2\Lambda(t)).$$

Thus, the unconditional distribution of $X(t)$ can be obtained by integrating v out

$$\begin{aligned} f_{X(t)}(x) &= \int_0^{+\infty} \frac{1}{\sqrt{2\pi\kappa^2 v\Lambda(t)}}\exp\left(-\frac{(x-v\Lambda(t))^2}{2\kappa^2 v\Lambda(t)}\right)\sqrt{\frac{\zeta}{2\pi v^3}}\exp\left(-\frac{\zeta(v-\mu)^2}{2\mu^2 v}\right)dv \\ &= \frac{\sqrt{\zeta}}{\pi\kappa\sqrt{\Lambda(t)}}\exp\left(\frac{x}{\kappa^2}+\frac{\zeta}{\mu}\right)\mathcal{K}_{-1}(\sqrt{a(t)b(t)})\sqrt{\frac{a(t)}{b(t)}}, \end{aligned}$$

where

$$a(t) = \frac{\zeta}{\mu^2} + \frac{\Lambda(t)}{\kappa^2},\ b(t) = \zeta + \frac{x^2}{\Lambda(t)\kappa^2},$$

and

$$\mathcal{K}_p(z) = \frac{1}{2}\int_0^{\infty} y^{p-1}\exp\left(-\frac{z}{2}(y+y^{-1})\right)dy$$

is the modified Bessel function of the second kind [1]. In addition, we have

$$\mathbb{E}[X(t)] = \mu\Lambda(t),\ \mathrm{Var}[X(t)] = \mu\kappa^2\Lambda(t) + \frac{\mu^3}{\zeta}\Lambda(t)^2.$$

2.2. Time to failure distribution

In practice, it is a routine task to make reliability inference of deteriorating products based on degradation modeling. Typically, the degradation $X(t)$ would correspond to a given failure threshold D_f, and the lifetime is defined as the first hitting time (FHT) of $X(t)$ to D_f: $T_f = \inf\{t: X(t) > D_f\}$. Conditional on v, it is known that the transformed FHT $\Lambda(T_f)$ follows an IG distribution $\mathcal{IG}(D_f/v, D_f^2/(v\kappa^2))$. Unconditional, the PDF of $\Lambda(T_f)$ can be derived as

$$f_{\Lambda(T_f)}(u) = \frac{D_f}{\pi u\kappa\mu}\sqrt{\frac{\zeta(\kappa^2\zeta+\mu^2 u)}{\left(\kappa^2\zeta u + D_f^2\right)}}\exp\left\{\frac{\zeta}{\mu}+\frac{D_f}{\kappa^2}\right\}\mathcal{K}_1\left(\sqrt{\left(\frac{\zeta}{\mu^2}+\frac{u}{\kappa^2}\right)\left(\zeta+\frac{D_f^2}{\kappa^2 u}\right)}\right). \tag{5}$$

Given that $\Lambda(t)$ is differentiable, the PDF of T_f under the calendar time t is

$$f_{T_f}(t) = f_{\Lambda(T_f)}(\Lambda(t))\frac{d\Lambda(t)}{dt}. \tag{6}$$

The reliability indices, e.g., the expected lifetime of the product can be derived based on $f_{T_f}(t)$. For instance, with the power law degradation trend $\Lambda(t) = t^\theta$, the mean and variance of T_f can be derived as

$$\begin{aligned} \mathbb{E}[T_f] &= \frac{2}{\pi}\sqrt{\frac{\zeta}{\kappa^2}}\left(\frac{D_f}{\mu}\right)^{\frac{1}{2}+\frac{1}{\theta}}\exp\left(\frac{D_f}{\kappa^2}+\frac{\zeta}{\mu}\right)\mathcal{K}_{\frac{1}{2}-\frac{1}{\theta}}\left(\frac{D_f}{\kappa^2}\right)\mathcal{K}_{\frac{1}{2}+\frac{1}{\theta}}\left(\frac{\zeta}{\mu}\right), \\ \mathrm{Var}[T_f] &= \frac{2}{\pi}\sqrt{\frac{\zeta}{\kappa^2}}\left(\frac{D_f}{\mu}\right)^{\frac{1}{2}+\frac{2}{\theta}}\exp\left(\frac{D_f}{\kappa^2}+\frac{\zeta}{\mu}\right)\mathcal{K}_{\frac{1}{2}-\frac{2}{\theta}}\left(\frac{D_f}{\kappa^2}\right)\mathcal{K}_{\frac{1}{2}+\frac{2}{\theta}}\left(\frac{\zeta}{\mu}\right) - \mathbb{E}[T_f]^2. \end{aligned}$$

Particularly, for the linear degradation case $\theta = 1$ and $\Lambda(t) = t$, we have

$$\begin{aligned} \mathbb{E}[T_f] &= D_f\left(\frac{1}{\mu}+\frac{1}{\zeta}\right), \\ \mathrm{Var}[T_f] &= (D_f\kappa^2 + D_f^2)\left(\frac{1}{\mu\zeta}+\frac{2}{\zeta^2}\right) + D_f\kappa^2\left(\frac{1}{\mu}+\frac{1}{\zeta}\right)^2. \end{aligned}$$

2.3. Inference from observations

Suppose that the degradation of a unit is observed at $\{t_1, ..., t_m\}$ with the observations $\boldsymbol{X} = (X_1, ..., X_m)^{\mathrm{T}}$, where the superscript "T" denotes matrix transposition. Define $X_0 = 0$ and $t_0 = 0$. Then, the increments $\Delta X_j = X_j - X_{j-1}$ are independent and Gaussian conditional on the drift v:

$$(\Delta X_j|v) \sim \mathcal{N}(v\Delta\Lambda_j, v\kappa^2\Delta\Lambda_j),$$

where $\Delta\Lambda_j = \Lambda(t_j) - \Lambda(t_{j-1})$. The joint distribution of v and $\boldsymbol{X}$ is readily obtained as

$$p(v, \boldsymbol{X}) = \sqrt{\frac{\zeta}{2\pi v^3}}\exp\left\{-\frac{\zeta(v-\mu)^2}{2\mu^2 v}\right\}\cdot\prod_{j=1}^{m}\frac{1}{\sqrt{2\pi v\kappa^2\Delta\Lambda_j}}\exp\left\{-\frac{(\Delta X_j - v\Delta\Lambda_j)^2}{2v\kappa^2\Delta\Lambda_j}\right\}.$$

By integrating v out, we can obtain the joint distribution of $\boldsymbol{X}$ as

$$p(\boldsymbol{X}) = \sqrt{\frac{\zeta}{(2\pi)^{m+1}\kappa^{2m}}}\prod_{j=1}^{m}\frac{1}{\sqrt{\Delta\Lambda_j}}\exp\left(\frac{X_m}{\kappa^2}+\frac{\zeta}{\mu}\right)2\mathcal{K}_p(\sqrt{ab})(b/a)^{p/2},$$

where

$$a = \frac{\zeta}{\mu^2} + \frac{\Lambda_m}{\kappa^2},\ b = \zeta + \frac{1}{\kappa^2}\sum_{j=1}^{m}\frac{\Delta X_j^2}{\Delta\Lambda_j},\ p = -\frac{m+1}{2}.$$

In many cases, it is of interest to estimate v based on the degradation observations. For instance, in the prognostics and health management of complex systems, the in-service measured degradation signals can be used to update the conditional distribution of v to monitor the system health status. Based on the above analysis, the conditional distribution of v conditional on the degradation observation $\boldsymbol{X}$ is

$$p(v|\boldsymbol{X}) = \frac{(a/b)^{p/2}}{2\mathcal{K}_p(\sqrt{ab})}v^{p-1}\exp\left(-\frac{1}{2}(av+bv^{-1})\right). \tag{7}$$

This indicates that v, conditional on the observations, follows a generalized IG distribution $\mathcal{GIG}(a, b, p)$ [6].

Suppose we have the degradation observations $\mathbf{X} = \{\boldsymbol{X}_1, ..., \boldsymbol{X}_n\}$ of n units, where $\boldsymbol{X}_i = (X_{i,1}, ..., X_{i,m_i})^{\mathrm{T}}$ are obtained at time points $t_{i,1}, ..., t_{i,m_i}$. Then, the unknown model parameters $\boldsymbol{\Theta} = (\mu, \zeta, \kappa)$ and the parameters $\boldsymbol{\theta}$ in $\Lambda(\cdot)$ can be estimated by maximizing the log-likelihood function

$$\ell(\boldsymbol{\Theta}, \boldsymbol{\theta}|\mathbf{X}) = \sum_{i=1}^{n}\ln p(\boldsymbol{X}_i|\boldsymbol{\Theta}, \boldsymbol{\theta}). \tag{8}$$

Because the model parameters are involved in the modified Bessel function $\mathcal{K}$, it brings difficulties in straightforward maximization of the log-likelihood function. On the other hand, the unit-specific degradation rate can be seen as missing data, where an EM algorithm can be used to find the ML estimates. The EM algorithm is a powerful method for ML estimation, and has been widely used in degradation problems. In the following, we detail the EM algorithm for parameter inference of the proposed model.

3. EM algorithm for ML estimation

3.1. Point estimation

We first assume the parameters $\boldsymbol{\theta}$ in $\Lambda(\cdot)$ are determined and focus on the estimation of $\boldsymbol{\Theta}$. For the proposed model, we consider the unobservable drift parameter v_i for each unit as missing data. Denote $\mathbf{V} = (v_1, ..., v_n)^{\mathrm{T}}$. Then, we can obtain the following log-likelihood function for the complete data $\{\mathbf{V}, \mathbf{X}\}$

$$\ell_c(\boldsymbol{\Theta}|\mathbf{V}, \mathbf{X}) = \ell_v + \ell_X, \tag{9}$$

where

$$\ell_v = -\frac{n}{2}\ln(2\pi) + \frac{n}{2}\ln\zeta - \frac{3}{2}\sum_{i=1}^{n}\ln v_i - \sum_{i=1}^{n}\frac{\zeta(v_i-\mu)^2}{2\mu^2 v_i},$$

and

$$\ell_X = \sum_{i=1}^{n}\left\{-\frac{m_i}{2}\ln(2\pi) - \frac{m_i}{2}\ln\kappa^2 v_i - \frac{1}{2}\sum_{j=1}^{m_i}\ln\Delta\Lambda_{i,j} - \frac{1}{2\kappa^2 v_i}\sum_{j=1}^{m_i}\frac{(\Delta X_{i,j} - v_i\Delta\Lambda_{i,j})^2}{\Delta\Lambda_{i,j}}\right\}.$$

With the complete-data log-likelihood, the EM algorithm iteratively implementing the E-step and the M-step can be carried out. Suppose we have the estimates $\boldsymbol{\Theta}^{(s)}$ for the model parameters $\boldsymbol{\Theta}$ in the sth iteration, the following Q-function needs to be calculated:

$$Q(\boldsymbol{\Theta}|\boldsymbol{\Theta}^{(s)}) = \mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta}^{(s)}}[\ell_c(\boldsymbol{\Theta}|\mathbf{V}, \mathbf{X})], \tag{10}$$

where the expectation is taken with respect to the conditional distribution of $(\mathbf{V}|\mathbf{X}, \boldsymbol{\Theta}^{(s)})$. Note that the degradation observations and the drift rate for different units are statistically independent from each other. Therefore,

$$p(\mathbf{V}|\mathbf{X}) = \prod_{i=1}^{n} p(v_i|\boldsymbol{X}_i),$$

where $p(v_i|\boldsymbol{X}_i)$ follows a generalized IG distribution with parameters (a_i, b_i, p_i), as given in (7). According to the properties of the generalized IG distribution, it can be readily obtained that

$$\begin{aligned}
\mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta}^{(s)}}[v_i] &= \frac{\sqrt{b_i}\mathcal{K}_{p_i+1}(\sqrt{a_i b_i})}{\sqrt{a_i}\mathcal{K}_{p_i}(\sqrt{a_i b_i})},\\
\mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta}^{(s)}}[v_i^{-1}] &= \frac{\sqrt{a_i}\mathcal{K}_{p_i+1}(\sqrt{a_i b_i})}{\sqrt{b_i}\mathcal{K}_{p_i}(\sqrt{a_i b_i})} - \frac{2p_i}{b_i}.
\end{aligned}$$

Once the Q-function is obtained, the estimates for the model parameters are updated by the maximization of the Q-function (M-step). Deriving the first order partial derivatives of the Q-function with respect to $\boldsymbol{\Theta}$ and setting them to zero, we obtain the following results:

$$\begin{aligned}
\hat{\mu}^{(s+1)} &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta}^{(s)}}[v_i],\\
\hat{\zeta}^{(s+1)} &= \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta}^{(s)}}[v_i^{-1}] - 1/\hat{\mu}^{(s+1)}},\\
\hat{\omega}^{(s+1)} &= \frac{1}{\sum_{i=1}^{n} m_i}\sum_{i=1}^{n}\left\{\Lambda_{i,m_i}\mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta}^{(s)}}[v_i] - 2X_{i,m_i}\right.\\
&\quad \left. + \left[\sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta\Lambda_{i,j}}\right]\mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta}^{(s)}}[v_i^{-1}]\right\},
\end{aligned}$$

where $\omega \triangleq \kappa^2$ for notation simplicity. The iteration is implemented until the difference between the estimates in two consecutive iterations is smaller than a given tolerance. Then, the ML estimates for $\boldsymbol{\Theta}$ can be obtained.

In the case that $\boldsymbol{\theta}$ in function $\Lambda(\cdot)$ should be estimated from the degradation data, a profile likelihood method can be utilized. Specifically, for any given $\boldsymbol{\theta}$, we can first estimate the parameters $\boldsymbol{\Theta}$ using the proposed EM algorithm. Here, the estimates $\widehat{\boldsymbol{\Theta}}$ are functions of $\boldsymbol{\theta}$, $\widehat{\boldsymbol{\Theta}} = \widehat{\boldsymbol{\Theta}}(\theta)$. Substituting them into Eq. (8), we can obtain the log-likelihood value as a function of $\boldsymbol{\theta}$: $\ell(\theta) = \ell(\widehat{\boldsymbol{\Theta}}(\theta), \theta)$. Then, the estimate for $\boldsymbol{\theta}$ can be obtained by the maximization of the profile log-likelihood.

3.2. Interval estimation

Suppose we get the ML estimates $\{\widehat{\boldsymbol{\Theta}}, \hat{\theta}\}$ with the EM algorithm. The confidence intervals for the model parameters can be obtained through normal asymptotics. Specifically, the observed information matrix can be obtained as a byproduct of the EM algorithm. Subsequently, the asymptotic covariance at the estimates can be obtained as the inverse of the observed information matrix and the relevant interval estimates can be derived. According to Louis [10], the observed information matrix at $\{\widehat{\boldsymbol{\Theta}}, \hat{\theta}\}$ can be expressed as

$$\begin{aligned}
\mathcal{I}(\widehat{\boldsymbol{\Theta}}, \hat{\theta}) &= -\begin{pmatrix}
\frac{\partial^2}{\partial\mu^2} & \frac{\partial^2}{\partial\mu\partial\zeta} & \frac{\partial^2}{\partial\mu\partial\omega} & \frac{\partial^2}{\partial\mu\partial\theta^{\mathrm{T}}}\\
\frac{\partial^2}{\partial\mu\partial\zeta} & \frac{\partial^2}{\partial\zeta^2} & \frac{\partial^2}{\partial\zeta\partial\omega} & \frac{\partial^2}{\partial\zeta\partial\theta^{\mathrm{T}}}\\
\frac{\partial^2}{\partial\mu\partial\omega} & \frac{\partial^2}{\partial\zeta\partial\omega} & \frac{\partial^2}{\partial\omega^2} & \frac{\partial^2}{\partial\omega\partial\theta^{\mathrm{T}}}\\
\frac{\partial^2}{\partial\mu\partial\theta} & \frac{\partial^2}{\partial\zeta\partial\theta} & \frac{\partial^2}{\partial\omega\partial\theta} & \frac{\partial^2}{\partial\theta\partial\theta^{\mathrm{T}}}
\end{pmatrix}\ell(\boldsymbol{\Theta}, \theta)|_{\boldsymbol{\Theta}=\widehat{\boldsymbol{\Theta}},\theta=\hat{\theta}}\\
&= \{-\mathbb{E}_{\mathbf{V}|\mathbf{X},\widehat{\boldsymbol{\Theta}},\hat{\theta}}[\nabla\nabla^{\mathrm{T}}\ell_c(\boldsymbol{\Theta}, \theta|\mathbf{V}, \mathbf{X})] - \mathbb{E}_{\mathbf{V}|\mathbf{X},\widehat{\boldsymbol{\Theta}},\hat{\theta}}[\nabla\ell_c(\boldsymbol{\Theta}, \theta|\mathbf{V}, \mathbf{X})\nabla^{\mathrm{T}}\ell_c(\boldsymbol{\Theta}, \theta|\mathbf{V}, \mathbf{X})]\\
&\quad + \mathbb{E}_{\mathbf{V}|\mathbf{X},\widehat{\boldsymbol{\Theta}},\hat{\theta}}[\nabla\ell_c(\boldsymbol{\Theta}, \theta|\mathbf{V}, \mathbf{X})]\mathbb{E}_{\mathbf{V}|\mathbf{X},\widehat{\boldsymbol{\Theta}},\hat{\theta}}[\nabla^{\mathrm{T}}\ell_c(\boldsymbol{\Theta}, \theta|\mathbf{V}, \mathbf{X})]\}_{\boldsymbol{\Theta}=\widehat{\boldsymbol{\Theta}},\theta=\hat{\theta}},
\end{aligned} \tag{11}$$

where ∇ is the vector differential operator.

Based on Eq. (11) and the complete log-likelihood in Eq. (9), we can derive the following results after some algebra:

$$\begin{aligned}
\frac{\partial^2\ell}{\partial\mu^2} &= \frac{\zeta}{\mu^4}\sum_{i=1}^{n}(-3\mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta},\theta}[v_i] + 2\mu) + \frac{\zeta^2}{\mu^6}\sum_{i=1}^{n}\mathrm{Var}[v_i|\mathbf{X}, \boldsymbol{\Theta}, \theta],\\
\frac{\partial^2\ell}{\partial\mu\partial\zeta} &= \frac{1}{\mu^3}\sum_{i=1}^{n}(\mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta},\theta}[v_i] - \mu) - \frac{\zeta}{2\mu^3}\sum_{i=1}^{n}\left(\frac{1}{\mu^2}\mathrm{Var}[v_i|\mathbf{X}, \boldsymbol{\Theta}, \theta] + \mathrm{Cov}[v_i, v_i^{-1}|\mathbf{X}, \boldsymbol{\Theta}, \theta]\right)\\
\frac{\partial^2\ell}{\partial\mu\partial\omega} &= \frac{\zeta}{2\mu^3\omega^2}\sum_{i=1}^{n}\left\{\Lambda_{i,m_i}\mathrm{Var}[v_i|\mathbf{X}, \boldsymbol{\Theta}, \theta] + \sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta\Lambda_{i,j}}\mathrm{Cov}[v_i, v_i^{-1}|\mathbf{X}, \boldsymbol{\Theta}, \theta]\right\},\\
\frac{\partial^2\ell}{\partial\zeta^2} &= -\frac{n}{2\zeta^2} + \frac{1}{4\mu^4}\sum_{i=1}^{n}(\mathrm{Var}[v_i|\mathbf{X}, \boldsymbol{\Theta}, \theta] + 2\mu^2\mathrm{Cov}[v_i, v_i^{-1}|\mathbf{X}, \boldsymbol{\Theta}, \theta] + \mu^4\mathrm{Var}[v_i^{-1}|\mathbf{X}, \boldsymbol{\Theta}, \theta]),\\
\frac{\partial^2\ell}{\partial\zeta\partial\omega} &= -\frac{1}{4\mu^2\omega^2}\sum_{i=1}^{n}\left(\Lambda_{i,m_i}\mathrm{Var}[v_i|\mathbf{X}, \boldsymbol{\Theta}, \theta] + \mu^2\sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta\Lambda_{i,j}}\mathrm{Var}[v_i^{-1}|\mathbf{X}, \boldsymbol{\Theta}, \theta]\right.\\
&\quad \left. + \left(\mu^2\Lambda_{i,m_i} + \sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta\Lambda_{i,j}}\right)\mathrm{Cov}[v_i, v_i^{-1}|\mathbf{X}, \boldsymbol{\Theta}, \theta]\right),\\
\frac{\partial^2\ell}{\partial\omega^2} &= \frac{1}{2\omega^2}\sum_{i=1}^{n}m_i - \frac{1}{\omega^3}\sum_{i=1}^{n}\left\{\Lambda_{i,m_i}\mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta},\theta}[v_i] - 2X_{i,m_i} + \sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta\Lambda_{i,j}}\mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta},\theta}[v_i^{-1}]\right\}\\
&\quad + \frac{1}{4\omega^4}\sum_{i=1}^{n}\left\{\Lambda_{i,m_i}^2\mathrm{Var}[v_i|\mathbf{X}, \boldsymbol{\Theta}, \theta] + \left[\sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta\Lambda_{i,j}}\right]^2\mathrm{Var}[v_i^{-1}|\mathbf{X}, \boldsymbol{\Theta}, \theta]\right.\\
&\quad \left. + 2\Lambda_{i,m_i}\sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta\Lambda_{i,j}}\mathrm{Cov}[v_i, v_i^{-1}|\mathbf{X}, \boldsymbol{\Theta}, \theta]\right\},
\end{aligned}$$

where

$$\begin{aligned}
\mathrm{Var}[v_i|\mathbf{X}, \boldsymbol{\Theta}, \theta] &= \mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta},\theta}[v_i^2] - \mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta},\theta}[v_i]^2,\\
\mathrm{Var}[v_i^{-1}|\mathbf{X}, \boldsymbol{\Theta}, \theta] &= \mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta},\theta}[v_i^{-2}] - \mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta},\theta}[v_i^{-1}]^2,\\
\mathrm{Cov}[v_i, v_i^{-1}|\mathbf{X}, \boldsymbol{\Theta}, \theta] &= 1 - \mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta},\theta}[v_i]\mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta},\theta}[v_i^{-1}].
\end{aligned}$$

According to the properties of the generalized IG distribution, we have

$$\begin{aligned}
\mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta},\theta}[v_i^2] &= \frac{b_i\mathcal{K}_{p_i+2}(\sqrt{a_i b_i})}{a_i\mathcal{K}_{p_i}(\sqrt{a_i b_i})},\\
\mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta},\theta}[v_i^{-2}] &= \frac{a_i\mathcal{K}_{p_i-2}(\sqrt{a_i b_i})}{b_i\mathcal{K}_{p_i}(\sqrt{a_i b_i})}.
\end{aligned}$$

In addition, the entries involving the derivatives of $\boldsymbol{\theta}$ are

$$\frac{\partial^2 \ell}{\partial \mu \partial \theta} = -\frac{\zeta}{2\mu^3\omega}\sum_{i=1}^{n}\left(\frac{\partial \Lambda_{i,m_i}}{\partial \theta}\mathrm{Var}[v_i|\mathbf{X},\boldsymbol{\Theta},\boldsymbol{\theta}] - \sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta\Lambda_{i,j}^2}\frac{\partial \Delta\Lambda_{i,j}}{\partial \theta}\mathrm{Cov}[v_i, v_i^{-1}|\mathbf{X},\boldsymbol{\Theta},\boldsymbol{\theta}]\right),$$

$$\frac{\partial^2 \ell}{\partial \zeta \partial \theta} = \frac{1}{4\mu^2\omega}\sum_{i=1}^{n}\left(\frac{\partial \Lambda_{i,m_i}}{\partial \theta}\mathrm{Var}[v_i|\mathbf{X},\boldsymbol{\Theta},\boldsymbol{\theta}] - \mu^2\sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta\Lambda_{i,j}^2}\frac{\partial \Delta\Lambda_{i,j}}{\partial \theta}\mathrm{Var}[v_i^{-1}|\mathbf{X},\boldsymbol{\Theta},\boldsymbol{\theta}] + \left(\mu^2\frac{\partial \Lambda_{i,m_i}}{\partial \theta} - \sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta\Lambda_{i,j}^2}\frac{\partial \Delta\Lambda_{i,j}}{\partial \theta}\right)\mathrm{Cov}[v_i, v_i^{-1}|\mathbf{X},\boldsymbol{\Theta},\boldsymbol{\theta}]\right),$$

$$\frac{\partial^2 \ell}{\partial \omega \partial \theta} = \frac{1}{2\omega^2}\sum_{i=1}^{n}\left(\frac{\partial \Lambda_{i,m_i}}{\partial \theta}\mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta},\boldsymbol{\theta}}[v_i] - \sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta\Lambda_{i,j}^2}\frac{\partial \Delta\Lambda_{i,j}}{\partial \theta}\mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta},\boldsymbol{\theta}}[v_i^{-1}]\right) - \frac{1}{4\omega^3}\sum_{i=1}^{n}\left(\Lambda_{i,m_i}\frac{\partial \Lambda_{i,m_i}}{\partial \theta}\mathrm{Var}[v_i|\mathbf{X},\boldsymbol{\Theta},\boldsymbol{\theta}] - \sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta\Lambda_{i,j}}\sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta\Lambda_{i,j}^2}\frac{\partial \Delta\Lambda_{i,j}}{\partial \theta}\mathrm{Var}[v_i^{-1}|\mathbf{X},\boldsymbol{\Theta},\boldsymbol{\theta}] + \left(\frac{\partial \Lambda_{i,m_i}}{\partial \theta}\sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta\Lambda_{i,j}} - \Lambda_{i,m_i}\sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta\Lambda_{i,j}^2}\frac{\partial \Delta\Lambda_{i,j}}{\partial \theta}\right)\mathrm{Cov}[v_i, v_i^{-1}|\mathbf{X},\boldsymbol{\Theta},\boldsymbol{\theta}]\right),$$

$$\frac{\partial^2 \ell}{\partial \theta \partial \theta^{\mathrm{T}}} = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\left(\frac{1}{\Delta\Lambda_{i,j}}\frac{\partial^2 \Delta\Lambda_{i,j}}{\partial \theta \partial \theta^{\mathrm{T}}} - \frac{1}{\Delta\Lambda_{i,j}^2}\frac{\partial \Delta\Lambda_{i,j}}{\partial \theta}\frac{\partial \Delta\Lambda_{i,j}}{\partial \theta^{\mathrm{T}}}\right) + \frac{1}{2\omega}\sum_{i=1}^{n}\mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta},\boldsymbol{\theta}}[v_i^{-1}]\sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta\Lambda_{i,j}^3}\left(\Delta\Lambda_{i,j}\frac{\partial^2 \Delta\Lambda_{i,j}}{\partial \theta \partial \theta^{\mathrm{T}}} - 2\frac{\partial \Delta\Lambda_{i,j}}{\partial \theta}\frac{\partial \Delta\Lambda_{i,j}}{\partial \theta^{\mathrm{T}}}\right) - \frac{1}{2\omega}\sum_{i=1}^{n}\mathbb{E}_{\mathbf{V}|\mathbf{X},\boldsymbol{\Theta},\boldsymbol{\theta}}[v_i]\frac{\partial^2 \Lambda_{i,m_i}}{\partial \theta \partial \theta^{\mathrm{T}}} + \frac{1}{4\omega^2}\sum_{i=1}^{n}\left\{\frac{\partial \Lambda_{i,m_i}}{\partial \theta}\frac{\partial \Lambda_{i,m_i}}{\partial \theta^{\mathrm{T}}}\mathrm{Var}[v_i|\mathbf{X},\boldsymbol{\Theta},\boldsymbol{\theta}] - 2\frac{\partial \Lambda_{i,m_i}}{\partial \theta}\sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta\Lambda_{i,j}^2}\frac{\partial \Delta\Lambda_{i,j}}{\partial \theta^{\mathrm{T}}}\mathrm{Cov}[v_i, v_i^{-1}|\mathbf{X},\boldsymbol{\Theta},\boldsymbol{\theta}] + \sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta\Lambda_{i,j}^2}\frac{\partial \Delta\Lambda_{i,j}}{\partial \theta}\sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta\Lambda_{i,j}^2}\frac{\partial \Delta\Lambda_{i,j}}{\partial \theta^{\mathrm{T}}}\mathrm{Var}[v_i^{-1}|\mathbf{X},\boldsymbol{\Theta},\boldsymbol{\theta}]\right\},$$

where

$$\frac{\partial \Delta\Lambda_{i,j}}{\partial \theta} = \begin{cases}\dfrac{\partial \Lambda(t_{i,1};\boldsymbol{\theta})}{\partial \theta}, & j=1\\ \dfrac{\partial \Lambda(t_{i,j};\boldsymbol{\theta})}{\partial \theta} - \dfrac{\partial \Lambda(t_{i,j-1};\boldsymbol{\theta})}{\partial \theta}, & j>1\end{cases}$$

$$\frac{\partial^2 \Delta\Lambda_{i,j}}{\partial \theta \partial \theta^{\mathrm{T}}} = \begin{cases}\dfrac{\partial^2 \Lambda(t_{i,1};\boldsymbol{\theta})}{\partial \theta \partial \theta^{\mathrm{T}}}, & j=1\\ \dfrac{\partial^2 \Lambda(t_{i,j};\boldsymbol{\theta})}{\partial \theta \partial \theta^{\mathrm{T}}} - \dfrac{\partial^2 \Lambda(t_{i,j-1};\boldsymbol{\theta})}{\partial \theta \partial \theta^{\mathrm{T}}}, & j>1\end{cases}$$

$$\frac{\partial \Lambda_{i,m_i}}{\partial \theta} = \frac{\partial \Lambda(t_{i,m_i};\boldsymbol{\theta})}{\partial \theta},\ \frac{\partial^2 \Lambda_{i,m_i}}{\partial \theta \partial \theta^{\mathrm{T}}} = \frac{\partial^2 \Lambda(t_{i,m_i};\boldsymbol{\theta})}{\partial \theta \partial \theta^{\mathrm{T}}}.$$

Clearly, when the parameters $\boldsymbol{\theta}$ in $\Lambda(\cdot)$ are given, we simply have

$$\mathcal{I}(\widehat{\boldsymbol{\Theta}}) = -\begin{pmatrix}\dfrac{\partial^2}{\partial \mu^2} & \dfrac{\partial^2}{\partial \mu \partial \zeta} & \dfrac{\partial^2}{\partial \mu \partial \omega}\\ \dfrac{\partial^2}{\partial \mu \partial \zeta} & \dfrac{\partial^2}{\partial \zeta^2} & \dfrac{\partial^2}{\partial \zeta \partial \omega}\\ \dfrac{\partial^2}{\partial \mu \partial \omega} & \dfrac{\partial^2}{\partial \zeta \partial \omega} & \dfrac{\partial^2}{\partial \omega^2}\end{pmatrix}\ell(\boldsymbol{\Theta})\Bigg|_{\boldsymbol{\Theta}=\widehat{\boldsymbol{\Theta}}}.$$

Thus, with the ML estimates at the end of the EM algorithm, the observed information matrix and the asymptotic covariance can be readily obtained. Based on the asymptotic covariance for the ML estimators, the interval estimates for other quantities, such as the degradation at time t and the p quantile of the failure time can be further obtained by the delta method.

4. Extension to ADT data

Consider a constant-stress ADT under K stress levels $\xi_1, ..., \xi_K$. Suppose n_k units are allocated in stress ξ_k for degradation testing. The degradation in each stress level can be modeled by the proposed Wiener process model with IG drift. From the AFT point of view, we assume the degradation of an unit in the kth stress level can be modeled as

$$X_{k,i}(t) = \mu\cdot r_{k,i}h(\xi_k)\Lambda(t) + \sigma\mathcal{B}(r_{k,i}h(\xi_k)\Lambda(t)).$$

Here, $h(\xi_k)$ is a determined acceleration factor that depends on the acceleration variables. In such a formulation, we essentially assume that the virtual age of a unit under elevated stresses is equivalent to that under the normal use stress by properly scaling the time axis by a factor $h(\xi_k)$. Reformulating the model in terms of the observable degradation rate, we consider the following model for ADT

$$X_{k,i}(t) = v_{k,i}\Lambda(t) + \kappa\mathcal{B}(v_{k,i}\Lambda(t)),$$

where $v_{k,i} = \mu r_{k,i}h(\xi_k) \sim \mathcal{IG}(\mu h(\xi_k), \zeta h(\xi_k))$ denotes the degradation rate of unit i under the kth stress level.

We assume the stress ξ_k has been standardized with $\xi_0 = 0$, and $h(\xi_k)$ has an exponential form $h(\xi_k) = \exp(\beta\xi_k)$, where β is a proper parameter to be determined [9]. Therefore, the model parameters in the ADT setting are $\boldsymbol{\Theta} = \{\mu, \zeta, \kappa, \beta\}$ and the possible unknown parameters $\boldsymbol{\theta}$ in $\Lambda(\cdot)$.

Suppose we have the degradation observations $\boldsymbol{X}_{k,1}, ..., \boldsymbol{X}_{k,n_k}$ for the n_k units in the kth level, where $\boldsymbol{X}_{k,i} = (X_{k,i,1}, ..., X_{k,i,m_{k,i}})^{\mathrm{T}}$ are measured at $t_{k,i,1}, ..., t_{k,i,m_{k,i}}$. The model parameters can be estimated by a similar EM algorithm as that in the single-stress case. In the following, we briefly present the extension of the proposed EM algorithm to the ADT data.

4.1. Point estimation using EM algorithm

As in the single-stress level case, the unobserved degradation drift $v_{k,i}$ for the units under different stress levels are viewed as the missing data. Accordingly, the complete-data log-likelihood can be obtained as

$$\begin{aligned}\ell_c &= \mathrm{const} + \frac{1}{2}\sum_{k=1}^{K}\left\{n_k\ln(\zeta h(\xi_k)) - \sum_{i=1}^{n_k}\frac{\zeta h(\xi_k)(v_{k,i}-\mu h(\xi_k))^2}{(\mu h(\xi_k))^2 v_{k,i}} - \sum_{i=1}^{n_k}m_{k,i}\ln\omega - \sum_{i=1}^{n_k}\sum_{j=1}^{m_{k,i}}\ln\Delta\Lambda_{k,i,j} - \frac{1}{\omega}\sum_{i=1}^{n_k}\sum_{j=1}^{m_{k,i}}\frac{(\Delta X_{k,i,j}-v_{k,i}\Delta\Lambda_{k,i,j})^2}{v_{k,i}\Delta\Lambda_{k,i,j}}\right\},\\ &= \mathrm{const} + \frac{1}{2}\sum_{k=1}^{K}\left\{n_k\ln\zeta + \beta n_k\xi_k - \sum_{i=1}^{n_k}\frac{\zeta(v_{k,i}-\mu\exp(\beta\xi_k))^2}{\mu^2\exp(\beta\xi_k)v_{k,i}} - \sum_{i=1}^{n_k}m_{k,i}\ln\omega - \sum_{i=1}^{n_k}\sum_{j=1}^{m_{k,i}}\ln\Delta\Lambda_{k,i,j} - \frac{1}{\omega}\sum_{i=1}^{n_k}\sum_{j=1}^{m_{k,i}}\frac{(\Delta X_{k,i,j}-v_{k,i}\Delta\Lambda_{k,i,j})^2}{v_{k,i}\Delta\Lambda_{k,i,j}}\right\}.\end{aligned}$$

where $\mathrm{const} = -\frac{1}{2}\sum_{k=1}^{K}\sum_{i=1}^{n_k}(1+m_{k,i})\ln(2\pi)$ and $\omega \triangleq \kappa^2$.

The conditional distribution of $v_{k,i}$ for $i = 1, ..., n_k$, $k = 1, ..., K$ is only dependent on the data of the ith unit in the kth stress level, i.e.,

$$p(v_{1,1},...,v_{1,n_1},...v_{K,1},...,v_{K,n_K}|\boldsymbol{X}_{1,1},...,\boldsymbol{X}_{1,n_1},...,\boldsymbol{X}_{K,1},...,\boldsymbol{X}_{K,n_K}) = \prod_{k=1}^{K}\prod_{i=1}^{n_k}p(v_{k,i}|\boldsymbol{X}_{k,i}).$$

Therefore, the conditional expectation of $v_{k,i}$ and $v_{k,i}^{-1}$, which are required in the E-step, can be obtained similarly as in Section 3. The M-step can be achieved by deriving the first order derivatives of the Q-function with respect to the unknown parameters $\boldsymbol{\Theta}$ and setting them to zero. After some algebra, we obtain the following equations

$$\mu = \frac{1}{\sum_{k=1}^{K}n_k}\sum_{k=1}^{K}\exp(-\beta\xi_k)\sum_{i=1}^{n_k}\mathbb{E}_{v_{k,i}|\boldsymbol{X}_{k,i},\boldsymbol{\Theta},\boldsymbol{\theta}}[v_{k,i}], \tag{12}$$

$$\frac{1}{\zeta} = \frac{1}{\sum_{k=1}^{K}n_k}\sum_{k=1}^{K}\exp(\beta\xi_k)\sum_{i=1}^{n_k}\mathbb{E}_{v_{k,i}|\boldsymbol{X}_{k,i},\boldsymbol{\Theta},\boldsymbol{\theta}}[v_{k,i}^{-1}] - \frac{1}{\mu}, \tag{13}$$

$$\frac{1}{\zeta} = \frac{1}{\sum_{k=1}^{K}n_k\xi_k}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\left(\xi_k\exp(\beta\xi_k)\mathbb{E}_{v_{k,i}|\boldsymbol{X}_{k,i},\boldsymbol{\Theta},\boldsymbol{\theta}}[v_{k,i}^{-1}] - \frac{1}{\mu^2}\xi_k\exp(-\beta\xi_k)\mathbb{E}_{v_{k,i}|\boldsymbol{X}_{k,i},\boldsymbol{\Theta},\boldsymbol{\theta}}[v_{k,i}]\right). \tag{14}$$

Substituting Eq. (12) into (13) and (14), β and ζ can be solved by combining the resulting two equations. Then, the estimates for μ can be obtained subsequently. The parameter $\omega = \kappa^2$ can be updated as

$$\omega^{(s+1)} = \frac{1}{\sum_{k=1}^{K}\sum_{i=1}^{n_k} m_{k,i}} \sum_{k=1}^{K}\sum_{i=1}^{n_k}\left\{\left[\sum_{j=1}^{m_i}\frac{\Delta X_{k,i,j}^2}{\Delta\Lambda_{k,i,j}}\right]\mathbb{E}_{v_{k,i}|\boldsymbol{X}_{k,i},\boldsymbol{\Theta}^{(s)}}[v_{k,i}^{-1}] - 2X_{k,i,m_i} + \Lambda_{k,i,m_i}\mathbb{E}_{v_{k,i}|\boldsymbol{X}_{k,i},\boldsymbol{\Theta}^{(s)}}[v_{k,i}]\right\}. \tag{15}$$

In the case that the parameters $\boldsymbol{\theta}$ in $\Lambda(\cdot)$ should be estimated, the estimates can be obtained by the profile likelihood method, as described in Section 3.

4.2. Interval estimation

The confidence intervals can be constructed similarly as in the single-stress case by the normal asymptotics. Nevertheless, the number of stress levels is generally small in practice, in which case the normal approximation is not accurate. Therefore, we prefer the bootstrap method to derive the interval estimates [25].

We implement the parametric bootstrap based on the estimated $\widehat{\boldsymbol{\Theta}}$ (or $\{\widehat{\boldsymbol{\Theta}}, \widehat{\theta}\}$), as outlined in Algorithm 1. With the bootstrap, B analogues of the ML estimates can be obtained. Consequently, the confidence intervals can be readily constructed by using the percentiles of these B analogues.

5. Illustrative examples

5.1. Laser degradation data

To illustrate the application of the proposed model, we apply it to analyze the GaAs laser degradation data from Meeker and Escobar [11]. The operating current of the laser device increases over time, and the laser device would fail when the operating current exceeds a specified threshold. The dataset consists of the degradation data of 15 testing samples, where each unit is measured at times {250, 500, ...,4000}. The increase of the operating current of the 15 units are illustrated in Fig. 2.

From the figure, we see that the degradation rates of the 15 units have a significant dispersion. Thus, we tentatively fit each degradation path by a basic Wiener process model given in Eq. (1), where the estimates for v_i and σ_i can be easily obtained as

$$\hat{v}_i = \frac{X_{i,m_i}}{t_{i,m_i}},\ \hat{\sigma}_i^2 = \frac{1}{m_i}\sum_{j=1}^{m_i}\frac{(\Delta X_{i,j} - \hat{v}_i\Delta t_{i,j})^2}{\Delta t_{i,j}}.$$

Here, an linear degradation trend $\Lambda(t) = t$ is employed because the operating current appear to increase linearly with time. As shown in Fig. 1, the estimated drift rates of the 15 units demonstrate some right skewness. This indicates that the proposed model is potentially promising.

With a simple calculation, it is found that the correlation ratio between the estimated $\hat{v}_i$ and $\hat{\sigma}_i^2$ is 0.5, implying a positive correlation between $\hat{v}_i$ and $\hat{\sigma}_i^2$. Thus, we impose the relationship $\sigma^2 = \kappa_{\text{ts}} v$, and fit the data by the following fixed-effect Wiener process model

$$X_i(t) = v_i t + \kappa_{\text{ts}}\mathcal{B}(v_i t).$$

The ML estimates for v_i, $i = 1, ...,15$ and κ_{ts}^2 can be derived after some algebra:

$$\hat{\kappa}_{\text{ts}}^2 = \frac{1}{\sum_{i=1}^{n} m_i}\sum_{i=1}^{n}\left(\sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta t_{i,j}}\hat{v}_i^{-1} - 2X_{i,m_i} + \hat{v}_i t_{i,m_i}\right),$$

$$\hat{v}_i = \frac{1}{2t_{i,m_i}}\left(\sqrt{m_i^2\hat{\kappa}_{\text{ts}}^4 + 4t_{i,m_i}\sum_{j=1}^{m_i}\frac{\Delta X_{i,j}^2}{\Delta t_{i,j}}} - m_i\hat{\kappa}_{\text{ts}}^2\right),\ i = 1,...,n.$$

The estimate for κ_{ts}^2 is 5.281×10^{-3}, and the empirical CDF for $\hat{v}_i$, $i = 1, ...,15$ is given in Fig. 3. After a check of the empirical CDF, we found that an IG distribution $\mathcal{IG}(\mu_{\text{ts}}, \zeta_{\text{ts}})$ could provide a good fit to $\hat{v}_i$, $i = 1, ...,15$. The parameters μ_{ts} and ζ_{ts} can be estimated as

Input: The ML estimates $\hat{\boldsymbol{\Theta}}$ (or $\{\hat{\boldsymbol{\Theta}}, \hat{\boldsymbol{\theta}}\}$).
Repeat
1. For $k = 1, \ldots, K$, generate n_k samples $\tilde{v}_{k,i}, i = 1, \ldots, n_k$ from the IG distribution $\mathcal{IG}\left(\hat{\mu}\exp(\hat{\beta}\xi_k), \hat{\zeta}\exp(\hat{\beta}\xi_k)\right)$ as the realizations of n_k heterogeneous drift rates in level k.
2. For $i = 1, \ldots, n_k$, generate the observations $\tilde{\boldsymbol{X}}_{k,i} = (\tilde{X}_{k,i,1}, \ldots, \tilde{X}_{k,i,m_i})^{\mathrm{T}}$ at $t_{k,i,1}, \ldots, t_{k,i,m_i}$ from the Wiener process
$$X_{k,i}(t) = \tilde{v}_{k,i}\hat{\Lambda}(t) + \hat{\kappa}\mathcal{B}\left(\tilde{v}_{k,i}\hat{\Lambda}(t)\right).$$
3. Obtain the estimate $\hat{\boldsymbol{\Theta}}^*$ (or $\{\hat{\boldsymbol{\Theta}}^*, \hat{\boldsymbol{\theta}}^*\}$) based on the sample $\{\tilde{\boldsymbol{X}}_1, \ldots, \tilde{\boldsymbol{X}}_K\}$ using the proposed EM algorithm.

Output: B analogues of the ML estimates $\hat{\boldsymbol{\Theta}}_1^*, \ldots, \hat{\boldsymbol{\Theta}}_B^*$ (or $\{\hat{\boldsymbol{\Theta}}_1^*, \hat{\boldsymbol{\theta}}_1^*\}, \ldots, \{\hat{\boldsymbol{\Theta}}_B^*, \hat{\boldsymbol{\theta}}_B^*\}$).

Algorithm 1. Parametric bootstrap for interval estimation of the ADT model.

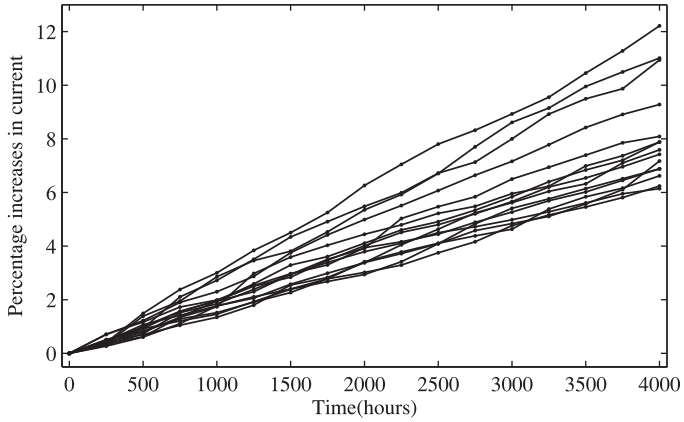


Fig. 2. Degradation paths of 15 laser devices.

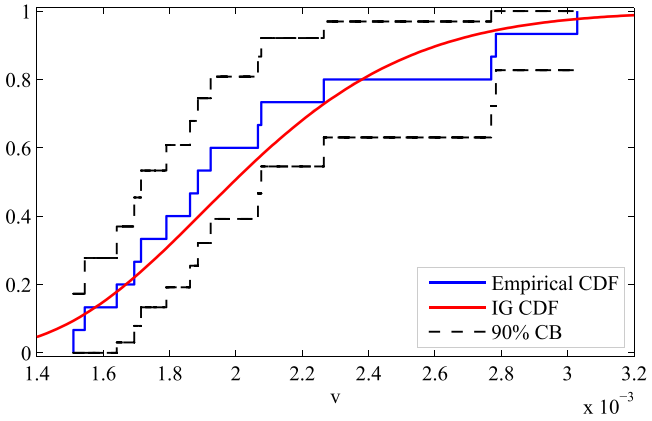


Fig. 3. Empirical CDF with 90% confidence band (CB) and the estimated IG CDF for $\hat{v}_i$, $i = 1, \ldots, 15$.

Table 1
ML estimates and standard deviations for the parameters in the proposed model when fitting to the laser data.

		μ	ζ	κ^2	θ
Linear	MLE	2.037×10^{-3}	5.094×10^{-2}	5.6080×10^{-2}	\
	SD	1.139×10^{-4}	2.154×10^{-4}	5.4053×10^{-3}	\
Power law	MLE	1.991×10^{-3}	4.982×10^{-2}	5.6078×10^{-2}	1.003
	SD	3.829×10^{-4}	2.286×10^{-4}	5.4051×10^{-3}	2.219×10^{-2}

$$\hat{\mu}_{\text{ts}} = \frac{1}{n}\sum_{i=1}^{n} \hat{v}_i = 2.037 \times 10^{-3},$$
$$\hat{\zeta}_{\text{ts}} = \left(\sum_{i=1}^{n} \hat{v}_i^{-1} - \hat{\mu}_{\text{ts}}^{-1}\right)^{-1} = 3.010 \times 10^{-3}.$$

As shown in Fig. 3, the IG distribution provides a good fit for the heterogeneous degradation rates.

Based on the above analysis, we fit the laser data using the proposed random-effects Wiener process model. The estimates for the model parameters are listed in Table 1, where the standard deviations (SD) are obtained by normal asymptotics. Note that the estimates $(\hat{\mu}, \hat{\zeta}, \hat{\kappa}^2)$ here are quite close to $(\hat{\mu}_{\text{ts}}, \hat{\zeta}_{\text{ts}}, \hat{\kappa}^2_{\text{ts}})$. Indeed, we can view $(\hat{\mu}_{\text{ts}}, \hat{\zeta}_{\text{ts}}, \hat{\kappa}^2_{\text{ts}})$ obtained in the above two-step manner as an approximate estimator. The estimates $(\hat{\mu}_{\text{ts}}, \hat{\zeta}_{\text{ts}}, \hat{\kappa}^2_{\text{ts}})$ can also serve as an initial guess for the EM algorithm.

Fig. 4(a) plots the estimate and the 90% confidence band for the expected degradation $\mathbb{E}[X(t)] = \mu t$. Assume the failure threshold is $D_f = 6$, i.e., the laser device is deemed failed if the operating current increases 6%. Then the lifetime distribution can be obtained according to Eqs. (5) and (6). Fig. 4(b) gives the failure time distribution along with the 90% pointwise confidence band for the laser device.

For comparison, we also fit the data by the following models: (i) the commonly used random-effects Wiener process model with Gaussian drift, see, e.g., [29]; (ii) the Wiener process model with skew-normal drift [16]; (iii) the Wiener process model with normal-gamma drift-volatility [24]. The estimation results are listed in Table 2. From the log-likelihood values and AIC values, we can see that the proposed model leads to a better fit compared to these existing models.

We also try the power law $\Lambda(\cdot)$, i.e., $\Lambda(t) = t^{\theta}$ to capture any possible non-linear degradation trend in the laser degradation, and the corresponding estimates are listed in Table 1. It is noted that the estimated θ is quite close to 1 with a relatively small SD. Moreover, the log-likelihood with power law $\Lambda(\cdot)$ is 74.10, which is only slightly larger than that in the linear case. According to the AIC criterion, it turns out that the linear model is better. Therefore, the linear Wiener model with IG drift is a good model for the laser degradation data.

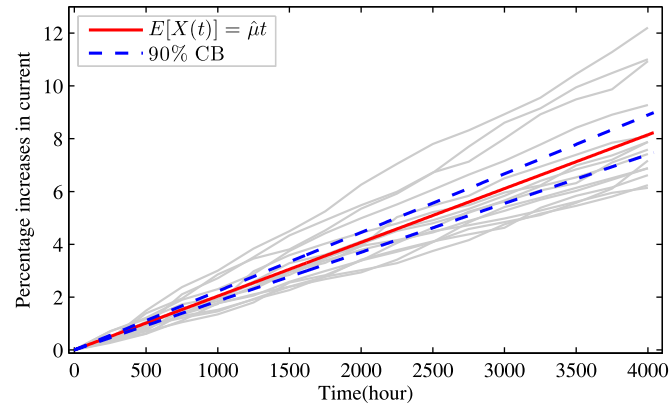
5.2. IRLED ADT data

To illustrate the application of the proposed model in analyzing ADT data, the constant-stress ADT data of infrared light-emitting diodes (IRLED) in [27, Example 8.10] are considered. IRLEDs are highly-reliable optoelectronic devices used widely in communication systems. The performance of IRLEDs is mainly characterized by the variation ratio of luminous power. To estimate the reliability at the operating current of 50 mA, 40 units were sampled and divided into two groups. A group of 25 units was tested at 170 mA and the group of 15 units was tested at 320 mA. The degradation data are listed in Tables 8.9 and 8.10 in [27], and the degradation paths of the 40 units are illustrated in Fig. 5.

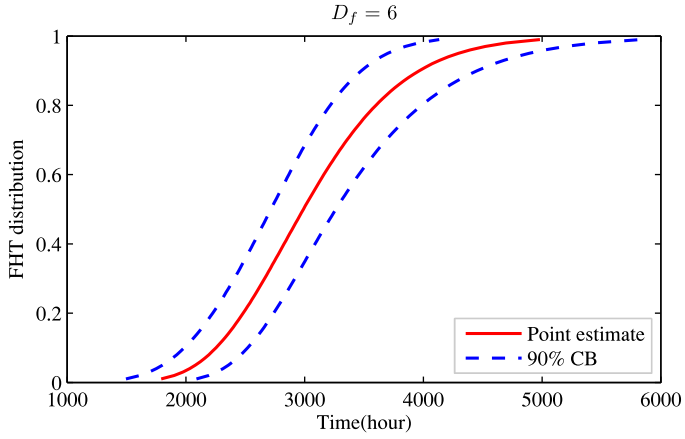
As can be noticed, the degradation of units, even under the same stress level, shows a considerable heterogeneity. To examine the random effects, we first apply the proposed model in Section 3 to fit the degradation data in each level individually. Table 3 presents the estimation results, where the power law degradation trend $\Lambda(t; \theta) = t^{\theta}$ is used following [27]. It can be noted that $\hat{\theta}_k$ under two different levels are close to each other, implying that the degradation mechanism is unchanged in elevated stresses. Moreover, the estimates $\hat{\kappa}_k^2$ under the two levels are close, justifying the assumption that κ and θ are constant across different stress levels.

As expected, the estimate $\hat{\mu}_k$ increases with the operating current. On the other hand, $\hat{\zeta}_k$ is also increasing with the stress level. A simple calculation shows that $\hat{\mu}_2/\hat{\mu}_1$ and $\hat{\zeta}_2/\hat{\zeta}_1$ are of the same magnitude. This justifies our model assumption based on the accelerated failure time principles. Hence, we apply the proposed model to fit the IRLED data and obtain the following ML estimates using the EM algorithm described in Section 4:

$\hat{\mu} = 5.457 \times 10^{-4}$, $\hat{\zeta} = 1.492 \times 10^{-3}$, $\hat{\beta} = 6.592$, $\hat{\kappa}^2 = 1.450$, and $\hat{\theta} = 0.741$.



(a)



(b)

Fig. 4. Left: Estimate for the expected degradation trend $E[X(t)]$; Right: Estimated life time distribution and 90% pointwise confidence band for $D_f = 6$.

Table 2
Performance comparison of different models when fitting to the laser data.

	Log-likelihood	AIC
Proposed IG drift-Linear	74.09	−142.18
Proposed IG drift-Power law	74.10	−140.20
Gaussian drift	69.19	−132.38
Skew-normal drift	71.11	−134.22
Normal-Gamma drift-volatility	72.86	−137.73

Table 3
Individual estimation results for the degradation under two stress levels.

level	$\hat{\mu}_k$	$\hat{\zeta}_k$	$\hat{\kappa}_k^2$	$\hat{\theta}_k$
170 mA	0.03928	0.069891	1.4165	0.7575
320 mA	0.3896	2.4029	1.517	0.7316

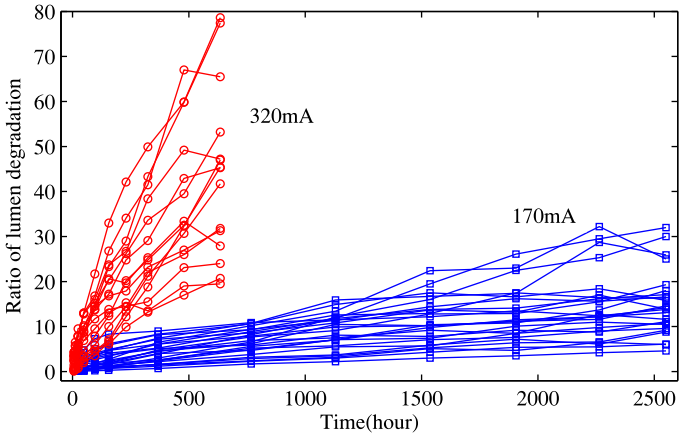


Fig. 5. Degradation paths of 40 IRLEDs under two levels of operating current.

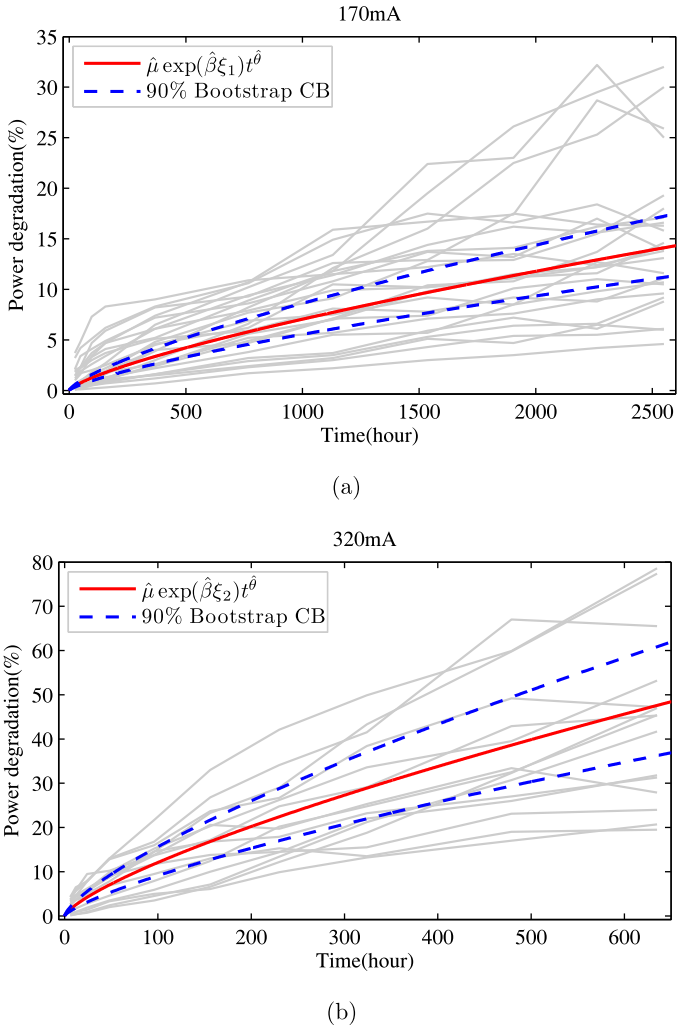


Fig. 6. Expected degradation trend $\mathbb{E}[X(t)]$ and 90% confidence bands under different operating current.

Note that we have standardized the stresses according to [9], where the degradation rate is linked to the operating current using a power law relationship [27].

Based on the estimated model, we get the pointwise point estimates together with the 90% confidence bands for the expected degradation paths under different stress levels, as shown in Fig. 6. The confidence bands are constructed by parametric bootstrap with $B = 5000$ re-samples. From the figures, we see that the estimated paths fit the degradation observations properly in both stress levels. With the estimated model, the degradation law can be extrapolated into the normal use stress, based on which the failure time distribution of the IRLED can be derived. Fig. 7 gives the CDF of the lifetime of the IRLED under two thresholds: $D_f = 10$ and $D_f = 20$. As indicated by the upper confidence bound of the CDF, the IRLED has a quite long life at the use stress level. Indeed, the lower end of 90% interval estimates for the expected lifetime under the failure threshold $D_f = 10$ is 3.195×10^5 hours (36.5 years). This implies the IRLED has an ultra-high reliability and its failure probability would be negligible in practical use.

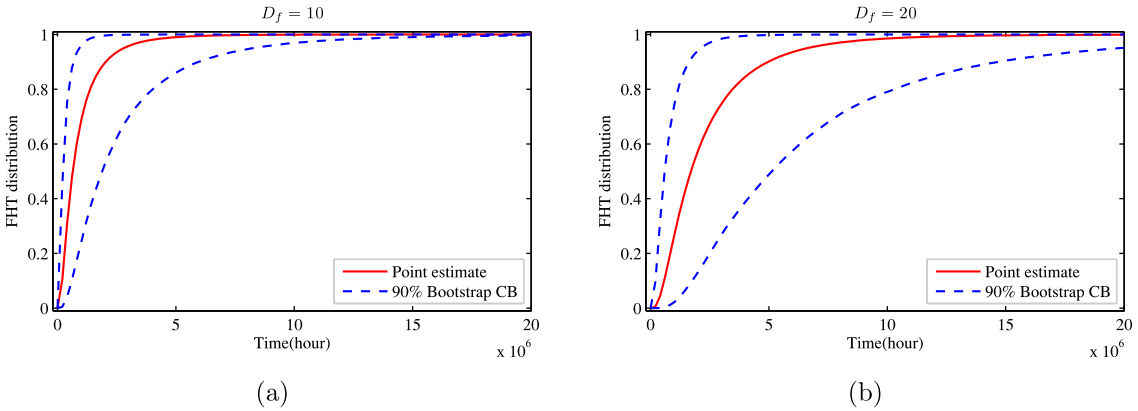


Fig. 7. Lifetime distribution (solid) and 90% confidence bands (dashed) of the IRLED under different thresholds.

6. Conclusions

In this study, we proposed a novel random-effects Wiener process model for degradation data based on linear accelerated failure time principle. The new Wiener process model exploits an IG distribution to describe the heterogeneous degradation rates in the population and has the diffusion coefficient dependent on the drift rate. Compared with the existing models, the proposed Wiener process model with IG drift has the following features:

- The IG distribution has a support on $(0, +\infty)$ and a flexible shape, which is more suitable to model the heterogeneous drift rates than normal distributions;
- The diffusion coefficient is positively correlated with the drift rate, which explains the phenomenon that a unit has a larger degradation rate often has a larger variability in the degradation path.

For the Wiener process model with IG drift, an EM algorithm was developed to obtain the point estimates as well as the interval estimates of the model parameters. The proposed model was further extended to the constant-stress ADT data. By applying to a laser degradation dataset and an IRLED dataset, the efficiency and applicability of the Wiener process model with IG drift in modeling single-stress degradation and multi-level ADT were properly demonstrated.

References

[1] Handbook of mathematical functions with formulas, graphs, and mathematical tables, chap. 9.6 Modified BesselFunctions I and K. In: Abramowitz M, Stegun IA, editors. New York: Dover; 1972. p. 374–7.
[2] Fan M, Zeng Z, Zio E, Kang R. Modeling dependent competing failure processes with degradation-shock dependence. Reliab Eng Syst Saf 2017;165:422–30.
[3] Feng Q, Peng H, Coit D. A degradation-based model for joint optimization of burn-in, quality inspection, and maintenance: a light display device application. Int J Adv ManufTechnol 2010;50(5–8):801–8.
[4] Hu C-H, Lee M-Y, Tang J. Optimum step-stress accelerated degradation test for wiener degradation process under constraints. Eur J Oper Res 2015;241(2):412–21.
[5] Huang Z, Xu Z, Wang W, Sun Y. Remaining useful life prediction for a nonlinear heterogeneous wiener process model with an adaptive drift. IEEE Trans Reliab 2015;64(2):687–700. https://doi.org/10.1109/TR.2015.2403433.
[6] Jørgensen B. Statistical properties of the generalized inverse gaussian distribution. 9. Springer Science & Business Media; 1982.
[7] Kong Y, Ye Z-S. A cumulative-exposure-based algorithm for failure data from a load-sharing system. IEEE Trans Reliab 2016;65(2):1001–13.
[8] Liao H, Elsayed EA. Reliability inference for field conditions from accelerated degradation testing. Nav Res Logist 2006;53(6):576–87.
[9] Lim H, Yum B-J. Optimal design of accelerated degradation tests based on wiener process models. J Appl Stat 2011;38(2):309–25.
[10] Louis TA. Finding the observed information matrix when using the em algorithm. J R Stat Soc Ser B (Methodological) 1982;44(2):226–33. http://www.jstor.org/stable/2345828
[11] Meeker WQ, Escobar LA. Statistical methods for reliability data. New York: John Wiley & Sons, INC.; 1998.
[12] Meeker WQ, Escobar LA, Lu CJ. Accelerated degradation tests: modeling and analysis. Technometrics 1998;40(2):89–99.
[13] Nelson W. Accelerated life testing-step-stress models and data analyses. IEEE Trans Reliab 1980;R-29(2):103–8. https://doi.org/10.1109/TR.1980.5220742.
[14] Newby M. Accelerated failure time models for reliability data analysis. Reliab Eng Syst Saf 1988;20(3):187–97.
[15] Park JI, Bae SJ. Direct prediction methods on lifetime distribution of organic light-emitting diodes from accelerated degradation tests. IEEE Trans Reliab 2010;59(1):74–90. https://doi.org/10.1109/TR.2010.2040761.
[16] Peng CY, Tseng ST. Statistical lifetime inference with skew-wiener linear degradation models. IEEE Trans Reliab 2013;62(2):338–50. https://doi.org/10.1109/TR.2013.2257055.
[17] Peng W, Li YF, Mi J, Yu L, Huang HZ. Reliability of complex systems under dynamic conditions: a bayesian multivariate degradation perspective. Reliab Eng Syst Saf 2016;153:75–87.
[18] Peng W, Li YF, Yang YJ, Mi J, Huang HZ. Bayesian degradation analysis with inverse gaussian process models under time-varying degradation rates. IEEE Trans Reliab 2017;66(1):84–96. https://doi.org/10.1109/TR.2016.2635149.
[19] Seshadri V. The inverse gaussian distribution: statistical theory and applications. 137. Springer Science & Business Media; 2012.
[20] Si X-S, Chen M-Y, Wang W, Hu C-H, Zhou D-H. Specifying measurement errors for required lifetime estimation performance. Eur J Oper Res 2013;231(3):631–44.
[21] Si X-S, Wang W, Hu C-H, Zhou D-H, Pecht MG. Remaining useful life estimation based on a nonlinear diffusion degradation process. IEEE Trans Reliab 2012;61(1):50–67.
[22] Tang S, Yu C, Wang X, Guo X, Si X. Remaining useful life prediction of lithium-ion batteries based on the wiener process with measurement error. Energies 2014;7(2):520–47.
[23] Tseng S-T, Wen Z-C. Step-stress accelerated degradation analysis for highly reliable products. J Qual Technol 2000;32(3):209.
[24] Wang X. Wiener processes with random effects for degradation data. J Multivariate Anal 2010;101(2):340–51.
[25] Wang X, Xu D. An inverse gaussian process model for degradation data. Technometrics 2010;52(2):188–97.
[26] Whitmore G, Schenkelberg F. Modelling accelerated degradation data using wiener diffusion with a time scale transformation. Lifetime Data Anal 1997;3(1):27–45.
[27] Yang G. Life cycle reliability engineering. John Wiley & Sons; 2007.
[28] Ye Z-S, Chen N, Shen Y. A new class of wiener process models for degradation analysis. Reliab Eng Syst Saf 2015;139:58–67.
[29] Ye Z-S, Wang Y, Tsui K-L, Pecht M. Degradation data analysis using wiener processes with measurement errors. IEEE Trans Reliab 2013;62(4):772–80.
[30] Zhai Q, Ye Z-S, Yang J, Zhao Y. Measurement errors in degradation-based burn-in. Reliab Eng Syst Saf 2016;150:126–35. https://doi.org/10.1016/j.ress.2016.01.015.