Accelerated degradation tests with inspection effects

Zhao, Xiujie; Chen, Piao; Gaudoin, Olivier; Doyen, Laurent

DOI
10.1016/j.ejor.2020.11.041

Publication date
2021

Document Version
Final published version

Published in
European Journal of Operational Research

European Journal of Operational Research 292 (2021) 1099–1114

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Innovative Applications of O.R.

Accelerated degradation tests with inspection effects

Xiujie Zhao[a], Piao Chen[b,*], Olivier Gaudoin[c], Laurent Doyen[c]

[a] College of Management and Economics, Tianjin University, Tianjin, China
[b] Department of Applied Mathematics, Technische Universiteit Delft, Delft, the Netherlands
[c] Laboratoire Jean Kuntzmann, Université Grenoble Alpes, Grenoble, France

ARTICLE INFO

Article history:
Received 6 January 2020
Accepted 27 November 2020
Available online 15 December 2020

Keywords:
Reliability
Accelerated degradation tests
Confidence density
Degradation modeling
Wiener process

ABSTRACT

This study proposes a framework to analyze accelerated degradation testing (ADT) data in the presence of inspection effects. Motivated by a real dataset from the electric industry, we study two types of effects induced by inspections. After each inspection, the system degradation level instantaneously reduces by a random value. Meanwhile, the degrading rate is elevated afterwards. Considering the absence of observations due to practical reasons, we employ the expectation–maximization (EM) algorithm to analytically estimate the unknown parameters in a stepwise Wiener degradation process with covariates. Moreover, to maintain the level of generality for the adaption of the method in various scenarios, a confidence density approach is utilized to hierarchically estimate the parameters in the acceleration link function. The proposed methods can provide efficient parameter estimation under complex link functions with multiple stress factors. Further, confidence intervals are derived based on the large-sample approximation. Simulation studies and a case study from Schneider Electric are used to illustrate the proposed methods. The results show that the proposed model yields a remarkably better fit to the Schneider data in comparison to the conventional Wiener ADT model.

1. Introduction and motivation

Reliability tests are widely used to predict product lifetime in various industries. A successfully planned and conducted reliability test can provide important information supporting managerial decisions under a reasonable test budget, thereby reducing both prospective costs and risks. In order to shorten the test duration, conventional life tests are commonly conducted under elevated stresses to accelerate the failures of test units. In recent decades, with the advances in sensors and monitoring technologies, degradation tests become preferable to life tests in the sense that they can predict reliability characteristics over time without the concern of censoring (Meeker, Escobar, & Lu, 1998). In a typical degradation test, discrete degradation measurements are taken and the observed degradation paths are then employed to make inferences about the product reliability. Degradation paths are usually modeled based on a quality characteristic (QC), such as the brightness of displays and battery life of electronic devices (Wang, Tang, Bae, & Xu, 2018).

When utilized for lifetime prediction, the structure and pattern of degradation data are desired to be as simplistic as possible to relieve the modeling complexity and computational burden. However, under practical usage or even controlled experimental conditions, degradation paths may inevitably behave atypical patterns from time to time. In degradation tests, since the stress levels are strictly controlled, the degradation paths of test systems are usually deemed to be stable during most of the test time. Nevertheless, some interventions to the test systems may be unavoidable due to various practical concerns. A common exercise is that engineers have to alter the test conditions temporarily to obtain degradation measurements. For example, many reliability tests are conducted under certain combinations of temperature and humidity, where test chambers that provide such environments are employed. Typical degradation tests of this kind can be found in Meeker and Escobar (1998, Chapter 21) and references therein. To take effective degradation measurements, the involvement of manual inspection or/and precise instruments are mandatory, yet the exercise is difficult, if not impossible under the test environment. In such cases, the test units have to be removed from the chamber temporarily, which results in the change of test environments. Although the duration of measurement is usually short or even negligible compared to the whole test duration, the drastic change of test environments may still cause substantial changes

* Corresponding author.
E-mail address: p.chen-6@tudelft.nl (P. Chen).

https://doi.org/10.1016/j.ejor.2020.11.041

on system degradation levels. Another reason of atypical degradation paths lies in the intervening nature of inspections. In other words, certain types of inspections may inherently influence the system degradation (Zhao, Gaudoin, Doyen, & Xie, 2019). One of the examples is the destructive test, where inspections can cause directly destructive effects to the test systems (Shi, Escobar, & Meeker, 2009). Considering the aforementioned issues, we propose to model the effects brought by inspections in degradation tests and afterwards investigate parameter estimation upon such testing data.

The research to be proposed is motivated by a real experiment carried out by Schneider Electric with the objective to reveal the degradation characteristic of an electrical distribution device. As a key part of the device, a mechanical linkage corrodes over time, which is a dominant cause of performance degradation. To quantify the degradation level of the device, engineers measure the torque that is needed to separate the linkage. A higher torque implies a more severe condition of corrosion. On the one hand, since the inspection separates the linkage, the grown corrosion is physically disassembled, leading to a reduction in the degradation level during the inspection. On the other hand, the inspection causes damage to the integrity of surface treatments in the linkage, which leads to a higher rate of corrosion. Obviously, the two types of effects are opposite with respect to the system health. Another concern of the problem is the difficulty in revealing the accurate degradation reduction during the inspection. To follow a basic rule to inspect systems, engineers tend to minimize the influence of inspection and therefore only measure the torque that separates the linkage. Once a measurement is obtained, the inspection is terminated immediately. Consequently, the measurement process may only capture the degradation level before inspections yet fail to observe degradation reduction. Although the engineers can give an approximation of the degradation reduction from the prior knowledge or other experiments, the accurate value can never be known. Thus, the reduction effect is a hidden variable that cannot be directly utilized for statistical inference.

By employing the experiment from Schneider Electric as an illustrative example, the study aims at establishing a framework to analyze accelerated degradation tests with complex inspection effects. In general, the proposed method can be applied to model degradation data in the presence of environmental covariates and interventions that exert both positive and negative effects.

2. Literature review

With the fast emergence of system monitoring technologies, modeling and inference of degradation data now play a vital role in the research area of reliability engineering and its interfaces with other areas, such as mechanical engineering (Wang & Tsui, 2017), energy (Lin et al., 2017), electrical engineering (Si, 2015) etc. Degradation analysis not only subsumes approaches to model relevant data, but also creates alternative planning methods of reliability tests for life prediction.

Interest in accelerated degradation test (ADT) has grown in recent decades owing to its successful applications to various products and systems, such as LED lamps, lithium-ion batteries and rail tracks (Ye & Xie, 2015). Initiated by Meeker et al. (1998), regression-based general path models are widely used for degradation modeling in ADT (Hong, Duan, Meeker, Stanley, & Gu, 2015). Further, due to clear physical explanations and appealing mathematical tractability, stochastic processes such as Wiener processes (Hu, Lee, & Tang, 2015), gamma processes (Tsai, Sung, Lio, Chang, & Lu, 2016) and inverse Gaussian processes (Ye, Chen, Tang, & Xie, 2014) start to play an influential role in ADT modeling and planning. For a more detailed overview, one can be referred to Limon, Yadav, and Liao (2017). To date, several new considerations on ADT planning and analysis emerged in the literature to tackle more practical issues. To name a few, Tseng and Lee (2016) proposed a general exponential-dispersion model to characterize degradation and gave the optimal ADT plans in analytical forms. In Li, Wu, Ma, Li, and Kang (2018), random fuzzy theory was adopted to model the uncertainty in ADT data. Wang and Tsui (2017) considered multiple stresses in ADT for rubber sealed O-rings. With regard to model uncertainty, Liu, Li, Zio, Kang, and Jiang (2017) applied the Bayesian modeling averaging approach to model ADT data.

In real problems, it is common that a degradation path cannot be characterized by models in regular forms such as linear or typical nonlinear ones (e.g., polynomial, logarithm and exponential models). The reasons behind an atypical degradation path can be rather complex. For example, Hong et al. (2015) proposed a degradation modeling approach by utilizing dynamic weathering covariates to characterize irregular degradation paths. In some other studies (Bae, Yuan, Ning, & Kuo, 2015; Wang et al., 2018), change-point detection and modeling were discussed for degradation data. Adaptive and dynamic methods for online degradation modeling have also prevailed in the literature (Si, 2015; Zhai & Ye, 2018).

Apart from these, human interventions to industrial systems can also be a pivotal cause of atypical degradation path, and a common example is imperfect maintenance (Mercier & Castro, 2019). Surprisingly, despite plentiful extant works on atypical degradation paths, we can only find very scant research in the literature that addressed similar issues in ADT problems. Xiao and Ye (2016) discussed the ADT planning problem with random initial degradation levels. In Ye, Hu, and Yu (2019), the initial performance of test units were considered to allocate units to stress levels. Nevertheless, these works did not incorporate the effect of inspection as described in Section 1.

The rest of the paper is organized as follows. In Section 3, a systematic approach to model construction and parameter estimation is established for ADT data with inspection effects. Section 4 discusses the uncertainty quantification of parameter estimators. Simulation studies are carried out in Section 5. Section 6 presents the case study from Schneider Electric. Finally, conclusions are drawn in Section 7.

3. Degradation models with inspection effects

3.1. Preliminaries

Consider a degradation test with M stress levels and N_i test units are allocated to stress level i, where $i = 1, \ldots, M$. For the jth test unit under stress i, as which we call unit (i, j) for simplicity, a total of O_{ij} inspections are carried out at time epochs $\tau_{ij1}, \ldots, \tau_{ijO_{ij}}$. Since an instant degradation reduction occurs upon each inspection, we denote the degradation level before the reductive effect by y^-_{ijk} for the kth inspection for unit (i, j), with $k = 0$ representing the initial inspection prior to the test. Meanwhile, the degradation level after the reductive effect is denoted by y^+_{ijk}. We introduce a variable z_{ijk} to model the proportion of degradation reduction with respect to y^-_{ijk} for the kth inspection of unit (i, j) as follows:

$$z_{ijk} = \frac{y^-_{ijk} - y^+_{ijk}}{y^-_{ijk}}, \quad 0 \le z_{ijk} \le 1. \tag{1}$$

The notational details are illustrated in Fig. 1. As mentioned in Section 1, y^-_{ijk} is usually observable whereas y^+_{ijk} is relatively difficult to reveal. Thus, for notational simplicity, we subset the data to $\boldsymbol{y}^-$ and $\boldsymbol{z}$ that are given by

$$\boldsymbol{y}^- = \{y^-_{ijk}, i = 1, \ldots, M, j = 1, \ldots, N_i, k = 0, \ldots, O_{ij}\},$$

$$\boldsymbol{z} = \{z_{ijk}, i = 1, \ldots, M, j = 1, \ldots, N_i, k = 0, \ldots, O_{ij}\}.$$

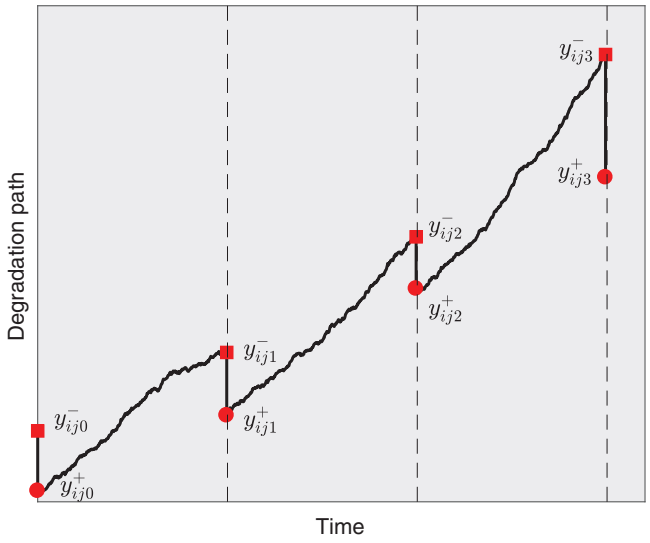

Fig. 1. Illustration for notations by a possible sample path of degradation for unit j tested under environment i.

To further facilitate degradation modeling, we let $\Delta \boldsymbol{y} = \{\Delta y_{ijk}, i = 1, \ldots, M, j = 1, \ldots, N_i, k = 1, \ldots, O_{ij}\}$, where $\Delta y_{ijk} = y^-_{ijk} - y^+_{ij(k-1)}$. Considering the notation in (1), we can rewrite Δy_{ijk} as

$$\Delta y_{ijk} = y^-_{ijk} - y^-_{ij(k-1)}(1 - z_{ij(k-1)}).$$

3.2. Wiener degradation model and inspection effect

We employ the Wiener process as the baseline model to characterize the inherent generation of degradation for system of interest. More concretely, if we assume that the system operates without any intervention, the degradation path of the system can be modeled by a drifted Wiener process. The Wiener process features independent Gaussian increments over non-overlapping time periods, which enables its wide application in degradation modeling. A drifted Wiener process $\{W(t); t \geq 0\}$ can be characterized by drift and diffusion parameters, denoted by μ and σ, respectively. In this manner, it gives $W(t) = \mu t + \sigma \mathcal{B}(t)$, where $\mathcal{B}(\cdot)$ is a standard Brownian motion. The increment $\Delta W(t-s) = W(t) - W(s)$ for any $t > s$ follows a normal distribution with mean $\mu(t-s)$ and variance $\sigma^2(t-s)$, i.e., $\Delta W(t-s) \sim N(\mu(t-s), \sigma^2(t-s))$. In the presence of inspection effects, i.e., nonzero z_{ijk}'s, the degradation path no longer follows the conventional Wiener process. Nevertheless, according to the aforementioned properties, given z_{ijk}'s, Δy_{ijk} are independent increments of a Wiener process and they follow:

$$\Delta y_{ijk} \sim N(\mu_{ijk}\Delta\tau_{ijk}, \sigma^2\Delta\tau_{ijk}). \tag{2}$$

where μ_{ijk} is the degradation rate between the $(k-1)$th and kth inspection for test unit (i, j) and $\Delta\tau_{ijk} = \tau_{ijk} - \tau_{ij(k-1)}$. Note that if the inspection has no effect on the system degradation, then z_{ijk}'s are 0 for all i, j and k, and the degradation path can be modeled by a conventional drifted Wiener process. Note that in the proposed model, we use a flexible notational convention in the sense that τ_{ijk} and O_{ij} can be different for different test units, implying that the proposed methods can be well applied to unbalanced data set in terms of time and number of observations.

Next, experimental factors as covariates are introduced into the model. The majority of research on degradation tests has suggested to use parametric models to link μ and covariates $\boldsymbol{x}_i$ (Jakob, Kimmelmann, & Bertsche, 2017). We use $f_{\text{acc}}(\boldsymbol{x}_i)$ to denote the baseline degradation rate under stress $\boldsymbol{x}_i$ and the acceleration model can be formulated in either parametric or nonparametric manners. The parametric acceleration model can take various forms based on the physical mechanism of degradation and factors involved in the test. The selection of $f_{\text{acc}}(\boldsymbol{x}_i)$ is not the main focus of the study, and some brief discussions regarding the Schneider example are given in Appendix A. Moreover, the complexity in $f_{\text{acc}}(\boldsymbol{x}_i)$ may impede the inferential efficiency under various specific models. Under different forms of $f_{\text{acc}}(\boldsymbol{x}_i)$, we wish to maximize the generalizability of the proposed model. Towards this end, we treat $f_{\text{acc}}(\boldsymbol{x}_i)$ as separate parameters first and then employ a hierarchical analysis in Section 3.6 for further inferences.

To capture the effect of degradation rate increase after the kth inspection, a function $g(k; \boldsymbol{\omega})$ is introduced, where $\boldsymbol{\omega}$ is a vector of unknown parameters. To benchmark the effect, we use $g(k; \boldsymbol{\omega})$ as an added term to the baseline degradation rate and therefore we assume $g(0; \boldsymbol{\omega}) \equiv 0$. Then, the degradation rate between the $(k-1)$th and kth inspection for test unit (i, j) is given by

$$\mu_{ijk} = f_{\text{acc}}(\boldsymbol{x}_i) + g(k; \boldsymbol{\omega}). \tag{3}$$

One implicit assumption from (3) is that the increase effect of degradation rate brought by inspections and stress variables $\boldsymbol{x}_i$ are independent. In other words, we assume that the inspection effect on degradation rate only depends on k and does not interact with environmental stresses. Since the degradation rate increases with the number of inspections, $g(k; \boldsymbol{\omega})$ is a non-decreasing function. The simplest forms are polynomial, e.g., the first and second order polynomial models are given as follows:

$$g(k; \omega) = \omega k, \quad g(k; \boldsymbol{\omega}) = \omega_1 k^2 + \omega_2 k.$$

Without loss of generality, we assume $g(k; \omega) = \omega k$ for analytical simplicity in the following analysis.

Recall that we have introduced z_{ijk} earlier to describe the proportion of degradation reduction. A beta distribution is employed to model the unconditioned z_{ijk}, for any realization of z_{ijk}, denoted by z, the density function is given by

$$f_{\text{Beta}}(z; u, v) = \frac{\Gamma(u+v)}{\Gamma(u)\Gamma(v)} z^{u-1}(1-z)^{v-1}, \quad 0 \leq z \leq 1, \tag{4}$$

where u and v are the shape parameters of the beta distribution. It is worth mentioning that beta distribution has been widely used to model the effect of the imperfect repair (Zhang, Gaudoin, & Xie, 2015).

Remark. If y^+_{ijk} is unobservable, u and v cannot be estimated from the model due to the absence of $\boldsymbol{z}$ under the frequentist setting. However, engineers may manage to approximately quantify the reductive effect from domain expertise or preliminary experiments. For the example from the Schneider Electric, engineers can carry out a different type of experiment to measure the degradation levels before and after the first inspection at $t = 0$ and model the reductive effects, though it is inapplicable during the ADT as it takes much longer time to obtain the reductive measurements. In the presence of the described available data at $t = 0$, it is easy to fit a beta distribution to the observed reductive measurements. Other types of models to characterize $\boldsymbol{z}$ can also be used after uncomplicated modifications to the likelihoods in the following contents in the section and density functions in Appendix B. It is noteworthy that the Bayesian method can be an appealing alternative, where u and v can be characterized by some prior distributions at first. The computational burden can be an important issue in Bayesian inference, especially when $\boldsymbol{z}$ is unobservable (Bernardo et al., 2003). Further, in this paper, we presume that z_{ijk}'s are independent random variables. The assumption is valid if the inspection effects of degradation reduction are instantaneous and do not interact over time. In realistic applications, the effects can interact over time under stress environments. In such cases, a joint distribution can be employed to characterize the interdependence of

z_{ijk}'s. However, the introduction of interdependence leads to more complicated likelihoods, which hinders the tractability of estimators. Monte Carlo EM algorithm can be useful in implementing the aforementioned extensions (Levine & Casella, 2001).

In Sections 3.3 and 3.4, we will discuss the modeling of degradation data under two different scenarios. In the first one, the degradation levels before and after the inspection are observable so that all the parameters in the model can be estimated through a complete likelihood. In the second one, only the degradation levels before the inspection can be obtained for inference, under which case parameters u and v are given a priori to facilitate the estimation of other unknown parameters.

3.3. Inspection effect modeling with complete observations

As preliminaries for the following analyses, this part aims at establishing data modeling framework via maximum likelihood estimation (MLE). The complete log-likelihood function of $\boldsymbol{\theta}_c$ under data $(\boldsymbol{y}^-, \boldsymbol{z})$ can be represented by

$$\begin{aligned}\log L(\boldsymbol{\theta}_c|\boldsymbol{y}^-,\boldsymbol{z}) &= \log p\left(\boldsymbol{y}^-|\boldsymbol{z},\boldsymbol{\theta}_c\right) + \log p(\boldsymbol{z}|\boldsymbol{\theta}_c)\\ &= \sum_{i=1}^{M}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}\left(-\frac{1}{2}\log 2\pi - \frac{1}{2}\log\Delta\tau_{ijk} - \log\sigma\right)\\ &\quad - \sum_{i=1}^{M}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}\left\{\frac{\left[y^-_{ijk} - y^-_{ij(k-1)}\left(1-z_{ij(k-1)}\right) - \mu_{ijk}\Delta\tau_{ijk}\right]^2}{2\sigma^2\Delta\tau_{ijk}}\right\}\\ &\quad + \sum_{i=1}^{M}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}+1}\left[\log\Gamma(u+v) - \log\Gamma(u) - \log\Gamma(v)\right.\\ &\quad \left. + (u-1)\log z_{ij(k-1)} + (v-1)\log\left(1-z_{ij(k-1)}\right)\right], \end{aligned} \tag{5}$$

where $\mu_{ijk} = f_{\text{acc}}(\boldsymbol{x}_i) + \omega k$. Note again that we assume that z_{ijk}'s are mutually independent and they are also independent of $\boldsymbol{y}$ and the environmental factors. Under the current model setting, the unknown parameters in the model can be summarized by $\boldsymbol{\theta}_c = (f_{\text{acc}}(\boldsymbol{x}_i), i = 1,\ldots,M, \omega, \sigma^2, u, v)^T$. If the effect of degradation reduction can be observed, the problem is simplified to a standard MLE problem with all observations available in (5). Further, since parameters u and v are independent of other parameters in the model, the likelihood involving u and v can be independently maximized via the observations z_{ijk}. The remaining part of the likelihood can also be maximized by numerical methods.

3.4. Inspection effect modeling with hidden effect observations

When y^+_{ijk} is unobservable, the effect of degradation reduction cannot be captured, which makes u and v in the model inestimable. This kind of information can also be quantified by a beta distribution as described in (4). As discussed before, we assume that the pilot distribution of z is known and characterized by u_0 and v_0, respectively. By holding the property of independence of z_{ijk}'s. the following log-likelihood is to be maximized to estimate $\boldsymbol{\theta} = (f_{\text{acc}}(\boldsymbol{x}_i), i = 1,\ldots,M, \omega, \sigma^2)$:

$$\begin{aligned}\log L(\boldsymbol{\theta}|\boldsymbol{y}^-,\boldsymbol{z}) &= \log p\left(\boldsymbol{y}^-|\boldsymbol{z},\boldsymbol{\theta}\right)\\ &= \sum_{i=1}^{M}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}\left(-\frac{1}{2}\log 2\pi - \frac{1}{2}\log\Delta\tau_{ijk} - \log\sigma\right)\\ &\quad - \sum_{i=1}^{M}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}\left\{\frac{\left[y^-_{ijk} - y^-_{ij(k-1)}\left(1-z_{ij(k-1)}\right) - \mu_{ijk}\Delta\tau_{ijk}\right]^2}{2\sigma^2\Delta\tau_{ijk}}\right\}. \end{aligned} \tag{6}$$

Due to the absence of z_{ijk}, the log-likelihood cannot be maximized in its current form. Alternatively, we resort to the expectation–maximization (EM) algorithm to obtain the parameter estimates. The EM algorithm is an iterative method to find the MLE for statistical models with latent variables. Following its first introduction by Dempster, Laird, and Rubin (1977), the EM algorithm has been studied extensively from both theoretical and practical perspectives (McLachlan & Krishnan, 2007). Specifically, in reliability engineering, the EM algorithm is commonly used to capture latent random effect in life and degradation models (Chen & Ye, 2017; Duan & Wang, 2018). Apart from the EM algorithm, the hidden semi-Markov model was used for health diagnosis and prognosis with latent effects in Dong and He (2007). A two-stage approach was proposed in Lee, Hu, and Tang (2017) to estimate the model from time-censored ADT data. The Kalman filtering technique has also been prevailingly employed in the remaining life estimation based on degradation models (Si, Wang, Hu, & Zhou, 2014). The EM algorithm consists of two steps: (1) the E-step in which the conditional expectation of the complete log-likelihood with respect to incomplete data is completed; (2) the M-step in which the expected log-likelihood is maximized to generate parameter estimation at the current iteration. Denote the conditional expectation of the complete log-likelihood at iteration n by $Q\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(n)}\right)$, we have

$$\begin{aligned}Q\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(n)}\right) &= \sum_{i=1}^{M}\sum_{j=1}^{N_i} O_{ij}\left(-\frac{1}{2}\log 2\pi - \frac{1}{2}\log\Delta\tau_{ijk} - \log\sigma\right)\\ &\quad - \frac{1}{2\sigma^2}\sum_{i=1}^{M}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}\frac{1}{\Delta\tau_{ijk}}\left[\left(y^-_{ijk} - y^-_{ij(k-1)}\right.\right.\\ &\quad \left. -f_{\text{acc}}(\boldsymbol{x}_i)\Delta\tau_{ijk} - \omega k\Delta\tau_{ijk}\right)^2 + 2y^-_{ij(k-1)}\left(y^-_{ijk} - y^-_{ij(k-1)}\right.\\ &\quad \left. -f_{\text{acc}}(\boldsymbol{x}_i)\Delta\tau_{ijk} - \omega k\Delta\tau_{ijk}\right)E\left(z_{ij(k-1)}|\boldsymbol{y}^-,\boldsymbol{\theta}^{(n)}\right)\\ &\quad \left. + \left(y^-_{ij(k-1)}\right)^2 E\left(z^2_{ij(k-1)}|\boldsymbol{y}^-,\boldsymbol{\theta}^{(n)}\right)\right]. \end{aligned} \tag{7}$$

In the expectation step, we need to compute $E\left(z_{ij(k-1)}|\boldsymbol{y}^-,\boldsymbol{\theta}^{(n)}\right)$ and $E\left(z^2_{ij(k-1)}|\boldsymbol{y}^-,\boldsymbol{\theta}^{(n)}\right)$, Unfortunately, the conditional distribution of $z_{ij(k-1)}$ cannot be identified as a known random distribution. We have to resort to numerical methods to evaluate the expected values with respect to $z_{ij(k-1)}$. Related details are given in Appendix B. Next, the first order derivatives of $Q\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(n)}\right)$ are provided for its maximization. Following previous analyses, we also treat $f_{\text{acc}}(\boldsymbol{x}_i)$ as unknown parameters. Thus,

$$\begin{aligned}\frac{\partial Q\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(n)}\right)}{\partial f_{\text{acc}}(\boldsymbol{x}_i)} &= -\frac{1}{\sigma^2}\left[f_{\text{acc}}(\boldsymbol{x}_i)\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}\Delta\tau_{ijk}\right.\\ &\quad - \sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}\left[y^-_{ijk} - y^-_{ij(k-1)} + y^-_{ij(k-1)}E\left(z_{ij(k-1)}|\boldsymbol{y}^-,\boldsymbol{\theta}^{(n)}\right)\right]\\ &\quad \left. + \omega\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}k\Delta\tau_{ijk}\right]. \end{aligned} \tag{8}$$

By solving $\partial Q\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(n)}\right)/\partial f_{\text{acc}}(\boldsymbol{x}_i) = 0$, we obtain the following equation:

$$f^{(n+1)}_{\text{acc}}(\boldsymbol{x}_i) = A^{(n)}_i - B_i\omega^{(n)}, \tag{9}$$

where

$$A^{(n)}_i = \frac{\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}\left[y^-_{ijk} - y^-_{ij(k-1)} + y^-_{ij(k-1)}E\left(z_{ij(k-1)}|\boldsymbol{y}^-,\boldsymbol{\theta}^{(n)}\right)\right]}{\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}\Delta\tau_{ijk}},$$

$$B_i = \frac{\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}} k\Delta\tau_{ijk}}{\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}} \Delta\tau_{ijk}}.$$

Further for ω, we have

$$\frac{\partial Q\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(n)}\right)}{\partial\omega} = -\frac{1}{\sigma^2}\Bigg[-\sum_{i=1}^{M}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}} k\Big(y^-_{ijk} - y^-_{ij(k-1)} + y^-_{ij(k-1)}E\left(z_{ij(k-1)}|\boldsymbol{y}^-,\boldsymbol{\theta}^{(n)}\right)\Big) + \sum_{i=1}^{M}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}} k\Delta\tau_{ijk}f_{\mathrm{acc}}(\boldsymbol{x}_i) + \omega\sum_{i=1}^{M}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}} k^2\Delta\tau_{ijk}\Bigg], \tag{10}$$

of which the solution is followed by

$$C\omega^{(n)} = D^{(n)} - \sum_{i=1}^{M} f_{\mathrm{acc}}(\boldsymbol{x}_i)F_i, \tag{11}$$

where

$$C = \sum_{i=1}^{M}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}} k^2\Delta\tau_{ijk},$$

$$D^{(n)} = \sum_{i=1}^{M}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}} k\Big[y^-_{ijk} - y^-_{ij(k-1)} + y^-_{ij(k-1)}E\left(z_{ij(k-1)}|\boldsymbol{y}^-,\boldsymbol{\theta}^{(n)}\right)\Big],$$

$$F_i = \sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}} k\Delta\tau_{ijk}.$$

To plug Eq. (9) to Eq. (11), the following equation is obtained

$$C\omega^{(n)} = D^{(n)} - \sum_{i=1}^{M}\left(A_i^{(n)} - B_i\omega\right)F_i, \tag{12}$$

which yields the estimates of ω and $f_{\mathrm{acc}}(\boldsymbol{x}_i;\boldsymbol{\delta})$ given by

$$\omega^{(n+1)} = \frac{D^{(n)} - \sum_{i=1}^{M} A_i^{(n)}F_i}{C - \sum_{i=1}^{M} B_iF_i}, \quad f_{\mathrm{acc}}^{(n+1)}(\boldsymbol{x}_i) = A_i^{(n)} - B_i\omega^{(n+1)}. \tag{13}$$

Then for σ^2, likewise we have

$$\frac{\partial Q\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(n)}\right)}{\partial\sigma^2} = -\frac{1}{2\sigma^2}\sum_{i=1}^{M}\sum_{j=1}^{N_i} O_{ij} + \frac{1}{2\sigma^4}\sum_{i=1}^{M}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}\frac{1}{\Delta\tau_{ijk}}\Big[\left(y^-_{ijk} - y^-_{ij(k-1)} - f_{\mathrm{acc}}(\boldsymbol{x}_i)\Delta\tau_{ijk} - \omega k\Delta\tau_{ijk}\right)^2 + 2y^-_{ij(k-1)}\left(y^-_{ijk} - y^-_{ij(k-1)} - f_{\mathrm{acc}}(\boldsymbol{x}_i)\Delta\tau_{ijk} - \omega k\Delta\tau_{ijk}\right)E\left(z_{ij(k-1)}|\boldsymbol{y}^-,\boldsymbol{\theta}^{(n)}\right) + \left(y^-_{ij(k-1)}\right)^2E\left(z^2_{ij(k-1)}|\boldsymbol{y}^-,\boldsymbol{\theta}^{(n)}\right)\Big]. \tag{14}$$

The root of the equation can be easily obtained by solving

$$\left(\sigma^2\right)^{(n+1)}\sum_{i=1}^{M}\sum_{j=1}^{N_i} O_{ij} = \sum_{i=1}^{M}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}\frac{1}{\Delta\tau_{ijk}}\Big[\left(y^-_{ijk} - y^-_{ij(k-1)} - f_{\mathrm{acc}}(\boldsymbol{x}_i)\Delta\tau_{ijk} - \omega k\Delta\tau_{ijk}\right)^2 + 2y^-_{ij(k-1)} \times \left(y^-_{ijk} - y^-_{ij(k-1)} - f_{\mathrm{acc}}(\boldsymbol{x}_i)\Delta\tau_{ijk} - \omega k\Delta\tau_{ijk}\right)E\left(z_{ij(k-1)}|\boldsymbol{y}^-,\boldsymbol{\theta}^{(n)}\right) + \left(y^-_{ij(k-1)}\right)^2E\left(z^2_{ij(k-1)}|\boldsymbol{y}^-,\boldsymbol{\theta}^{(n)}\right)\Big]. \tag{15}$$

Through (9)–(15), the current iteration of the EM algorithm is realized. The iterations are continued until the convergence of parameter estimators.

3.5. Guess of initial estimates and ending of iterations

To start the aforementioned EM algorithm, starting estimates $\boldsymbol{\theta}^{(0)}$ are needed. The convergence speed of the algorithm hinges on the selection of $\boldsymbol{\theta}^{(0)}$. Here, we utilize the mean of the z_{ijk} to approximately obtain $\boldsymbol{\theta}^{(0)}$. It is obvious that $E(z_{ijk}) = u_0/(u_0+v_0)$. Therefore, to use $E(z_{ijk})$ rather than unobservable z_{ijk}, we have

$$\left(y^-_{ijk} - \frac{v_0y^-_{ij(k-1)}}{u_0+v_0}\right) \sim N\left(f_{\mathrm{acc}}(\boldsymbol{x}_i)\Delta\tau_{ijk} + \omega k\Delta\tau_{ijk}, \sigma^2\Delta\tau_{ijk}\right), \tag{16}$$

where the left-hand-side term is completely observable and after the manipulation of a typical MLE, $\boldsymbol{\theta}^{(0)}$ can be given as in Appendix C.

Another issue of the EM algorithm is when to terminate the iterations. The question poses a tradeoff between the estimating precision and computational efficiency. Generally, it is a common criterion to terminate the iterations when the proportions of changes in absolute values of parameter estimators are smaller than critical values $\boldsymbol{\varepsilon}$. It is a vector because different parameters may have different critical values. For the problem described in the paper, parameters play different roles depending on how decision makers would utilize the estimates. For example, in terms of life prediction, $f_{\mathrm{acc}}(\boldsymbol{x}_i)$'s are important for the extrapolation to understand the degradation rate under normal usage conditions, especially in the presence of condition fluctuations, while ω is more useful if the device is frequently inspected. Regarding these reliability issues, $\boldsymbol{\varepsilon}$ can be properly determined to satisfy the required estimating accuracy.

3.6. A hierarchical analysis to estimate $f_{\mathrm{acc}}(\mathbf{x}_i)$

In aforementioned analyses, $f_{\mathrm{acc}}(\boldsymbol{x}_i), i = 1,\ldots,M$ are treated as unknown parameters for estimation. As one of the main objectives of the study, the estimation of degradation rate under normal usage condition is realized by a hierarchical method. Due to the possible complexity in $f_{\mathrm{acc}}(\boldsymbol{x}_i)$, it could be onerous to derive analytical iterative solutions to the parameters herein to the EM algorithm. In view of this, the hierarchical method can provide reasonable estimation and meanwhile keep the mathematical derivations in the paper directly adaptable in various scenarios. Liu, Liu, and Xie (2015) reported a method to conduct meta-analysis of independent studies via a confidence density (CD) approach. In the paper, we employ a revised version of the confidence density to estimate the parameters $\boldsymbol{\delta}$ in $f_{\mathrm{acc}}(\boldsymbol{x}_i)$, and we denote $f_{\mathrm{acc}}(\boldsymbol{x}_i)$ by $f_{\mathrm{acc}}(\boldsymbol{x}_i;\boldsymbol{\delta})$ in the following context. As discussed in Appendix A, the form of $f_{\mathrm{acc}}(\boldsymbol{x}_i;\boldsymbol{\delta})$ depends on certain known physical mechanisms. To ensure $f_{\mathrm{acc}}(\boldsymbol{x}_i;\boldsymbol{\delta})$ to be greater than zero, we take natural logarithm on it. Due to the invariance property of MLE, the MLE of $\log f_{\mathrm{acc}}(\boldsymbol{x}_i;\boldsymbol{\delta})$ is readily given by $\log\hat{f}_{\mathrm{acc}}(\boldsymbol{x}_i;\boldsymbol{\delta})$. Three times differentiable mapping functions $\boldsymbol{\mathcal{M}} = (\mathcal{M}_1,\ldots,\mathcal{M}_M)'$ is used to link $\log f_{\mathrm{acc}}(\boldsymbol{x}_i;\boldsymbol{\delta})$ and the unknown parameter vector $\boldsymbol{\delta}$:

$$\log f_{\mathrm{acc}}(\boldsymbol{x}_i;\boldsymbol{\delta}) = \mathcal{M}_i(\boldsymbol{\delta}). \tag{17}$$

Further, following Xie and Singh (2013), we can construct a CD for $\log f_{\mathrm{acc}}(\boldsymbol{x}_i;\boldsymbol{\delta}), i = 1,\ldots,M$, which is a multivariate normal (MN) distribution with mean $\left(\log\hat{f}_{\mathrm{acc}}(\boldsymbol{x}_1;\boldsymbol{\delta}),\ldots,\log\hat{f}_{\mathrm{acc}}(\boldsymbol{x}_M;\boldsymbol{\delta})\right)'$ and covariance matrix $\{[\mathbf{I}(\hat{\boldsymbol{\theta}})]^{-1}\oslash\hat{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}}^T\}_{1:M}$, where $\{\cdot\}_{1:M}$ denotes the $M\times M$ square matrix partitioned from the upper left in the original matrix and $\oslash$ denotes an element-wise division. Note that $\boldsymbol{\delta}$ is only involved in $f_{\mathrm{acc}}(\boldsymbol{x}_i;\boldsymbol{\delta})$, thus we only use the first M rows and columns in the original covariance matrix in the model described in Section 3.4, and the parameters ω and σ^2 are assumed known in the subsection. The covariance matrix is approximated by the delta methods. For notational convenience,

we let $\hat{\mathbf{V}} = \{[\mathbf{I}(\hat{\boldsymbol{\theta}})]^{-1} \oslash \hat{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}}^T\}_{1:M}$. Additionally, we let $\mathbf{V}$ be the covariance matrix under true parameters. Then the CD for $\boldsymbol{\delta}$ is given in a form of MN distribution by

$$h(\boldsymbol{\delta}) = (2\pi)^{-M/2} \det(\hat{\mathbf{V}})^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{a}^T\hat{\mathbf{V}}^{-1}\boldsymbol{a}\right), \tag{18}$$

where $\boldsymbol{a}$ is a M-dimensional column vector with each element given by

$$a_i = \log f_{\text{acc}}(\boldsymbol{x}_i; \boldsymbol{\delta}) - \log \hat{f}_{\text{acc}}(\boldsymbol{x}_i; \boldsymbol{\delta}), \quad i = 1, \ldots, M.$$

By maximizing (18), we obtain the point estimator of $\boldsymbol{\delta}$ under CD, which we denote by $\hat{\boldsymbol{\delta}}_{\text{CD}}$:

$$\hat{\boldsymbol{\delta}}_{\text{CD}} = \arg\max_{\boldsymbol{\delta}} h(\boldsymbol{\delta}). \tag{19}$$

The reason for using the CD estimation is twofold. First, as mentioned previously, it facilitates the derivation of closed-form estimators under the EM framework. Second, the efficiency of estimation is not compromised by the hierarchical operations by CD estimation. The following analyses are presented to justify the latter statement. Under a conventional MLE framework, if $\boldsymbol{\delta}$ is directly used to maximize the likelihood function in (6), we can obtain the MLE of $\boldsymbol{\delta}$ given by

$$\hat{\boldsymbol{\delta}}_{\text{DIR}} = \arg\max_{\boldsymbol{\delta}} L(\boldsymbol{\delta}; \boldsymbol{y}^-, \boldsymbol{z}). \tag{20}$$

For notational convenience, we let $\boldsymbol{\zeta} = \left(\log f_{\text{acc}}(\boldsymbol{x}_1; \boldsymbol{\delta}), \ldots, \log f_{\text{acc}}(\boldsymbol{x}_M; \boldsymbol{\delta})\right)'$ and use $L(\boldsymbol{\delta})$ and $L(\boldsymbol{\zeta})$ to respectively represent the likelihood functions $L(\boldsymbol{\delta}; \boldsymbol{y}^-, \boldsymbol{z})$ and $L(\boldsymbol{\zeta}; \boldsymbol{y}^-, \boldsymbol{z})$, where $L(\boldsymbol{\zeta}; \boldsymbol{y}^-, \boldsymbol{z})$ denotes the likelihood under $\boldsymbol{\zeta}$. Thus, we have $\boldsymbol{\zeta} = \mathcal{M}(\boldsymbol{\delta})$. Moreover, let $n = \sum_{i=1}^{M}\sum_{j=1}^{N_i} O_{ij}$ be the total number of observations in the test.

Lemma 1. *As $n \to \infty$, the direct estimator $\hat{\boldsymbol{\delta}}_{\text{DIR}}$ obtained from (20) is consistent and normally distributed. Specifically,*

$$n^{1/2}\left(\hat{\boldsymbol{\delta}}_{\text{DIR}} - \boldsymbol{\delta}\right) \xrightarrow{d} \text{MN}\left(\mathbf{0}, \left[\mathbf{J}(\boldsymbol{\delta})^T\tilde{\mathbf{I}}(\boldsymbol{\zeta})\mathbf{J}(\boldsymbol{\delta})\right]^{-1}\right), \tag{21}$$

where $\tilde{\mathbf{I}}(\boldsymbol{\zeta}) = \mathbf{V}^{-1}$ and $\mathbf{J}(\boldsymbol{\delta}) = \partial\mathcal{M}(\boldsymbol{\delta})/\partial\boldsymbol{\delta}$ is the Jacobian of $\mathcal{M}$ with respect to $\boldsymbol{\delta}$.

The proofs of the lemma and the following results are provided in Appendix D. Lemma 1 implies the asymptotic properties of the direct estimators $\hat{\boldsymbol{\delta}}_{\text{DIR}}$ via the delta method. Next, we focus on the CD estimators $\hat{\boldsymbol{\delta}}_{\text{CD}}$ with the following lemma.

Lemma 2. *The first-order derivative of the log-confidence density function $\log h(\boldsymbol{\zeta}) \equiv \log h(\boldsymbol{\delta})$ with respect to $\boldsymbol{\zeta}$, is asymptotically equivalent to the score function $s(\boldsymbol{\delta}) = \partial \log L(\boldsymbol{\delta})/\partial\boldsymbol{\delta}$.*

According to Lemma 2, the CD estimator $\hat{\boldsymbol{\delta}}_{\text{CD}}$ and direct estimator $\hat{\boldsymbol{\delta}}_{\text{DIR}}$ share exactly identical asymptotic properties. Therefore, we can introduce Theorem 1 in analogy to Lemma 1 immediately following Lemma 2.

Theorem 1. *As $n \to \infty$, the CD estimator $\hat{\boldsymbol{\delta}}_{\text{CD}}$ is consistent and normally distributed. Specifically,*

$$n^{1/2}\left(\hat{\boldsymbol{\delta}}_{\text{CD}} - \boldsymbol{\delta}\right) \xrightarrow{d} \text{MN}\left(\mathbf{0}, \left[\mathbf{J}(\boldsymbol{\delta})^T\tilde{\mathbf{I}}(\boldsymbol{\zeta})\mathbf{J}(\boldsymbol{\delta})\right]^{-1}\right), \tag{22}$$

where $\tilde{\mathbf{I}}(\boldsymbol{\zeta}) = \mathbf{V}^{-1}$ and $\mathbf{J}(\boldsymbol{\delta}) = \partial\mathcal{M}(\boldsymbol{\delta})/\partial\boldsymbol{\delta}$.

The statements in the theorem show that the CD approach is asymptotically as efficient as the direct estimation approach. To quantify the uncertainty in the CD estimators, we put forward Corollary 1 based on Lemma 2 and Theorem 1.

Corollary 1. *The covariance matrix of $n^{1/2}\left(\hat{\boldsymbol{\delta}}_{\text{CD}} - \boldsymbol{\delta}\right)$ can be consistently estimated by $n\hat{\Sigma}_{\text{CD}}$, where*

$$\hat{\Sigma}_{\text{CD}} = \left[-\frac{\partial^2}{\partial\boldsymbol{\delta}\partial\boldsymbol{\delta}^T}\log h\left(\hat{\boldsymbol{\delta}}_{\text{CD}}\right)\right]^{-1}, \tag{23}$$

Remark. On the one hand, the confidence density based methods can address the aforementioned problem to hierarchically estimate parameters without imposing difficulties in the EM algorithm. On the other hand, as was advised in Liu et al. (2015), the information from independent studies can be well combined via the confidence density. In the problem we have been focusing on in the paper, the proposed hierarchical analysis can be applied to degradation tests under different acceleration functions and finally yield integrated results for the parameters of interest, which could be a subset or transformation of parameters that are already involved in those tests.

In light of the previous analyses on the hierarchical estimation, we have justified the efficiency of the CD approach. The estimation of $\boldsymbol{\delta}$ can be readily obtained via (19). Due to the nonlinear and non-additive properties of the Peck model, it is not intuitive to compare the effects of environmental factors under the usage environments, which could be of great interest to reliability engineers and decision makers. The following proposition is given under the Peck model to compare the effects brought by a single-unit change in temperature and relative humidity.

Proposition 1. *(Intuitive comparison of effects under the Peck model) The baseline degradation rate at the reference environment $f_{\text{acc}}(\boldsymbol{x}_{\text{ref}}; \boldsymbol{\delta})$ under parameters $\boldsymbol{\delta} = (\delta_0, E_a, \delta_1)$ is more sensitive to temperature if*

$$\frac{E_a}{11605} \cdot \frac{1}{TK_{\text{ref}}^2} \cdot \frac{RH_{\text{ref}}}{\delta_1} > 1,$$

and is more sensitive to relative humidity otherwise.

The proposition can be straightforwardly justified by taking the first-order derivative on $f_{\text{acc}}(\boldsymbol{x}; \boldsymbol{\delta})$ with respect to TK and RH, thus the proof is omitted. The proposition will be used for illustration in Section 6.2. If acceleration models other than the Peck model are used, similar statements can also be entertained for effect comparison.

4. Uncertainty quantification of the estimated parameters

To quantify the uncertainties in the parameter estimators is a vital task to enable and justify the adoption of the estimators in knowledge creation and decision making. Compared to point estimation, interval estimation is usually preferable in real problems. In this section, we discuss the large-sample based method to construct confidence intervals for estimated parameters.

The assumptions of large-sample approximation are commonly utilized to provide asymptotic covariance matrix of parameter estimators from which confidence intervals are obtained. For complete datasets, a routine practice is to compute the Fisher information from the log-likelihood directly, the elements in the Fisher information matrix $\mathbf{I}(\boldsymbol{\theta})$ are given by

$$\left[\mathbf{I}(\boldsymbol{\theta})\right]_{ij} = E\left[-\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log L(\boldsymbol{\theta}|\boldsymbol{y}^-, \boldsymbol{z})|\boldsymbol{\theta}\right]. \tag{24}$$

Asymptotically, $\hat{\boldsymbol{\theta}}$ follows a MN distribution, i.e., $\hat{\boldsymbol{\theta}} \sim N(\boldsymbol{\theta}, \left[\mathbf{I}(\boldsymbol{\theta})\right]^{-1})$. Since $\boldsymbol{\theta}$ is also unknown, we can alternatively employ the observed information $\mathbf{I}(\hat{\boldsymbol{\theta}})$ to compute the asymptotic covariance matrix of $\boldsymbol{\theta}$.

For incomplete data sets with hidden observations as discussed in Section 3.4, we adopt the approach from Oakes (1999), where the Fisher information can be directly calculated via function Q.

Table 1
Experimental design $\boldsymbol{\pi}$ for the test under treatment 1–9 (numbered in parenthesis).

RH level	Temperature level			Total
	303.15 Kelvin (30 degree Celsius)	318.15 Kelvin (45 degree Celsius)	333.15 Kelvin (60 degree Celsius)	
60	(1) 16/49	(4) 8/49	(7) 4/49	4/7
75	(2) 8/49	(5) 4/49	(8) 2/49	2/7
90	(3) 4/49	(6) 2/49	(9) 1/49	1/7
Total	4/7	2/7	1/7	1

Table 2
Parameters as input to simulation study.

Model	Parameter	Value
Peck model	δ_0	1
	E_a	2×10^7
	δ_1	3
Increased degradation rate	ω	1
Diffusion parameter	σ^2	0.3
Degradation reduction (known)	u_0	2
	v_0	3

Accordingly, the observed Fisher information can be computed by

$$\mathbf{I}\left(\hat{\boldsymbol{\theta}}\right) = -\left[\frac{\partial^2 Q\left(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}\right)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T} + \frac{\partial^2 Q\left(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}\right)}{\partial\boldsymbol{\theta}\partial\hat{\boldsymbol{\theta}}^T}\right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}. \tag{25}$$

Note that the second term in the r.h.s. of the equation is viewed as the missing information due to the absence of $\boldsymbol{z}$. Likewise, the asymptotic confidence intervals for unknown parameters can be constructed by the Fisher information. The derivation of (25) requires manipulations based on Appendix B, and the analytical details of (25) are given in Appendix E. As a side note, for the parameter σ^2, the normal approximated confidence intervals are usually inappropriate. Alternatively,we build the confidence intervals based on $\log\sigma^2$ via the delta method. Further, the confidence intervals of $\log\sigma^2$ are transformed by an exponential operation to quantify the uncertainty in σ^2. More approaches for uncertainty quantification based on EM algorithm can be found in Louis (1982) and Meng and Rubin (1991).

5. Simulation study

To facilitate the simulation, we need to specify an experimental design of the degradation test. By letting $N = \sum_i N_i$ be the total test units and $\pi_i = N_i/N$ be the proportion of test units allocated to stress i, we suppose that $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_M)$ is a pre-specified experimental design for the simulation study. Without loss of generality, the test is assumed to allow for the change in temperature and humidity under the Peck model introduced in Appendix A. Additionally, three levels for each factor are specified and we follow a 4:2:1 allocation rule for each factor (Meeker & Escobar, 1998; Meeker & Hahn, 1977). To be specific, the test plan allocates more test units to lower stress levels to avoid overwhelming extrapolation. The detailed plan is shown in Table 1.

Parameters are set as shown in Table 2 for the purpose of illustration. As a side note, the reference temperature and relative humidity are fixed at $TK_{\text{ref}} = 293.15$ Kelvin (20 degree Celsius) and $RH_{\text{ref}} = 50\%$, respectively.

To explore the effect of sample sizes, we will show results under $N = 49,\ 98$ and 147 in the simulation studies. Test units are assumed to be inspected for 3 times. The following four sub-studies constitute this section to explore the effectiveness of the proposed model.

5.1. Estimates from the EM algorithm and confidence intervals

First, point estimates are obtained from the EM algorithm under 1000 simulation replicates for each sample size of interest. The convergence criterion is set as $\varepsilon = 0.001$ (1‰) for all parameters. In Table 3, the mean bias and root mean squared error (RMSE) are shown under $N = 49,\ 98$ and 147. By observing the results, we can imply that the EM algorithm can accurately estimate the unknown parameters, with rather low mean bias and RMSE even under moderate sample sizes. Moreover, the accuracy of the estimation enhances with the increase in sample size. Specifically, compared to other parameters, the estimation accuracy of σ^2 is relatively low but drastically improves over the sample size. A possible reason behind this is that the estimation of σ^2 involves both $E(z_{ij(k-1)}|\boldsymbol{y}^-, \boldsymbol{\theta}^{(n)})$ and $E(z^2_{ij(k-1)}|\boldsymbol{y}^-, \boldsymbol{\theta}^{(n)})$ as indicated in (15), where more uncertainty of hidden variables are brought into the estimators.

Further, with the extant point estimates, the confidence intervals are constructed via the method proposed in Section 4. Specifically, the $(1-\alpha)\times 100\%$ confidence interval is given by

$$\hat{\theta}_i \pm z_{\alpha/2}\left\{[\mathbf{I}(\hat{\boldsymbol{\theta}})]_{ii}\right\}^{1/2},$$

where $[\cdot]_{ii}$ denotes the ith diagonal element of a matrix and $z_{\alpha/2}$ is the $1-\alpha/2$ quantile of the standard normal distribution. In Table 4, the coverage probabilities and the average lengths are listed under the simulated datasets under three sample sizes. We compute the 95% confidence intervals under large-sample approximation. As the sample size increases, the coverage probability becomes closer to 0.95 and the average length is shortened. As seen, even under a sample of 49, the confidence intervals perform well and for most parameters over 90% of them can cover the true values. Again, influenced by the relatively large bias, the coverage probability of σ^2 is moderately lower under small sample sizes. Recall that we use the $\log\sigma^2$ to construct confidence intervals for σ^2. The trick is proven to benefit the performance. A supporting example is that under $N = 49$, we obtain a coverage probability of 0.841 comparing to 0.750 where σ^2 is directly used to quantify the uncertainty.

5.2. Hierarchical analysis

We now consider the estimation of $\boldsymbol{\delta}$ in the Peck model by means of the proposed hierarchical analysis. Likewise, the performance of point estimation and uncertainty quantification are listed in Table 5 and Table 6, respectively.

The results imply good performances with low mean bias and RMSE for the point estimators as well as coverage probabilities that are close enough to 0.95. It is worth noting that the hierarchical analysis consumes limited computational efforts. For example, point estimation together with uncertainty quantification under $N = 147$ only takes less than 1 seconds on a single Intel i5 core. If hierarchical analysis is not employed, the complexity of $f_{\text{acc}}(\cdot)$ will hinder the derivation of analytical results in the EM algorithm, which could introduce enormous computational burdens to the problem.

5.3. Sensitivity analysis with respect to the degradation reduction effect

As discussed in Section 3, the choice of u_0 and v_0 leans on the experience of engineers. Misspecification of the distribution for z_0 may occur and lead to higher bias of parameter estimation. Here, by assuming the true values of u_0 and v_0 to be 2 and 3, respectively, we change the assumed values to explore the influence of misspecification under sample size $N = 49$. Note that $E(z_0) = 0.4$

Table 3
Mean bias and RMSE of unknown parameters under $N = 49, 98$ and 147.

Parameter $\boldsymbol{\theta}$	True value	$N = 49$		$N = 98$		$N = 147$	
		Bias	RMSE	Bias	RMSE	Bias	RMSE
$f_{\text{acc}}(\boldsymbol{x}_1)$	2.104	0.011	0.264	−0.001	0.190	0.008	0.155
$f_{\text{acc}}(\boldsymbol{x}_2)$	4.110	0.010	0.283	0.005	0.202	0.004	0.162
$f_{\text{acc}}(\boldsymbol{x}_3)$	7.101	−0.005	0.344	0.004	0.250	0.006	0.200
$f_{\text{acc}}(\boldsymbol{x}_4)$	2.751	−0.001	0.288	0.003	0.206	0.001	0.173
$f_{\text{acc}}(\boldsymbol{x}_5)$	5.373	0.003	0.344	0.006	0.248	0.009	0.196
$f_{\text{acc}}(\boldsymbol{x}_6)$	9.284	−0.002	0.439	0.007	0.313	0.006	0.266
$f_{\text{acc}}(\boldsymbol{x}_7)$	3.511	0.013	0.343	0.003	0.247	0.011	0.204
$f_{\text{acc}}(\boldsymbol{x}_8)$	6.857	0.006	0.444	−0.006	0.316	0.012	0.256
$f_{\text{acc}}(\boldsymbol{x}_9)$	11.849	−0.001	0.647	−0.012	0.425	−0.004	0.358
ω	1.000	−0.009	0.178	−0.003	0.130	−0.008	0.106
σ^2	0.300	−0.044	0.075	−0.020	0.049	−0.009	0.037

Table 4
Coverage probability and average length of the 95% confidence intervals of $\boldsymbol{\theta}$ under $N = 49, 98$ and 147.

Parameter $\boldsymbol{\theta}$	$N = 49$		$N = 98$		$N = 147$	
	Cov. prob.	Avg. len.	Cov. prob.	Avg. len.	Cov. prob.	Avg. len.
$f_{\text{acc}}(\boldsymbol{x}_1)$	0.916	0.944	0.922	0.693	0.929	0.572
$f_{\text{acc}}(\boldsymbol{x}_2)$	0.915	1.032	0.930	0.758	0.944	0.627
$f_{\text{acc}}(\boldsymbol{x}_3)$	0.919	1.267	0.938	0.924	0.954	0.766
$f_{\text{acc}}(\boldsymbol{x}_4)$	0.918	1.044	0.933	0.766	0.948	0.633
$f_{\text{acc}}(\boldsymbol{x}_5)$	0.926	1.254	0.933	0.922	0.948	0.762
$f_{\text{acc}}(\boldsymbol{x}_6)$	0.927	1.665	0.951	1.215	0.949	1.002
$f_{\text{acc}}(\boldsymbol{x}_7)$	0.933	1.249	0.933	0.917	0.948	0.757
$f_{\text{acc}}(\boldsymbol{x}_8)$	0.927	1.647	0.946	1.201	0.941	0.993
$f_{\text{acc}}(\boldsymbol{x}_9)$	0.933	2.380	0.949	1.676	0.946	1.367
ω	0.913	0.638	0.925	0.465	0.937	0.383
σ^2	0.841	0.220	0.901	0.167	0.936	0.140

Table 5
Mean bias and RMSE of estimated $\boldsymbol{\delta}$ under $N = 49, 98$ and 147.

Parameter $\boldsymbol{\delta}$	True value	$N = 49$		$N = 98$		$N = 147$	
		Bias	RMSE	Bias	RMSE	Bias	RMSE
δ_0	1	0.050	0.176	0.025	0.112	0.017	0.092
E_a $(\times 10^7)$	2	−0.039	0.2481	−0.020	0.160	−0.012	0.138
δ_1	3	−0.081	0.315	−0.031	0.187	−0.027	0.171

Table 6
Coverage probability and average length of the 95% confidence intervals of $\boldsymbol{\delta}$ under $N = 49, 98$ and 147.

Parameter $\boldsymbol{\delta}$	$N = 49$		$N = 98$		$N = 147$	
	Cov. prob.	Avg. len.	Cov. prob.	Avg. len.	Cov. prob.	Avg. len.
δ_0	0.933	0.542	0.944	0.389	0.957	0.319
E_a $(\times 10^7)$	0.944	0.692	0.941	0.508	0.958	0.419
δ_1	0.905	0.686	0.938	0.506	0.948	0.417

and $\text{var}(z_0) = 0.04$ hold under true values. In Table 7, we list the bias and RMSE under four settings of misspecfied values of u_0 and v_0 as follows:

- High Mean: $u_0 = 3,\ v_0 = 2$, i.e., $E(z_0) = 0.6, \text{var}(z_0) = 0.04$;
- Low Mean: $u_0 = 0.6,\ v_0 = 2.4$, i.e., $E(z_0) = 0.2, \text{var}(z_0) = 0.04$;
- High Variance: $u_0 = 1.2,\ v_0 = 1.8$, i.e., $E(z_0) = 0.4, \text{var}(z_0) = 0.06$;
- Low Variance: $u_0 = 9.2,\ v_0 = 13.8$, i.e., $E(z_0) = 0.4, \text{var}(z_0) = 0.01$.

As seen from the result, the misspecification of mean of z_0 brings considerable bias to the estimators of $f_{\text{acc}}(\boldsymbol{x}_i)$ and ω, while the estimation of σ^2 suffers more when the variance is misspecified. For extrapolating analysis based on the ADT data, $f_{\text{acc}}(\boldsymbol{x}_i)$ and ω play more important roles. For this purpose, engineers should focus on evaluating the mean effect of degradation reduction for a better estimation accuracy. To explore the influence of misspecification on the estimation of parameters in the Peck model, we carry out the hierarchical methods to estimate $\boldsymbol{\delta}$ and show the mean bias and RMSE in Table 8. The estimated $\boldsymbol{\delta}$ suffers a considerable bias when the mean of z_0 is misspecified, while the influence of misspecified variance exerts relatively smaller influence on the estimation accuracy.

6. Case study

6.1. Data from Schneider Electric

Schneider Electric conducted a degradation test for a type of electrical distribution device. A total of 104 test units underwent the test under four different settings of temperature and humidity. Specifically, two levels of temperature (313.15 Kelvin and 333.15

Table 7
Mean bias and RMSE of unknown parameters $\boldsymbol{\theta}$ under misspecified u_0 and v_0.

Parameter $\boldsymbol{\theta}$	High mean		Low mean		High variance		Low variance	
	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
$f_{\text{acc}}(\boldsymbol{x}_1)$	−0.662	0.712	0.644	0.704	0.065	0.273	−0.008	0.287
$f_{\text{acc}}(\boldsymbol{x}_2)$	−0.566	0.642	0.539	0.625	0.052	0.305	−0.008	0.310
$f_{\text{acc}}(\boldsymbol{x}_3)$	−0.531	0.640	0.465	0.595	0.032	0.352	0.009	0.397
$f_{\text{acc}}(\boldsymbol{x}_4)$	−0.614	0.676	0.600	0.674	0.063	0.294	−0.010	0.312
$f_{\text{acc}}(\boldsymbol{x}_5)$	−0.535	0.639	0.487	0.608	0.039	0.346	−0.008	0.366
$f_{\text{acc}}(\boldsymbol{x}_6)$	−0.498	0.700	0.458	0.684	0.041	0.486	0.007	0.556
$f_{\text{acc}}(\boldsymbol{x}_7)$	−0.574	0.671	0.550	0.662	0.054	0.352	−0.006	0.377
$f_{\text{acc}}(\boldsymbol{x}_8)$	−0.529	0.699	0.441	0.654	0.025	0.450	−0.005	0.528
$f_{\text{acc}}(\boldsymbol{x}_9)$	−0.471	0.835	0.394	0.811	−0.006	0.647	0.001	0.762
ω	0.755	0.776	−0.733	0.761	−0.058	0.192	0.008	0.191
σ^2	−0.025	0.071	0.005	0.071	−0.078	0.096	0.213	0.239

Table 8
Mean bias and RMSE of unknown parameters $\boldsymbol{\delta}$ under misspecified u_0 and v_0.

Parameter $\boldsymbol{\delta}$	High mean		Low mean		High variance		High variance	
	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
δ_0	−0.181	0.221	0.178	0.202	0.096	0.176	0.035	0.162
E_a (10^7)	0.152	0.262	−0.146	0.233	−0.052	0.187	−0.009	0.221
δ_1	0.252	0.328	−0.234	0.321	−0.105	0.210	−0.023	0.208

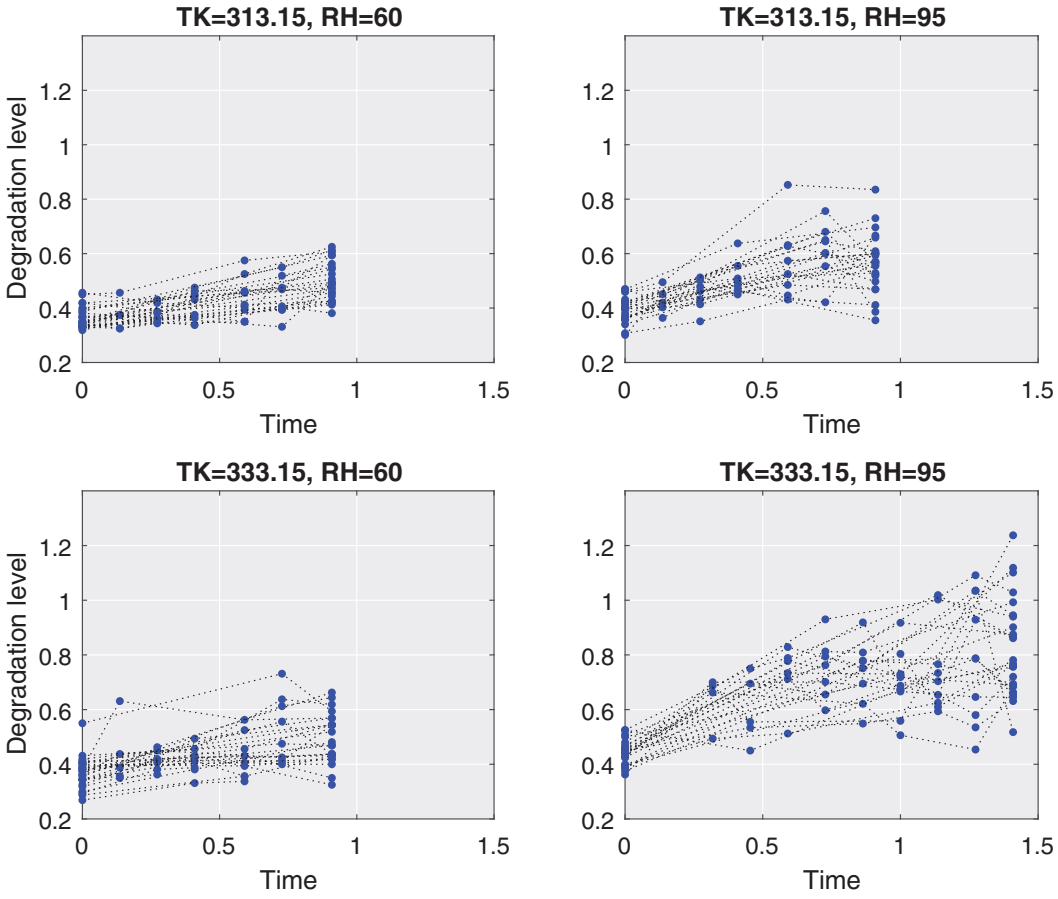

Fig. 2. Degradation test data under 4 stress levels.

Kelvin) and two levels of relative humidity (60% and 90%) are considered. Note that 313.15 Kelvin and 333.15 Kelvin are equivalent to 40 degree Celsius and 60 degree Celsius, respectively. The test chambers provide generally higher stresses than the usage condition of the device, thus the test can be regarded as an ADT. The test data under each stress conditions are plotted in Fig. 2. It can be observed that the observation epochs (τ_{ijk}) and number of observations (O_{ijk}) vary across the test units in the test. The specifications of the test are shown in Table 9. Based on the discussions with engineers from Schneider, we propose to model the degradation reduction effect by a beta distribution with parameters $u_0 = 1$ and $v_0 = 3$.

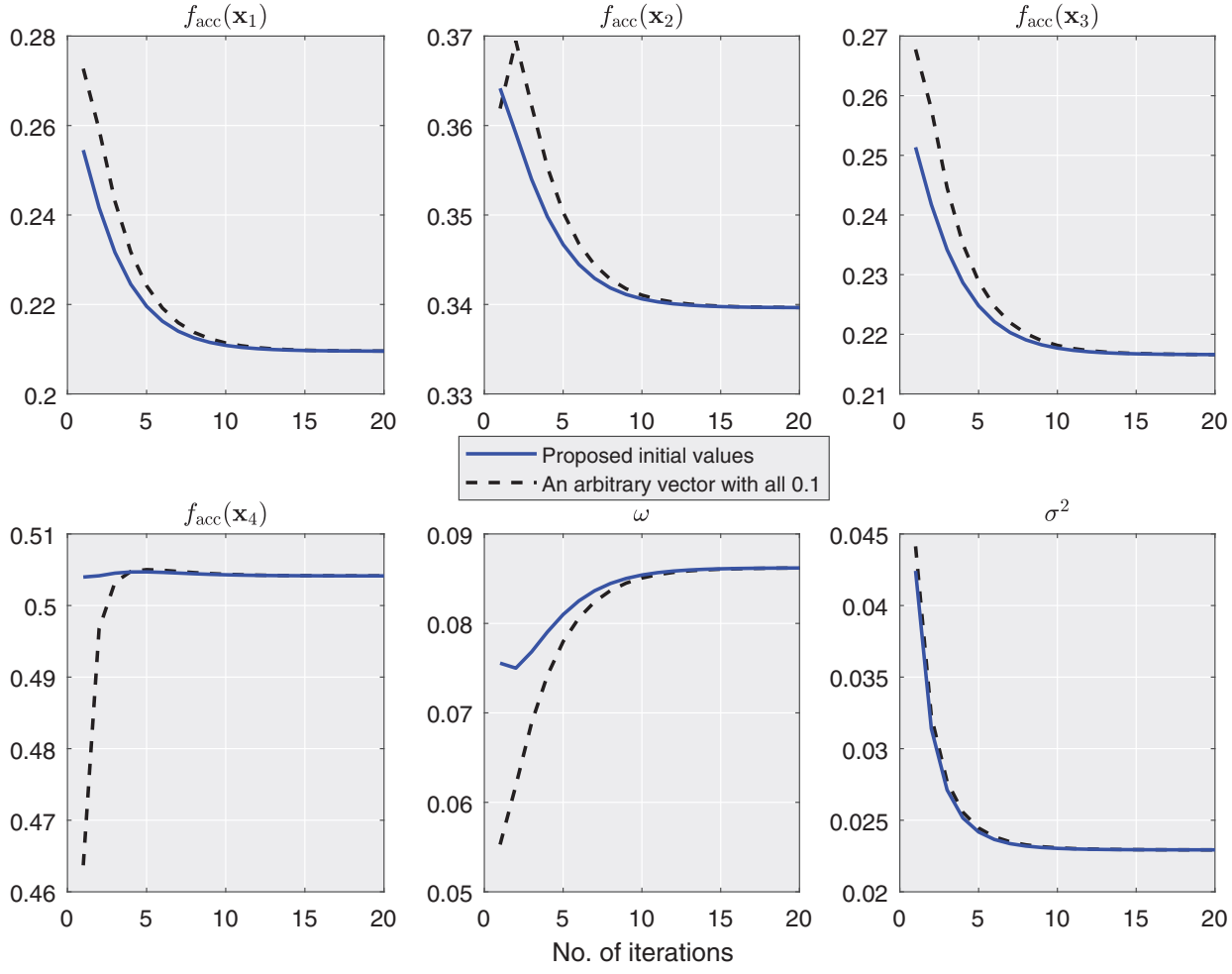

Fig. 3. Convergence of the EM algorithm.

Table 9
Overview of the test from Schneider Electric.

i	TK_i (in absolute Kelvin)	RH_i (in %)	# of test units	Duration
1	313.15	60	28	0.9091
2	313.15	95	24	0.9091
3	333.15	60	28	0.9091
4	333.15	95	24	1.4091

6.2. Model estimation

We proceed to the parameter estimation and uncertainty quantification. First, we are interested in how the parameter estimators converge through the iterations. In Fig. 3, the estimates at each iteration are plotted. We compare results under proposed initial estimates (see Section 3.5) against initial values of setting 0.1 for all parameters. As shown, the proposed initial values can speed up the convergence of EM algorithm by selecting the initial values close to the MLE. Table 10 reports the MLE, asymptotic standard deviation and 95% confidence intervals of $\boldsymbol{\theta}$ (LB and UB represents the lower bound and upper bound of a confidence interval, respectively). We can observe that the lower bound 95% confidence interval of ω is positive, which implies that the effect of degradation rate increase is significant.

Next, the hierarchical analysis is carried out and the results are shown in Table 11. It can be seen that both $\hat{E}_a$ and $\hat{\delta}_1$ are significantly positive. Thus, temperature and relative humidity both exert accelerating influences on the degradation.

At the current stage, it is not easy to tell how the environmental factors influence the degradation rate in intuitive senses, thus we utilize Proposition 1 evaluated at the CD estimate of $\boldsymbol{\delta}$ as follows:

$$\frac{\hat{E}_a}{11,605}\cdot\frac{1}{TK_{\mathrm{ref}}^2}\cdot\frac{RH_{\mathrm{ref}}}{\hat{\delta}_1}=0.6189,$$

which implies that under the usage environments, the baseline degradation rate is more sensitive to the change in relative humidity. From the practical point of view, it is of more significance to prevent the relative humidity from becoming overwhelmingly high.

To demonstrate the advantages of the proposed model, we perform a comparison with the conventional ADT model that overlooks the inspection effects. Specifically, the conventional model assumes linear Wiener degradation paths. Table 12 lists parameter estimates as well as the Bayesian information criterion (BIC) values under various combinations of assumed u and v as well as under the conventional Wiener degradation model. We can observe a remarkable improvement by considering the inspection effects. Under the BIC, the advantage of the proposed model is overwhelming. Thus there is no further need for the comparisons under criteria with smaller penalty term of additional parameters, such as the Akaike information criterion (AIC). Specifically, we can observe that the estimates of σ^2 under the proposed models (around 0.02 to 0.03) are considerably smaller than that under the conventional model (0.052). This implies that the proposed model can capture more uncertainty systematically via the inspection effects, which is ignored by the conventional model.

Table 10
MLE and confidence intervals of unknown parameters.

	Parameters					
	$f_{acc}(\boldsymbol{x}_1)$	$f_{acc}(\boldsymbol{x}_2)$	$f_{acc}(\boldsymbol{x}_3)$	$f_{acc}(\boldsymbol{x}_4)$	ω	σ^2
MLE	0.2104	0.3403	0.2173	0.5042	0.0857	0.0230
Asymptotic std	0.0573	0.0625	0.0569	0.0583	0.0264	0.0031
LB of 95 % CI	0.0981	0.2177	0.1057	0.3899	0.0340	0.0168
UB of 95% CI	0.3228	0.4629	0.3289	0.6185	0.1373	0.0292

Table 11
CD estimates and confidence intervals of parameters in $f_{acc}(\cdot)$.

	Parameters		
	δ_0	E_a $(\times 10^7)$	δ_1
CD estimate	0.1014	1.8154	1.4721
Asymptotic std	0.0394	0.8454	0.3812
LB of 95% CI	0.0214	0.1584	0.7251
UB of 95% CI	0.1787	3.4723	2.2191

6.3. Analyses of reliability characteristics

The ultimate goal of degradation data modeling is to better understand the reliability and support related decision making. Towards this end, reliability engineers usually make use of the results of the estimated model to infer reliability characteristics of interest. Some commonly used characteristics include mean time to failure (MTTF) and life quantiles. By taking into account both accelerated environmental factors and inspection effects, the proposed methods can generate some interesting results that differ from those under conventional approaches.

First, we are curious about the reliability under different environmental conditions, which is one of the main purposes of the ADT. Four combinations of temperature and relative humidity are considered. An appealing property of the Wiener degradation model is that the lifetime follows an inverse Gaussian distribution, under which the reliability function is analytical. By assuming the threshold to be 0.5, the reliability curves and 80% pointwise confidence bands subject to estimation uncertainty are shown in Fig. 4. High temperature ($TK = 303.15$ Kelvin or 30 degree Celsius) or high humidity ($RH = 70\%$) shortens the lifetime. Further, the ef-

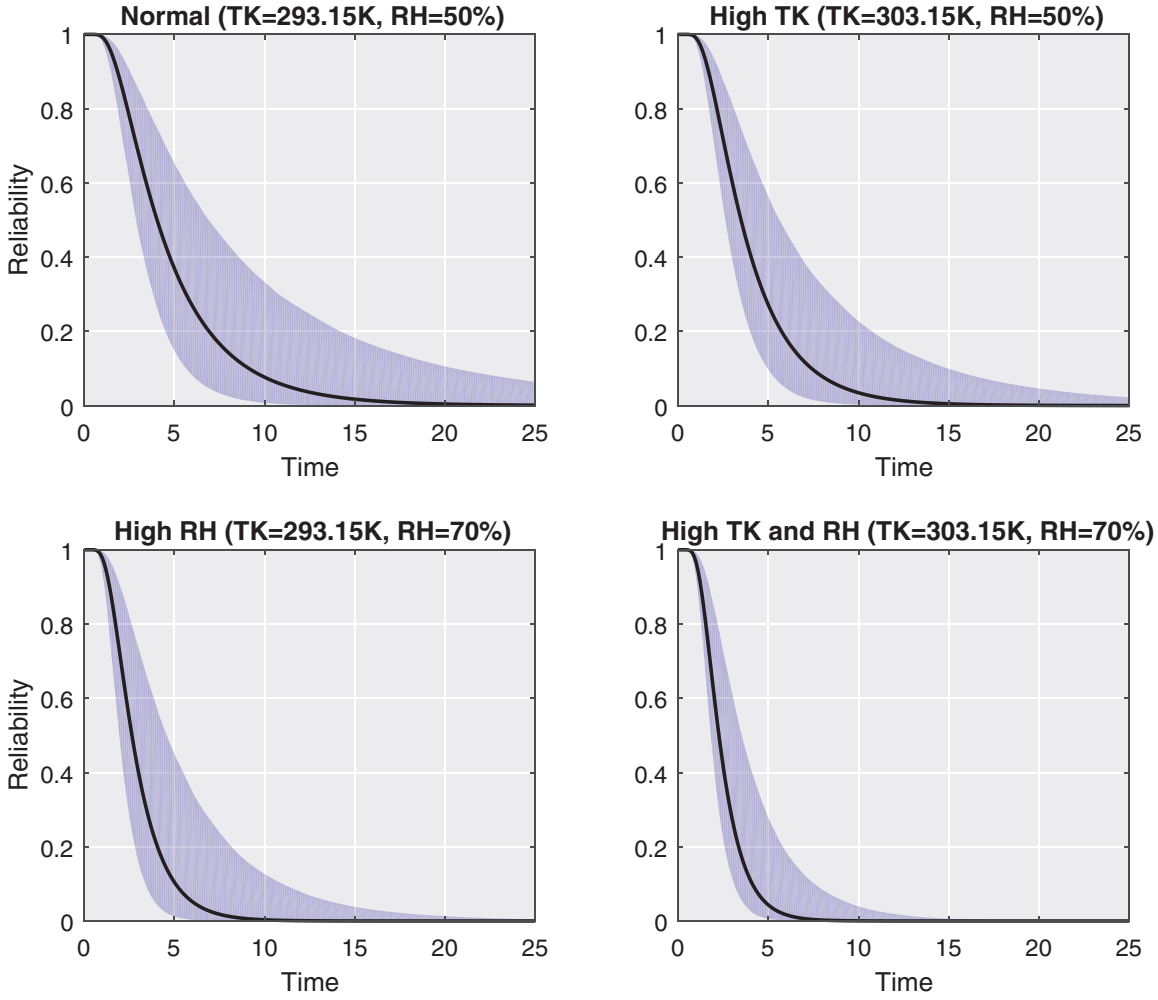

Fig. 4. Reliability curves and pointwise 80% confidence bands under different environmental conditions with no inspection.

Table 12
Performance comparison between the proposed model and the conventional model.

Model under (u, v)	Parameters						BIC
	$f_{\text{acc}}(\boldsymbol{x}_1)$	$f_{\text{acc}}(\boldsymbol{x}_2)$	$f_{\text{acc}}(\boldsymbol{x}_3)$	$f_{\text{acc}}(\boldsymbol{x}_4)$	ω	σ^2	
(0.5,0.5)	0.194	0.368	0.257	0.570	0.187	0.022	−521
(0.5,1)	0.190	0.331	0.211	0.525	0.102	0.023	−511
(0.5,3.5)	0.197	0.300	0.191	0.455	0.028	0.026	−480
(1,0.5)	0.353	0.552	0.386	0.708	0.216	0.028	−461
(1,1)	0.255	0.427	0.290	0.595	0.183	0.020	−543
(1,3)	**0.210**	**0.340**	**0.217**	**0.504**	**0.086**	**0.023**	**−512**
(2,1)	0.350	0.532	0.373	0.694	0.245	0.040	−363
(2,2)	0.289	0.452	0.308	0.609	0.186	0.028	−466
(2,8)	0.227	0.339	0.224	0.485	0.064	0.026	−478
Conventional	**0.145**	**0.188**	**0.124**	**0.272**	**NA**	**0.052**	**−299**

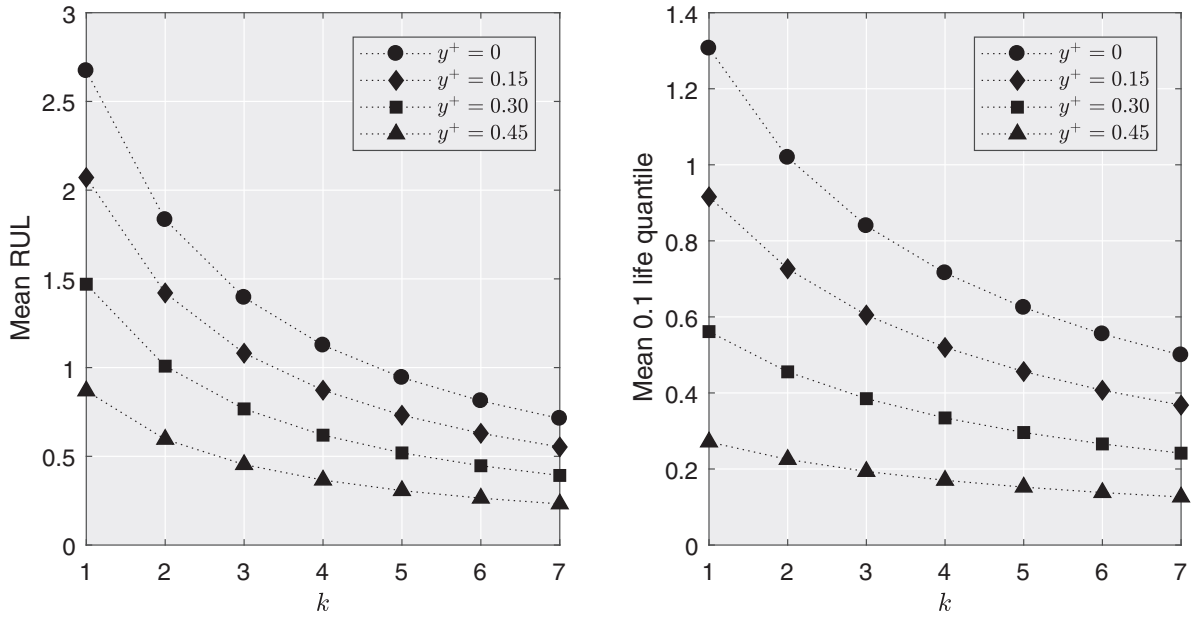

Fig. 5. ML estimated mean and 0.1 quantile of RUL under different k and y^+.

fect of relative humidity seems more significant, which validates the statement from Proposition 1. Practically, managers should keep the device working in cool and dry environments to prolong its lifetime. Additionally, due to the extrapolation of factors in ADTs, there exists considerable variability in the reliability curves, especially under lower *TK* and *RH*, of which decision makers need to beware. More conservative decisions are usually suggested in the presence of larger model uncertainty.

As the company's another objective is to predict the remaining useful life (RUL) for devices of different age and inspection history, under the estimated model, we plot the mean RUL and 0.1 RUL quantile under different combinations of current degradation level (y^+) and historical times of inspection (k) in Fig. 5. As is in accordance with intuitions, a higher y^+ or k leads to smaller RUL. To be more specific, the mean RUL decreases in proportion to the increase in y^+, while the influence of k is more significant when it is smaller. The behavior of 0.1 RUL quantile is similar except a faster decrease in the quantile when y^+ becomes higher. It is noteworthy that, in the presence of inspections, the reliability function in a general form over a time horizon from zero to infinity is onerous to derive due to the randomness in both degradation levels and the effect of inspections. Decision makers can employ the RUL distribution with given inspected degradation levels, which is tractable, to characterize the reliability characteristics and support future decisions.

For a better exposition of degradation paths under different inspection intervals, we simulate and plot the expected degradation path and confidence bands in Fig. 6. We set the same scale for x- and y-axes for comparison. Obviously, more frequent inspections lead to faster degradation despite the existence of degradation reduction effect. Interestingly, we observe that the uncertainty in the degradation paths is the smallest when $\Delta\tau = 0.1$. It is probably a result of tradeoff between the uncertainties in the two types of inspection effects. For the studied device, it may not be wise to inspect the system too frequently. Nevertheless, the decision maker should consider the risk of unexpected failures under limited information.

7. Concluding remarks

We focus on an atypical degradation modeling problem motivated by a real example. As a first attempt to deal with partially observable data with manifold inspection effects, the study utilizes the EM algorithm to iteratively derive the parameter estimators. Further, a hierarchical confidence density is used to accurately estimate the acceleration model. Uncertainty quantification is carried out via large-sample approximation. The manipulation of the proposed model can be widely applied to different types of ADT data. Simulation studies validate the proposed methods and show satisfactory estimating performance under different sample sizes. The

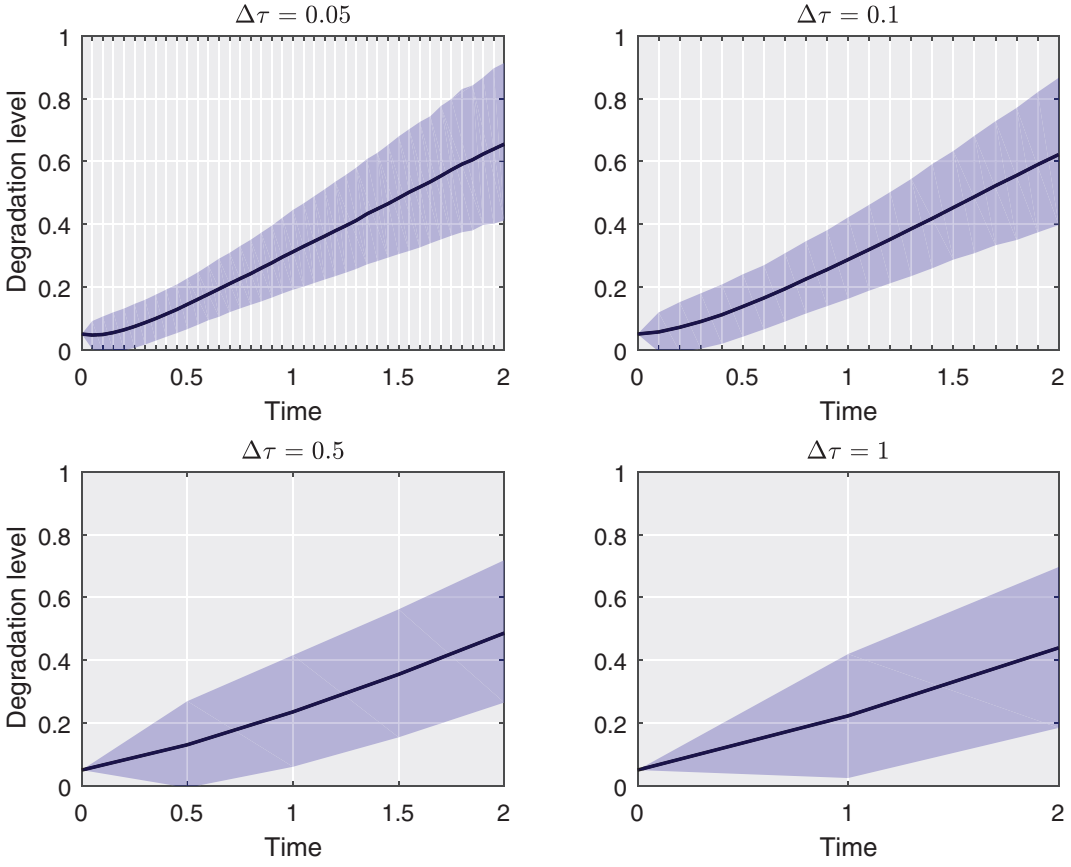

Fig. 6. Expected degradation paths and pointwise 90% confidence bands under different inspection intervals.

computational burden is significantly reduced by the hierarchical analysis. A case study from Schneider Electric demonstrates the advantage of the model over conventional degradation models. A series of reliability analyses based on the estimation results are provided to explore managerial insights from practical perspectives.

The involvement of inspection effects can be further investigated for reliability modeling and management. Bayesian methods can be explored to characterize the hidden inspection effects. Dynamic approaches to system inspection and maintenance optimization based on the current model are of interest to effectively manage industrial systems. As is mentioned in Section 3.6, it is also useful to combine the information from different tests, which puts forward some opportunities for future research. For more detailed extensions to solve real industrial problems, relevant discussions of other degradation models such as the gamma processes and inverse Gaussian processes and more models for inspection effects can be further explored.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under grant 72002149, 71971181, 72032005 and 71802145.

Appendix A. Acceleration mechanism and link functions

The appendix discusses the acceleration mechanism under two common experimental factors: temperature and humidity. Note that we only consider the parametric form of link function $f_{\text{acc}}(\boldsymbol{x}_i)$ in (3) here. To establish parametric models for $f_{\text{acc}}(\boldsymbol{x}_i)$, some domain knowledge on the physical failure mechanism is required. Otherwise, it is suggested to employ nonparametric extrapolating methods to predict the degradation rate at use conditions. By denoting the parametric form of $f_{\text{acc}}(\boldsymbol{x}_i)$ by $f_{\text{acc}}(\boldsymbol{x}_i;\boldsymbol{\delta})$, a Peck model is used for the joint modeling of temperature and relative humidity (Peck, 1986). To be specific,

$$f_{\text{acc}}(\boldsymbol{x}_i;\boldsymbol{\delta}) = \delta_0 \exp\left(\frac{E_a}{11605}\left(\frac{1}{TK_{\text{ref}}} - \frac{1}{TK_i}\right)\right)\left(\frac{RH_i}{RH_{\text{ref}}}\right)^{\delta_1}, \tag{A.1}$$

where $\boldsymbol{x}_i = (TK_i, RH_i)$ is the vector of absolute temperature and relative humidity at stress level i, and $x_{\text{ref}} = (TK_{\text{ref}}, RH_{\text{ref}})$ is covariate vector at reference stress level, which is often set at the covariate values under the use condition. The unknown parameters are $\boldsymbol{\delta} = (\delta_0, E_a, \delta_1)^T$, where E_a is the activation energy that is constant to a material and δ_1 characterizes the effect of humidity.

Appendix B. Density function of $z_{ij(k-1)}$

Given observable $\boldsymbol{y}^-$ and $\boldsymbol{\theta}^{(n)}$, we have

$$f_{z_{ij(k-1)}}\left(z|\boldsymbol{y}^-,\boldsymbol{\theta}^{(n)}\right) = \frac{f\left(\boldsymbol{y}^-;\boldsymbol{\theta}^{(n)},z\right)f\left(z;\boldsymbol{\theta}^{(n)}\right)}{f\left(\boldsymbol{y}^-;\boldsymbol{\theta}^{(n)}\right)} = \frac{\phi\left(\frac{y^-_{ijk}-y^-_{ij(k-1)}(1-z)-\mu^{(n)}_{ijk}\Delta\tau_{ijk}}{\sigma\sqrt{\Delta\tau_{ijk}}}\right)f_{\text{Beta}}(z;u_0,v_0)}{\int_0^1\left(\frac{y^-_{ijk}-y^-_{ij(k-1)}(1-x)-\mu^{(n)}_{ijk}\Delta\tau_{ijk}}{\sigma\sqrt{\Delta\tau_{ijk}}}\right)f_{\text{Beta}}(x;u_0,v_0)\mathrm{d}x}.$$

To facilitate further analysis in Appendix E, the equation can be elaborated into the following form:

$$f_{z_{ij(k-1)}}\left(z|\boldsymbol{y}^-,\boldsymbol{\theta}^{(n)}\right) = \frac{\mathcal{F}(z)f_{\text{Beta}}(z;u_0,v_0)}{\int_0^1\mathcal{F}(z)f_{\text{Beta}}(x;u_0,v_0)\mathrm{d}x}, \tag{B.1}$$

where

$$\mathcal{F}(z)=\exp\left[-\frac{\left(y^{-}_{ij(k-1)}z\right)^2+2y^{-}_{ij(k-1)}z\left(y^{-}_{ijk}-y^{-}_{ij(k-1)}-\mu^{(n)}_{ijk}\Delta\tau_{ijk}\right)}{2\left(\sigma^2\right)^{(n)}\Delta\tau_{ijk}}\right]. \tag{B.2}$$

Moreover, the first-order derivatives of f with respect to $\boldsymbol{\theta}^{(n)}$ are given by

$$\frac{\partial f_{z_{ij(k-1)}}\left(z|\boldsymbol{y}^{-},\boldsymbol{\theta}^{(n)}\right)}{\partial\theta_i^{(n)}}=\frac{\mathcal{F}'_{\theta_i}(z)f_{\text{Beta}}(z;u_0,v_0)\int_0^1\mathcal{F}(x)f_{\text{Beta}}(x;u_0,v_0)\mathrm{d}x}{\left[\int_0^1\mathcal{F}(x)f_{\text{Beta}}(x;u_0,v_0)\mathrm{d}x\right]^2}-\frac{\mathcal{F}(z)f_{\text{Beta}}(z;u_0,v_0)\int_0^1\mathcal{F}'_{\theta_i}(x)f_{\text{Beta}}(x;u_0,v_0)\mathrm{d}x}{\left[\int_0^1\mathcal{F}(z)f_{\text{Beta}}(x;u_0,v_0)\mathrm{d}x\right]^2}, \tag{B.3}$$

where $\mathcal{F}'_{\theta_i}(z)$'s are calculated by

$$\begin{aligned}
\mathcal{F}'_{f_{\text{acc}}(\boldsymbol{x}_i)}(z)&=\frac{\mathcal{F}y^{-}_{ij(k-1)}z}{\left(\sigma^2\right)^{(n)}},\\
\mathcal{F}'_{\omega}(z)&=\frac{\mathcal{F}y^{-}_{ij(k-1)}zk}{\left(\sigma^2\right)^{(n)}},\\
\mathcal{F}'_{\sigma^2}(z)&=\mathcal{F}\frac{\left(y^{-}_{ij(k-1)}z\right)^2+2y^{-}_{ij(k-1)}z\left(y^{-}_{ijk}-y^{-}_{ij(k-1)}-\mu^{(n)}_{ijk}\Delta\tau_{ijk}\right)}{2\left[\left(\sigma^2\right)^{(n)}\right]^2\Delta\tau_{ijk}}.
\end{aligned} \tag{B.4}$$

Appendix C. Expression for initial guesses of starting points

Following the analytical results in Section 3.4, we can calculate the initial values of parameter estimates by the following equations:

$$\begin{cases}
\omega^{(0)}=\frac{D-\sum_{i=1}^{M}A_iF_i}{C-\sum_{i=1}^{M}B_iF_i},\\
f^{(0)}_{\text{acc}}(\boldsymbol{x}_i)=A_i-B_i\omega^{(0)},\\
\left(\sigma^2\right)^{(0)}\sum_{i=1}^{M}\sum_{j=1}^{N_i}O_{ij}=G,
\end{cases}$$

where

$$\begin{aligned}
A_i&=\frac{\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}\left[y^{-}_{ijk}-y^{-}_{ij(k-1)}+y^{-}_{ij(k-1)}u_0/(u_0+v_0)\right]}{\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}\Delta\tau_{ijk}},\\
D&=\sum_{i=1}^{M}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}k\left(y^{-}_{ijk}-y^{-}_{ij(k-1)}+y^{-}_{ij(k-1)}u_0/(u_0+v_0)\right),\\
G&=\sum_{i=1}^{M}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}\frac{1}{\Delta\tau_{ijk}}\left\{\left(y^{-}_{ijk}-y^{-}_{ij(k-1)}-f_{\text{acc}}(\boldsymbol{x}_i)\Delta\tau_{ijk}-\omega k\Delta\tau_{ijk}\right)^2\right.\\
&\quad+2y^{-}_{ij(k-1)}\left(y^{-}_{ijk}-y^{-}_{ij(k-1)}-f_{\text{acc}}(\boldsymbol{x}_i)\Delta\tau_{ijk}-\omega k\Delta\tau_{ijk}\right)u_0/(u_0+v_0)\\
&\quad\left.+\left(y^{-}_{ij(k-1)}\right)^2\left[u_0^2/(u_0+v_0)^2+u_0v_0/\left[(u_0+v_0)^2(u_0+v_0+1)\right]\right]\right\}.
\end{aligned}$$

Appendix D. Technical proofs in Section 3.6

D.1. Proof of Lemma 1

In the following proof, we assume that the regularity conditions described in Appendix B in Liu et al. (2015) hold for the likelihood contributions and $\boldsymbol{\zeta}$ is identifiable in the form of likelihood under the current model. As a side note, we are using the denominator-layout notations. Further, let $L^*(\boldsymbol{\zeta})\equiv L(\boldsymbol{\delta})$ under $\boldsymbol{\zeta}=\boldsymbol{\mathcal{M}}(\boldsymbol{\delta})$. We apply the Taylor approximation as follows:

$$\frac{\partial}{\partial\boldsymbol{\delta}}\log L(\hat{\boldsymbol{\delta}}_{\text{DIR}})\approx\frac{\partial}{\partial\boldsymbol{\delta}}\log L(\boldsymbol{\delta})+\frac{\partial}{\partial\boldsymbol{\delta}\partial\boldsymbol{\delta}^T}\log L(\boldsymbol{\delta})(\hat{\boldsymbol{\delta}}_{\text{DIR}}-\boldsymbol{\delta}). \tag{D.1}$$

The l.h.s. of the equation is zero and further we obtain

$$\frac{\partial}{\partial\boldsymbol{\delta}}\log L(\boldsymbol{\delta})=\mathbf{J}(\boldsymbol{\delta})^T\frac{\partial}{\partial\boldsymbol{\zeta}}\log L^*(\boldsymbol{\zeta}) \tag{D.2}$$

and

$$\frac{\partial}{\partial\boldsymbol{\delta}\partial\boldsymbol{\delta}^T}\log L(\boldsymbol{\delta})=\mathbf{J}(\boldsymbol{\delta})^T\left[\frac{\partial}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\zeta}^T}\log L^*(\boldsymbol{\zeta})\right]\mathbf{J}(\boldsymbol{\delta})=-\mathbf{J}(\boldsymbol{\delta})^T\tilde{\boldsymbol{\Gamma}}(\boldsymbol{\zeta})\mathbf{J}(\boldsymbol{\delta}), \tag{D.3}$$

where $\tilde{\boldsymbol{\Gamma}}(\boldsymbol{\zeta})$ is the observed Fisher information matrix. By plugging (D.2) and (D.3) to (D.1), the form of $n^{1/2}\left(\hat{\boldsymbol{\delta}}_{\text{DIR}}-\boldsymbol{\delta}\right)$ is rewritten by

$$n^{1/2}\left(\hat{\boldsymbol{\delta}}_{\text{DIR}}-\boldsymbol{\delta}\right)=\left[\mathbf{J}(\boldsymbol{\delta})^T\frac{\tilde{\boldsymbol{\Gamma}}(\boldsymbol{\zeta})}{n}\mathbf{J}(\boldsymbol{\delta})\right]^{-1}\times\mathbf{J}(\boldsymbol{\delta})^T\left[n^{-1/2}\frac{\partial}{\partial\boldsymbol{\zeta}}\log L^*(\boldsymbol{\zeta})\right].$$

It is well known that as $n\to\infty$, $\tilde{\boldsymbol{\Gamma}}(\boldsymbol{\zeta})/n$ converges to $\tilde{\mathbf{I}}(\boldsymbol{\zeta})$ in probability. As implied in Liu et al. (2015), the random part of the equation follows MN distribution:

$$n^{-1/2}\frac{\partial}{\partial\boldsymbol{\zeta}}\log L^*(\boldsymbol{\zeta})\xrightarrow{d}\text{MN}\left(\mathbf{0},\tilde{\mathbf{I}}(\boldsymbol{\zeta})\right),$$

which can immediately imply that $n^{1/2}\left(\hat{\boldsymbol{\delta}}_{\text{DIR}}-\boldsymbol{\delta}\right)$ follows a MN distribution with zero mean, indicating the consistency of $\hat{\boldsymbol{\delta}}_{\text{DIR}}$. The covariance matrix of $\hat{\boldsymbol{\delta}}_{\text{DIR}}$ is then obtained by

$$\begin{aligned}
\text{var}\left[n^{1/2}\left(\hat{\boldsymbol{\delta}}_{\text{DIR}}-\boldsymbol{\delta}\right)\right]&=\underbrace{\left[\mathbf{J}(\boldsymbol{\delta})^T\tilde{\mathbf{I}}(\boldsymbol{\zeta})\mathbf{J}(\boldsymbol{\delta})\right]^{-1}\times\mathbf{J}(\boldsymbol{\delta})^T\tilde{\mathbf{I}}(\boldsymbol{\zeta})\mathbf{J}(\boldsymbol{\delta})}_{\text{identity}}\times\left[\mathbf{J}(\boldsymbol{\delta})^T\tilde{\mathbf{I}}(\boldsymbol{\zeta})\mathbf{J}(\boldsymbol{\delta})\right]^{-1}\\
&=\left[\mathbf{J}(\boldsymbol{\delta})^T\tilde{\mathbf{I}}(\boldsymbol{\zeta})\mathbf{J}(\boldsymbol{\delta})\right]^{-1}.
\end{aligned}$$

This completes the proof of Lemma 1.

D.2. Proof of Lemma 2

We can rewrite $\log h(\boldsymbol{\delta})/\partial\boldsymbol{\delta}$ as

$$\frac{\partial\log h(\boldsymbol{\delta})}{\partial\boldsymbol{\delta}}=\mathbf{J}(\boldsymbol{\delta})^T\hat{\mathbf{V}}^{-1}(\hat{\boldsymbol{\zeta}}-\boldsymbol{M}(\boldsymbol{\delta})). \tag{D.4}$$

For $\partial\log L(\boldsymbol{\delta})/\partial\boldsymbol{\delta}$, a Taylor approximation is employed and we obtain

$$\begin{aligned}
\frac{\partial\log L(\boldsymbol{\delta})}{\partial\boldsymbol{\delta}}&=\mathbf{J}(\boldsymbol{\delta})^T\frac{\partial}{\partial\boldsymbol{\zeta}}\log L^*(\boldsymbol{\zeta})\\
&\approx\mathbf{J}(\boldsymbol{\delta})^T\left[\frac{\partial}{\partial\boldsymbol{\zeta}}\log L^*(\hat{\boldsymbol{\zeta}})+\frac{\partial^2}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\zeta}^T}\log L^*(\hat{\boldsymbol{\zeta}})(\boldsymbol{\zeta}-\hat{\boldsymbol{\zeta}})\right]\\
&=\mathbf{J}(\boldsymbol{\delta})^T\left[0-\hat{\mathbf{V}}^{-1}(\boldsymbol{\zeta}-\hat{\boldsymbol{\zeta}})\right]=\mathbf{J}(\boldsymbol{\delta})^T\hat{\mathbf{V}}^{-1}(\hat{\boldsymbol{\zeta}}-\boldsymbol{M}(\boldsymbol{\delta})).
\end{aligned} \tag{D.5}$$

As seen, (D.4) and (D.5) are equal under the Taylor approximation. Thus, they have the same asymptotic properties, which completes the proof.

Appendix E. The observed Fisher information from the method in Oakes (1999)

The following second-order derivatives of Q are given and used for the construction of the Fisher information in (25):

$$\frac{\partial^2Q}{\partial f^2_{\text{acc}}(x_i;\boldsymbol{\delta})}|(\boldsymbol{\theta}=\hat{\boldsymbol{\theta}})=-\frac{1}{\hat{\sigma}^2}\sum_{j=1}^{N_i}\sum_{k=1}^{o_{ij}}\Delta\tau_{ijk},\ \frac{\partial^2Q}{\partial f_{\text{acc}}(\boldsymbol{x}_i;\boldsymbol{\delta})\partial\sigma^2}|(\boldsymbol{\theta}=\hat{\boldsymbol{\theta}})$$

$$= \frac{1}{\hat{\sigma}^4}\left[\hat{f}_{\mathrm{acc}}(\boldsymbol{x}_i)\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}\Delta\tau_{ijk} - \sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}\left[y^-_{ijk} - y^-_{ij(k-1)} + y^-_{ij(k-1)}E\left(z_{ij(k-1)}|\boldsymbol{y}^-,\hat{\boldsymbol{\theta}}\right)\right] + \hat{\omega}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}k\Delta\tau_{ijk}\right],$$

$$\frac{\partial^2 Q}{\partial f_{\mathrm{acc}}(\boldsymbol{x}_i;\boldsymbol{\delta})\partial\omega}|(\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}) = -\frac{1}{\hat{\sigma}^2}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}k\Delta\tau_{ijk},$$

$$\frac{\partial^2 Q}{\partial\omega^2}|(\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}) = -\frac{1}{\hat{\sigma}^2}\sum_{i=1}^{M}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}k^2\Delta\tau_{ijk}.$$

$$\frac{\partial^2 Q}{\partial\omega\partial\sigma^2}|(\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}) = \frac{1}{\hat{\sigma}^4}\left[\sum_{i=1}^{M}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}k\left(y^-_{ijk} - y^-_{ij(k-1)} + y^-_{ij(k-1)}E\left(z_{ij(k-1)}|\boldsymbol{y}^-,\hat{\boldsymbol{\theta}}\right)\right) + \sum_{i=1}^{M}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}k\Delta\tau_{ijk}\hat{f}_{\mathrm{acc}}(\boldsymbol{x}_i) + \hat{\omega}\sum_{i=1}^{M}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}k^2\Delta\tau_{ijk}\right],$$

$$\frac{\partial Q\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(n)}\right)}{\partial\left(\sigma^2\right)^2}|(\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}) = \frac{1}{2\hat{\sigma}^4}\sum_{i=1}^{M}\sum_{j=1}^{N_i}O_{ij} + \frac{1}{\hat{\sigma}^6}\sum_{i=1}^{M}\sum_{j=1}^{N_i}\sum_{k=1}^{O_{ij}}\frac{1}{\Delta\tau_{ijk}}\left[\left(y^-_{ijk} - y^-_{ij(k-1)} - \hat{f}_{\mathrm{acc}}(\boldsymbol{x}_i)\Delta\tau_{ijk} - \hat{\omega}k\Delta\tau_{ijk}\right)^2 + 2y^-_{ij(k-1)}\left(y^-_{ijk} - y^-_{ij(k-1)} - \hat{f}_{\mathrm{acc}}(\boldsymbol{x}_i)\Delta\tau_{ijk} - \hat{\omega}k\Delta\tau_{ijk}\right)E\left(z_{ij(k-1)}|\boldsymbol{y}^-,\hat{\boldsymbol{\theta}}\right) + \left(y^-_{ij(k-1)}\right)^2E\left(z^2_{ij(k-1)}|\boldsymbol{y}^-,\hat{\boldsymbol{\theta}}\right)\right].$$

The second-order derivatives directly provide the first term in the r.h.s. of (25). For the second term, derivatives can be directly taken on (8) to (15) with respect to $\boldsymbol{\theta}^{(n)}$ and replace $\boldsymbol{\theta}^{(n)}$ with $\hat{\boldsymbol{\theta}}$. Specifically, we need to compute $\partial E\left(z_{ij(k-1)}|\boldsymbol{y}^-,\hat{\boldsymbol{\theta}}\right)/\partial\hat{\boldsymbol{\theta}}$ and $\partial E\left(z^2_{ij(k-1)}|\boldsymbol{y}^-,\hat{\boldsymbol{\theta}}\right)/\partial\hat{\boldsymbol{\theta}}$. Combining the results from (B.1) to (B.4) in Appendix B, analytical results can be obtained to finally compute (25) from:

$$\frac{\partial E\left(z_{ij(k-1)}|\boldsymbol{y}^-,\hat{\boldsymbol{\theta}}\right)}{\partial\hat{\theta}_i} = \int_0^1\frac{\partial f_{z_{ij(k-1)}}\left(z|\boldsymbol{y}^-,\hat{\boldsymbol{\theta}}\right)}{\partial\hat{\theta}_i}z\mathrm{d}z,$$

$$\frac{\partial E\left(z^2_{ij(k-1)}|\boldsymbol{y}^-,\hat{\boldsymbol{\theta}}\right)}{\partial\hat{\theta}_i} = \int_0^1\frac{\partial f_{z_{ij(k-1)}}\left(z|\boldsymbol{y}^-,\hat{\boldsymbol{\theta}}\right)}{\partial\hat{\theta}_i}z^2\mathrm{d}z.$$

References

Bae, S. J., Yuan, T., Ning, S., & Kuo, W. (2015). A Bayesian approach to modeling two-phase degradation using change-point regression. *Reliability Engineering & System Safety, 134*, 66–74. https://doi.org/10.1016/j.ress.2014.10.009.

Bernardo, J., Bayarri, M., Berger, J., Dawid, A., Heckerman, D., Smith, A., West, M., et al. (2003). The variational Bayesian em algorithm for incomplete data: With application to scoring graphical model structures. *Bayesian Statistics, 7*(453–464), 210.

Chen, P., & Ye, Z.-S. (2017). Random effects models for aggregate lifetime data. *IEEE Transactions on Reliability, 66*(1), 76–83. https://doi.org/10.1109/TR.2016.2611625.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological), 39*(1), 1–38.

Dong, M., & He, D. (2007). Hidden semi-Markov model-based methodology for multi-sensor equipment health diagnosis and prognosis. *European Journal of Operational Research, 178*(3), 858–878. https://doi.org/10.1016/j.ejor.2006.01.041.

Duan, F., & Wang, G. (2018). Exponential-dispersion degradation process models with random effects and covariates. *IEEE Transactions on Reliability, 67*(3), 1128–1142. https://doi.org/10.1109/TR.2018.2849087.

Hong, Y., Duan, Y., Meeker, W. Q., Stanley, D. L., & Gu, X. (2015). Statistical methods for degradation data with dynamic covariates information and an application to outdoor weathering data. *Technometrics, 57*(2), 180–193. https://doi.org/10.1080/00401706.2014.915891.

Hu, C.-H., Lee, M.-Y., & Tang, J. (2015). Optimum step-stress accelerated degradation test for Wiener degradation process under constraints. *European Journal of Operational Research, 241*(2), 412–421. https://doi.org/10.1016/j.ejor.2014.09.003.

Jakob, F., Kimmelmann, M., & Bertsche, B. (2017). Selection of acceleration models for test planning and model usage. *IEEE Transactions on Reliability, 66*(2), 298–308.

Lee, M. Y., Hu, C. H., & Tang, J. (2017). A two-stage latent variable estimation procedure for time-censored accelerated degradation tests. *IEEE Transactions on Reliability, 66*(4), 1266–1279. https://doi.org/10.1109/TR.2017.2731680.

Levine, R. A., & Casella, G. (2001). Implementations of the Monte Carlo EM algorithm. *Journal of Computational and Graphical Statistics, 10*(3), 422–439. https://doi.org/10.1198/106186001317115045.

Li, X.-Y., Wu, J.-P., Ma, H.-G., Li, X., & Kang, R. (2018). A random fuzzy accelerated degradation model and statistical analysis. *IEEE Transactions on Fuzzy Systems, 26*(3), 1638–1650. https://doi.org/10.1109/TFUZZ.2017.2738607.

Limon, S., Yadav, O. P., & Liao, H. (2017). A literature review on planning and analysis of accelerated testing for reliability assessment. *Quality and Reliability Engineering International, 33*(8), 2361–2383. https://doi.org/10.1002/qre.2195.

Lin, Y., Su, Y., Cheng, Y., Tao, L., Noktehdan, A., Chong, J., … Jin, H. (2017). Lithium-ion battery capacity fading dynamics modelling for formulation optimization: A stochastic approach to accelerate the design process. *Applied Energy, 202*, 138–152. https://doi.org/10.1016/j.apenergy.2017.04.027.

Liu, D., Liu, R. Y., & Xie, M. (2015). Multivariate meta-analysis of heterogeneous studies using only summary statistics: efficiency and robustness. *Journal of the American Statistical Association, 110*(509), 326–340. https://doi.org/10.1080/01621459.2014.899235.

Liu, L., Li, X.-y., Zio, E., Kang, R., & Jiang, T.-m. (2017). Model uncertainty in accelerated degradation testing Analysis. *IEEE Transactions on Reliability, 66*(3), 603–615. https://doi.org/10.1109/TR.2017.2696341.

Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological), 44*(2), 226–233.

McLachlan, G., & Krishnan, T. (2007). *The EM algorithm and extensions*: 382. John Wiley & Sons.

Meeker, W. Q., & Escobar, L. A. (1998). *Statistical methods for reliability data*. John Wiley & Sons.

Meeker, W. Q., Escobar, L. A., & Lu, C. J. (1998). Accelerated degradation tests: modeling and analysis. *Technometrics, 40*(2), 89–99. https://doi.org/10.1080/00401706.1998.10485191.

Meeker, W. Q., & Hahn, G. J. (1977). Asymptotically optimum over-stress tests to estimate the survival probability at a condition with a low expected failure probability. *Technometrics, 19*(4), 381–399.

Meng, X.-L., & Rubin, D. B. (1991). Using EM to obtain asymptotic variance-covariance matrices: The SEM algorithm. *Journal of the American Statistical Association, 86*(416), 899–909. https://doi.org/10.1080/01621459.1991.10475130.

Mercier, S., & Castro, I. (2019). Stochastic comparisons of imperfect maintenance models for a gamma deteriorating system. *European Journal of Operational Research, 273*(1), 237–248. https://doi.org/10.1016/j.ejor.2018.06.020.

Oakes, D. (1999). Direct calculation of the information matrix via the EM. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 61*(2), 479–482. https://doi.org/10.1111/1467-9868.00188.

Peck, D. S. (1986). Comprehensive model for humidity testing correlation. In *Proceedings of the 24th international reliability physics symposium* (pp. 44–50). https://doi.org/10.1109/IRPS.1986.362110.

Shi, Y., Escobar, L. A., & Meeker, W. Q. (2009). Accelerated destructive degradation test planning. *Technometrics, 51*(1), 1–13.

Si, X. S. (2015). An adaptive prognostic approach via nonlinear degradation modeling: Application to battery data. *IEEE Transactions on Industrial Electronics, 62*(8), 5082–5096. https://doi.org/10.1109/TIE.2015.2393840.

Si, X.-S., Wang, W., Hu, C.-H., & Zhou, D.-H. (2014). Estimating remaining useful life with three-source variability in degradation modeling. *IEEE Transactions on Reliability, 63*(1), 167–190.

Tsai, T.-R., Sung, W.-Y., Lio, Y. L., Chang, S. I., & Lu, J.-C. (2016). Optimal two-variable accelerated degradation test plan for gamma degradation processes. *IEEE Transactions on Reliability, 65*(1), 459–468. https://doi.org/10.1109/TR.2015.2435774.

Tseng, S.-T., & Lee, I.-C. (2016). Optimum allocation rule for accelerated degradation tests with a class of exponential-dispersion degradation models. *Technometrics, 58*(2), 244–254. https://doi.org/10.1080/00401706.2015.1033109.

Wang, D., & Tsui, K.-L. (2017). Statistical modeling of bearing degradation signals. *IEEE Transactions on Reliability, 66*(4), 1331–1344. https://doi.org/10.1109/TR.2017.2739126.

Wang, P., Tang, Y., Bae, S. J., & Xu, A. (2018). Bayesian approach for two-phase degradation data based on change-point Wiener process with measurement errors. *IEEE Transactions on Reliability, 67*(2), 688–700. https://doi.org/10.1109/TR.2017.2785978.

Xiao, X., & Ye, Z. (2016). Optimal design for destructive degradation tests with random initial degradation values using the Wiener process. *IEEE Transactions on Reliability, 65*(3), 1327–1342. https://doi.org/10.1109/TR.2016.2575442.

Xie, M.-g., & Singh, K. (2013). Confidence distribution, the frequentist distribution estimator of a parameter: A review. *International Statistical Review, 81*(1), 3–39. https://doi.org/10.1111/insr.12000.

Ye, Z.-S., Chen, L.-P., Tang, L. C., & Xie, M. (2014). Accelerated degradation test planning using the inverse Gaussian process. *IEEE Transactions on Reliability, 63*(3), 750–763. https://doi.org/10.1109/TR.2014.2315773.

Ye, Z.-S., Hu, Q., & Yu, D. (2019). Strategic allocation of test units in an accelerated degradation test plan. *Journal of Quality Technology, 51*(1), 64–80. https://doi.org/10.1080/00224065.2018.1545495.

Ye, Z.-S., & Xie, M. (2015). Stochastic modelling and analysis of degradation for highly reliable products. *Applied Stochastic Models in Business and Industry, 31*(1), 16–32.

Zhai, Q., & Ye, Z.-S. (2018). Degradation in common dynamic environments. *Technometrics, 60*(4), 461–471. https://doi.org/10.1080/00401706.2017.1375994.

Zhang, M., Gaudoin, O., & Xie, M. (2015). Degradation-based maintenance decision using stochastic filtering for systems under imperfect maintenance. *European Journal of Operational Research, 245*(2), 531–541. https://doi.org/10.1016/j.ejor.2015.02.050.

Zhao, X., Gaudoin, O., Doyen, L., & Xie, M. (2019). Optimal inspection and replacement policy based on experimental degradation data with covariates. *IISE Transactions, 51*(3), 322–336. https://doi.org/10.1080/24725854.2018.1488308.