Statistical Theory and Related Fields

ISSN: 2475-4269 (Print) 2475-4277 (Online) Journal homepage: www.tandfonline.com/journals/tstf20

Reliability analysis of two-stage nonlinear phase-dependent Wiener degradation process

Binxian Zhuang, Wanping Lv, Shuqi Fan, Cheng Li, Jiaxing Huang, Qiang Guan & Yongxian Wen

To cite this article: Binxian Zhuang, Wanping Lv, Shuqi Fan, Cheng Li, Jiaxing Huang, Qiang Guan & Yongxian Wen (04 Mar 2026): Reliability analysis of two-stage nonlinear phase-dependent Wiener degradation process, Statistical Theory and Related Fields, DOI: 10.1080/24754269.2026.2633809

To link to this article: https://doi.org/10.1080/24754269.2026.2633809

Published online: 04 Mar 2026.

STATISTICAL THEORY AND RELATED FIELDS
https://doi.org/10.1080/24754269.2026.2633809

Taylor & Francis
Taylor & Francis Group

Reliability analysis of two-stage nonlinear phase-dependent Wiener degradation process

Binxian Zhuang[a,b], Wanping Lv[a,b], Shuqi Fan[a,b], Cheng Li[a,b], Jiaxing Huang[a], Qiang Guan[c] and Yongxian Wen[a,b]

[a]College of Computer and Information Science, Fujian Agriculture and Forestry University, Fuzhou, People's Republic of China; [b]Institute of Statistics and Application, Fujian Agriculture and Forestry University, Fuzhou, People's Republic of China; [c]College of Information Engineering, Sanming University, Sanming, People's Republic of China

ABSTRACT
Accurately assessing reliability and predicting the life of products is an important part of Prediction and Prognostics and Health Management. According to historical degradation data of products, the degradation model of two-stage nonlinear phase-dependent Wiener process is proposed to evaluate the dependability and forecast the lifetime of individual products. Firstly, a two-stage phase-dependent degradation model based on a nonlinear Wiener process was developed to characterize nonlinear nature of the degradation of product performance by simultaneously taking the correlation of degradation stages and the degradation heterogeneity of products into account; Secondly, the model parameters are estimated using EM algorithm based on historical degradation data of products; Finally, simulation studies are used to demonstrate the superiority of the approach for assessing offline reliability, and the validity of the method proposed in this paper is verified through an example of high-voltage pulse capacitors. The results show that compared with the degradation model based on phase-independent nonlinear Wiener processes, the two-stage nonlinear phase-dependent Wiener process degradation model can better capture actual degradation paths of products and provide a more reasonable assessment of lifetime reliability of products.

ARTICLE HISTORY
Received 10 April 2025
Revised 22 August 2025
Accepted 12 February 2026

KEYWORDS
Phase-dependent degradation model; nonlinear; Wiener process; two-stage degradation; EM algorithm; Bayes method; reliability assessment

1. Introduction

Accurately predicting products (systems) life in engineering practice has considerable research and practical significance, attracting extensive attention from scholars over the past decades. Reliability analysis and life prediction have made significant theoretical progress and widespread application over decades of research. In general, methods for reliability analysis and life prediction are divided into three overall categories: physical failure approaches, data-driven methods, and methods that combine both of the above methods (Cai et al., 2019;

CONTACT Yongxian Wen wenyx9681@fafu.edu.cn, wen9681@sina.com; Qiang Guan sxtjgq@126.com
College of Computer and Information Science, Fujian Agriculture and Forestry University, Fuzhou 350002, People's Republic of China; Institute of Statistics and Application, Fujian Agriculture and Forestry University, Fuzhou 350002, People's Republic of China; College of Information Engineering, Sanming University, Sanming 365004, People's Republic of China

Y. Wang et al., 2016; Yuan et al., 2020; Zhuang et al., 2023). Among data-driven methods, the statistical data-driven assesses the reliability of products and predicts life in a probabilistic way, based on useful monitoring data and statistical models. This method has strong mathematical properties and is widely used in a variety of application scenarios. It is one of the most popular life prediction methods at present (Si et al., 2011). The stochastic process approach, one of the statistical data-driven approaches, offers a better description of the degradation process of products. By modeling the monitoring degradation data, the probability distribution of lifetime can be obtained, which facilitates the quantification of lifetime uncertainty (Dong et al., 2022; J. X. Yang et al., 2024). Therefore, the degradation reliability analysis and life prediction methods based on stochastic processes have become one of the mainstream methods in relevant researches (Pandey et al., 2009).

In the stochastic process approach, Wiener process is a class of diffusion processes with drift terms, which can describe the monotonous or non-monotonous degradation process that shows a tendency to increase or decrease due to a large number of small losses, and possesses good mathematical properties. Therefore, Wiener-based methods have been widely used in the degradation model for reliability analysis and lifetime prediction (Ding et al., 2024; Z. S. Ye et al., 2015; J. Yang et al., 2025).

Recently, the degradation model based on the Wiener process generally assumes that products follow a single stochastic degradation phase in the process of degradation, while in the engineering practice due to the complexity of degradation mechanism, the influence of environmental factors, and the change of system's operating state, the degradation rate and the degree of data fluctuation of some products show a two-phase or even multi-phase characteristic (Wu et al., 2024; Zhou et al., 2014). J. Yang et al. (2025) proposed an ensemble multi-stage Wiener modeling paradigm that combines temporal phase mutation patterns and dynamic volatility analysis to significantly enhance the accuracy of state-of-health estimation and remaining useful life (RUL) prediction for lead-acid batteries. P. Zhang et al. (2019) for the problem of RUL prediction of a liquid coupling device with the characteristic of two-phase in degradation process, introduced random effects to describe differences among products, established a two-stage linear Wiener process degradation model, and adopted EM algorithm and Bayesian theory to achieve the offline identification and online updates of model parameters. P. P. Wang et al. (2018) established a two-stage linear Wiener process degradation model by considering degradation heterogeneity of products, used a hierarchical Bayesian approach and maximum likelihood method to estimate model parameters, and validated the way on organic light emitting diodes. Bae et al. (2015) proposed a hierarchical Bayesian change-point regression model to fit a two-stage degradation process, derived the failure time distribution of randomly selected products from the population, and developed a Gibbs sampling algorithm to estimate the change point and predict the failure time distribution of randomly selected products, which was validated on plasma display panels.

However, the degradation model based on the two-stage linear Wiener process generally assumes that the degradation stages are mutually independent. H. Wang et al. (2018) were the first to use the degradation process of lasers and bearings to argue that the assumption is not reasonable, and proposed that there is a correlation between the degradation stages. Thus, they proposed a two-stage linear Wiener process model with stage correlation and applied Bayesian method for life prediction. Based on their research, H. Wang et al. (2020) simultaneously considered the correlation of degradation stages and degradation heterogeneity of products, and proposed a mixed-effects model based on the linear Wiener process to capture two-stage degradation of products.

However, changes in the external environment, load, internal state and other factors can either strengthen or weaken the degradation of product performance, resulting in nonlinear degradation characteristics. Therefore, the degradation model based on nonlinear Wiener processes can capture the degradation process of product performance more accurately and realistically. Ding et al. (2024) proposed an improved adaptive nonlinear Wiener degradation model that effectively solved the limitations of Markovianity in the traditional Wiener process by integrating Kalman filtering, EM algorithm, and RTS smoothing technology, and was verified on C-MAPSS dataset. Huang et al. (2018) proposed a degradation model based on a multi-stage nonlinear Wiener process and predicted the aircraft engine lifetime. A. Zhang et al. (2020) proposed a two-stage nonlinear degradation model based on Wiener process and verified the validity of model on high-voltage pulse capacitors.

The above research status on the degradation process of products indicates that the reliability analysis and lifetime prediction of high-reliability and long-lifespan products have received significant attention in related fields, and the Wiener process has a wide range of applications in degradation modeling and life prediction, but the existing research results mainly focus on two-stage linear or nonlinear Wiener process degradation models that take degradation heterogeneity of products into account. Moreover, the degradation models considering phase-dependent are established only based on two-stage linear Wiener process.

To overcome this shortcoming, a two-stage nonlinear phase-dependent Wiener process degradation model is proposed, which simultaneously considers the correlation of degradation stages and degradation heterogeneity of products. Additionally, it assesses the reliability and predicts lifetime of individual products based on their historical degradation data.

2. Model formulation

2.1. Single-stage nonlinear Wiener process degradation model

Denote the Wiener process with a constant rate as $\{X(t), t \geq 0\}$. The degradation model based on the Wiener process is expressed as

$$X(t) = \alpha_0 + \int_0^t \lambda(t, \theta)\, \mathrm{d}t + \sigma B(t), \tag{1}$$

where $X(t)$ denotes the performance degradation amount at moment t, α_0 denotes the initial degradation amount, $\lambda(t, \theta)$ denotes the degradation rate, $\sigma > 0$ denotes the diffusion coefficient, and $\{B(t), t \geq 0\}$ denotes the standard Brownian motion.

When $\lambda(t, \theta) = \beta r t^{(r-1)}$, the model is defined as a nonlinear Wiener process, namely,

$$X(t) = \alpha_0 + \beta t^r + \sigma B(t), \tag{2}$$

where $\beta > 0$ denotes the drift coefficient, and r is the unknown parameter in nonlinear function of drift coefficient.

According to degradation reliability theory, if $X(t)$ is incremental, given a failure threshold D, if $X(t) \leq D$, the product is considered to work properly; if $X(t) > D$, the product is considered to have failed.

Based on the conception of the first reaching time of a stochastic process, the lifetime T of product is defined as the first time $X(t)$ that exceeds the failure threshold D, namely,

$$T = \inf\{t : X(t) \geq D \mid \alpha_0 < D\}. \tag{3}$$

Thus, the lifetime distribution of product (Deng et al., 2006) can be expressed as

$$F_T(t;\alpha_0,\beta,\sigma) = P(X(t) \geq D) = 1 - \Phi\left[\frac{D-\alpha_0-\beta t^r}{\sqrt{\sigma^2 t}}\right], \tag{4}$$

where $\Phi(\cdot)$ is the distribution function of standard normal distribution. Therefore, the reliability function of the product can be expressed as

$$R(t;\alpha_0,\beta,\sigma) = 1 - F_T(t;\alpha_0,\beta,\sigma) = \Phi\left(\frac{D-\alpha_0-\beta^r}{\sqrt{\sigma^2 t}}\right). \tag{5}$$

2.2. Two-stage nonlinear Wiener process degradation models

Assuming that there is degradation data for n products, let $X_i = (X_{i,1}, X_{i,2}, \ldots, X_{i,m_i})^\top$ denote the degradation data collected from the ith product at time $t_{i,1} < t_{i,2} < \cdots < t_{i,m_i}$, where $i = 1, 2, \ldots, n$. Thus, the two-stage nonlinear phase dependent Wiener process degradation model is denoted as

$$\begin{aligned} X_{i,j} &= \begin{cases} \alpha_{1,i} + \beta_{1,i} t_{i,j}^{r_1} + \sigma_1 B(t_{i,j}), & t_{i,j} \leq \tau_i, \\ \alpha_{2,i} + \beta_{2,i}(t_{i,j}-\tau_i)^{r_2} + \sigma_2 B(t_{i,j}-\tau_i), & t_{i,j} > \tau_i, \end{cases} \\ \theta_i &= (\alpha_{1,i}, \beta_{1,i}, \alpha_{2,i}, \beta_{2,i})^\top \sim \mathrm{MVN}(\mu_\theta, \Sigma_\theta), \end{aligned} \tag{6}$$

where $j = 1, 2, \ldots, m_i$; $\alpha_{1,i}$ and $\alpha_{2,i}$ denote the initial values of two phases in degradation process of the ith product, respectively; $\beta_{1,i}$ and $\beta_{2,i}$ are drift parameters, which denote the degradation rates of two phases of the ith product; σ_1 and σ_2 are constants denoting the diffusion coefficient of all units (assumed constant); τ_i is the change point in degradation process of the ith product; $B(\cdot)$ is the standard Brownian motion. Therefore, at this point $X_{i,j}$ follows a normal distribution with mean in Equation (7) and covariance in Equation (8). The probability density function (PDF) of θ_i is shown in Equation (9).

$$E(X_{i,j}) = \begin{cases} \alpha_{1,i} + \beta_{1,i} t_{i,j}^{r_1}, & t_{i,j} \leq \tau_i, \\ \alpha_{2,i} + \beta_{2,i}(t_{i,j}-\tau_i)^{r_2}, & t_{i,j} > \tau_i; \end{cases} \tag{7}$$

$$\mathrm{Var}(X_{i,j}) = \begin{cases} \sigma_1^2 t_{i,j}, & t_{i,j} \leq \tau_i, \\ \sigma_2^2 (t_{i,j}-\tau_i), & t_{i,j} > \tau_i; \end{cases} \tag{8}$$

$$p(\theta_i|\mu_\theta, \Sigma_\theta) = (2\pi)^{-\frac{k}{2}} |\Sigma_\theta|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\theta_i-\mu_\theta)^\top \Sigma_\theta^{-1}(\theta_i-\mu_\theta)\right], \tag{9}$$

where k denotes the dimension of θ_i.

In the model of Equation (6), a two-stage nonlinear phase-dependent Wiener degradation process is used to describe the degradation data X_i of the ith product with the parameters θ_i, $r_1, r_2, \sigma_1, \sigma_2, \tau_i$. The parameter vector θ_i follows a multivariate normal distribution with mean vector μ_θ and covariance matrix Σ_θ. Among them, μ_θ represents the common characteristics, namely, the fixed effects related to overall average performance, while Σ_θ represents random effects due to degradation heterogeneity of products. Moreover, the covariance between the parameters of different degradation stages reflects the correlation between the stages (H. Wang et al., 2020). For simplicity of representation, we denote the change point as $\tau = (\tau_1, \tau_2, \ldots, \tau_n)$. Accounting for the individual variability of change points, we make

τ follow the normal distribution $N(\mu_\tau, \sigma_\tau^2)$ (Li, 2009; Ng, 2008; P. P. Wang et al., 2018), and the model parameters are denoted as $\Theta = \{\mu_\theta, \Sigma_\theta, \sigma_1, \sigma_2, r_1, r_2, \tau\}$.

The above two-stage nonlinear phase-dependent Wiener process degradation model can be extended into a multi-stage nonlinear phase-dependent Wiener process degradation model, namely

$$\begin{aligned} X_{i,j} &= [\alpha_{1,i} + \beta_{1,i} t_{i,j}^{r_1} + \sigma_1 B(t_{i,j})] I_{(0,\tau_{i,1})}(t) \\ &\quad + [\alpha_{2,i} + \beta_{2,i}(t_{i,j} - \tau_{i,1})^{r_2} + \sigma_2 B(t_{i,j} - \tau_{i,1})] I_{(\tau_{i,1},\tau_{i,2})}(t) \\ &\quad + \cdots + [\alpha_{p,i} + \beta_{p,i}(t_{i,j} - \tau_{i,p-1})^{r_p} + \sigma_p B(t_{i,j} - \tau_{i,p-1})] I_{(\tau_{i,p-1},\infty)}(t), \\ \theta_i &= (\alpha_{1,i}, \beta_{1,i}, \alpha_{2,i}, \beta_{2,i}, \dots, \alpha_{p,i}, \beta_{p,i})^\top \sim \mathrm{MVN}(\mu_\theta, \Sigma_\theta), \end{aligned} \tag{10}$$

where

$$I_A(t) = \begin{cases} 1, & t \in A, \\ 0, & \text{else}. \end{cases}$$

The model of Equation (10) allows reliability analysis and lifetime prediction of multi-stage nonlinear degradation processes. When θ_i is known, the model of Equation (10) is a multi-stage nonlinear degradation model with independent degradation stages (Huang et al., 2018), namely,

$$\begin{aligned} X(t) &= [\alpha_0 + \beta_1 t^{r_1} + \sigma_1 B(t)] I_{(0,\tau_1)}(t) \\ &\quad + [\alpha_{\tau_1} + \beta_2 (t - \tau_1)^{r_2} + \sigma_2 B(t - \tau_1)] I_{(\tau_1,\tau_2)}(t) \\ &\quad + \cdots + [\alpha_{\tau_{p-1}} + \beta_p (t - \tau_{p-1})^{r_p} + \sigma_p B(t - \tau_{p-1})] I_{(\tau_{p-1},\infty)}(t), \end{aligned} \tag{11}$$

where α_0 represents the initial degradation amount, α_{τ_q} is the performance degradation value of the product at moment $\tau_q (q = 1, 2, \dots, p)$, and τ_q is the time to reach each performance degradation cut-off value, that is, the change point.

2.3. Lifetime distribution

A two-stage example will be used for degradation modeling study and case analysis.

Firstly, the two-stage independent nonlinear Wiener process degradation model is shown in Equation (12):

$$X(t) = [\alpha_0 + \beta_1 t^{r_1} + \sigma_1 B(t)] I_{(0,\tau)}(t) + [\alpha_\tau + \beta_2 (t - \tau)^{r_2} + \sigma_2 B(t - \tau)] I_{(\tau,\infty)}(t). \tag{12}$$

We can obtain the life distribution and reliability function of product performance degradation (A. Zhang et al., 2020):

$$F_T(t; \beta_1, \beta_2, \sigma_1, \sigma_2, \mu_\tau) = P(X(t) \geq D) = 1 - \Phi\left[\frac{D - \beta_1 \mu_\tau^{r_1} - \beta_2 (t - \mu_\tau)^{r_2}}{\sqrt{\sigma_1^2 \mu_\tau + \sigma_2^2 (t - \mu_\tau)}}\right], \tag{13}$$

$$R_T(t; \beta_1, \beta_2, \sigma_1, \sigma_2, \mu_\tau) = 1 - F_T(t; \beta_1, \beta_2, \sigma_1, \sigma_2, \mu_\tau). \tag{14}$$

Secondly, based on lifetime distributions and reliability functions of the single-stage nonlinear Wiener process degradation model and the two-stage independent nonlinear Wiener process degradation model, and according to the established model of Equation (6), we will derive the lifetime distribution of product: if the product fails in the first phase, i.e., $T \leq \tau$, the lifetime distribution is only related to the first phase of the degradation process; if the product fails in the second phase, namely, $T > \tau$, the lifetime distribution needs to fulfill the condition that the product does not fail in the first phase. Therefore, the lifetime distribution and reliability function of product can be obtained according to Equations (4) and (13) (Deng et al., 2006; A. Zhang et al., 2020):

$$F_T(t;\alpha_{1,i},\beta_{1,i},\beta_{2,i},\sigma_1,\sigma_2,r_1,r_2,\mu_\tau) = \begin{cases} 1-\Phi\left(\dfrac{D-\alpha_{1,i}-\beta_{1,i}t^{r_1}}{\sqrt{\sigma_1^2 t}}\right), & t \leq \tau_i, \\ 1-\Phi\left(\dfrac{D-\beta_{1,i}\mu_\tau^{r_1}-\beta_{2,i}(t-\mu_\tau)^{r_2}}{\sqrt{\sigma_1^2\mu_\tau+\sigma_2^2(t-\mu_\tau)}}\right), & t > \tau_i, \end{cases} \tag{15}$$

$$R_T(t;\alpha_{1,i},\beta_{1,i},\beta_{2,i},\sigma_1,\sigma_2,r_1,r_2,\mu_\tau) = 1 - F_T(t;\alpha_{1,i},\beta_{1,i},\beta_{2,i},\sigma_1,\sigma_2,r_1,r_2,\mu_\tau), \tag{16}$$

where D is the failure threshold.

Furthermore, to ensure the rigour of the model established in Equation (6) for RUL prediction, the conditional survival probability of the product is based on the ratio of the first-passage time distribution and the survival function:

$$S(t \mid \tau) = P(T > t \mid T > \tau) = \frac{S(t)}{S(\tau)} = \frac{P(\inf_{\beta \leq t} X(\beta) < D)}{P(\inf_{\beta \leq \tau} X(\beta) < D)}, \quad t \geq \tau, \tag{17}$$

where $S(t)$ is the survival function.

When the estimated values of parameters $\Theta = \{\mu_\theta, \Sigma_\theta, \sigma_1, \sigma_2, r_1, r_2, \tau\}$ of the model of Equation (6) and μ_τ are derived, we can obtain the lifetime distribution and reliability function of product.

3. Parameter estimation for two-stage nonlinear phase-dependent Wiener process degradation model

Based on degradation data of products, the parameter estimation of two-stage nonlinear phase dependent Wiener process degradation model is mainly introduced in this section. To facilitate the following analysis, the transformed degradation data $Y_i = (Y_{i,1}, Y_{i,2}, \ldots, Y_{i,m_i})^\top$ of the ith product ($i = 1, 2, \ldots, n$) will be introduced, in which,

$$Y_{i,j} = \begin{cases} X_{i,j}, & j = 1 \quad \text{or } c_i + 1, \\ X_{i,j} - X_{i,j-1}, & \text{else}, \end{cases} \tag{18}$$

where $j = 1, 2, \ldots, m_i$; $c_i \in \{1, 2, \ldots, m_i\}$, and $t_{i,c_i} = \tau_i$, namely, t_{i,c_i} is change-point moment of the ith product; define $t_{i,0} = 0$. Given θ_i, σ_1, σ_2, r_1, r_2, τ_i, $Y_{i,j}$ follows a normal distribution with mean in Equation (19) and covariance in Equation (20):

$$E(Y_{i,j}|\theta_i, r_1, r_2, \tau_i) = \begin{cases} \alpha_{1,i}I(j=1) + \beta_{1,i}(t_{i,j}^{r_1} - t_{i,j-1}^{r_1}), & j \leq c_i, \\ \alpha_{2,i}I(j=c_i+1) + \beta_{2,i}[(t_{i,j}-t_{i,c_i})^{r_2} - (t_{i,j-1}-t_{i,c_i})^{r_2}], & j > c_i, \end{cases} \tag{19}$$

$$\mathrm{Cov}(Y_{i,j}, Y_{i,k}|\sigma_1, \sigma_2, \tau_i) = \begin{cases} \sigma_1^2(t_{i,j} - t_{i,j-1}), & j = k \le c_i, \\ \sigma_2^2(t_{i,j} - t_{i,j-1}), & j = k > c_i, \\ 0, & \text{else}, \end{cases} \tag{20}$$

where $j = k = 1, 2, \ldots, m_i$.

Thus, when $\theta_i, \sigma_1, \sigma_2, r_1, r_2, \tau_i$ are known, for convenience, we directly introduce the vector form of the joint probability density function of Y_i, as shown in Equation (21):

$$\begin{aligned} \rho(Y_i|\theta_i, r_1, r_2, \sigma_1, \sigma_2, \tau_i) = {}& (2\pi)^{-\frac{m_i}{2}} \sigma_1^{-c_i} \sigma_2^{-(m_i-c_i)} |T_i|^{-\frac{1}{2}} \\ & \times \exp\left[-\frac{1}{2}(Y_i - \Lambda_i\theta_i)^\top \Delta_i T_i^{-1} \Delta_i (Y_i - \Lambda_i\theta_i)\right], \end{aligned} \tag{21}$$

where

$$\begin{aligned} \Lambda_i &= \begin{bmatrix} \Lambda_{i,1} & \\ & \Lambda_{i,2} \end{bmatrix}, \\ \Lambda_{i,1} &= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ t_{i,1}^{r_1} & t_{i,2}^{r_1} - t_{i,1}^{r_1} & \cdots & t_{i,c_i}^{r_1} - t_{i,c_i-1}^{r_1} \end{pmatrix}^\top, \\ \Lambda_{i,2} &= \left(\begin{matrix} 1 & 0 & \cdots \\ (t_{i,c_i+1} - t_{i,c_i})^{r_2} & (t_{i,c_i+2} - t_{i,c_i})^{r_2} - (t_{i,c_i+1} - t_{i,c_i})^{r_2} & \cdots \end{matrix}\right. \\ &\qquad \left.\begin{matrix} 0 \\ (t_{i,m_i} - t_{i,c_i})^{r_2} - (t_{i,m_i-1} - t_{i,c_i})^{r_2} \end{matrix}\right)^\top, \\ \Delta_i &= \sigma_1^{-1}\Delta_{i,1} + \sigma_2^{-1}\Delta_{i,2}, \\ \Delta_{i,1} &= \mathrm{diag}(I(t_{i,1} \le \tau_i), I(t_{i,2} \le \tau_i), \ldots, I(t_{i,m_i} \le \tau_i)), \\ \Delta_{i,2} &= \mathrm{diag}(I(t_{i,1} > \tau_i), I(t_{i,2} > \tau_i), \ldots, I(t_{i,m_i} > \tau_i)), \end{aligned}$$

and

$$T_i = \mathrm{diag}(t_{i,1}, t_{i,2} - t_{i,1}, \ldots, t_{i,m_i} - t_{i,m_i-1}).$$

3.1. Parameter estimation based on EM algorithm

Assuming that transformed degradation data $\mathrm{Da} = \{Y_1, Y_2, \ldots, Y_n\}$ of products is given, the model parameter Θ is estimated by maximizing the marginal likelihood (Chen & Tsui, 2013), where

$$\widehat{\Theta} = \arg\max_{\Theta} \prod_{i=1}^{n} \int \rho(Y_i|\theta_i, r_1, r_2, \sigma_1, \sigma_2, \tau_i)\rho(\theta_i|\mu_\theta, \Sigma_\theta)\,\mathrm{d}\theta_i. \tag{22}$$

Since the explicit expression of $\widehat{\Theta}$ is difficult to obtain, we adopt the EM algorithm to find the maximum likelihood estimates of the model parameters. First, we gain the observed data $\mathrm{Da} = \{Y_1, Y_2, \ldots, Y_n\}$, followed by the introduction of hidden parameters $\psi = \{\theta_1, \theta_2, \ldots, \theta_n\}$. At this point, the full log-likelihood function can be expressed as

$$l(\Theta|\mathrm{Da}, \psi) = l_1(\psi, r_1, r_2, \sigma_1, \sigma_2, \tau|\mathrm{Da}) + l_2(\mu_\theta, \Sigma_\theta|\psi), \tag{23}$$

where

$$l_1(\psi, r_1, r_2, \sigma_1, \sigma_2, \tau|\mathrm{Da})$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} \ln \rho(Y_i|\theta_i, r_1, r_2, \sigma_1, \sigma_2, \tau_i) \\
&= -\sum_{i=1}^{n} \frac{m_i}{2} \ln(2\pi) - \frac{1}{2}\sum_{i=1}^{n} \ln|T_i| - \sum_{i=1}^{n} c_i \ln\sigma_1 - \sum_{i=1}^{n}(m_i - c_i)\ln\sigma_2 \\
&\quad - \frac{1}{2}\sum_{i=1}^{n}(Y_i - \Lambda_i\theta_i)^\top \Delta_i T_i^{-1}\Delta_i(Y_i - \Lambda_i\theta_i),
\end{aligned} \tag{24}
$$

and

$$
\begin{aligned}
l_2(\mu_\theta, \Sigma_\theta|\psi) &= \sum_{i=1}^{n} \ln \rho(\theta_i|\mu_\theta, \Sigma_\theta) \\
&= -2n\ln(2\pi) - \frac{1}{2}n\ln|\Sigma_\theta| - \frac{1}{2}\sum_{i=1}^{n}(\theta_i - \mu_\theta)^\top \Sigma_\theta^{-1}(\theta_i - \mu_\theta).
\end{aligned} \tag{25}
$$

3.1.1. E-Step

$\Theta^{(s)} = \{\mu_\theta^{(s)}, \Sigma_\theta^{(s)}, \sigma_1^{(s)}, \sigma_2^{(s)}, r_1^{(s)}, r_2^{(s)}, \tau^{(s)}\}$ denotes the sth estimated parameters, and $Q(\Theta \mid \Theta^{(s)}, \mathrm{Da})$ denotes the conditional expectation of the full log-likelihood function $l(\Theta \mid \mathrm{Da}, \psi)$, which can be obtained by applying matrix derivation (see Appendix A for the proof):

$$
\begin{aligned}
Q(\Theta|\Theta^{(s)}, \mathrm{Da}) &= E_{\theta_i|Y_i,\Theta^{(s)}}[l(\Theta|\mathrm{Da})] \\
&= E_{\theta_i|Y_i,\Theta^{(s)}}[l_1(\psi, r_1, r_2, \sigma_1, \sigma_2, \tau|\mathrm{Da}) + l_2(\mu_\theta, \Sigma_\theta|\psi)] \\
&= -\sum_{i=1}^{n}\frac{m_i}{2}\ln(2\pi) - \frac{1}{2}\sum_{i=1}^{n}\ln|T_i| \\
&\quad - \sum_{i=1}^{n} c_i \ln\sigma_1 - \sum_{i=1}^{n}(m_i - c_i)\ln\sigma_2 \\
&\quad - \frac{1}{2}\sum_{i=1}^{n}(Y_i - \Lambda_i\nu_{\theta,i}^{(s)})^\top \Delta_i T_i^{-1}\Delta_i(Y_i - \Lambda_i\nu_{\theta,i}^{(s)}) \\
&\quad - \frac{1}{2}\sum_{i=1}^{n}\mathrm{tr}(\Lambda_i^\top \Delta_i T_i^{-1}\Delta_i\Lambda_i C_{\theta,i}^{(s)}) \\
&\quad - 2n\ln(2\pi) - \frac{1}{2}n\ln|\Sigma_\theta| \\
&\quad - \frac{1}{2}\sum_{i=1}^{n}(\nu_{\theta,i}^{(s)} - \mu_\theta)^\top \Sigma_\theta^{-1}(\nu_{\theta,i}^{(s)} - \mu_\theta) - \frac{1}{2}\sum_{i=1}^{n}\mathrm{tr}(\Sigma_\theta^{-1}C_{\theta,i}^{(s)}),
\end{aligned} \tag{26}
$$

in which

$$
\begin{aligned}
\nu_{\theta,i}^{(s)} &= E(\theta_i|Y_i, \Theta^{(s)}) \\
&= [(\Lambda_i^{(s)})^\top \Delta_i^{(s)} T_i^{-1}\Delta_i^{(s)}\Lambda_i^{(s)} + (\Sigma_\theta^{(s)})^{-1}]^{-1}
\end{aligned}
$$

$$\times\,[(Y_i^{(s)})^\top \Delta_i^{(s)} T_i^{-1} \Delta_i^{(s)} \Lambda_i^{(s)} + (\mu_\theta^{(s)})^\top (\Sigma_\theta^{(s)})^{-1}]^\top, \tag{27}$$

$$C_{\theta,i}^{(s)} = \mathrm{Var}(\theta_i | Y_i, \Theta^{(s)}) = [(\Lambda_i^{(s)})^\top \Delta_i^{(s)} T_i^{-1} \Delta_i^{(s)} \Lambda_i^{(s)} + (\Sigma_\theta^{(s)})^{-1}]^{-1}, \tag{28}$$

and $\mathrm{tr}(\cdot)$ denotes the trace of a matrix, where

$$\Lambda_i^{(s)} = \begin{bmatrix} \Lambda_{i,1}^{(s)} & \\ & \Lambda_{i,2}^{(s)} \end{bmatrix}, \tag{29}$$

$$\Lambda_{i,1}^{(s)} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ t_{i,1}^{r_1^{(s)}} & t_{i,2}^{r_1^{(s)}} - t_{i,1}^{r_1^{(s)}} & \cdots & t_{i,c_i^{(s)}}^{r_1^{(s)}} - t_{i,c_i^{(s)}-1}^{r_1^{(s)}} \end{pmatrix}^\top, \tag{30}$$

$$\Lambda_{i,2}^{(s)} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ (t_{i,c_i^{(s)}+1} - t_{i,c_i^{(s)}})^{r_2^{(s)}} & (t_{i,c_i^{(s)}+2} - t_{i,c_i^{(s)}})^{r_2^{(s)}} - (t_{i,c_i^{(s)}+1} - t_{i,c_i^{(s)}})^{r_2^{(s)}} & \cdots & (t_{i,m_i} - t_{i,c_i^{(s)}})^{r_2^{(s)}} - (t_{i,m_i-1} - t_{i,c_i^{(s)}})^{r_2^{(s)}} \end{pmatrix}^\top, \tag{31}$$

$$\Delta_i^{(s)} = (\sigma_1^{(s)})^{-1} \Delta_{i,1}^{(s)} + (\sigma_2^{(s)})^{-1} \Delta_{i,2}^{(s)}, \tag{32}$$

$$\Delta_{i,1}^{(s)} = \mathrm{diag}(I(t_{i,1} \le \tau_i^{(s)}), I(t_{i,2} \le \tau_i^{(s)}), \ldots, I(t_{i,m_i} \le \tau_i^{(s)})), \tag{33}$$

$$\Delta_{i,2}^{(s)} = \mathrm{diag}(I(t_{i,1} > \tau_i^{(s)}), I(t_{i,2} > \tau_i^{(s)}), \ldots, I(t_{i,m_i} > \tau_i^{(s)})), \tag{34}$$

$C_i^{(s)}$ is the subscript and $t_{i,c_i^{(s)}} = \tau_i^{(s)}$.

3.1.2. M-Step

The M-Step solves $\Theta^{(s+1)}$ by maximizing $Q(\Theta|\Theta^{(s)}, \mathrm{Da})$. Since $\tau = (\tau_1, \tau_2, \ldots, \tau_n)^\top$ in model parameter $\Theta = \{\mu_\theta, \Sigma_\theta, \sigma_1, \sigma_2, r_1, r_2, \tau\}$ is discrete and r_1 and r_2 are exponential parts of the model, it is difficult to obtain $\tau^{(s+1)}$, $r_1^{(s+1)}$ and $r_2^{(s+1)}$ directly by maximizing $Q(\Theta \mid \Theta^{(s)}, \mathrm{Da})$. Therefore, we first derive the explicit expressions for $\mu_\theta^{(s+1)}$, $\Sigma_\theta^{(s+1)}$, $\sigma_1^{(s+1)}$, $\sigma_2^{(s+1)}$ as shown in Equations (35)–(38) below:

$$\mu_\theta^{(s+1)} = \frac{1}{n} \sum_{i=1}^n \nu_{\theta,i}^{(s)}, \tag{35}$$

$$\Sigma_\theta^{(s+1)} = \frac{1}{n} \left[\sum_{i=1}^n (\nu_{\theta,i}^{(s)} - \mu_\theta^{(s+1)})(\nu_{\theta,i}^{(s)} - \mu_\theta^{(s+1)})^\top + \sum_{i=1}^n C_{\theta,i}^{(s)} \right], \tag{36}$$

$$\begin{aligned} \sigma_1^{(s+1)} = \frac{1}{\sum_{i=1}^n c_i^{(s+1)}} &\left[\sum_{i=1}^n (Y_i - \Lambda_i^{(s+1)} \nu_{\theta,i}^{(s)})^\top \Delta_{i,1}^{(s+1)} T_i^{-1} \Delta_i^{(s)} (Y_i - \Lambda_i^{(s+1)} \nu_{\theta,i}^{(s)}) \right. \\ &\left. + \sum_{i=1}^n \mathrm{tr}((\Lambda_i^{(s+1)})^\top \Delta_{i,1}^{(s+1)} T_i^{-1} \Delta_i^{(s)} \Lambda_i^{(s+1)} C_{\theta,i}^{(s)}) \right], \end{aligned} \tag{37}$$

$$\begin{aligned} \sigma_2^{(s+1)} = \frac{1}{\sum_{i=1}^n (m_i - c_i^{(s+1)})} &\left[\sum_{i=1}^n (Y_i - \Lambda_i^{(s+1)} \nu_{\theta,i}^{(s)})^\top \Delta_{i,2}^{(s+1)} T_i^{-1} \Delta_i^{(s)} (Y_i - \Lambda_i^{(s+1)} \nu_{\theta,i}^{(s)}) \right. \\ &\left. + \sum_{i=1}^n \mathrm{tr}((\Lambda_i^{(s+1)})^\top \Delta_{i,2}^{(s+1)} T_i^{-1} \Delta_i^{(s)} \Lambda_i^{(s+1)} C_{\theta,i}^{(s)}) \right], \end{aligned} \tag{38}$$

where $c_i^{(s+1)}$, $\Lambda_i^{(s+1)}$, $\Delta_i^{(s)}$, $\Delta_{i,1}^{(s+1)}$ and $\Delta_{i,2}^{(s+1)}$ are defined in the same way as Equation (26).

From Equations (35) and (36), it can be seen that the values of $\mu_\theta^{(s+1)}$, $\Sigma_\theta^{(s+1)}$ are independent of $\tau^{(s+1)}$, $r_1^{(s+1)}$ and $r_2^{(s+1)}$. Therefore, the values of $\mu_\theta^{(s+1)}$, $\Sigma_\theta^{(s+1)}$ can be obtained directly.

Then, σ_1 and σ_2 in Equation (26) are substituted with Equations (37) and (38). Thus, Equation (26) is a function of τ, r_1 and r_2, which is denoted as $Q(\tau, r_1, r_2 \mid \sigma_1^{(s+1)}, \sigma_2^{(s+1)}, \mathrm{Da})$, as shown in Equation (39):

$$\begin{aligned} &Q(\tau, r_1, r_2 \mid \sigma_1^{(s+1)}, \sigma_2^{(s+1)}, \mathrm{Da}) \\ &\propto -\sum_{i=1}^{n} c_i^{(s)} \ln\left(\frac{\sum_{i=1}^{n} (Y_i - \Lambda_i^{(s+1)} \nu_{\theta,i}^{(s)})^\top \Delta_{i,1}^{(s)} T_i^{-1} \Delta_i^{(s+1)} (Y_i - \Lambda_i^{(s+1)} \nu_{\theta,i}^{(s)})}{\sum_{i=1}^{n} c_i^{(s)}} \right. \\ &\quad \left. + \frac{\sum_{i=1}^{n} \mathrm{tr}((\Lambda_i^{(s+1)})^\top \Delta_{i,1}^{(s+1)} T_i^{-1} \Delta_i^{(s+1)} \Lambda_i^{(s+1)} C_{\theta,i}^{(s)})}{\sum_{i=1}^{n} c_i^{(s)}} \right) \\ &\quad - \sum_{i=1}^{n} (m_i - c_i^{(s)}) \ln\left(\frac{\sum_{i=1}^{n} (Y_i - \Lambda_i^{(s+1)} \nu_{\theta,i}^{(s)})^\top \Delta_{i,2}^{(s+1)} T_i^{-1} \Delta_i^{(s+1)} (Y_i - \Lambda_i^{(s+1)} \nu_{\theta,i}^{(s)})}{\sum_{i=1}^{n} (m_i - c_i^{(s)})} \right. \\ &\quad \left. + \frac{\sum_{i=1}^{n} \mathrm{tr}((\Lambda_i^{(s+1)})^\top \Delta_{i,2}^{(s+1)} T_i^{-1} \Delta_i^{(s+1)} \Lambda_i^{(s+1)} C_{\theta,i}^{(s)})}{\sum_{i=1}^{n} (m_i - c_i^{(s)})} \right). \end{aligned} \tag{39}$$

Let $h = (\tau, r_1, r_2)^\top$, and then denote $Q(\tau, r_1, r_2 \mid \sigma_1^{(s+1)}, \sigma_2^{(s+1)}, \mathrm{Da})$ as $Q(h \mid \sigma_1^{(s+1)}, \sigma_2^{(s+1)}, \mathrm{Da})$. At this point, the solution for $h^{(s+1)}$ can be obtained by maximizing $Q(h \mid \sigma_1^{(s+1)}, \sigma_2^{(s+1)}, \mathrm{Da})$, that is

$$h^{(s+1)} = \arg\max_h Q(h \mid \sigma_1^{(s+1)}, \sigma_2^{(s+1)}, \mathrm{Da}). \tag{40}$$

After the solutions of $h^{(s+1)}$ are obtained, bringing them into Equations (37) and (38), the values of $\sigma_1^{(s+1)}$ and $\sigma_2^{(s+1)}$ are obtained.

The EM algorithm will be converged as meeting one of the following two conditions: (1) it iterates between E-step and M-step until the difference between $\Theta^{(s)}$ and $\Theta^{(s+1)}$ is smaller than a pre-defined value δ, i.e., $\|\Theta^{(s+1)} - \Theta^{(s)}\| < \delta$, such that the relative change value for the conditional expectation of the full log-likelihood function $l(\Theta \mid \mathrm{Da}, \psi)$ between two consecutive iterations is less than the threshold value (10^{-6}), i.e., $\frac{|Q(\Theta|\Theta^{(s+1)},\mathrm{Da}) - Q(\Theta|\Theta^{(s)},\mathrm{Da})|}{Q(\Theta|\Theta^{(s)},\mathrm{Da})} < 10^{-6}$; (2) $s_{\max}(800)$ means that the maximum iteration number is reached, at which point the algorithm stops iterating.

Based on the iterative process described above, the mean and variance estimates of the change point can be obtained from Equation (41) (A. Zhang et al., 2020):

$$\begin{cases} \hat{\mu}_\tau = \frac{\sum_{i=1}^{n} \tau_i}{n}, \\ \hat{\sigma}_\tau^2 = \frac{\sum_{i=1}^{n} (\tau_i - \hat{\mu}_\tau)^2}{n-1}. \end{cases} \tag{41}$$

In addition, to improve the accuracy of parameter estimation of the model of Equation (6), the initial values of the parameters were predicted by the heuristic methods of the literature (H. Wang et al., 2020; Z. S. Ye et al., 2013; A. Zhang et al., 2020).

Table 1. The specified model parameters for simulation.

Model parameters	Parameter settings
μ_θ	$[377.27 \quad 3.6005\times10^{-5} \quad 378.40 \quad 3.2246\times10^{-3}]$
Σ_θ	$\begin{bmatrix} 0.9537\times10^{2} & 1.4290\times10^{-4} & 0.9197\times10^{2} & 2.6833\times10^{-4} \\ 1.4290\times10^{-4} & 2.1805\times10^{-10} & 1.3776\times10^{-4} & 1.8624\times10^{-9} \\ 0.9197\times10^{2} & 1.3776\times10^{-4} & 8.8698\times10^{1} & 2.4155\times10^{-4} \\ 2.6833\times10^{-4} & 1.8624\times10^{-9} & 2.4155\times10^{-4} & 5.4396\times10^{-7} \end{bmatrix}$
σ_1	7.4544×10^{-2}
σ_2	11.8469×10^{-2}
r_1	4.6185
r_2	2.5114

4. Simulation

A simulation is carried out in this section. Through the simulation, we focus on the performance of parameter estimation using the EM algorithm.

Firstly, a simulation is provided to investigate the effect of sample size on the performance of the proposed method. We consider nine experiments with sample sizes ranging from 2 to 10, respectively, and parameters of the true models are specified in Table 1. In this case, the change point is generated from $N(6.6090, 0.2907)$, and the failure threshold D of the product is 440.

The simulation was repeated 100 times for each experiment, and the average degradation paths are plotted for comparison in Figure 1. The average degradation path curves obtained by the proposed method exhibit high consistency under different sample sizes, indicating that sample size has a minimal effect on the performance of the proposed method. Then the proposed method has better stability for the sample size. It is worth noting that the differences of curves under different sample sizes are primarily concentrated in the transition region between the two degradation stages (i.e., near the change point), with these differences primarily manifesting as fluctuations amplitude of change-point position. This phenomenon showed that accurate identification of a change point is critical to the model performance, possibly because the degradation mechanism transition process near change point is more sensitive to the sample size. Additionally, the curves exhibit high consistency in other regions, further validating that the proposed method has better robustness during the main degradation phase.

Secondly, to investigate the robustness of the proposed method, we set up four different sets of model parameter truth values (as shown from Tables 2– 5) for simulation. The first and second groups both describe upward-trending degradation paths; while the third and fourth groups both show downward-trending degradation paths. Meanwhile, the change points of the four simulation experiments are generated from $N(6, 0.5)$. Then, the degradation rate of the first stage is faster than that of the second stage in the degradation paths of both the first and third groups of simulation experiments; the opposite is true for the second and fourth groups of simulation experiments.

The detailed steps are as follows.

(1) We randomly generate degradation data of products from the two-stage nonlinear phase-dependent Wiener process degradation model of Equation (6) (Per simulation: 5 products with 14 degradation data each).

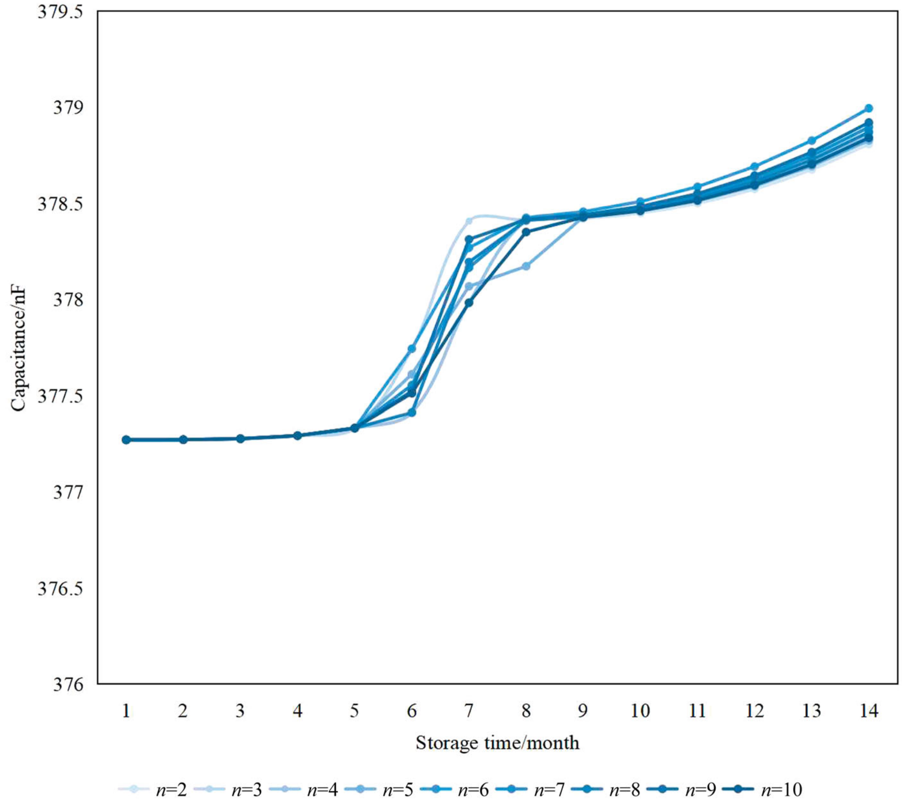

Figure 1. Comparison results of fitted mean degradation paths under different sample sizes.

Table 2. The specified model parameters for simulation (Group 1).

Model parameters	Parameter settings
μ_θ	$[300 \quad 9.6\times10^{-5} \quad 302 \quad 8.2\times10^{-3}]$
Σ_θ	$\begin{bmatrix} 0.9531\times10^{-7} & -4.0452\times10^{-12} & 1.0616\times10^{-6} & -7.3737\times10^{-10} \\ -4.0452\times10^{-12} & 9.7689\times10^{-15} & -2.6860\times10^{-11} & 2.1815\times10^{-14} \\ 1.0616\times10^{-6} & -2.6860\times10^{-11} & 6.0471\times10^{-4} & -6.3461\times10^{-8} \\ -7.3737\times10^{-10} & 2.1815\times10^{-14} & -6.3461\times10^{-8} & 1.0223\times10^{-10} \end{bmatrix}$
σ_1	1.0×10^{-3}
σ_2	0.5
r_1	2
r_2	5.5
τ	[6.7613 6.6502 5.8223 6.0287 5.9956]

(2) The model parameters are estimated based on the generated degradation data using the proposed methodology.
(3) The first two steps were repeated 1000 times to obtain 1000 estimates of the model parameters.
(4) Seek the average and mean-squared error of the parameter estimation values for these 1000 times.

The mean values and mean square errors of the 1000 estimates of model parameters in the four sets of simulation experiments are shown in Tables 6– 9. Among them, we denote

Table 3. The specified model parameters for simulation (Group 2).

Model parameters	Parameter settings
μ_θ	$[300 \quad 6.2 \times 10^{-3} \quad 302 \quad 4.0 \times 10^{-5}]$
Σ_θ	$\begin{bmatrix} 1.3539 \times 10^{-6} & 2.3441 \times 10^{-10} & 9.8690 \times 10^{-6} & -6.5125 \times 10^{-9} \\ 2.3441 \times 10^{-10} & 6.1595 \times 10^{-11} & 2.7662 \times 10^{-8} & 2.3970 \times 10^{-12} \\ 9.8690 \times 10^{-6} & 2.7662 \times 10^{-8} & 1.4057 \times 10^{-4} & -7.6496 \times 10^{-8} \\ -6.5125 \times 10^{-9} & 2.3970 \times 10^{-12} & -7.6496 \times 10^{-8} & 7.2042 \times 10^{-11} \end{bmatrix}$
σ_1	1.0×10^{-3}
σ_2	2.8×10^{-1}
r_1	3
r_2	6
τ	[6.4613 6.5502 4.9223 5.0587 4.8956]

Table 4. The specified model parameters for simulation (Group 3).

Model parameters	Parameter settings
μ_θ	$[302 \quad -6.2 \times 10^{-3} \quad 300 \quad -6.5 \times 10^{-5}]$
Σ_θ	$\begin{bmatrix} 5.5531 \times 10^{-7} & 6.3134 \times 10^{-10} & 7.9892 \times 10^{-8} & 5.8950 \times 10^{-9} \\ 6.3134 \times 10^{-10} & 1.0851 \times 10^{-11} & 3.7257 \times 10^{-9} & 7.5503 \times 10^{-12} \\ 7.9892 \times 10^{-8} & 3.7257 \times 10^{-9} & 5.0229 \times 10^{-4} & -8.2662 \times 10^{-9} \\ 5.8950 \times 10^{-9} & 7.5503 \times 10^{-12} & -8.2662 \times 10^{-9} & 8.3696 \times 10^{-11} \end{bmatrix}$
σ_1	1.0×10^{-3}
σ_2	0.2
r_1	3
r_2	5.5
τ	[6.6613 6.5502 5.6223 5.0587 5.8956]

Table 5. The specified model parameters for simulation (Group 4).

Model parameters	Parameter settings
μ_θ	$[302 \quad -6.5 \times 10^{-5} \quad 300 \quad -6.2 \times 10^{-3}]$
Σ_θ	$\begin{bmatrix} 0.7465 \times 10^{-6} & -4.7503 \times 10^{-13} & 6.8090 \times 10^{-7} & -4.6559 \times 10^{-10} \\ -4.7503 \times 10^{-13} & 6.1416 \times 10^{-16} & -3.0631 \times 10^{-12} & 2.7247 \times 10^{-15} \\ 6.8090 \times 10^{-7} & -3.0631 \times 10^{-12} & 6.1578 \times 10^{-4} & -6.3074 \times 10^{-8} \\ -4.6559 \times 10^{-10} & 2.7247 \times 10^{-15} & -6.3074 \times 10^{-8} & 1.0114 \times 10^{-10} \end{bmatrix}$
σ_1	1.0×10^{-3}
σ_2	1.3
r_1	6
r_2	2
τ	[6.7613 6.6502 5.7223 4.7287 5.8956]

the components in μ_θ and τ as $\mu_{\theta 1} \sim \mu_{\theta 4}$ and $\tau_1 \sim \tau_5$, respectively, and the determinant of Σ_θ is calculated to enable comprehensive evaluation.

From Tables 6–9, both the model parameter estimation deviations and mean-squared errors remain within 10% of the true values for all four simulation experiments. The simulation results indicate minimal deviations between the parameter estimates and their true values. In addition, we note that the mean square errors of the model parameters are also very small, indicating the estimation of the EM algorithm has good accuracy and robustness in this paper.

Table 6. Comparison of the results of the first group of simulation experiments.

Model parameter	True value	Average value	Deviation from the mean	Mean square error	Percentage of mean square error
$\mu_{\theta 1}$	300	300.0000087	0.000002890%	$1.22711542 \times 10^{-7}$	0.00000004090%
$\mu_{\theta 2}$	9.6×10^{-5}	$9.60555661 \times 10^{-5}$	0.057881354%	$1.09817947 \times 10^{-12}$	0.00000114394%
$\mu_{\theta 3}$	302	302.00013837	0.000045818%	$5.00350586 \times 10^{-6}$	0.00000165679%
$\mu_{\theta 4}$	8.2×10^{-3}	$8.20000943 \times 10^{-3}$	0.000115000%	$3.56666839 \times 10^{-12}$	0.00000004350%
$\lvert\Sigma_\theta\rvert$	4.9551×10^{-35}	$4.84652959 \times 10^{-35}$	2.191084136%	$5.01392879 \times 10^{-69}$	0%
σ_1	1.0×10^{-3}	$0.90210685 \times 10^{-3}$	9.789315000%	$3.35385773 \times 10^{-8}$	0.00335385773%
σ_2	0.5	0.50345901	0.691802000%	1.5871×10^{-4}	0.03174200000%
r_1	5.5	5.49969319	0.005578364%	$3.93618893 \times 10^{-5}$	0.00071567071%
r_2	2	2.0102702	0.513510000%	1.1682×10^{-4}	0.00584100000%
τ_1	6.5613	6.7783456	3.307966409%	0.002558357	0.03899161345%
τ_2	6.5502	6.78082367	3.520864554%	0.02851826	0.43537998840%
τ_3	5.8223	5.84539435	0.396653384%	0.00307359	0.05278996273%
τ_4	6.0587	5.98933229	1.144927295%	0.38430019	6.34294799214%
τ_5	5.9956	5.93185709	1.063161485%	0.00619933	0.10339799186%

Table 7. Comparison of the results of the second group of simulation experiments.

Model parameter	True value	Average value	Deviation from the mean	Mean square error	Percentage of mean square error
$\mu_{\theta 1}$	300	299.99997643	0.00000786%	$1.51873503 \times 10^{-7}$	0.00000005%
$\mu_{\theta 2}$	6.2×10^{-3}	$6.20185935 \times 10^{-3}$	0.02998948%	$2.40235240 \times 10^{-10}$	0.00000387%
$\mu_{\theta 3}$	302	301.99859563	0.00046502%	$7.18584738 \times 10^{-5}$	0.00002379%
$\mu_{\theta 4}$	4.0×10^{-5}	$4.12735902 \times 10^{-5}$	3.18397561%	$1.63797271 \times 10^{-12}$	0.00000409%
$\lvert\Sigma_\theta\rvert$	1.17×10^{-31}	$1.08911723 \times 10^{-31}$	7.06397869%	$3.84829548 \times 10^{-62}$	0%
σ_1	1.0×10^{-3}	$1.00906848 \times 10^{-3}$	0.90684844%	$4.65992192 \times 10^{-8}$	0.00465992%
σ_2	0.28	0.29334795	4.76712507%	$2.08376419 \times 10^{-4}$	0.07442015%
r_1	3	2.99982693	0.00576904%	$3.85736335 \times 10^{-6}$	0.00012858%
r_2	6	6	0%	0	0%
τ_1	6.4613	6.61573366	2.39013290%	0.57342085	8.87469774%
τ_2	6.5502	6.65944358	1.66778998%	0.31646614	4.83139665%
τ_3	4.9223	4.86354885	1.19357116%	0.00500687	0.10171819%
τ_4	5.0587	4.69434095	7.20262229%	0.41876488	8.27811252%
τ_5	4.8956	4.86147730	0.69700748%	0.00236622	0.04833359%

Table 8. Comparison of the results of the third group of simulation experiments.

Model parameter	True value	Average value	Deviation from the mean	Mean square error	Percentage of mean square error
$\mu_{\theta 1}$	302	302.00001085	0.00000359%	$1.41606403 \times 10^{-7}$	0.00000005%
$\mu_{\theta 2}$	-6.2×10^{-3}	$-6.21102429 \times 10^{-3}$	0.17781118%	$4.01085669 \times 10^{-10}$	0.00000647%
$\mu_{\theta 3}$	300	299.99509718	0.00163427%	$1.19861148 \times 10^{-4}$	0.00003995%
$\mu_{\theta 4}$	-6.5×10^{-5}	$-6.66117892 \times 10^{-5}$	2.47967570%	$2.06271133 \times 10^{-10}$	0.00031734%
$\lvert\Sigma_\theta\rvert$	5.883×10^{-32}	$5.65083699 \times 10^{-32}$	3.94633714%	$3.13972894 \times 10^{-63}$	0%
σ_1	1.0×10^{-3}	$0.91562009 \times 10^{-3}$	8.43799093%	$3.68264812 \times 10^{-8}$	0.00368265%
σ_2	0.2	0.18716317	6.41841521%	$2.76650220 \times 10^{-4}$	0.13832511%
r_1	3	2.99897897	0.03403428%	$3.92699916 \times 10^{-6}$	0.00013090%
r_2	5.5	5.57413459	1.34790162%	$1.53239148 \times 10^{-2}$	0.27861663%
τ_1	6.6613	6.76726759	1.59079450%	0.11526876	1.73042441%
τ_2	6.5502	6.76318055	3.25151219%	0.18514301	2.82652460%
τ_3	5.6223	5.77730910	2.75704075%	0.09872383	1.75593315%
τ_4	5.0587	4.86802267	3.76929505%	0.09505762	1.87909193%
τ_5	5.8956	5.88021949	0.26088110%	0.00132794	0.02252423%

Table 9. Comparison of the results of the fourth group of simulation experiments.

Model parameter	True value	Average value	Deviation from the mean	Mean square error	Percentage of mean square error
$\mu_{\theta 1}$	302	301.9999987	0.000000420%	$1.36739414 \times 10^{-7}$	0.00000004528%
$\mu_{\theta 2}$	-6.5×10^{-5}	$-6.50009761 \times 10^{-5}$	0.001501746%	$3.03271919 \times 10^{-13}$	0.00000046657%
$\mu_{\theta 3}$	300	299.9999777	0.000007449%	$5.80764691 \times 10^{-6}$	0.00000193588%
$\mu_{\theta 4}$	-6.2×10^{-3}	$-6.19999419 \times 10^{-3}$	0.000093759%	$2.98766535 \times 10^{-12}$	0.00000004819%
$\lvert\Sigma_\theta\rvert$	2.6629×10^{-35}	$2.68522671 \times 10^{-35}$	0.838435827%	$9.53806107 \times 10^{-69}$	0%
σ_1	1.0×10^{-3}	$0.91379511 \times 10^{-3}$	8.620488687%	$3.54171701 \times 10^{-8}$	0.00354171701%
σ_2	1.3	1.303131604	0.240892604%	$2.76650220 \times 10^{-4}$	0.00495770661%
r_1	6	6.000017477	0.000291285%	$6.44501859 \times 10^{-5}$	0.00040865060%
r_2	2	2.018395269	0.919763470%	$2.45190362 \times 10^{-5}$	0.02232750292%
τ_1	6.7613	6.749046209	0.181234238%	0.002441371	0.03610800594%
τ_2	6.6502	6.751522313	1.523597978%	0.012657813	0.19033732287%
τ_3	5.7223	5.870126472	2.583340125%	0.023681579	0.41384721082%
τ_4	4.7287	4.876015219	3.115342884%	0.022663242	0.47927003352%
τ_5	5.8956	5.899354057	0.063675579%	0.000194184	0.00329370962%

5. Real data analysis

In this section, the degradation data of high-voltage pulse capacitors validate the rationality and effectiveness of the proposed model and are compared with the study in the literature (A. Zhang et al., 2020).

Prolonged storage leads to progressive capacitance increase in high-voltage pulse capacitors, ultimately causing failure. The capacitance of five units of a certain type of high-voltage pulse capacitor was measured at one-month intervals in the literature (H. F. Ye et al., 2007) shown in Table 10.

The relative change in the capacitance of the high-voltage pulse capacitor over time is given in Figure 2. The degradation data in Figure 2 were modelled and analysed using a two-stage nonlinear model based on phase independence in the literature (A. Zhang et al., 2020). For further comparative analysis and validation, the degradation model in literature (A. Zhang et al., 2020) is used as a reference method for comparative study, denoted as M_2, while the proposed method in this paper is denoted as M_1.

From Figure 2, the degradation path of high-voltage pulse capacitors presents a distinct two-stage degradation characteristic. In the first stage, the degradation rate is relatively slow, exhibiting a nonlinear acceleration trend; in the second stage, the degradation rate accelerates significantly, presenting a steeper nonlinear growth curve. This two-stage nonlinear degradation characteristic may be closely related to the gradual aging of internal insulating materials of the capacitor and the intensification of partial discharge phenomena. It is noteworthy that the change point between the two degradation stages exhibits the individual variability, reflecting minor differences in material properties or manufacturing processes among different capacitor samples. This two-stage nonlinear degradation pattern provides an important theoretical basis for accurately predicting the RUL of high-voltage pulse capacitors. Therefore, the degradation path of the high-voltage pulse capacitors is described by the two-stage nonlinear phase-dependent Wiener process degradation model proposed in this paper.

To compare the goodness-of-fit of M_1 and M_2, we measure them by the Akaike information criterion (AIC), which is defined as

$$\mathrm{AIC} = 2k - 2\ln(L),$$

Table 10. Capacitance of 5 high-voltage pulse capacitors (unit: nF).

Sample number	Number of measurements													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
C1	411.37	411.55	411.36	411.34	411.51	411.75	412.03	413.06	412.95	413.05	413.26	413.47	413.74	414.09
C2	412.45	412.57	412.41	412.41	412.56	412.79	413.06	414.01	413.85	413.87	414.16	414.39	414.63	414.46
C3	408.87	408.96	408.72	408.74	408.89	409.13	409.44	410.31	410.20	410.28	410.45	410.68	410.86	411.24
C4	429.80	429.72	429.78	429.81	430.05	430.32	430.54	430.61	430.71	430.85	430.93	431.14	431.39	431.75
C5	433.48	433.20	433.29	433.33	433.52	433.90	434.31	434.28	434.30	434.45	434.71	434.90	435.18	435.60

Note: Unit of all time points in the table is month.

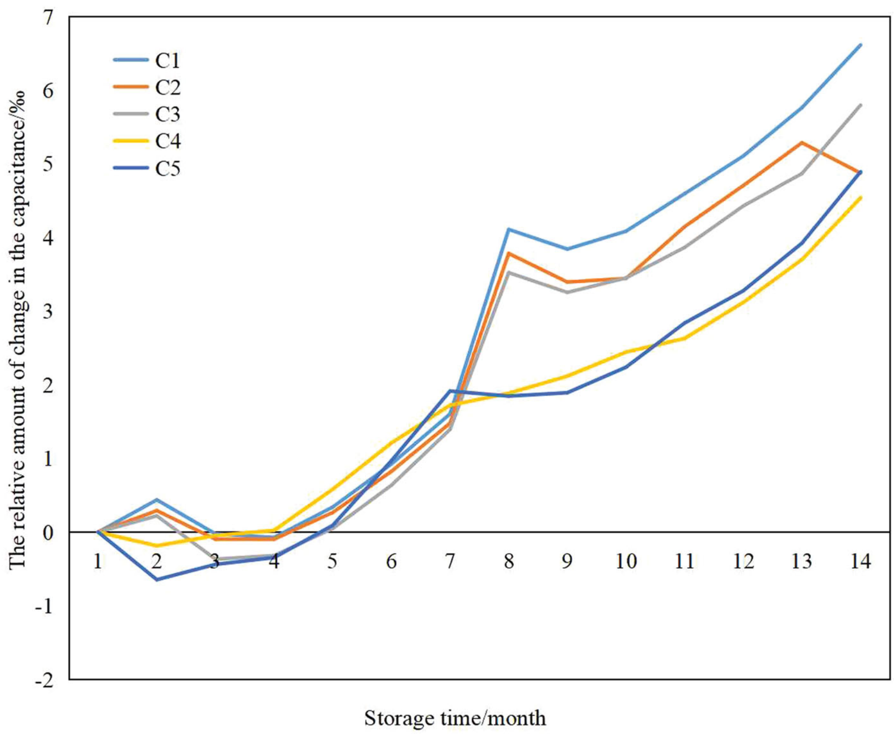

Figure 2. Degradation paths of the relative amount of change in the capacitance.

Table 11. Comparison of AIC values of M_1 and M_2.

Model	M_1	M_2
k	23	13
$\ln(L)$	161.209	−7.164
AIC	−276.418	40.328

Note: M_1 is the proposed method. M_2 is the method of literature (A. Zhang et al., 2020).

where k denotes the number of unknown parameters in the model, and $\ln(L)$ represents the log-likelihood function value.

The values of AIC of the two models in Table 11 show that M_1 has a smaller AIC value and can fit the degradation path of the high-voltage impulse capacitors more reasonably and efficiently.

5.1. Parameter estimation

The degradation data from five high-voltage impulse capacitors was used to produce parameter estimates using M_1 for modeling analysis. The estimated model parameters of Methods M_1 and M_2 are shown in Table 12 (note: for the convenience of comparison, we have replaced the model parameters of Method M_2, which have the same meaning in Method M_1, accordingly).

Table 12. Estimated parameters of our model.

Model parameters	Method M_1	Model parameters	Method M_2
μ_θ	$[419.1879 \quad 4.0006 \times 10^{-5} \quad 420.4440 \quad 3.5829 \times 10^{-3}]$	$\beta_{1,i}$	1.857×10^{-4}
Σ_θ	$\begin{bmatrix} 1.0597 \times 10^{2} & 1.5878 \times 10^{-4} & 1.0219 \times 10^{2} & 2.9814 \times 10^{-4} \\ 1.5878 \times 10^{-4} & 2.4228 \times 10^{-10} & 1.5307 \times 10^{-4} & 2.0693 \times 10^{-9} \\ 1.0219 \times 10^{2} & 1.5307 \times 10^{-4} & 9.8553 \times 10^{1} & 2.6839 \times 10^{-4} \\ 2.9814 \times 10^{-4} & 2.0693 \times 10^{-9} & 2.6839 \times 10^{-4} & 6.0440 \times 10^{-7} \end{bmatrix}$	$\beta_{2,j}$	5.242×10^{-2}
σ_1	8.2827×10^{-2}	σ_1	40.1995×10^{-2}
σ_2	13.1632×10^{-2}	σ_2	22.5344×10^{-2}
r_1	5.1317	r_1	5.125
r_2	2.7904	r_2	2.048
τ	[7.0134 6.9958 6.9990 6.0128 6.0242]	τ	[7.004 7.000 6.978 5.583 5.879]

Note: M_1 is the proposed method. M_2 is the method of literature (A. Zhang et al., 2020).

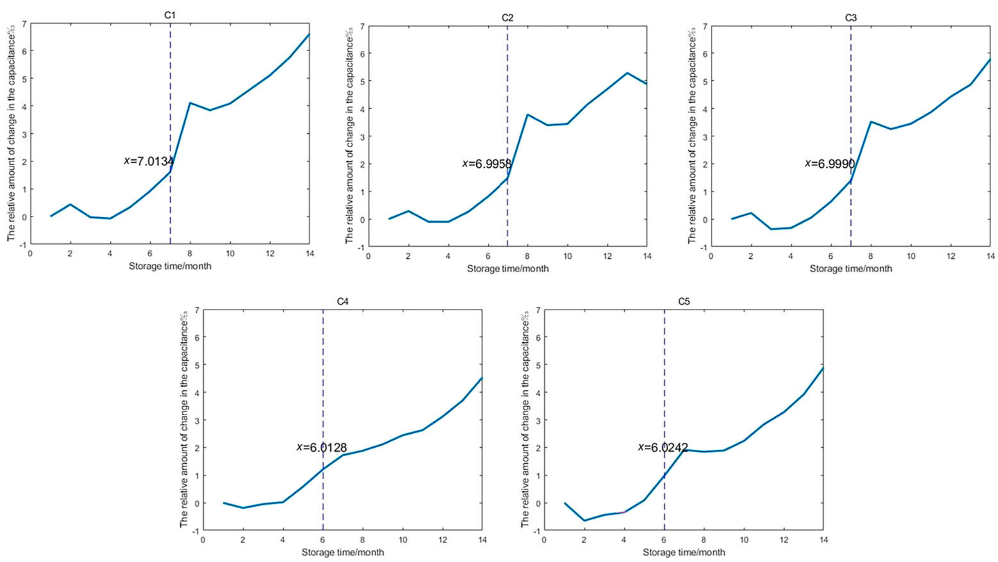

Figure 3. The estimates of the change points.

First, the location of the change point is determined according to the method proposed in the literature (H. Wang et al., 2020), which is to compare the goodness-of-fit of the two models of Equation (2) and Equation (6) by using AIC and Bayesian Information Criterion (BIC). If the model of Equation (6) is superior to the model of Equation (2), the degradation is considered to have entered the second stage and the change point is detected; conversely, the degradation is considered to be still in the first stage. In addition, if the change point has been identified in historical monitoring of degradation data, no further localization is required for subsequent monitoring. As shown in Table 12, the proposed method yields change-point estimates of (7.0134, 6.9958, 6.9990, 6.0128, and 6.0242) months

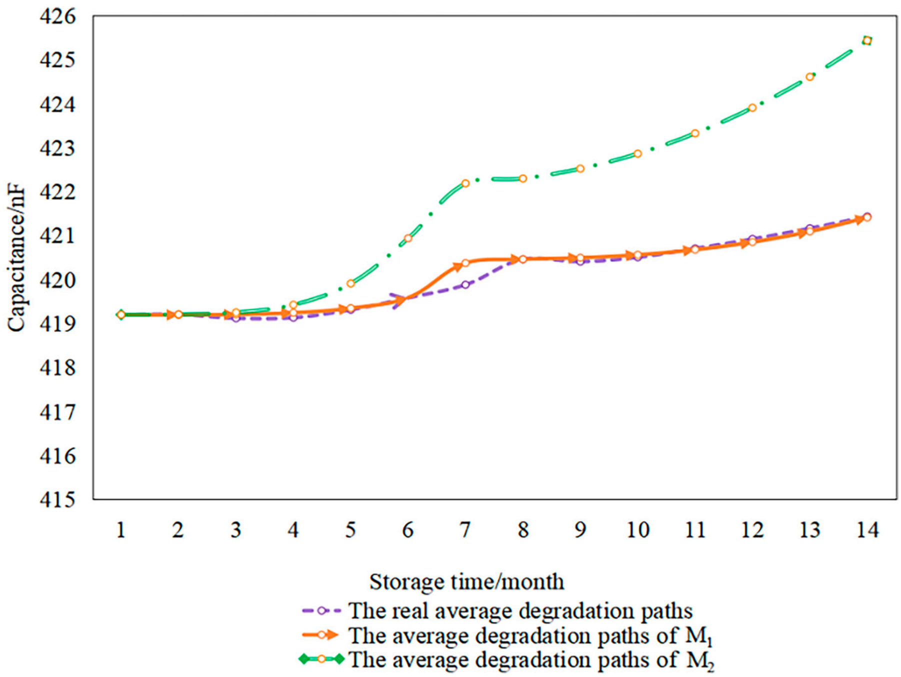

Figure 4. Comparative results of mean degradation curves.

for the five high-voltage pulse capacitors, with corresponding visualizations provided in Figure 3.

From the aforementioned change-point results, it can be concluded that the degradation change points of high-voltage pulse capacitors are all located between measurement time points. Moreover, there are obvious differences in the positions of change points among different high-voltage pulse capacitors, with significant variations in the model parameters of both stages, which is also consistent with the characteristics of nonlinear degradation. Among them, the drift parameter of the first stage is smaller than that of the second stage, and $r_1 > r_2$, which corresponds to the actual situation that the degradation rate of the high-voltage pulse capacitors slows down with the extension of storage time.

In addition, from Table 12, the estimates of model parameters obtained by Methods M_1 and M_2 are generally trending relatively the same, indicating that the proposed method has a good accuracy. Moreover, there is a large difference between the obtained estimates of the five parameters $\beta_{1,i}$, $\beta_{2,i}$, σ_1, σ_2 and r_2 and that of Method M_2, because the proposed method takes the correlation between the two degradation stages of the high-voltage pulse capacitors into account.

To further verify the rationality and effectiveness of the proposed method, the average degradation paths of M_1 and M_2 are given in Figure 4 and compared with the real average degradation paths of five high-voltage pulse capacitors. From Figure 4, the average degradation path of M_1 is closer to the real path, while that of M_2 exhibits a more pronounced deviation, which indicates that the consideration of the correlation between degradation phases is more in line with the actual degradation of high-voltage impulse capacitors, and demonstrates that the proposed method can better capture the characteristics of degradation of product.

Table 13. Comparative results of lifetime prediction.

Model	M_1	M_2
$t_{0.5}$	73.24	34.21
$t_{0.1}$	73.31	33.59

Note: M_1 is the proposed method. M_2 is the method of literature (A. Zhang et al., 2020).

5.2. Lifetime prediction

In this subsection, the degradation data of high-voltage pulse capacitors is used for lifetime prediction. The five high-voltage pulse capacitors used in this paper were considered failed when their capacitance drift exceeds 5% (Feng et al., 2012). Here, the median lifetime $t_{0.5}$ and quartile point lifetime $t_{0.1}$ of the high-voltage pulse capacitors are predicted using Method M_1 and Method M_2, respectively, and the results are shown in Table 13.

It is evident from Table 13 that Method M_2 is more conservative than Method M_1 in predicting both the median lifetime $t_{0.5}$ and the quartile lifetime $t_{0.1}$. Because the estimates of $\beta_{1,i}$ and $\beta_{2,i}$ in Method M_2 differ significantly from that in Method M_1, while the estimates of r_1 and r_2 differ less compared to Method M_1, what results in the degradation rate of the high-voltage impulse capacitors in Method M_2 will be faster that in Method M_1. Therefore, the degradation profile of Method M_2 will reach the failure threshold sooner. However, combined with the goodness-of-fit and the average degradation path, Method M_1 is consistent with the actual degradation of high-voltage pulse capacitors.

Therefore, disregarding the correlation between degradation stages can lead to more conservative predictions, which can result in overusing of materials, energy, and other resources in design and production. In contrast, it is possible to predict products' life more rationally by considering the correlation between degradation stages in the proposed method, which neither leads to hazardous predictions nor excessively wastes resources.

6. Conclusion

In this study, a two-stage nonlinear phase-dependent Wiener process degradation model is proposed to address the nonlinear characteristics of two-stage degradation process of products. Superior to existing studies, the proposed model considers both the degradation phase correlation and degradation heterogeneity of products in solving the problem of the nonlinear degradation process. Based on the historical degradation data of products, the EM algorithm is utilized to estimate the unknown parameters in the phase-dependent degradation model and to derive the lifetime distribution as well as the reliability function of products in this paper. The effectiveness and superiority of the proposed method are verified through simulation and examples of degradation of high-voltage pulse capacitors, in which, simulations show that the EM algorithm has good accuracy and robustness in estimating the model parameters. Case studies show that compared with the literature (A. Zhang et al., 2020), the proposed model exhibits significant advantages in fitting the degradation path of high-voltage pulse capacitors. Specifically, by comparing AIC values of the two models in Table 11, the proposed model (M_1) performs more reasonably and effectively to fit the degradation path of high-voltage pulse capacitors due to the smaller AIC value of M_1. Furthermore, as shown in Figure 4, the average degradation path predicted by M_1 highly aligns with the actual observed average degradation path, while the comparison model (M_2) exhibits significant deviation. This result clearly demonstrates that the proposed method by considering the correlation between degradation stages, more accurately reflects the actual degradation of high-voltage pulse capacitors, thereby better capturing the degradation characteristics of product.

This study not only provides a new perspective on degradation path modeling at the theoretical level, but also offers more reliable technical support for the health management, lifespan prediction, and maintenance decision-making of high-voltage pulse capacitors in practice.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Science and Technology Innovation Special Foundation of Fujian Agriculture and Forestry University of China under [grant number KFB25041A]; the National Natural Science Foundation of Fujian under [grant number 2025J01393]; the 2023 Ministry of Education for Employment and Education Project under [grant number 20230104862].

References

Bae, S. J., Yuan, T., Ning, S. L., & 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

Cai, Z. Y., Chen, Y. X., Guo, J. S., Wang, Z. Z., & Deng, L. (2019). Remaining lifetime prediction for device with measurement error and random effect (in Chinese). *Systems Engineering and Electronics*, *41*(7), 1658–1664.

Chen, N., & Tsui, K. L. (2013). Condition monitoring and remaining useful life prediction using degradation signals: Revisited. *IIE Transactions*, *45*(9), 939–952. https://doi.org/10.1080/0740817X.2012.706376

Deng, A. M., Chen, X., Zhang, C. H., & Wang, Y. S. (2006). Reliability assessment based on performance degradation data (in Chinese). *Journal of Astronautics*, *27*(3), 546–552.

Ding, C. B., Zhao, X., Zhu, Z. H., Zhao, H., & Zhang, T. L. (2024). Remaining life-span prediction based on adaptive nonlinear Wiener process (in Chinese). *Fire Control & Command Control*, *49*(8), 105–113.

Dong, Q., Zheng, J. F., Hu, C. H., Li, B., & Mou, H. X. (2022). Remaining useful life prognostic method based on two-stage adaptive Wiener process (in Chinese). *Acta Automatica Sinica*, *48*(2), 539–553.

Feng, J., Sun, Q., & Jin, T. D. (2012). Storage life prediction for a high-performance capacitor using multi-phase Wiener degradation model. *Communications in Statistics – Simulation and Computation*, *41*(8), 1317–1335. https://doi.org/10.1080/03610918.2011.624241

Huang, L., Liu, J. Q., & Gong, Y. J. (2018). Multi-phase residual life prediction of engines based on Wiener process (in Chinese). *Journal of Beijing University of Aeronautics and Astronautics*, *44*(5), 1081–1087.

Li, M.. (2009). *Research on the Degradation-Failure Modeling and Analysis Methods of Degradation Path with Random Change-Points (in Chinese)* [Unpublished doctoral dissertation]. National University of Defense Technology.

Ng, T. S. (2008). An application of the EM algorithm to degradation modeling. *IEEE Transactions on Reliability*, *57*(1), 2–13. https://doi.org/10.1109/TR.2008.916867

Pandey, M. D., Yuan, X. X., & Van Noortwijk, J. M. (2009). The influence of temporal uncertainty of deterioration on life-cycle management of structures. *Structure and Infrastructure Engineering*, *5*(2), 145–156. https://doi.org/10.1080/15732470601012154

Si, X. S., Wang, W., Hu, C. H., & D. H. Zhou (2011). Remaining useful life estimation–A review on the statistical data driven approaches. *European Journal of Operational Research*, *213*(1), 1–14. https://doi.org/10.1016/j.ejor.2010.11.018
Wang, H., Ma, X., & Zhao, Y. (2020). A mixed-effects model of two-phase degradation process for reliability assessment and RUL prediction. *Microelectronics Reliability*, *107*,113622https://doi.org/10.1016/j.microrel.2020.113622
Wang, H., Zhao, Y., & Ma, X. (2018). Remaining useful life prediction using a novel two-stage Wiener process with stage correlation. *IEEE Access*, *6*, 65227–65238. https://doi.org/10.1109/ACCESS.2018.2877630
Wang, P. P., Tang, Y. C., Bae, S. J., & He, Y. (2018). Bayesian analysis of two-phase degradation data based on change-point Wiener process. *Reliability Engineering & System Safety*, *170*, 244–256. https://doi.org/10.1016/j.ress.2017.09.027
Wang, Y., Peng, Y. Z., Zi, Y. Y., Jin, X. H., & Tsui, K. L. (2016). A two-stage data-driven-based prognostic approach for bearing degradation problem. *IEEE Transactions on Industrial Informatics*, *12*(3), 924–932. https://doi.org/10.1109/TII.2016.2535368
Wu, J. P., Wang, L. Y., Huang, X. Z., Yang, J., & Du, M. L. (2024). The RUL prediction based on improved Wiener degradation model for wet friction components. *Measurement Science and Technology*, *35*(7),076126. https://doi.org/10.1088/1361-6501/ad3ea0
Yang, J., Hong, Y. M., Wang, W., & Wu, G. H. (2025). Remaining useful life prediction of lead-acid battery using multi-phase Wiener process-based degradation model. *Process Safety and Environmental Protection*, *197*,106974. https://doi.org/10.1016/j.psep.2025.106974
Yang, J. X., Tang, S. J., Li, L., Sun, X. Y., Qi, S., & Si, X. S. (2024). Remaining useful life prediction based on implicit nonlinear Wiener degradation process (in Chinese). *Journal of Beijing University of Aeronautics and Astronautics*, *50*(1), 328–340.
Ye, H. F., Yao, Y. H., & Yu, C. L. (2007). Capacitance variability of high voltage pulse capacitors (in Chinese). *High Voltage Engineering*, *33*(6), 37–41.
Ye, Z. S., Chen, N., & Shen, Y. (2015). A new class of Wiener process models for degradation analysis. *Reliability Engineering & System Safety*, *139*, 58–67. https://doi.org/10.1016/j.ress.2015.02.005
Ye, Z. S., Wang, Y., Tsui, K. L., & Pecht, M. (2013). Degradation data analysis using Wiener processes with measurement errors. *IEEE Transactions on Reliability*, *62*(4), 772–780.
Yuan, Y., Zhang, Y., & Ding, H. (2020). Research on key technology of industrial artificial intelligence and its application in predictive maintenance (in Chinese). *Acta Automatica Sinica*, *46*(10), 2013–2030.
Zhang, A., Wang, Z. H., Zhao, Z. P., Qin, X. G., Wu, Q., & Liu, C. R. (2020). Two-stage nonlinear Wiener process for degradation modeling and reliability analysis (in Chinese). *Systems Engineering and Electronics*, *42*(4), 954–959.
Zhang, P., Hu, C. H., Bai, C., Zhang, Y., & Zhang, J. X. (2019). Remaining useful life estimation method for two-phase degradation system with random effects (in Chinese). *China Measurement & Test*, *45*(1), 1–7.
Zhou, R., Serban, N., & Gebraeel, N. (2014). Degradation-based residual life prediction under different environments. *The Annals of Applied Statistics*, *8*(3), 1671–1689. https://doi.org/10.1214/14-AOAS749
Zhuang, L. L., Xu, A. C., & Wang, X. L. (2023). A prognostic driven predictive maintenance framework based on Bayesian deep learning. *Reliability Engineering & System Safety*, *234*,109181https://doi.org/10.1016/j.ress.2023.109181

Appendix. Proof of Equation (26)

Proposition 1: *For $\theta_i \sim N(\mu_\theta, \Sigma_\theta)$, using Bayes' theorem, $\theta_s | Y_i, \Theta^{(i)}$ obeys a multivariate normal distribution with mean $v_{\theta,i}$ and variance $C_{\theta,i}$ (H. Wang et al., 2018), where,*

$$
\begin{aligned}
v_{\theta,i} &= [\Lambda_i^\top \Delta_i T_i^{-1} \Delta_i \Lambda_i + \Sigma_\theta^{-1}]^{-1} [Y_i^\top \Delta_i T_i^{-1} \Delta_i \Lambda_i + \mu_\theta^\top \Sigma_\theta^{-1}]^\top, \\
C_{\theta,i} &= [\Lambda_i^\top \Delta_i T_i^{-1} \Delta_i \Lambda_i + \Sigma_\theta^{-1}]^{-1}.
\end{aligned}
$$

Proof: The proof process of Equation (26) is as follows.

$$
\begin{aligned}
Q(\Theta|\Theta^{(s)}, \mathrm{Da}) &= E_{\Theta_i|Y_i,\Theta^{(s)}}[l(\Theta|\mathrm{Da})] \\
&= E_{\Theta_i|Y_i,\Theta^{(s)}}[l_1(\psi, r_1, r_2, \sigma_1, \sigma_2, \tau|\mathrm{Da}) + l_2(\mu_\theta, \Sigma_\theta|\psi)] \\
&= E_{\Theta_i|Y_i,\Theta^{(s)}}\Bigg[-\sum_{i=1}^n \frac{m_i}{2}\ln(2\pi) - \frac{1}{2}\sum_{i=1}^n \ln|T_i| - \sum_{i=1}^n c_i \ln\sigma_1 \\
&\quad - \sum_{i=1}^n (m_i - c_i)\ln\sigma_2 - \frac{1}{2}\sum_{i=1}^n (Y_i - \Lambda_i\theta_i)^\top \Delta_i T_i^{-1}\Delta_i(Y_i - \Lambda_i\theta_i) \\
&\quad - 2n\ln(2\pi) - \frac{1}{2}n\ln|\Sigma_\theta| + \frac{1}{2}\sum_{i=1}^n (\theta_i - \mu_\theta)^\top \Sigma_\theta^{-1}(\theta_i - \mu_\theta)\Bigg] \\
&= -\sum_{i=1}^n \frac{m_i}{2}\ln(2\pi) - \frac{1}{2}\sum_{i=1}^n \ln|T_i| - \sum_{i=1}^n c_i\ln\sigma_1 - \sum_{i=1}^n (m_i - c_i)\ln\sigma_2 \\
&\quad - 2n\ln(2\pi) - \frac{1}{2}n\ln|\Sigma_\theta| \\
&\quad + E_{\Theta_i|Y_i,\Theta^{(s)}}\left[-\frac{1}{2}\sum_{i=1}^n (Y_i - \Lambda_i\theta_i)^\top \Delta_i T_i^{-1}\Delta_i(Y_i - \Lambda_i\theta_i)\right] \\
&\quad + E_{\Theta_i|Y_i,\Theta^{(s)}}\left[-\frac{1}{2}\sum_{i=1}^n (\theta_i - \mu_\theta)^\top \Sigma_\theta^{-1}(\theta_i - \mu_\theta)\right].
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
&E_{\theta_i|Y_i,\Theta^{(s)}}[(Y_i - \Lambda_i\theta_i)^\top \Delta_i T_i^{-1}\Delta_i(Y_i - \Lambda_i\theta_i)] \\
&\quad = \mathrm{tr}\{E_{\theta_i|Y_i,\Theta^{(s)}}[(Y_i - \Lambda_i\theta_i)^\top \Delta_i T_i^{-1}\Delta_i(Y_i - \Lambda_i\theta_i)]\} \\
&\quad = E_{\theta_i|Y_i,\Theta^{(s)}}\{\mathrm{tr}[(Y_i - \Lambda_i\theta_i)^\top \Delta_i T_i^{-1}\Delta_i(Y_i - \Lambda_i\theta_i)]\} \\
&\quad = E_{\theta_i|Y_i,\Theta^{(s)}}\{\mathrm{tr}[\Delta_i T_i^{-1}\Delta_i(Y_i - \Lambda_i\theta_i)(Y_i - \Lambda_i\theta_i)^\top]\} \\
&\quad = \mathrm{tr}\{E_{\theta_i|Y_i,\Theta^{(s)}}[\Delta_i T_i^{-1}\Delta_i(Y_i - \Lambda_i\theta_i)(Y_i - \Lambda_i\theta_i)^\top]\} \\
&\quad = \mathrm{tr}\{\Delta_i T_i^{-1}\Delta_i E_{\theta_i|Y_i,\Theta^{(s)}}[(Y_i - \Lambda_i\theta_i)(Y_i - \Lambda_i\theta_i)^\top]\} \\
&\quad = \mathrm{tr}\{\Delta_i T_i^{-1}\Delta_i[\mathrm{Var}_{\theta_i|Y_i,\Theta^{(s)}}(Y_i - \Lambda_i\theta_i) + E_{\theta_i|Y_i,\Theta^{(s)}}(Y_i - \Lambda_i\theta_i)^\top E_{\theta_i|Y_i,\Theta^{(s)}}(Y_i - \Lambda_i\theta_i)]\} \\
&\quad = \mathrm{tr}[\Delta_i T_i^{-1}\Delta_i \mathrm{Var}_{\theta_i|Y_i,\Theta^{(s)}}(Y_i - \Lambda_i\theta_i)] \\
&\qquad + \mathrm{tr}[E_{\theta_i|Y_i,\Theta^{(s)}}(Y_i - \Lambda_i\theta_i)^\top \Delta_i T_i^{-1}\Delta_i E_{\theta_i|Y_i,\Theta^{(s)}}(Y_i - \Lambda_i\theta_i)] \\
&\quad = \mathrm{tr}(\Lambda_i^\top \Delta_i T_i^{-1}\Delta_i\Lambda_i C_{\theta,i}^{(s)}) + E_{\theta_i|Y_i,\Theta^{(s)}}(Y_i - \Lambda_i\theta_i)^\top \Delta_i T_i^{-1}\Delta_i E_{\theta_i|Y_i,\Theta^{(s)}}(Y_i - \Lambda_i\theta_i) \\
&\quad = \mathrm{tr}(\Lambda_i^\top \Delta_i T_i^{-1}\Delta_i\Lambda_i C_{\theta,i}^{(s)}) + (Y_i - \Lambda_i\nu_{\theta,i}^{(s)})^\top \Delta_i T_i^{-1}\Delta_i(Y_i - \Lambda_i\nu_{\theta,i}^{(s)}),
\end{aligned}
$$

where

$$
\begin{aligned}
E_{\theta_i|Y_i,\Theta^{(s)}}(Y_i - \Lambda_i\theta_i) &= (Y_i - \Lambda_i\nu_{\theta,i}^{(s)}), \\
\mathrm{Var}_{\theta_i|Y_i,\Theta^{(s)}}(Y_i - \Lambda_i\theta_i) &= \Lambda_i^\top \Lambda_i C_{\theta,i}^{(s)}.
\end{aligned}
$$

The same reasoning can be used to prove that

$$
\begin{aligned}
&E_{\theta_i|Y_i,\Theta^{(s)}}[(\theta_i - \mu_\theta)^\top \Sigma_\theta^{-1}(\theta_i - \mu_\theta)] \\
&\quad = (\nu_{\theta,i}^{(s)} - \mu_\theta)^\top \Sigma_\theta^{-1}(\nu_{\theta,i}^{(s)} - \mu_\theta) + \mathrm{tr}(\Sigma_\theta^{-1} C_{\theta,i}^{(s)}).
\end{aligned}
$$

Thus, Equation (26) is proved. ∎