Parameter estimation for Load-sharing system subject to Wiener degradation process using E-M algorithm

Abstract- In practice, many systems exhibit load-sharing behavior, where the surviving components share the total load imposed on the system. Different from general systems, the components of load-sharing systems are interdependent in nature, in such a way that when one component fails, the system load has to be shared by the remaining components, which increases the failure rate or degradation rate of the remaining components. Due to the load-sharing mechanism among components, parameter estimation and reliability assessment are usually complicated for load-sharing systems. Although load-sharing systems with components subject to sudden failures have been intensely studied in literatures with detailed estimation and analysis approaches, those with components subject to degradation are rarely investigated. In this paper, we propose the parameter estimation method for load-sharing systems subject to continuous degradation with a constant load. Likelihood function based on degradation data of components is established as a first step. The maximum likelihood estimators for unknown parameters are deduced and obtained via EM algorithm considering the non-closed form of the likelihood function. Numerical examples are used to illustrate the effectiveness of the proposed method.

Key words-Load-sharing system, continuous degradation, Wiener degradation, constant load, EM algorithm

1 Introduction

The assumption of independence among components in a system is commonly adopted in reliability engineering for the convenience of analysis. However, more often than not, systems are subject to various types of component dependence, which increasingly challenges the assumption of independence. There are several models to describe such dependence in practice. In some cases, components in a system share the total load on the system. Consider a system consisting of parallel connected components, when the components fail one by one, system load is redistributed among the surviving components. Generally, this is referred to as load-sharing [1].

Early studies on load-sharing systems can be found in Daniels[2], Birnbaum and Saunders[3], Harlow and Phoenix[4] and Lee et al.[5]. While Durham and Lynch[6], Yang and Younis[7] and Singh et al. [8] are among the more recent investigations on load-sharing systems. Meanwhile, applications and implementations of load-sharing systems in different areas are also investigated such as computer systems[9], power grids[10] and gear systems[11]. Most research from literatures on load-sharing systems investigated the characteristics of reliability with a certain known load-sharing rule. The assumption that model parameters are known *a priori* is always involved in related literatures, whereas few has studied the parameter estimations of the model on the basis of real data. Exceptions can be found in the following existing works: Kim and Kvam [12] estimated the parameters under the assumption of an exponential underlying lifetime distribution and equal load-sharing rule; Singh et al. [8] introduced Bayesian approach to estimate the parameters; Deshpande et al. [13] modeled load-sharing system with a family of semiparametric distributions and suggested the estimates for the parameters; Park[14,15] detailed the maximum likelihood estimates (MLE) of the parameters for load-sharing systems with components that have exponential, Weibull, normal and lognormal distributed life in these two works.

Note that an assumption of parametric or semiparametric lifetime distributions is commonly introduced in previous literatures on load-sharing systems. This implies that the functioning condition of the components are supposed to be invariant and the component failures are supposed to be sudden and catastrophic. However, this is not always a proper assumption, considering the varying working conditions and the slow but cumulative system damage. Recently, with advances of modern technologies, systems are becoming increasingly more reliable, which may go through a long period of deterioration in terms of performance before eventually fail. In such situations, degradation model has been pointed out to be more convenient and flexible in Singpurwalla[16]. In a degradation model, performance degradation of a certain product is concerned rather than a mere failure time. The system is supposed to fail when the performance degrades below a specific threshold, which means that the system lifetime can also be modeled with respect to the corresponding degradation model. On the basis of a proper degradation model, a series of works on reliability assessment, maintenance[17], forecasting[18] and warranty policies can be further studied.

In previous literatures on load-sharing systems, several conventional parametric models have been investigated. As first proposed in Whitmore and Schenkelberg[19], Wiener degradation process can effectively characterize the scenario where the degradation increment is supposed to be a cumulation of additive tiny independent effects in infinitesimal time interval. This independent increment characteristic of Wiener degradation process provides multiple convenient properties in application[20]. Meanwhile, a closed form expression of lifetime distribution induced by degradation endows the Wiener degradation process many facilities in practice. Except for Wiener process, Gamma process[21,22] and Inverse Gaussian degradation process[23] can also be implemented in degradation model. For more details of stochastic models characterizing degradation, readers can refer to[24,25].

As mentioned above, when failure models are not adequate for a load-sharing system, degradation model can be incorporated. However, there are few literatures focusing on load-sharing systems subject to degradation. As an exception, Liu et al.[26] considered a load-sharing system with components subject to continuous degradation processes, where system reliability and maintenance policies are investigated under constant and cumulative load respectively. Yet the parameters of the degradation processes are assumed to be constant and known and in advance. In this paper, we will assess the load-sharing system subject to degradation with unknown parameters. Likelihood function is established based on observed degradation data of components. As the exact failure time of a degrading component is supposed to be unobservable, the closed form of likelihood function directly based on observed data cannot be attained. Numerical methods are essential to obtain the MLEs for the parameters. Different from Park[15], where a Newton type method is used for parameter estimation, we introduce an EM algorithm to calculate the parameters from MLEs.

The rest of the paper is organized as follows. In section 2, we propose the assumptions and interpretations for load-sharing systems subject to Wiener degradation. Section 3 gives a direct establishment of likelihood function based on the censored data set. In section 4, we introduce and implement the EM algorithm for the MLE. In section 5, we give detailed process to solve the MLEs through EM algorithm. Numerical examples and simulation study are illustrated in section 6. Concluding remarks are given in section 7.

2 Load-sharing systems subject to degradation

2.1 General model assumptions.

For a load-sharing system subject to component degradation, the degradation level of a component is concerned instead of conventional failure time. The basic assumptions are listed as follows for the model formulation in this paper:

1) The system is composed of J components. The components in the system are functioning independently and follow an identical mean degradation path.
2) Each component is subject to a continuous degradation path. Degradation level at time t of the j-th component is denoted by $X_j(t;\ \theta_j)$, with parameter vector θ_j.
3) A component fails immediately when its degradation level exceeds a pre-specified threshold ω.
4) Total load of the system is redistributed among the remaining components with an instant and additive impact on each component after a component failure.
5) The system fails instantly when and only when all the components fail.

In this paper, we consider a commonly used type of continuous degradation process which has independent increment property: Wiener process. We denote the increment of the degradation as $\Delta X_j(t) = X_j\left(t + \Delta t;\ \theta_j\right) - X_j(t;\ \theta_j)$. The probability density function (PDF) for the degradation increment of component j is defined as $f_j\left(\Delta X_j(t)|\theta_j, \Delta t\right)$.

2.2 Wiener degradation process

Wiener process is commonly employed to characterize a continuous degradation process[27]. In this paper, we confine the Wiener process to a linear shift form. Under the assumption of Wiener degradation process with linear shift, the performance of component j at time t can be modeled as:

$$X_j\left(t; \theta_j\right) = \mu_j t + \sigma_j B(t),$$

where $\theta_j = \left(\mu_j, \sigma_j\right)$ is the parameter vector for component j, μ_j is the drift parameter and σ_j is the diffusion parameter. $B(t)$ is the standard Brownian motion. The corresponding PDF for the increment of component j in time Δt can be defined by:

$$f_j\left(\Delta X_j(t)|\theta_j, \Delta t\right) = \left(2\pi {\sigma_j}^2 \Delta t\right)^{-\frac{1}{2}} \exp\left\{-\frac{\left(\Delta X_j(t) - \mu_j \Delta t\right)^2}{2{\sigma_j}^2 \Delta t}\right\}, \tag{1}$$

under a pre-specified critical threshold ω, the failure time of a component can be defined as:

$$T_j = \inf\{t: X_j(t;\ \mu_j,\ \sigma_j) \geq \omega |\ \mu_j,\ \sigma_j\}. \tag{2}$$

The corresponding distribution of T_i is an inverse Gaussian distribution with the respective PDF and the cumulative density function (CDF) as follows:

$$f_{IG}(T_j|\omega, \mu_j, \sigma_j) = \frac{1}{\sqrt{2\pi\ \sigma_j{}^2 T_j{}^3}} exp\left\{-\frac{(\omega - \mu_j T_j)^2}{2\ \sigma_j{}^2 T_j}\right\}, \tag{3}$$

$$F_{IG}(T_j|\omega, \mu_j, \sigma_j) = \Phi\left(\frac{\mu_j T_j - \omega}{\sigma_j\sqrt{T_j}}\right) + exp\left\{\frac{2\mu_j T_j}{\sigma_j{}^2}\right\}\Phi\left(\frac{-\mu_j T_j - \omega}{\sigma_j\sqrt{T_j}}\right), \tag{4}$$

where Φ is the CDF of standard normal distribution.

2.3 Load-sharing rule

The mode of load redistribution among the remaining components is determined by load-sharing rules. Most of the existing work regarding load-sharing systems has assumed a known load-sharing rule. Conventional load-sharing rules investigated in literatures include equal load-sharing, tapered load-sharing, local load-sharing, nearest-neighbor load-sharing and hybrid load-sharing rule. For more details on this load-sharing rules, readers are referred to Durham et al. [28]

In this paper, we consider the widely applied equal load-sharing rule, under which the total load is redistributed among the remaining components equally. We also confine the total load on system $L(t)$ at time t to a constant load L. Denote the load for component j at time t by $L_j(t)$, then the total performance deterioration of a single component in the system at time t is the accumulation of cumulative degradation and the current load. We still denote the total performance for component j at time t by $X_j(t)$ for convenience, which can be represented as:

$$X_j(t;\ \theta_j) = \mu_j t + \sigma_j B(t) + L_j(t), \tag{5}$$

under the equal load-sharing rule,

$$L_j(t) = \frac{L}{J - N(t)}, \tag{6}$$

and $X_j(t;\ \theta_j)$ can be further represented as:

$$X_j(t;\ \theta_j) = \mu_j t + \sigma_j B(t) + \frac{L}{J - N(t)}, \tag{7}$$

3 The construction of likelihood for load-sharing systems with degrading components

In this section we will construct the likelihood function for load-sharing systems subject to degradation. Although Park [14,15] detailed the likelihood function for load-sharing systems with exponential, Weibull, normal and log-normal components, load-sharing systems with degrading components cannot be constructed through the existing steps in these literatures. First, we provide a brief description of the available data for components during a test of such systems. Let $N(t)$ be the number of failed components in the system by time t.

Consider a single system sample for degradation test. As detailed in significant number of literatures, in such a test, degradation performance of each component is measured at time $t_i, i = 1, \cdots, N$, and let $t_0 = 0$. Denote $X_j(t_i)$ as the degradation performance of component j at time t_i . Without loss of generality, we label all the components in the order of their failure times, i.e., X_j denote the degradation path of the j-th failing component. After a component fails, its degradation performance of this component can no longer be observed. Note that in equation (2), when the interval between t_i and $t_{i-1}, , i = 1, \cdots, N$ is small enough, simultaneous failures of multiple components in each interval can be neglected in the sense of probability. In this paper, we assume that there is at most one component failure during an inspection interval. Note that this assumption holds in the context of the article.

As the exact failure time of each component cannot be observed, the results through a single test can be listed as below:

$$\begin{array}{cccc} X_1(t_1) & , & \cdots & , & X_1\left(t_{i_1}\right) \\ X_2(t_1) & , & \cdots & , & X_1\left(t_{i_2}\right) \\ \vdots & & & & \vdots \\ X_J(t_1) & , & \cdots & , & X_J\left(t_{i_J}\right) \end{array}$$

where $t_{i_1} < t_{i_2} < \cdots < t_{i_J}$, and the j-th component fails in (t_{i_j}, t_{i_j+1}). In the last observation at time t_N, the whole system has failed in (t_{i_J}, t_{i_J+1}), so it is natural to terminate the test at t_{i_J+1} and denote $t_{i_J+1} = t_N$. Additionally, following the routine of degradation test in significant number of literatures, we consider a test with fixed inspection interval: $\Delta t = t_{k+1} - t_k$, where $k = 1, \cdots, t_N$. The parameter estimation in this paper is based on the observations listed above. When each component subject to a continuous degradation process, it is easy to verify

$$\mathcal{P}\{\Delta N(t) = N(t+\Delta t) - N(t) \geq 2\} = o(\Delta t), \tag{8}$$

which means the assumption of no simultaneous failures are observed within each time interval of Δt is acceptable.

For component j, let $\Delta X_j^k = X_j(t_k) - X_j(t_{k-1})$, where $X_j(t_0) = 0, j = 1, \cdots, J$; under the equal load-sharing rule, the following equation holds:

$$L_j(t_k) = \begin{cases} \frac{L}{J}, & k \in \{0, \cdots, i_1\}, \\ \frac{L}{J-1}, & k \in \{i_1 + 1, \cdots, i_2\}, \\ & \vdots \\ L, & k \in \{i_{J-1} + 1, \cdots, i_J\}. \end{cases} \tag{9}$$

Due to the independent increment property of the degradation process considered in this paper, ΔX_j^k is independently distributed. With respect to the drift of ΔX_j^k in the interval (t_{k-1}, t_{k-1}) induced by $\Delta L_j^k = L_j(t_k) - L_j(t_{k-1})$, the PDF of ΔX_j^k can be represented as:

$$\Delta X_j^k \sim f_j\left(\Delta X_j^k - \Delta L_j^k | \theta_j, \Delta t\right), j = 1, \cdots, J, k = 1, \cdots, i_j, \tag{10}$$

where $f_j\left(\cdot\, | \theta_j, \Delta t\right)$ is previously defined in section 2.1.

Meanwhile, the real failure time of component j is censored within the interval $\left(t_{i_j}, t_{i_j+1}\right), j \in \{1, \cdots, J-1\}$. The likelihood for the real lifetime $T^{(j)}$ contributed by component j within the interval $\left(t_{i_j}, t_{i_j+1}\right)$ is $F_{IG}\left(t_{i_j+1} - t_{i_j} \left| \omega - X_j\left(t_{i_j}\right), \theta_j\right.\right)$ under Wiener degradation. Thus, the complete likelihood function based on the data for θ_j can be represented as follows:

$$L\left(\theta_j\right) = \prod_{k=1}^{i_j} f_j\left(\Delta X_j^k - \Delta L_j^k | \theta_j, \Delta t\right) * F_j\left(t_{i_j+1} - t_{i_j}\right), \tag{11}$$

and the complete likelihood for $\theta = \left(\theta_1, \cdots, \theta_J\right)$ is

$$L(\theta) = \prod_{j=1}^{J} \prod_{k=1}^{i_j} f_j\left(\Delta X_j^k - \Delta L_j^k | \theta_j, \Delta t\right) * F_j\left(t_{i_j+1} - t_{i_j}\right), \tag{12}$$

where $F_j\left(t_{i_j+1} - t_{i_j}\right)$ is a consistent notation for the CDF of lifetime of component j instead of $F_{IG}\left(t_{i_j+1} - t_{i_j} \left| \omega - X_j\left(t_{i_j}\right), \theta_j\right.\right)$ under Wiener degradation through this paper.

Since $F_j\left(t_{i_j+1} - t_{i_j}\right)$ in the above likelihood function of the CDF is not in a closed form under

Wiener degradation pattern, we need to resort to numerical methods for the direct maximization of the likelihood function. One conventional and widely utilized method is the Newton-Raphson type of numerical methods. However, this type of methods can be very sensitive to the starting value or the results of each iteration, which leads to difficulty in convergence to the real solution. Also, when the number of parameters is large, the ineffectiveness of Newton-Raphson method becomes significant. A compromise to deal with such a problem is to ignore the censoring interval for each component, emitting the second item of each $F_j\left(t_{i_{j+1}} - t_{i_j}\right)$ in the right side of equation (12). This can cause a significant loss of information from the data especially when the number of components is large. In this paper, we introduce the EM algorithm to overcome the difficulty in pursuing the MLE for the parameters from complete likelihood function.

4 The EM algorithm and its application in the likelihood function

In this part, we give a brief review of EM algorithm. We transform the likelihood function into the form based on complete data instead of censored data and implement it into the EM algorithm for the numerical calculation of the MLE of parameters.

4.1 EM algorithm

EM algorithm is an iteration method to find the maximum likelihood or maximum posteriori of parameters for a statistical model when latent variables are involved. The EM algorithm was first introduced by Dempster et al. [29]. Further studies on the EM algorithm can be found in [30–32]. Although in some cases, especially when the target function is sharp around its maximum, the performance of EM algorithm may be not as fast as Newton methods. However, in cases of large number of parameters and a potential smooth character of the target function around the maximum, the EM algorithm can give a stable and faster performance with the aid of powerful computers.

We first give a brief introduction of the EM algorithm in the application for missing or censored data. The conventional EM algorithm can be divided in two steps: Expectation step (E-step) and Maximization step (M-step). In the E-step, expectation of the log-likelihood function is computed, with respect to the conditional distribution of the unobserved data given the observed data. The calculation is based on the current estimation of parameters; In the M-step, the quantity of the conditional expectation in E-step is maximized and the estimation of parameters for next iteration is calculated. To facilitate the application of the EM algorithm in this paper for our problem, some necessary details are provided in advance. Denote the set of observed data by X, the censored or missing data by Z and the vector of unknown parameters by θ. The likelihood function based on the complete data can be written as:

$$L(\theta|X,Z) = p(X,Z|\theta). \tag{13}$$

- Expectation-step:

$$Q\left(\theta\middle|\theta^{(n)}\right) = E_{Z|X,\theta}\left(\log L(\theta|X,Z)\right); \tag{14}$$

- Maximization-step:

$$\theta^{(n+1)} = \underset{\theta}{\operatorname{argmax}}\, Q\left(\theta\middle|\theta^{(n)}\right), \tag{15}$$

where $\theta^{(n)}$ denotes the parameter estimates in the nth iteration.

4.2 The implementation of E-M algorithm in likelihood function

To implement the EM algorithm, we first denote the real lifetime of component j in interval $\left(t_{i_j}, t_{i_j+1}\right)$ by Z_j . Note that Z_j is unobservable or latent, meanwhile, $\left\{X_j(t_1), \cdots, X_j\left(t_{i_j}\right)\right\}$ is the observed data set for component j. The distribution of Z_j can be represented by a truncated inverse Gaussian distribution under Wiener degradation with the PDF as follows:

$$\tilde{f}_j\left(Z_j|\omega - X_j\left(t_{i_j}\right), \theta_j\right) = \frac{f_{IG}\left(Z_j\middle|\omega - X_j\left(t_{i_j}\right), \theta_j\right)}{F_{IG}\left(t_{i_j+1} - t_{i_j}|\omega - X_j\left(t_{i_j}\right), \theta_j\right)}. \tag{16}$$

Likewise, as in the notation of $F_j\left(t_{i_j+1} - t_{i_j}\right)$, we replace $\tilde{f}_j\left(Z_j|\omega - X_j\left(t_{i_j}\right), \theta_j\right)$ by $\tilde{f}_j(Z_j)$ for simplicity. The complete likelihood function based on $\left\{X_j(t_1), \cdots, X_j\left(t_{i_j}\right)\right\}$ and Z_j can be constructed as follows:

$$L'(\theta) = \prod_{j=1}^{J}\prod_{k=1}^{i_j} f_j\left(\Delta X_j^k - \Delta L_j^k|\theta_j, \Delta t\right) * f_{IG}\left(Z_j\middle|\omega - X_j\left(t_{i_j}\right), \theta_j\right). \tag{17}$$

The notation of $L'(\theta)$ is to distinguish the likelihood based on pseudo complete data from $L(\theta)$ in the former sections. The corresponding log-likelihood function can be represented as

$$\log L'(\theta) = \sum_{j=1}^{J}\left(\sum_{k=1}^{i_j} \log f_j\left(\Delta X_j^k - \Delta L_j^k|\theta_j, \Delta t\right) + \log f_{IG}\left(Z_j\middle|\omega - X_j\left(t_{i_j}\right), \theta_j\right)\right). \tag{18}$$

Denoting the parameter estimation of θ in the t-th iteration by $\theta^{(t)} = \left(\theta_1^{(t)}, \cdots, \theta_J^{(t)}\right)$, the expectation of the second item in the log-likelihood function is therefore given by

$$E\left(f_{IG}\left(Z_j\middle|\omega - X_j\left(t_{i_j}\right), \theta_j\right)|\theta_j^{(t)}\right) = \int_0^{t_{i_j+1}} log\, f_{IG}\left(Z_j\middle|\omega - X_j\left(t_{i_j}\right), \theta_j\right) \cdot \tilde{f}_j(Z_j)dZ_j, \quad (19)$$

and the target expectation function in the E-step is

$$Q(\theta|\theta^{(n)}) = E(\log L'(\theta)|\theta^{(n)}) = \sum_{j=1}^{J}\sum_{k=1}^{i_j} \log f_j(\Delta X_j^k - \Delta L_j^k|\theta_j, \Delta t)$$

$$+\sum_{j=1}^{J} E\left(\log f_{IG}\left(Z_j|\omega - X_j\left(t_{i_j}\right), \mu_j, \sigma_j\right)|\theta_j^{(n)}\right). \quad (20)$$

Based on the calculation of the second item in the equation above, $Q(\theta|\theta^{(n)})$ in the E-step is available. Then in the M-step, the maximization of $Q(\theta|\theta^{(n)})$ is computed to obtain the parameter estimation $\theta^{(n+1)}$ of θ for the $(t+1)$-th iteration. When $\left|\theta^{(n+1)} - \theta^{(n)}\right| < \varepsilon$, where ε is a pre-specified small constant, the iteration is terminated and the associated result can be considered as an approximation of the parameter θ. Details under two types of degradation processes will be illustrated in the next section.

5 ML estimation of unknown parameters based on EM algorithm

In this section, we will illustrate the detailed access to the parameters of MLEs based on the likelihood function using EM algorithm. In the previous section, we illustrate the framework of EM algorithm for the likelihood function based on complete data. In this section, we will derive the explicit form of likelihood function under Wiener degradation processes, where the EM algorithm is implemented as introduced in the previous section.

As the expectation item in $Q(\theta|\theta^{(n)})$ is a in a summation form, it is equivalent to investigate the parameters for each component in the summation. In the case of Wiener degradation, by substituting equation (2), and denote $\omega - X_j\left(t_{i_j}\right)$ by ω_j, we have

$$\log f_{IG}\left(Z_j|\omega - X_j\left(t_{i_j}\right), \mu_j, \sigma_j\right) = \log f_{IG}(Z_j|\omega_j, \mu_j, \sigma_j)$$

$$= C_j - \log\sigma_j + \frac{\omega_j\mu_j}{\sigma_j{}^2} - \frac{3}{2}\log Z_j$$

$$-\frac{{\omega_j}^2}{2*{\sigma_j}^2*Z_j}-\frac{{\mu_j}^2*Z_j}{2{\sigma_j}^2}, \tag{21}$$

where C_j is a constant irrelevant to (μ_j, σ_j). The expectation of the log-likelihood function is formulated as

$$E\left[\log f_{IG}\left(Z_j|\omega_j,\mu_j,\sigma_j\right)\right]=C_j-\log\sigma_j+\frac{\omega_j\mu_j}{{\sigma_j}^2}-\frac{{\mu_j}^2}{2{\sigma_j}^2}E_{j1}^{(n)}-\frac{{\omega_j}^2}{2{\sigma_j}^2}E_{j2}^{(n)}-\frac{3}{2}E_{j3}^{(n)}, \tag{22}$$

where

$$\begin{cases} E_{j1}^{(n)}=E_{Z_j}\left(Z_j|\mu_j^{(n)},\sigma_j^{(n)}\right)=\int_0^{t_{i_j+1}} Z_j\cdot f_j^*\left(Z_j|\omega_j,\mu_j^{(n)},\sigma_j^{(n)}\right)dZ_j, \\ E_{j2}^{(n)}=E_{Z_j}\left(\frac{1}{Z_j}|\mu_j^{(n)},\sigma_j^{(n)}\right)=\int_0^{t_{i_j+1}}\frac{1}{Z_j}\cdot f_j^*\left(Z_j|\omega_j,\mu_j^{(n)},\sigma_j^{(n)}\right)dZ_j, \\ E_{j3}^{(n)}=E_{Z_j}\left(log\, Z_j\,|\mu_j^{(n)},\sigma_j^{(n)}\right)=\int_0^{t_{i_j+1}} log\, Z_j\cdot f_j^*\left(Z_j|\omega_j,\mu_j^{(n)},\sigma_j^{(n)}\right)dZ_j. \end{cases} \tag{23}$$

Note that all the three items above should be calculated in each iteration of the EM algorithm. Following the direct result of *Theorem 1* in Ye et al. [33], $E_{j3}^{(n)}$ can be represented as:

$$E_{j1}^{(n)}=\frac{\frac{\omega}{\mu_j^{(n)}}\left[1-F_{IG}\left(\frac{\omega^2}{\left(t_{i_j+1}-t_{i_j}\right){\mu_j^{(n)}}^2}|\omega-X_j\left(t_{i_j}\right),\mu_j^{(n)},\sigma_j^{(n)}\right)\right]}{F_{IG}\left(t_{i_j+1}-t_{i_j}|\omega-X_j\left(t_{i_j}\right),\mu_j^{(n)},\sigma_j^{(n)}\right)}. \tag{24}$$

$E_{j3}^{(n)}$ is a constant that does not depend on parameter $\left(\mu_j^{(n)},\sigma_j^{(n)}\right)$ in each iteration of the EM algorithm, thus in the M-step, the maximization of the parameters is free of $E_{j3}^{(n)}$ due to the derivation process. In other words, the numerical result of $E_{j3}^{(n)}$ is not concerned and will be integrated into C_j in the rest of the paper. For the calculation of $E_{j2}^{(n)}$, numerical methods are needed. As the integral process is on a bounded interval $(0, t_{i_j+1})$, so numerical methods for such a problem are very fast and robust. By substituting the results and the equal load-sharing rule into the aforementioned framework, the EM algorithm can be implemented as follows

- E-step:

$$Q\left(\theta|\theta^{(n)}\right)=E\left(\log L'(\theta)|\theta^{(n)}\right)=\sum_{j=1}^J\sum_{k=1}^{i_j} log\, f_j\left(\Delta X_j^k-\Delta L_j^k|\theta_j,\Delta t\right)$$
$$+\sum_{j=1}^J E\left(\log\tilde{f}_j\left(Z_j|\omega-X_j\left(t_{i_j}\right),\mu_j,\sigma_j\right)|\theta_j^{(n)}\right)$$

$$= \sum_{j=1}^{J}\sum_{k=1}^{i_j}\left(C_{jk}^{*} - \log\sigma_j - \frac{1}{2{\sigma_j}^2}\left(\Delta X_j^k - \Delta L_j^k - \Delta t\mu_j\right)^2\right)$$

$$+ \sum_{j=1}^{J}\left(C_j - \log\sigma_j + \frac{\omega_j\mu_j}{{\sigma_j}^2} - \frac{{\mu_j}^2}{2{\sigma_j}^2}E_{j1}^{(n)} - \frac{{\omega_j}^2}{2{\sigma_j}^2}E_{j2}^{(n)}\right)$$

$$= C + \sum_{j=1}^{J}\left(\begin{array}{c} -i_j\log\sigma_j - \\ \frac{1}{2{\sigma_j}^2}\left(X_j\left(t_{i_j}\right) - L\left(\frac{1}{J-j+1} - \frac{1}{J}\right) - t_{i_j}\mu_j\right)^2 \end{array}\right)$$

$$+ \sum_{j=1}^{J}\left(-\log\sigma_j + \frac{1}{{\sigma_j}^2}\left(\omega_j\mu_j - \frac{\mu_j}{2}E_{j1}^{(n)} - \frac{{\omega_j}^2}{2}E_{j2}^{(n)}\right)\right), \qquad (25)$$

where C_{jk}^{*}, C_j and C are all constants free of unknown parameters.

- M-step:

By differentiating $Q\left(\theta|\theta^{(n)}\right)$ in the E-step with respect to each μ_j, we obtain

$$\frac{\partial Q\left(\theta|\theta^{(n)}\right)}{\partial\mu_j} = \frac{\Delta t}{{\sigma_j}^2}\sum_{k=1}^{i_j}\left(\Delta X_j^k - \Delta L_j^k - \Delta t\mu_j\right) + \frac{1}{{\sigma_j}^2}\left(\omega_j - E_{j1}^{(n)}\mu_j\right)$$

$$= \frac{\Delta t}{{\sigma_j}^2}\left(X_j\left(t_{i_j}\right) - L\left(\frac{1}{J-j+1} - \frac{1}{J}\right) - t_{i_j}\mu_j\right) + \frac{1}{{\sigma_j}^2}\left(\omega_j - E_{j1}^{(n)}\mu_j\right). \qquad (26)$$

By differentiating $Q\left(\theta|\theta^{(n)}\right)$ in the E-step in terms of each σ_j, it follows

$$\frac{\partial Q\left(\theta|\theta^{(t)}\right)}{\partial\sigma_j} = \sum_{k=1}^{i_j}\left(-\frac{1}{\sigma_j} + \frac{1}{{\sigma_j}^3}\left(\Delta X_j^k - \Delta L_j^k - \Delta t\mu_j\right)^2\right)$$

$$+ \left(-\frac{1}{\sigma_j} + \frac{1}{{\sigma_j}^3}\left({\mu_j}^2E_{j1}^{(t)} + {\omega_j}^2E_{j2}^{(t)} - 2\omega_j\mu_j\right)\right)$$

$$= -\frac{i_j+1}{\sigma_j} + \frac{1}{{\sigma_j}^3}\left(\begin{array}{c} \sum_{k=1}^{i_j}\left(\Delta X_j^k - \Delta L_j^k - \Delta t\mu_j\right)^2 \\ +\left({\mu_j}^2E_{j1}^{(t)} + {\omega_j}^2E_{j2}^{(t)} - 2\omega_j\mu_j\right) \end{array}\right). \qquad (27)$$

By solving $\frac{\partial Q\left(\theta|\theta^{(t)}\right)}{\partial\mu_j} = 0$ and $\frac{\partial Q\left(\theta|\theta^{(t)}\right)}{\partial\sigma_j} = 0$, the MLEs for the $(t+1)$ iteration can be obtained as:

$$\mu_j^{(t+1)} = \frac{X_j\left(t_{i_j}\right) - L\left(\frac{1}{J-j+1} - \frac{1}{J}\right) + \omega_j}{\Delta t \cdot t_{i_j} + E_{j1}^{(t)}}, \tag{28}$$

and,

$$\begin{aligned}
\sigma_j^{(t+1)^2} &= \frac{\sum_{k=1}^{i_j}\left(\Delta X_j^k - \Delta L_j^k - \Delta t \mu_j^{(t+1)}\right)^2 + \left({\mu_j}^2 E_{j1}^{(t)} + {\omega_j}^2 E_{j2}^{(t)} - 2\omega_j \mu_j^{(t+1)}\right)}{i_j + 1} \\
&= \frac{i_j \Delta t \cdot t_{i_j} \mu_j^{(t+1)^2} - 2\Delta t \mu_j^{(t+1)} \sum_{k=1}^{i_j}\left(\Delta X_j^k - \Delta L_j^k\right) + \sum_{k=1}^{i_j}\left(\Delta X_j^k - \Delta L_j^k\right)^2}{i_j + 1} \\
&\quad + \frac{{\mu_j}^2 E_{j1}^{(t)} + {\omega_j}^2 E_{j2}^{(t)} - 2\omega_j \mu_j^{(t+1)}}{i_j + 1} \\
&= \frac{i_j \Delta t \cdot t_{i_j} \mu_j^{(t+1)^2} - 2\Delta t \mu_j^{(t+1)} \left(X_j\left(t_{i_j}\right) - L\left(\frac{1}{J-j+1} - \frac{1}{J}\right)\right)}{i_j + 1} \\
&\quad + \frac{\sum_{k=1}^{i_j}\left(\Delta X_j^k - \Delta L_j^k\right)^2}{i_j + 1} + \frac{{\mu_j}^2 E_{j1}^{(t)} + {\omega_j}^2 E_{j2}^{(t)} - 2\omega_j \mu_j^{(t+1)}}{i_j + 1} \\
&= \frac{i_j \Delta t \cdot t_{i_j} \mu_j^{(t+1)^2} - 2\Delta t \mu_j^{(t+1)} \left(\left(\Delta t \cdot t_{i_j} + E_{j1}^{(t)}\right)\mu_j^{(t+1)} - \omega_j\right)}{i_j + 1} \\
&\quad + \frac{\sum_{k=1}^{i_j}\left(\Delta X_j^k - \Delta L_j^k\right)^2}{i_j + 1} + \frac{{\mu_j}^2 E_{j1}^{(t)} + {\omega_j}^2 E_{j2}^{(t)} - 2\omega_j \mu_j^{(t+1)}}{i_j + 1} \\
&= \frac{1}{i_j + 1} \mu_j^{(n+1)^2} \left(E_{j1}^{(n)}(1 - 2\Delta t) + \Delta t \cdot t_{i_j}\left(i_j - 2\Delta t\right)\right) \\
&\quad + \frac{1}{i_j + 1}\left(2\omega_j \mu_j^{(n+1)}(\Delta t - 1) + {\omega_j}^2 E_{j2}^{(n)}\right) \\
&\quad + \frac{1}{i_j + 1} \sum_{k=1}^{i_j}\left(\left(\Delta X_j^k - \Delta L_j^k\right)^2\right).
\end{aligned} \tag{29}$$

Note that the third item in the equation $\frac{1}{i_j+1}\sum_{k=1}^{i_j}\left(\left(\Delta X_j^k - \Delta L_j^k\right)^2\right)$ does not depend on the results from the t-th iteration, but only related to j, we can calculate it before implementing algorithm and substitute the result into each iteration.

6 Simulation study

6.1 Numerical example

In this section, we use a numerical example to illustrate the effectiveness of the proposed parameter estimation approach. We generate artificial degradation data according to pre-specified parameters. Then we apply the EM algorithm on the simulated data. The number of components are set as 3. The degradation parameters and failure thresholds of the components are set as:

$$\mu_1 = 1, \sigma_1 = 2, \mu_2 = 2, \sigma_2 = 4, \mu_3 = 3, \sigma_3 = 3,$$
$$\omega_1 = 35, \omega_2 = 65, \omega_3 = 85$$

The simulated data set is illustrated in Table 6.1. When the degradation performance of any component exceeds the corresponding failure threshold, the observation is terminated and the exact failure time is censored.

Based on the data set in Table 1, we use the EM algorithm in our approach according to the preceding procedures to estimate the parameters. Initial values of the EM procedure are set as

$$\mu_1 = 0, \sigma_1 = 1, \mu_2 = 1, \sigma_2 = 1, \mu_3 = 1, \sigma_3 = 1$$

Table 1. Simulated data set for a load-sharing system with 3 components subject to Wiener degradation

Observation Time	Component 1	Component 2	Component 3
1	0	0	0
2	2.0734	4.3589	1.8328
3	6.7431	5.0391	2.4822
4	3.2254	6.0425	5.4578
5	5.9497	13.5013	6.1937
6	7.5872	20.6374	13.3039
7	5.9711	27.8062	11.1694
8	6.1047	31.9922	13.8626
9	7.7896	28.6627	16.1383
10	15.9467	33.0312	20.0959
11	22.4856	41.0521	24.0345
12	20.7858	44.5077	24.4391
13	27.8557	50.1461	27.3496
14	30.3065	54.5540	29.8552
15	32.1802	54.8402	34.7384

16	57.5157	41.0184
17	55.8664	47.3486
18	60.9203	47.7549
19	57.8319	50.9869
20	58.0564	50.3446
21	61.3184	50.0042
22		52.9835
23		60.5847
24		61.2747
25		65.3866
26		67.7085
27		74.0692
28		73.7973
29		76.8924
30		81.1523
31		83.5498

The estimation results are listed in Table 2. From Table 2, the final estimated values through EM algorithm after 17 iterations are

$$\hat{\mu}_1 = 0.9063,\ \hat{\sigma}_1 = 1.8904,$$
$$\hat{\mu}_2 = 2.1016,\ \hat{\sigma}_2 = 3.9605,$$
$$\hat{\mu}_3 = 3.0558,\ \hat{\sigma}_3 = 2.9397.$$

The deviation of the estimated values from the real ones are acceptable and the convergence of the iterations are fast (within 17 iterations).

Table 2. Iterations of EM algorithm sequences

Iteration	$\boldsymbol{\mu_1}$	$\boldsymbol{\sigma_1}$	$\boldsymbol{\mu_2}$	$\boldsymbol{\sigma_2}$	$\boldsymbol{\mu_3}$	$\boldsymbol{\sigma_3}$
0	0.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	0.6431	1.7233	2.0211	3.9187	3.1123	2.9433
2	0.7476	1.7647	2.0415	3.9354	3.0824	2.9416
3	0.8611	1.8213	2.0822	3.9462	3.0751	2.9408
4	0.8689	1.8515	2.0911	3.9518	3.0476	2.9404
5	0.8727	1.8721	2.0934	3.9533	3.0489	2.9402
6	0.8928	1.8813	2.0954	3.9546	3.0502	2.9399
7	0.9014	1.8844	2.0971	3.9558	3.0513	2.9398

8	0.9027	1.8864	2.0979	3.9570	3.0523	2.9398
9	0.9033	1.8871	2.0987	3.9581	3.0531	2.9397
10	0.9041	1.8878	2.0992	3.9588	3.0538	2.9397
11	0.9045	1.8883	2.0997	3.9593	3.0545	2.9397
12	0.9051	1.8888	2.1002	3.9598	3.0551	2.9397
13	0.9056	1.8893	2.1007	3.9601	3.0554	2.9397
14	0.9061	1.8897	2.1011	3.9603	3.0557	2.9397
15	0.9062	1.8901	2.1013	3.9604	3.0558	2.9397
16	0.9063	1.8893	2.1015	3.9605	3.0558	2.9397
17	0.9063	1.8904	2.1016	3.9605	3.0558	2.9397

6.2 Comparative study

We illustrate the properties of the EM algorithm in this case by comparing the estimated results with those using Newton type methods, i.e., the classical quasi Newton method in Fletcher[34]. The programs of Newton type methods can be found in existing literatures. In this part, we set component number at 2. The initial values of the parameters are:

$$\mu_1 = 1, \sigma_1 = 2, \mu_2 = 2, \sigma_2 = 4$$

The data used in this analysis is identical to that of Component 1 and 2 in Table 1.

The results are listed in Table 3. As the starting values have significant influence on the performance of Newton type methods, we choose different starting values for both the two methods around the real values of parameters.

Table 3. Estimated results of the parameters through the proposed EM method and the Newton type method

	$\boldsymbol{\mu_1}$	$\boldsymbol{\sigma_1}$	$\boldsymbol{\mu_2}$	$\boldsymbol{\sigma_2}$
Initial value	0.0000	0.5000	1.0000	1.0000
EM	0.9092	1.8904	2.1016	3.9605
Newton	0.2044	2.2513	0.7736	2.4982
Initial value	0.5000	1.0000	1.5000	2.0000
EM	0.9063	1.8904	2.1016	3.9605
Newton	0.2044	2.4541	1.2377	-0.8454
Initial value	0.5000	1.5000	2.0000	3.0000
EM	0.9063	1.8901	2.1016	3.9605
Newton	0.9066	1.8932	2.1016	3.9602

Initial value	1.0000	1.0000	1.5000	3.5000
EM	0.9063	1.8904	2.1016	3.9605
Newton	0.9066	1.8932	2.1016	3.9602
Initial value	2.0000	2.0000	2.5000	3.5000
EM	0.9063	1.8904	2.1016	3.9605
Newton	0.9066	1.8932	2.1016	3.9602

From Table 3, we generate initial values for both the two methods artificially. The proposed EM algorithm has a robust performance. However, the Newton type method is quite sensitive to the initial values. In the first two trials, the initial values are far from the real parameters and the Newton type method fails to converge. Pitfalls are significant for the Newton type method in this specific estimation problem. Although this may not hold in other problems, we can still conclude that our method provides a better and more robust performance than Newton type method. Actually EM algorithm generally outperform Newton type methods in cases under large scale of parameters or sparse prior parameter information.

7 Concluding remarks

In this paper, we deal with an estimation issue for load-sharing systems subject to degradation. As the actual failure time for each component cannot be observed, the closed form of the MLE cannot be derived. We introduce EM method for the parameter estimation in this case. Under the Wiener degradation model, we illustrate the procedure of the proposed EM algorithm and give the closed form parameter estimates for each step. Numerical results of the estimated parameters can be obtained through our method. Meanwhile, we compare the proposed EM algorithm with the Newton type method and show the advantage of EM method against the Newton type method through the numerical study.

In future research, the model assumption can be extended and more degradation models in addition to Wiener degradation model can be investigated for modeling and parameter estimation. Meanwhile, the application of EM algorithm is possible to be extended to more general systems for parameter estimation.

References

1. Singh B, Gupta P K. Load-sharing system model and its application to the real data set[J]. Mathematics and Computers in Simulation, 2012, 82(9): 1615-1629.

2. Daniels H E. The statistical theory of the strength of bundles of threads. I[C]//Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences. The Royal Society, 1945, 183(995): 405-435.

3. Birnbaum Z W, Saunders S C. A statistical model for life-length of materials[J]. Journal of the American Statistical Association, 1958, 53(281): 151-160.

4. Harlow D G, Phoenix S L. Probability distributions for the strength of fibrous materials under local load sharing I: two-level failure and edge effects[J]. Advances in Applied Probability, 1982, 14(1): 68-94.

5. Lee S, Durham S, Lynch J. On the calculation of the reliability of general load sharing systems[J]. Journal of Applied Probability, 1995, 32(3): 777-792.

6. Durham S D, Lynch J D. A threshold representation for the strength distribution of a complex load sharing system[J]. Journal of Statistical Planning and Inference, 2000, 83(1): 25-46.

7. Yang K, Younis H. A semi-analytical Monte Carlo simulation method for system's reliability with load sharing and damage accumulation[J]. Reliability Engineering & System Safety, 2005, 87(2): 191-200.

8. Singh B, Sharma K K, Kumar A. A classical and Bayesian estimation of a k-components load-sharing parallel system[J]. Computational Statistics & Data Analysis, 2008, 52(12): 5175-5185.

9. Levitin G, Dai Y S. Service reliability and performance in grid system with star topology[J]. Reliability Engineering & System Safety, 2007, 92(1): 40-46.

10. Basu A K, Bhattacharya A, Chowdhury S, et al. Planned scheduling for economic power sharing in a CHP-based micro-grid[J]. IEEE Transactions on Power systems, 2012, 27(1): 30-38.

11. Yu H, Eberhard P, Zhao Y, et al. Sharing behavior of load transmission on gear pair systems actuated by parallel arrangements of multiple pinions[J]. Mechanism and Machine Theory, 2013, 65: 58-70.

12. Kim H, Kvam PH. Reliability estimation based on system data with an unknown load share rule. Vol. 10, Lifetime Data Analysis. 2004. p. 83–94.

13. Deshpande J V, Dewan I, Naik-Nimbalkar U V. A family of distributions to model load sharing systems[J]. Journal of Statistical Planning and Inference, 2010, 140(6): 1441-1451.

14. Park C. Parameter estimation for the reliability of load-sharing systems[J]. IIE Transactions, 2010, 42(10): 753-765.

15. Park C. Parameter estimation from load-sharing system data using the expectation–maximization algorithm[J]. IIE Transactions, 2013, 45(2): 147-163.

16. Singpurwalla N D. Survival in dynamic environments[J]. Statistical Science, 1995: 86-103.

17. Jardine A K S, Lin D, Banjevic D. A review on machinery diagnostics and prognostics implementing condition-based maintenance[J]. Mechanical systems and signal processing, 2006, 20(7): 1483-1510.

18. Si X S, Wang W, Hu C H, et al. Remaining useful life estimation–a review on the statistical data driven approaches[J]. European journal of operational research, 2011, 213(1): 1-14.

19. Whitmore G A, Schenkelberg F. Modelling accelerated degradation data using Wiener diffusion with a time scale transformation[J]. Lifetime data analysis, 1997, 3(1): 27-45.

20. Zhai Q, Ye Z S. RUL Prediction of Deteriorating Products Using An Adaptive Wiener Process Model[J]. IEEE Transactions on Industrial Informatics, 2017.

21. Singpurwalla N. Gamma processes and their generalizations: an overview. In: Engineering probabilistic design and maintenance for flood protection. Springer; 1997. p. 67–75.

22. Van Noortwijk J M. A survey of the application of gamma processes in maintenance[J]. Reliability Engineering & System Safety, 2009, 94(1): 2-21.

23. Ye Z S, Chen N. The inverse Gaussian process as a degradation model[J]. Technometrics, 2014, 56(3): 302-311.

24. Ye Z S, Xie M. Stochastic modelling and analysis of degradation for highly reliable products[J]. Applied Stochastic Models in Business and Industry, 2015, 31(1): 16-32.

25. Zhai Q, Ye Z-S. Degradation in Common Dynamic Environments. Technometrics. 2017;(just-accepted).

26. Liu B, Xie M, Kuo W. Reliability modeling and preventive maintenance of load-sharing systemswith degrading components[J]. IIE Transactions, 2016, 48(8): 699-709.

27. Nikulin M S, Limnios N, Balakrishnan N, et al. Advances in degradation modeling: applications to reliability[J]. Survival Analysis, and Finance. Birkhauser Boston, 2010.

28. Durham S D, Lynch J D, Padgett W J, et al. Localized load-sharing rules and Markov-Weibull fibers: A comparison of microcomposite failure data with Monte Carlo simulations[J]. Journal of composite materials, 1997, 31(18): 1856-1882.

29. Dempster A P, Laird N M, Rubin D B. Maximum likelihood from incomplete data via the EM algorithm[J]. Journal of the royal statistical society. Series B (methodological), 1977: 1-38.

30. Wu C F J. On the convergence properties of the EM algorithm[J]. The Annals of statistics, 1983: 95-103.

31. Little RJA, Rubin DB. Statistical analysis with missing data. John Wiley & Sons; 2014.

32. Taghipour S, Banjevic D. Trend analysis of the power law process using Expectation–Maximization algorithm for data censored by inspection intervals[J]. Reliability Engineering & System Safety, 2011, 96(10): 1340-1348.

33. Ye Z S, Shen Y, Xie M. Degradation-based burn-in with preventive maintenance[J]. European Journal of Operational Research, 2012, 221(2): 360-367.

34. Fletcher R. A new approach to variable metric algorithms[J]. The computer journal, 1970, 13(3): 317-322.