Estimation of a Flexible Tampered Failure Rate Model Under Stage Life Testing Experiment

Dhrubasish Bhattacharyya[a], Debasis Kundu[b]

[a] *Department of Mathematical Sciences, Rajiv Gandhi Institute of Petroleum Technology, Jais, Amethi - 229304, Uttar Pradesh, India.*

[b] *Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India.*

Abstract

Life testing experiments are crucial for assessing product reliability, but traditional methods struggle with high product reliability, destructive testing, and time constraints. Stage Life Testing (SLT), which combines progressive censoring and step-stress testing, offers a flexible alternative by re-testing censored items under different stress levels. This article deals with the estimation methodology under SLT experiments using the tampered failure rate model, which assumes multiplicative effects on failure rates in different stress levels. The estimation procedure is proposed without making any specific assumptions about the shape of the hazard rate functions. Instead, it is assumed that the hazard functions in different stress levels belong to a class of distributions characterized by piecewise constant hazard rates. To address the challenges of complex maximum likelihood estimation, we propose an Expectation Conditional Maximization (ECM) algorithm, which efficiently handles parameter estimation. Simulation study is carried out to assess the performance of the proposed estimation procedure. Methodology for accommodating unknown cut-points is also proposed and illustrations using simulated data are presented. Extension to general k-step stage life testing is also considered. Finally, a computational algorithm is proposed to determine optimal censoring scheme based on D-optimality criteria.

Keywords: Stage Life Testing, Expectation-Maximization algorithm, Expectation Conditional Maximization algorithm, Maximum likelihood estimators, Bootstrap confidence interval.
AMS 2010 classification code : 62N01, 62N02, 62F10.

---

1. Introduction

In recent years, life testing experiments have become increasingly important for assessing the reliability of products and systems. The primary objective of such experiments is to evaluate key reliability characteristics of the test units. Traditionally, this involves subjecting a group of identical items to normal operating conditions and recording the time until each item fails. However, with modern advances in science and technology, many industrial products now exhibit high reliability with extended lifespans. As a result, conventional life testing methods often require long durations to collect meaningful failure time data. In addition, many life tests are destructive in nature, rendering the tested items unusable afterward. To overcome these limitations, researchers commonly adopt Accelerated Life Testing (ALT) techniques and various censoring strategies to obtain useful reliability information within a feasible time frame.

Email addresses: `dhrubasish018@gmail.com` (Dhrubasish Bhattacharyya), `kundu@iitk.ac.in` (Debasis Kundu)

In an ALT experiment, the experimental units are subjected to stress levels higher than those under normal operating conditions. This results in failure of items earlier than under normal conditions. For comprehensive surveys on accelerated life tests, we refer to Meeker and Hahn (1985), Meeker et al. (2021), Bagdonavičius and Nikulin (2001) and Nelson (2009). A special type of ALT experiment is the step-stress life test, which enables the experimenter to change the level of the stress factors in a sequential manner during the experiment; see DeGroot and Goel (1979), Nelson (1980), Balakrishnan et al. (2007), Balakrishnan and Xie (2007a), Balakrishnan and Xie (2007b), Han and Balakrishnan (2010) and others. For a comprehensive review on step-stress, see Kundu and Ganguly (2017). A step-stress life testing with two different stress levels is known as a simple step-stress life testing. The simple step-stress model is a special case of step-stress testing with only two stress levels and one stress-change time. Initially, n identical items are put on the initial stress level s_1. The stress level is changed to s_2 (generally $s_2 > s_1$) at the prefixed stress-change time T_1. If all n failures are observed before T_1 on stress level s_1 then the life test is terminated at the n-th failure time without changing the stress level. Otherwise, all surviving items are put on the increased stress level s_2 at T_1 until failure. The life test continues until all n items have failed.

In order to make life testing experiments cost-effective and less time-consuming, it is also a standard practice to employ different censoring schemes. The most commonly used censoring schemes are Type-I and Type-II. In Type-I censoring, the experiment is terminated at a pre-specified time. This has the possibility that no failure is observed during the time. To ensure a certain number of failures in a life testing experiment, Type-II censoring had been introduced in the literature where the experiment is terminated after a specific number of failures are observed. During the last two decades, the progressive censoring schemes have received widespread attention in statistical literature since these schemes add flexibility to other censoring schemes by allowing withdrawals of items during experiment. The most popular progressive censoring schemes are progressive Type-I and progressive Type-II schemes; see Cohen (1963), Basak et al. (2009), Chen and Lio (2010), Kundu and Joarder (2006), and Kundu (2008) in this regard. A comprehensive discussion of probabilistic properties and inferential results is provided by Balakrishnan and Aggarwala (2000) and Balakrishnan (2007).

It is imperative to observe that progressively censored items in a life testing experiment can further be applied in other life testing experiments or they can be further used on the test at a higher stress level. Therefore, if the study on the remaining life time of the censored units on the other stress level is conducted along with the original stress level, the loss of information due to censoring in the original life testing experiment can be compensated. Recently, in order to accommodate the further use of the progressively censored items, Laumen and Cramer (2019) have proposed a new life testing experiment where the censored items are further put on the life testing experiment with different (usually higher) stress level. This life testing is termed as 'Stage Life Testing' (SLT) and it connects both the ideas of progressive censoring and step-stress life testing. They adapted the idea of step-stress testing and assumed that the censored items are tested on a different stage whereas the remaining items are still tested under standard conditions. The basic situation with one censoring time T_1 and two stages s_1, s_2 is inspired by the simple step-stress testing. Laumen and Cramer (2019) also proposed the estimation procedure for two exponential populations in the context of the cumulative exposure model (introduced by Sedyakin, 1966). The extension to general k stages has also been considered by Laumen and Cramer (2021). The cumulative exposure (CE) model relates the lifetime distribution of items at one stress level to the lifetime distribution at the preceding stress level by shifting the lifetime distribution; see Nelson (1980) for details. More recently, Samanta and Kundu (2024a) explored both classical and

Bayesian inference for the Weibull model parameters in a two-stage life testing setup, assuming the distributions at the two stress levels are connected through the Khamis–Higgins model. Samanta and Kundu (2024b) considered both classical and Bayesian inference of a dependent competing risk model when data are collected from a simple step-stress stage life-test experiment. The dependency structures are explained through the Marshall–Olkin bivariate Weibull distribution. Furthermore, Samanta et al. (2022) proposed order-restricted estimation techniques for model parameters in a multi-step SLT framework, assuming Weibull lifetimes with common shape but varying scale parameters across stress levels. They also developed computational methods for determining optimal censoring schemes in such settings.

In this article, we consider the estimation of a flexible family of distributions in the context of a tampered failure rate model. The tampered failure rate model relies on the assumption that change of stress has a multiplicative effect on the failure rate function over the remaining life, and is widely used in step-stress life testing. It is pertinent to mention that the CE model provides no transparent physical motivation except for the special case of a scale family of distributions in different stress levels, whereas the tampered failure rate model is a plausible formulation when one views the stress to act in the form of a time acceleration; see Bhattacharyya and Soejoeti (1989) for detailed discussions regarding the same. This motivates us to consider the tampered failure rate model in the SLT setup. Further, we propose estimation procedures without making any specific assumptions about the shape of hazard rate functions. Instead, we assume that the hazard rate functions in different stress levels belong to a class of distributions characterized by piecewise constant hazard rates. This assumption introduces significant flexibility into our model, making it adaptable to a wide range of real-world scenarios. A key strength of our proposed model is its data-driven approach to the shape of the failure rate function. We consider the maximum likelihood estimates (MLEs) of the model parameters assuming that the number of cut points and the positions of the cut points are known. The complicated likelihood function in our model precludes the derivation of closed-form expressions for the maximum likelihood estimators (MLEs), thereby necessitating the use of high-dimensional numerical optimization techniques. However, standard optimization algorithms often struggle in such settings due to their sensitivity to initial values and the computational burden imposed by increased dimensionality. Additionally, the conventional Expectation-Maximization (EM) algorithm proves inadequate, as it does not yield a tractable or explicit M-step in our framework. To overcome these challenges, we adopt the Expectation Conditional Maximization (ECM) algorithm, an extension of the EM approach that decomposes the M step into a sequence of simpler, constrained maximization steps. This strategy allows for a more efficient and stable estimation of the model parameters. Furthermore, recognizing that the locations of the hazard rate cut points are typically unknown in practical applications, we extend our methodology to account for this uncertainty. This enhancement significantly broadens the applicability of our model, making it both versatile and well suited for real-world applications.

The EM algorithm, introduced by Dempster (1977), is a two-step iterative procedure used to obtain MLEs of unknown parameters of a likelihood function. The algorithm consists of two main steps: the E-step, where the expected value of the complete-data log-likelihood is calculated based on the observed data and current parameter estimates; and the M-step, where this expectation is maximized with respect to the parameters. These steps are repeated iteratively until convergence. A key advantage of the EM algorithm lies in its ability to handle missing or censored observations; see, for example, Ng et al. (2002) and Balakrishnan and Ling (2012). Despite its widespread applicability, the EM algorithm often leads to a computationally intensive or analytically intractable maximization step. To address this, Meng and Rubin (1993) proposed the Expectation Conditional Maximization (ECM) algorithm, a variant of EM in which the M step is decomposed into a sequence of simpler Conditional Maximization (CM) steps. In our context, where the M-step is also highly intricate, we adopt the ECM framework, which allows us to perform two successive constrained optimization steps more efficiently.

The paper is organized as follows. In Section 2, we provide a description of the model and a construction of the likelihood function. The detailed estimation methodology assuming known cut-points is presented in Section 3. An extensive simulation study is carried out and the results are presented in Section 4. Section 5 deals with a methodology to accommodate unknown cut-points. Illustrations using simulated data sets are also presented in Section 5. The extension of the methodology to general k-stage life testing is also proposed in Section 6. Finally, in Section 7, we present the computation of the optimal censoring scheme using a D-optimality based criteria.

2. Model description and the likelihood function

Here we present a brief description of the model and the construction of the likelihood function. Suppose that n identical objects are placed on a life test. The initial conditions are called stage s_1. At the prefixed stress-change time T_1, R_1^* items of the surviving items are randomly withdrawn and further tested on stage s_2. Suppose D_1 is the number of observed failures until time T_1 and D_2 is the number of observed failures after time T_1 in stage s_1. Let $Y_{1:M:n} \leq \cdots \leq Y_{D_1:M:n}$ be the ordered observations in stage s_1 before time T_1 and $Y_{D_1+1:M:n} \leq \cdots \leq Y_{D_1+D_2:M:n}$ be the ordered observations in stage s_1 from time T_1 onward. Here, $M = D_1 + D_2$ denotes the number of failures observed in stage s_1. The number of withdrawals R_1^* is chosen as $\rho(D_1)$, where $\rho(x)$ is typically defined either as $\rho(x) = \lfloor \pi \lfloor n - x \rfloor \rfloor$, with $\lfloor \cdot \rfloor$ denoting the floor function, or as $\rho(x) = \min\{ n - x, R_1 \}$. Here the proportion $\pi \in [0, 1]$ and the number R_1 are prespecified. Suppose D_1^* is the number of observed failures in stage s_2. Also, let $Z_{1:D_1^*,R_1^*} \leq \cdots \leq Z_{D_1^*:D_1^*,R_1^*}$ be the ordered failure times observed in stage s_2. We also assume that the experiment is terminated at prefixed time T_2. A schematic diagram of the stage life testing is presented in Figure 1 for illustration.

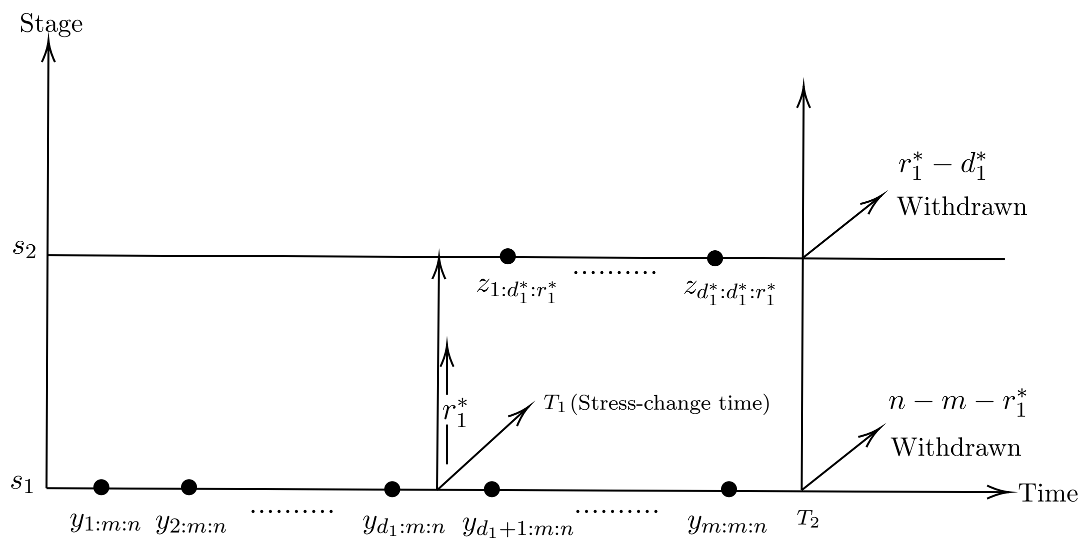

Figure 1: Illustration of SLT with stress-change time T_1 and test termination time T_2.

We assume that the failure rate function $r(t)$ of an item undergoing stress levels s_1 and s_2 is given by

$$h(t) = \begin{cases} \theta_1 h_0(t), & \text{under stress } s_1 \\ \theta_2 h_0(t), & \text{under stress } s_2. \end{cases}$$

Here we assume that the hazard rate $h_0(t)$ is piecewise constant. Hence $h_0(t)$ can be written as

$$h_0(t) = \sum_{j=1}^{\mathcal{M}} c_j I_{[\tau_{j-1}, \tau_j)}(t), \text{ where } I_{[\tau_{j-1}, \tau_j)}(t) = \begin{cases} 1, & \text{if } t \in [\tau_{j-1}, \tau_j) \\ 0, & \text{otherwise} \end{cases}$$

and $0 = \tau_0 < \tau_1 < \cdots < \tau_{\mathcal{M}-1} < \tau_{\mathcal{M}} = \infty$ are known cut points, and the c_j's are unknown quantities. In order to avoid identifiability issue, we assume $c_{\mathcal{M}} = 1$ without loss of generality. Observe that $\mathcal{M} = 1$ corresponds to the exponential model. We further assume that $\tau_{\mathcal{M}-1}$ is strictly less than the test termination time T_2. The cumulative hazard function $H_0(t)$ can be obtained as

$$H_0(t) = \int_0^t h_0(\tau) d\tau = \sum_{j=1}^{\mathcal{M}} c_j \left\{ \min\left(t, \tau_j\right) - \tau_{j-1} \right\} I_{[\tau_{j-1}, \tau_{\mathcal{M}})}(t).$$

The survival function $S_0(t)$ and the pdf $f_0(t)$ can then be obtained from the following relations

$$S_0(t) = e^{-H_0(t)} \text{ and } f_0(t) = h_0(t) e^{-H_0(t)}.$$

Hence it follows that the survival function $S_1(t)$ and the probability density function (pdf) $f_1(t)$ of an item in stage s_1 can be obtained as

$$S_1(t) = e^{-\theta_1 H_0(t)} \text{ and } f_1(t) = \theta_1 h_0(t) e^{-\theta_1 H_0(t)},$$

respectively. Now for $t \geq T_1$, the survival function under stage s_2 can be written as

$$S_2(t) = e^{-\int_0^t h(u)du} = e^{-\theta_1 \int_0^{T_1} h(u)du - \theta_2 \int_{T_1}^t h_0(u)du} = e^{-(\theta_1 - \theta_2) H_0(T_1)} e^{-\theta_2 H_0(t)}.$$

Hence the pdf of the lifetime of an item under stage s_2 can written as

$$f_2(t) = \theta_2 h_0(t) e^{-(\theta_1 - \theta_2) H_0(T_1)} e^{-\theta_2 H_0(t)}, \quad T_1 \leq t < T_2.$$

Then for given value of $R_1^* = r_1^*$, the likelihood function based on the observations $\mathbf{y} = (y_{1:m:n}, \ldots, y_{m:m:n})$ and $\mathbf{z} = \left(z_{1:d_1^*:r_1}, \ldots, z_{d_1^*:d_1^*:r_1^*}\right)$, can be written as

$$\begin{aligned}
\mathcal{L}(\mathbf{c}, \theta | \mathbf{y}, \mathbf{z}) &= C \prod_{i=1}^{m} f_1\left(y_{i:m:n}\right) \left\{S_1(T_2)\right\}^{n-m-r_1^*} \prod_{j=1}^{d_1^*} f_2\left(z_{j:d_1^*:r_1^*}\right) \left\{S_2(T_2)\right\}^{r_1^* - d_1^*} \\
&= C \theta_1^m \exp\left\{ -\theta_1 \sum_{i=1}^{m} H_0\left(y_{i:m:n}\right) - \theta_1 (n - m - r_1^*) H_0(T_2) \right\} \prod_{i=1}^{m} h_0\left(y_{i:m:n}\right) \prod_{j=1}^{d_1^*} h_0\left(z_{j:d_1^* r_1^*}\right) \\
&\times \theta_2^{d_1^*} \exp\left\{ -r_1^* (\theta_1 - \theta_2) H_0(T_1) - \theta_2 \sum_{j=1}^{d_1^*} H_0\left(z_{j:d_1^* r_1^*}\right) - \theta_2 (r_1^* - d_1^*) H_0(T_2) \right\},
\end{aligned}$$

where $C = \binom{n}{d_1}\binom{n-d_1-r_1^*}{d_2} d_1! d_2!$. Hence the log-likelihood function can be written as

$$
\begin{aligned}
\log \mathcal{L}(\mathbf{c}, \theta | \mathbf{y}, \mathbf{z}) \propto\ & m \log \theta_1 + \sum_{i=1}^{m} \log h_0\left(y_{i:m:n}\right) - \theta_1 \sum_{i=1}^{m} H_0\left(y_{i:m:n}\right) - \theta_1 (n - m - r_1^*) H_0(T_2) \\
& + d_1^* \log \theta_2 + \sum_{j=1}^{d_1^*} \log h_0\left(z_{j:d_1^* r_1^*}\right) - r_1^*(\theta_1 - \theta_2) H_0(T_1) - \theta_2 \sum_{j=1}^{d_1^*} H_0\left(z_{j:d_1^*:r_1^*}\right) \\
& - \theta_2 (r_1^* - d_1^*) H_0(T_2).
\end{aligned}
$$

Note that one needs to solve a $(\mathcal{M}+1)$-dimensional nonlinear optimization problem in order to obtain the maximum likelihood estimates of the parameters. Standard optimization methods such as the Newton-Raphson method may be used to solve the problem. However, these methods have some disadvantages, specifically, these methods are quite sensitive to initial guesses. Consequently, it is difficult to choose the initial guesses, particularly if $\mathcal{M}$ is large. In order to address this issue, we reformulate the problem as a missing data problem and propose an efficient Expectation Conditional Maximization (ECM) algorithm to compute the MLEs of the model parameters.

3. Missing data approach and the ECM algorithm

We now treat the estimation of model parameters under the SLT experiment as an incomplete data problem and propose an ECM algorithm to obtain the maximum likelihood estimators of the unknown model parameters. We denote the missing data by $\boldsymbol{M} = (\boldsymbol{M}_1, \boldsymbol{M}_2)$, where $\boldsymbol{M}_1 = \left(M_{1,1}, \ldots, M_{1,n-m-r_1^*}\right)$ is the vector of missing observations in stage s_1 and $\boldsymbol{M}_2 = \left(M_{2,1}, \ldots, M_{2,r_1^*-d_1^*}\right)$ is the vector of missing observations in stage s_2. The complete data can be obtained by combining $\boldsymbol{Y}$, $\boldsymbol{Z}$ and $\boldsymbol{M}$, i.e., is given by $\boldsymbol{X} = (\boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{M})$. The following theorem provides the conditional distribution of missing data $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ given the observations $(\boldsymbol{y}, \boldsymbol{z})$.

Theorem 3.1. *The random variables $M_{1,i}$'s ($M_{2,j}$'s) are independent and identically distributed for $i = 1, \ldots, n - m - r_1^*$ ($j = 1, \ldots, r_1^* - d_1^*$). Also, the conditional distributions of $M_{1:1}$ and $M_{2:1}$ are given by*

$$
f_{M_{1,1}|\boldsymbol{Y},\boldsymbol{Z}}\left(t|\boldsymbol{y}, \boldsymbol{z}\right) = \frac{f_1(t)}{S_1(T_2)},\ t > T_2
$$

and

$$
f_{M_{2,1}|\boldsymbol{Y},\boldsymbol{Z}}\left(t|\boldsymbol{y}, \boldsymbol{z}\right) = \frac{f_2(t)}{S_2(T_2)},\ t > T_2,
$$

respectively.

Proof: The proof closely follows the arguments used in Ng et al. (2002) and is therefore omitted for conciseness. □

Now we will briefly review the ECM algorithm which will be useful for further developments of the article. Suppose $\boldsymbol{X} = (X_1, \ldots, X_m)$ is a set of random observations from a distribution with pdf $f(x; \theta)$ where $\theta \in \mathbb{R}^d$ is a set of unknown parameters. In addition, suppose that $\boldsymbol{Z} = (Z_1, \ldots, Z_n)$ is another set of incomplete observations on an unobservable random variable Z and $\boldsymbol{Y} = (\boldsymbol{X}, \boldsymbol{Z})$ is the set of complete observations. Let $f(\boldsymbol{y}; \theta) = f(\boldsymbol{x}, \boldsymbol{z}; \theta)$ be the pdf governing $\boldsymbol{Y}$ and $f(\boldsymbol{z}|\boldsymbol{x}; \theta)$ be the conditional pdf of $\boldsymbol{Z}|\boldsymbol{X}$. Then the EM algorithm consists of iterations of the following two steps:

(i) The E-step computes the expected value of the log-likelihood function of the complete data with respect to the conditional pdf $f(\boldsymbol{z}|\boldsymbol{x};\theta^{(h)})$, i.e.,

$$Q\left(\theta|\boldsymbol{x};\theta^{(h)}\right) = \int \{\log f(\boldsymbol{x},\boldsymbol{z};\theta)\}\, f(\boldsymbol{z}|\boldsymbol{x};\theta^{(h)})d\boldsymbol{z}.$$

(ii) The M-step finds $\theta^{(h+1)} = \arg\max_{\theta} Q\left(\theta|\boldsymbol{x};\theta^{(h)}\right)$.

Here $\theta^{(1)},\ldots,\theta^{(h)}$ are the solutions obtained from h iterations and $\theta^{(0)}$ is the initial guess of θ. The ECM algorithm replaces the single M-step (ii) with a number of conditional M-steps (see Meng and Rubin, 1993 for details). Suppose $\{g_s(\theta) : s = 1,\ldots,S\}$ are S preselected vector-valued functions of θ. Let us define the constraint sets corresponding to the S constrained maximization problems as

$$\Theta_s\left(\theta^{(h+(s-1)/S)}\right) := \left\{\theta : g_s(\theta) = g_s\left(\theta^{(h+(s-1)/S)}\right)\right\},\ s = 1,\ldots,S;$$

where $\theta^{(h+(s-1)/S)}$, $s = 1,\ldots,S$ is a sequence of solutions to the maximization problem described as follows. Suppose we have solutions $\theta^{(1)},\ldots,\theta^{(h)}$ from h iterations. Then the $(h+1)$-th iteration consists of S sub-iterations described below:

(a) for $s = 1$, we solve

$$\arg\max_{\theta} Q\left(\theta|\boldsymbol{x};\theta^{(h)}\right) \text{ subject to } \theta \in \Theta_1\left(\theta^{(h)}\right)$$

and denote this solution by $\theta^{(h+1/S)}$.

(b) For $s = 2,\ldots,S-1$, we solve

$$\arg\max_{\theta} Q\left(\theta|\boldsymbol{x};\theta^{(h)}\right) \text{ subject to } \theta \in \Theta_s\left(\theta^{(h+(s-1)/S)}\right)$$

and denote this solution by $\theta^{(h+s/S)}$, $s = 2,\ldots,S-1$.

(c) For $s = S$, we solve

$$\arg\max_{\theta} Q\left(\theta|\boldsymbol{x};\theta^{(h)}\right) \text{ subject to } \theta \in \Theta_S\left(\theta^{(h+(S-1)/S)}\right)$$

and denote this solution by $\theta^{(h+1)}$.

This completes the $(h+1)$-th iteration of the algorithm. Note that

$$Q\left(\theta^{(h+1)}|\boldsymbol{x};\theta^{(h)}\right) \geq Q\left(\theta^{(h+(S-1)/S)}|\boldsymbol{x};\theta^{(h)}\right) \geq \cdots \geq Q\left(\theta^{(h+1/S)}|\boldsymbol{x};\theta^{(h)}\right) \geq Q\left(\theta^{(h)}|\boldsymbol{x};\theta^{(h)}\right).$$

Hence the ECM algorithm is a generalized EM algorithm, see Neal and Hinton (1998) for details. We now proceed to obtain the ECM algorithm in order to obtain the MLEs of the proposed model.

3.1. The Expectation (E)-Step

Suppose $\lambda(\boldsymbol{\theta},\boldsymbol{c}|\boldsymbol{X})$ denote the log-likelihood function corresponding to the complete data. Then the E-step of the ECM algorithm requires the computation of the expectation of $\lambda(\boldsymbol{\theta},\boldsymbol{c}|\boldsymbol{X})$, with respect to the conditional distribution of $\boldsymbol{M}$, given $\boldsymbol{Y},\boldsymbol{Z}$ and the current estimate of the parameters $\boldsymbol{\theta}^{(h)},\boldsymbol{c}^{(h)}$. Under the proposed model, the expectation of $\lambda(\boldsymbol{\theta},\boldsymbol{c}|\boldsymbol{X})$, given $\boldsymbol{y},\boldsymbol{z}$ and

$\boldsymbol{\theta^{(h)}}, \boldsymbol{c^{(h)}}$ can be written as

$$
\begin{aligned}
Q\left(\boldsymbol{\theta}, \boldsymbol{c}|\boldsymbol{\theta}^{(h)}, \boldsymbol{c}^{(h)}\right) = & (n-r_1^*)\log\theta_1 + \sum_{i=1}^{m}\log h_0\left(y_{i:m:n}\right) - \theta_1\sum_{i=1}^{m} H_0(y_{i:m:n}) + r_1^*\log\theta_2 \\
& - r_1^*(\theta_1-\theta_2)H_0(T_1) + \sum_{j=1}^{d_1^*}\log h_0(z_{j:d_1^*:r_1^*}) - \theta_2\sum_{j=1}^{d_1^*} H_0(z_{j:d_1^*:r_1^*}) \\
& + (n-m-r_1^*)\left\{E\left[\log h_0(M_{1:1})\right] - \theta_1 E\left[H_0(M_{1:1})\right]\right\} \\
& + (r_1^*-d_1^*)\left\{E\left[\log h_0(M_{2:1})\right] - \theta_2 E\left[H_0(M_{2:1})\right]\right\}. \qquad (1)
\end{aligned}
$$

Note that the conditional distributions of $M_{1:1}$ and $M_{2:1}$, given the current estimates of the parameters $\boldsymbol{c}^{(h)}, \boldsymbol{\theta}^{(h)}$, are given by

$$
f_{M_{1:1}|\boldsymbol{c}^{(h)},\boldsymbol{\theta}^{(h)}}(t) = \frac{\theta_1^{(h)} h_0^{(h)}(t)}{e^{-\theta_1^{(h)} H_0^{(h)}(T_2)}} e^{-\theta_1^{(h)} H_0^{(h)}(t)}, \ t > T_2
$$

and

$$
f_{M_{2:1}|\boldsymbol{c}^{(h)},\boldsymbol{\theta}^{(h)}}(t) = \frac{\theta_2^{(h)} h_0^{(h)}(t)}{e^{-\theta_2^{(h)} H_0^{(h)}(T_2)}} e^{-\theta_2^{(h)} H_0^{(h)}(t)}, \ t > T_2,
$$

respectively. We now proceed to obtain the expression for $Q\left(\boldsymbol{\theta}, \boldsymbol{c}|\boldsymbol{\theta}^{(h)}, \boldsymbol{c}^{(h)}\right)$. Note that since $T_2 > \tau_{M-1}$, $E\left[H_0(M_{1:1})\right]$ can be obtained as

$$
E\left[H_0(M_{1:1})\right] = \frac{\theta_1^{(h)}}{S_1^{(h)}(T_2)}\int_{T_2}^{\infty} H_0(x) e^{-\theta_1^{(h)} H_0^{(h)}(x)} dx = \frac{\theta_1^{(h)}}{S_1^{(h)}(T_2)} I, \text{ say,}
$$

where I can written as

$$
\begin{aligned}
I &= e^{-\theta_1^{(h)} H_0^{(h)}(\tau_{\mathcal{M}-1})}\int_{T_2}^{\infty}\left\{H_0(\tau_{\mathcal{M}-1}) + (x-\tau_{\mathcal{M}-1})\right\} e^{-\theta_1^{(h)}(x-\tau_{\mathcal{M}-1})} dx. \\
&= e^{-\theta_1^{(h)} H_0^{(h)}(\tau_{\mathcal{M}-1})}\left(I_1 + I_2\right), \text{ say.}
\end{aligned}
$$

Here $I_1 = H_0(\tau_{\mathcal{M}-1}) e^{-\theta_1^{(h)}(T_2-\tau_{\mathcal{M}-1})}/\theta_1^{(h)}$ and simplification followed by integration by parts yields

$$
I_2 = \frac{1}{\theta_1^{(h)}}\left[\left(T_2-\tau_{\mathcal{M}-1}\right) e^{-\theta_1^{(h)}(T_2-\tau_{\mathcal{M}-1})} + \frac{1}{\theta_1^{(h)}} e^{-\theta_1^{(h)}(T_2-\tau_{\mathcal{M}-1})}\right].
$$

Hence it follows that

$$
\begin{aligned}
I &= \frac{1}{\theta_1^{(h)}} e^{-\left\{\theta_1^{(h)} H_0^{(h)}(\tau_{\mathcal{M}-1}) + \theta_1^{(h)}(T_2-\tau_{\mathcal{M}-1})\right\}}\left[H_0(\tau_{\mathcal{M}-1}) + (T_2-\tau_{\mathcal{M}-1}) + \frac{1}{\theta_1^{(h)}}\right] \\
&= \frac{S_1^{(h)}(T_2)}{\theta_1^{(h)}}\left\{H_0(T_2) + \frac{1}{\theta_1^{(h)}}\right\}
\end{aligned}
$$

and consequently $E\left[H_0(M_{1:1})\right] = H_0(T_2) + 1/\theta_1^{(h)}$. Similar calculations show that $E\left[H_0(M_{2:1})\right] = H_0(T_2) + 1/\theta_2^{(h)}$. Since the test termination time $T_2 > \tau_{\mathcal{M}-1}$ and $h_0(t) = c_{\mathcal{M}} = 1$ for $t \in [\tau_{\mathcal{M}-1}, \tau_{\mathcal{M}})$, it follows that $E[\log h_0(M_{1:1})] = E[\log h_0(M_{2:1})] = 0$. Substituting these quantities in (1), one can obtain $Q\left(\boldsymbol{\theta}, \boldsymbol{c} | \boldsymbol{\theta}^{(h)}, \boldsymbol{c}^{(h)}\right)$ in the E-step of the proposed algorithm.

3.2. The Conditional Maximization (CM)-Steps

We now formalize the CM-steps for the proposed model. Let $\boldsymbol{\phi} = (\boldsymbol{\theta}, \boldsymbol{c})$ and suppose that we have the solutions $\boldsymbol{\phi}^{(0)}, \ldots \boldsymbol{\phi}^{(h)}$ from h iterations, where $\boldsymbol{\phi}^{(h)} = (\boldsymbol{\theta}^{(h)}, \boldsymbol{c}^{(h)})$. We consider the constraint sets Θ_s as given by

$$\Theta_s\left(\boldsymbol{\phi}^{(h+(s-1)/2)}\right) := \left\{\boldsymbol{\phi} : g_s(\boldsymbol{\phi}) = g_s\left(\boldsymbol{\phi}^{(h+(s-1)/2)}\right)\right\},\ s = 1, 2;$$

where $g_s(\boldsymbol{\phi})$, $s = 1, 2$ are preselected vector valued functions of ϕ. In our case, $g_1(\boldsymbol{\phi}) = \boldsymbol{\theta}$ and $g_2(\boldsymbol{\phi}) = \boldsymbol{c}$. Here $\boldsymbol{\phi}^{(h+(s-1)/2)}$, $s = 1, 2$ is a sequence of solutions to the maximization problem described as follows. The $(h+1)$-th iteration of the algorithm contains two sub-iterations as given below:

1) For $s = 1$, we solve

$$\arg\max_{\boldsymbol{\phi}} Q\left(\boldsymbol{\theta}, \boldsymbol{c} | \boldsymbol{\theta}^{(h)}, \boldsymbol{c}^{(h)}\right) \text{ subject to } \boldsymbol{\phi} \in \Theta_1\left(\boldsymbol{\phi}^{(h)}\right)$$

and denote this solution by $\boldsymbol{\phi}^{(h+1/2)}$.

2) For $s = 2$, we solve

$$\arg\max_{\boldsymbol{\phi}} Q\left(\boldsymbol{\theta}, \boldsymbol{c} | \boldsymbol{\theta}^{(h)}, \boldsymbol{c}^{(h)}\right) \text{ subject to } \boldsymbol{\phi} \in \Theta_2(\boldsymbol{\phi}^{(h+1/2)})$$

and denote this solution by $\boldsymbol{\phi}^{(h+1)}$.

This completes the $(h+1)$-th iteration of the algorithm. In order to facilitate the ECM algorithm, we now proceed to obtain the explicit expressions for the CM-steps. Note that the partial derivatives of $H_0(x)$ and $\log h_0(x)$ with respect to c_q are given by

$$\frac{\partial H_0(x)}{\partial c_q} = \left\{\min(x, \tau_q) - \tau_{q-1}\right\} I_{[\tau_{q-1}, \tau_{\mathcal{M}})}(x) = P(x, q),\ \text{say}$$

and $\frac{\partial \log h_0(x)}{\partial c_q} = I_{[\tau_{q-1}, \tau_q)}/c_q$, for $q = 1, \ldots, \mathcal{M} - 1$. Hence it follows that

$$\begin{aligned}
\frac{\partial Q}{\partial c_q} = \frac{1}{c_q}&\left\{\sum_{i=1}^{m} I_{[\tau_{q-1}, \tau_q)}(y_{i:m:n}) + \sum_{j=1}^{d_1^*} I_{[\tau_{q-1}, \tau_q)}(z_{j:d_1^*:r_1^*})\right\} \\
&- \theta_1^{(h)} \sum_{i=1}^{m} P(y_{i:m:n}, q) - \theta_2^{(h)} \sum_{j=1}^{d_1^*} P(z_{j:d_1^*:r_1^*}, q) - r_1^*(\theta_1^{(h)} - \theta_2^{(h)}) P(T_1, q) \\
&- (\tau_q - \tau_{q-1})\left\{\theta_1^{(h)}(n - m - r_1^*) + \theta_2^{(h)}(r_1^* - d_1^*)\right\}.
\end{aligned}$$

Further calculations show that

$$\frac{\partial Q}{\partial \theta_1} = \frac{n - r_1^*}{\theta_1} - \sum_{i=1}^{m} H_0^{(h)}(y_{i:m:n}) - r_1^* H_0^{(h)}(T_1) - (n - m - r_1^*)\left\{H_0^{(h)}(T_2) + 1/\theta_1^{(h)}\right\}$$

and

$$\frac{\partial Q}{\partial \theta_2} = \frac{r_1^*}{\theta_2} - \sum_{j=1}^{d_1^*} H_0^{(h)}(z_{j:d_1^*:r_1^*}) + r_1^* H_0^{(h)}(T_1) - (r_1^* - d_1^*)\left\{H_0^{(h)}(T_2) + 1/\theta_2^{(h)}\right\}.$$

Now using (1), the iterative CM-steps at $(h+1)$-th iteration of the algorithm can be computed as follows:

(i) For $s = 1$, the solution of the constrained maximization problem can be obtained as $\boldsymbol{\phi}^{(h+1/2)} = \left(\boldsymbol{\theta}^{(h+1/2)}, \boldsymbol{c}^{(h+1/2)}\right)$, where $\boldsymbol{\theta}^{(h+1/2)} = \left(\theta_1^{(h)}, \theta_2^{(h)}\right)$ and

$$c_q^{(h+1/2)} = \frac{A^{(h)}\left(\boldsymbol{y}, \boldsymbol{z}\right)}{B^{(h)}\left(\boldsymbol{y}, \boldsymbol{z}\right)}, \ q = 1, \ldots, M-1;$$

where

$$A^{(h)}\left(\boldsymbol{y}, \boldsymbol{z}\right) = \sum_{i=1}^{m} I_{[\tau_{q-1}, \tau_q)}(y_{i:m:n}) + \sum_{j=1}^{d_1^*} I_{[\tau_{q-1}, \tau_q)}(z_{j:d_1^*:r_1^*})$$

and

$$B^{(h)}\left(\boldsymbol{y}, \boldsymbol{z}\right) = \theta_1^{(h)} \sum_{i=1}^{m} P(y_{i:m:n}, q) + \theta_2^{(h)} \sum_{j=1}^{d_1^*} P(z_{j:d_1^*:r_1^*}, q) + r_1^*(\theta_1^{(h)} - \theta_2^{(h)}) P(T_1, q)$$
$$+ (\tau_q - \tau_{q-1})\left\{\theta_1^{(h)}(n - m - r_1^*) + \theta_2^{(h)}(r_1^* - d_1^*)\right\}.$$

(ii) For $s = 2$, the solution to the constrained maximization problem can be obtained as $\boldsymbol{\phi}^{(h+1)} = \left(\boldsymbol{\theta}^{(h+1)}, \boldsymbol{c}^{(h+1)}\right)$, where $\boldsymbol{c}^{(h+1)} = \boldsymbol{c}^{(h+1/2)}$. The solution $\boldsymbol{\theta}^{(h+1)}$ can be obtained as

$$\theta_1^{(h+1)} = \frac{n - r_1^*}{\sum_{i=1}^{m} H_0^{(h+1)}(y_{i:m:n}) + r_1^* H_0^{(h+1)}(T_1) + (n - m - r_1^*)\left\{H_0^{(h+1)}(T_2) + \frac{1}{\theta_1^{(h)}}\right\}}$$

and

$$\theta_2^{(h+1)} = \frac{r_1^*}{\sum_{j=1}^{d_1^*} H_0^{(h+1)}(z_{j:d_1^*:r_1^*}) - r_1^* H_0^{(h+1)}(T_1) + (r_1^* - d_1^*)\left\{H_0^{(h+1)}(T_2) + \frac{1}{\theta_2^{(h)}}\right\}}.$$

This completes the $(h+1)$-th iteration of the algorithm. We propose the obvious initialization of the parameters as $c_1 = \cdots = c_{\mathcal{M}-1} = \theta_1 = \theta_2 = 1$.

Remark 3.1. *Note that the existence and uniqueness of the solutions in the conditional maximization steps of the proposed algorithm are based on the fulfillment of the following condition:*

(i) $\sum_{i=1}^{m} I_{[\tau_{q-1}, \tau_q)}(y_{i:m:n}) + \sum_{j=1}^{d_1^*} I_{[\tau_{q-1}, \tau_q)}(z_{j:d_1^*:r_1^*}) > 0$, *i.e., each of the sub-intervals* $[\tau_{q-1}, \tau_q)$ *contains at least one of the observations irrespective of stress-levels,*

(ii) $0 < r_1^* < n$.

Remark 3.2. *The convergence behavior of the proposed ECM algorithm can be established by applying the theoretical results from Meng and Rubin (1993). According to their framework, convergence is guaranteed under two key conditions: (i) the existence of a unique solution at each* *conditional maximization (CM) step, and (ii) the so-called space-filling condition on the collection of constraint sets* $\{g_s\}_{s=1}^S$. *The second condition ensures that, from any interior point of the parameter space* Θ, *the algorithm can explore all possible directions during optimization and is not restricted by the geometry of the constraint sets.*

This condition can be verified when each constraint function g_s *is differentiable and the matrix of partial derivatives,* $\nabla g_s(\boldsymbol{\phi})$, *formed by the gradients of its components, is of full column rank for all* $\boldsymbol{\phi}$ *in the interior* Θ_0 *of* Θ. *Under these assumptions, the constraints are said to be space-filling at a point* $\boldsymbol{\phi}$ *if the intersection of the column spaces* $J_s(\boldsymbol{\phi}) = \{\nabla g_s(\boldsymbol{\phi})\lambda : \lambda \in \mathbb{R}^{d_s}\}$ *across all* s *yields only the zero vector; i.e.,* $J(\boldsymbol{\phi}) = \bigcap_{s=1}^S J_s(\boldsymbol{\phi}) = \{\mathbf{0}\}$.

In our setup, the gradient matrices $\nabla g_1(\boldsymbol{\phi})$ *and* $\nabla g_2(\boldsymbol{\phi})$ *are given by:*

$$\nabla g_1(\boldsymbol{\phi}) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix}, \quad \nabla g_2(\boldsymbol{\phi}) = \begin{pmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

It is straightforward to verify that both matrices are of full column rank. Moreover, their column spaces are orthogonal complements, implying $J_1(\boldsymbol{\phi}) \cap J_2(\boldsymbol{\phi}) = \{\mathbf{0}\}$. *Hence, the space-filling condition is satisfied.*

Additionally, uniqueness of the solution at each CM step is ensured under the condition that

$$\sum_{i=1}^{m} I_{[\tau_{q-1},\tau_q)}(y_{i:m:n}) + \sum_{j=1}^{d_1^*} I_{[\tau_{q-1},\tau_q)}(z_{j:d_1^*:r_1^*}) > 0, \quad \textit{for all } q = 1, \ldots, M,$$

provided $r_1^* > 0$. *Therefore, by invoking Theorem 3 of Meng and Rubin (1993), we conclude that the proposed ECM algorithm converges to a stationary point of the observed likelihood function* $\mathcal{L}(\mathbf{c}, \boldsymbol{\theta} \mid \mathbf{y}, \mathbf{z})$.

Remark 3.3. *Note that for the exponential model* ($\mathcal{M} = 1$), *we have* $f_0(t) = S_0(t) = e^{-t}$ *and* $H_0(t) = t$ *for all* $t \geq 0$. *Hence for observed value of* $R_1^* = r_1^*$, *the log-likelihood function based on the observations* $\mathbf{y} = (y_{1:m:n}, \ldots, y_{m:m:n})$ *and* $\mathbf{z} = \left(z_{1:d_1^*:r_1}, \ldots, z_{d_1^*:d_1^*:r_1^*}\right)$, *can be written as*

$$\log \mathcal{L}(\mathbf{c}, \theta|\mathbf{y}, \mathbf{z}) \propto m \log \theta_1 - \theta_1 \sum_{i=1}^{m} y_{i:m:n} - \theta_1(n - m - r_1^*)T_2 + d_1^* \log \theta_2 - r_1^*(\theta_1 - \theta_2)T_1$$
$$- \theta_2 \sum_{j=1}^{d_1^*} z_{j:d_1^*:r_1^*} - \theta_2(r_1^* - d_1^*)T_2.$$

Consequently, the MlEs of the model parameters can be obtained explicitly as

$$\hat{\theta}_1 = \frac{m}{\sum_{i=1}^m y_{i:m:n} + (n - m - r_1^*)T_2 + r_1^* T_1}$$

and

$$\hat{\theta}_2 = \frac{d_1^*}{\sum_{j=1}^{d_1^*} z_{j:d_1^*:r_1^*} + (r_1^* - d_1^*)T_2 - r_1^* T_1}.$$

4. Simulation study

In this section, we carry out a simulation study in order to assess the performance of the proposed estimation procedure. Several researchers have observed that M = 2 and M = 3 work quite well in practice, see, for example, Balakrishnan et al. (2016), Murphy (1996), Goodman et al. (2011), Samanta and Kundu (2023) and Samanta and Kundu (2025). Hence, in this case we restrict ourselves to M = 2 and M = 3 only. For certain choices of the parameters $\boldsymbol{c}$, $\boldsymbol{\theta}$ and cut points, we generate SLT order statistics corresponding to different values of n, stress-change time T_1, test termination time T_2 and different censoring schemes (proportion of withdrawals π). In each case, we compute the MLEs of the parameters along with associated mean squared errors (MSEs) based on 1000 replications. These results, corresponding to $\mathcal{M} = 2$ and $\mathcal{M} = 3$, are presented in Tables 1 and 2, respectively, where we report the Average Estimates (AEs) and associated MSEs.

The SLT ordered samples are generated using the following algorithm:

Step 1: Generate n random samples $U_1 \ldots, U_n$ from the standard uniform distribution. Then find the order statistics $U_{1:n} \leq \cdot \leq U_{n:n}$ based on the generated random samples.

Step 2: For $1 \leq j \leq n$, set $V^{(1)}_{j:n} = F_1^{-1}(U_{j:n})$ and $V^{(2)}_{j:n} = F_2^{-1}(U_{j:n})$, where $F_1(t) = e^{-\theta_1 H_0(t)}$ and $F_2(t) = e^{-\theta_2\{H_0(T_1+t)-H_0(T_1)\}}$.

Step 3: Define $\mathcal{P} = \left\{\alpha \in \{1, \ldots, n\} : 0 < V^{(1)}_{\alpha:n} < T_1\right\}$, $\mathcal{P}^c = \{1, \ldots, n\}\backslash\mathcal{P}$, $D_1 = |\mathcal{P}|$ and set $R_1^* = \lfloor \pi|\mathcal{P}^c|\rfloor$. Here $\lfloor\cdot\rfloor$ denotes the floor function and $|\cdot|$ denotes cardinality.

Step 4: Choose randomly a without-replacement sample $\mathcal{R} \subseteq \mathcal{P}^c$ with $|\mathcal{R}| = R_1^*$, set $\mathcal{Q} = \left\{\alpha \in \mathcal{P}^c\backslash\mathcal{R} : V^{(1)}_{\alpha:n} < T_2\right\}$, $\mathcal{S} = \left\{\alpha \in \mathcal{R} : V^{(2)}_{\alpha:n} + T_1 < T_2\right\}$ and define $D_2 = |\mathcal{Q}|$, $D_1^* = |S|$, $M = D_1 + D_2$.

Step 5: Set

$$(Y_{h:M:n} : h = 1, \ldots, D_1) = \text{sort}\left(V^{(1)}_{h:n} : h \in \mathcal{P}\right),$$

$$(Y_{D_1+i:M:n} : h = 1, \ldots, D_2) = \text{sort}\left(V^{(1)}_{i:n} : i \in \mathcal{Q}\right)$$

and

$$\left(Z_{j:D_1^*:R_1^*} : j = 1, \ldots, D_1^*\right) = \text{sort}\left(V^{(2)}_{j:n} + T_1 : j \in \mathcal{S}\right).$$

For each generated sample, we also evaluate 90% bootstrap confidence interval (see Efron, 1982). The parametric percentile bootstrap confidence intervals for $\mathcal{M} = 2$ are obtained using the following steps:

Step 1: For a generated SLT ordered sample, compute the maximum likelihood estimates $\hat{c}_1$, $\hat{\theta}_1$ and $\hat{\theta}_2$.

Step 2: Generate bootstrap SLT ordered samples $\boldsymbol{y}^* = (y^*_{1:m:n}, \ldots, y^*_{m:m:n})$ and $\boldsymbol{z}^* = \left(z^*_{1:d_1^*:r_1^*}, \ldots, z^*_{d_1^*:d_1^*:r_1^*}\right)$ corresponding to the parameters $\hat{c}_1$, $\hat{\theta}_1$ and $\hat{\theta}_2$ and the censoring scheme

π.

Step 3: Compute the MLEs $\hat{c}_1^*$, $\hat{\theta}_1^*$ and $\hat{\theta}_2^*$ based on the bootstrap sample.

Step 4: Repeat Steps 2 and 3 B times to obtain $\{\hat{c}_{11}^*, \ldots, \hat{c}_{1B}^*\}$, $\{\hat{\theta}_{11}^*, \ldots, \hat{\theta}_{1B}^*\}$ and $\{\hat{\theta}_{21}^*, \ldots, \hat{\theta}_{2B}^*\}$.

Step 5: Construct a $100(1-\alpha)\%$ confidence interval for c_1, θ_1 and θ_2 as $\left(\hat{c}_{1k}^*, \hat{c}_{1k'}^*\right)$, $\left(\hat{\theta}_{1k}^*, \hat{\theta}_{1k'}^*\right)$ and $\left(\hat{\theta}_{2k}^*, \hat{\theta}_{2k'}^*\right)$ respectively, where $k = \lfloor \frac{\alpha}{2} B \rfloor$, $k' = \lfloor 1 - \frac{\alpha}{2} B \rfloor$.

Similar procedure is carried out for obtaining confidence intervals corresponding to $\mathcal{M} = 3$. We evaluate average length (AL) of the bootstrap confidence intervals and the associated coverage probability (CP) based on 500 iterations with 500 bootstrap samples in each iteration. These results are reported in Tables 3 and 4 corresponding to $\mathcal{M} = 2$ and $\mathcal{M} = 3$, respectively.

From Tables 1 and 2, it can be observed that the estimates are close to the true values in all scenarios, which demonstrates the good finite-sample performance of the proposed ECM-based estimation procedure. The MSEs decrease as the sample size n increases, confirming the consistency of the estimators. Moreover, for fixed n, increasing the proportion of censored items re-tested at higher stress levels (π) reduces the MSEs further for θ_2, highlighting that information recovery is enhanced by incorporating censored observations in the SLT framework. Comparing $M = 2$ and $M = 3$, we note that although both models perform satisfactorily, the estimates for $M = 3$ tend to have slightly higher variability when n is small, which is expected due to the higher dimensionality of the parameter space. However, this effect diminishes with larger samples, indicating the stability of the methodology even under a more flexible hazard function.

The results presented in Tables 3 and 4 show that the lengths of the confidence intervals decrease as n increases, reflecting improved precision of the estimators with larger sample sizes. For a fixed n, higher censoring proportions (π) also lead to shorter confidence intervals for θ_2, which again indicates that including censored items at higher stress levels improves efficiency. The coverage probabilities remain close to the nominal 90% across all cases, thereby indicating the robustness of the bootstrap procedure in the SLT setup. We also observe that the intervals for $M = 3$ are slightly wider than those for $M = 2$, which is consistent with the fact that more parameters need to be estimated for $M = 3$. In general, these findings confirm that the proposed estimation procedure, coupled with bootstrap inference, provides accurate and reliable results in a variety of practical experimental settings.

5. Determination of cut points and data analysis

Throughout the article, we have assumed that both the number $\mathcal{M}$ and the positions of the cut points (τ_j's) in the hazard function are known. However, in most practical situations, it is more realistic to assume that the positions of the cut points are unknown. One common approach to handling unknown cut points is to choose equidistant cut points, though this method may not always be practical or accurate. In order to address this issue, we propose a methodology that is both easy to implement in practice and flexible enough to generalize the case of equidistant cut points. We are mainly concerned about the shape of the hazard rate function in $[0, \tau]$, where $\tau = \max(y_{m:m:n}, z_{d_1^*:d_1^*:r_1^*})$. Here we model the positions of the cut points as the arrival times of events in $[0, \tau]$, from a non-homogeneous Poisson process $\{N(t) : t \geq 0\}$, with intensity function $u(t) = \alpha t^{\alpha-1}$, $\alpha > 0$. Since the arrival times of the events from $N(t)$ are not observed, we treat them

Table 1: AE and MSE corresponding to $\mathcal{M} = 2$

n, T_1, T_2	π	Estimates					
		c_1		θ_1		θ_2	
		AE	MSE	AE	MSE	AE	MSE
$\tau_1 = 1, c_1 = 2, \theta_1 = 0.2, \theta_2 = 0.3$							
80, 1.5, 7	$\pi = 0.6$	2.0841	0.4537	0.2039	0.0022	0.3065	0.0044
	$\pi = 0.7$	2.1130	0.6039	0.2056	0.0032	0.3071	0.0036
	$\pi = 0.8$	2.1747	0.7444	0.2054	0.0039	0.3065	0.0032
100, 1.5, 7	$\pi = 0.6$	2.0556	0.3584	0.2041	0.0018	0.3072	0.0036
	$\pi = 0.7$	2.0813	0.4271	0.2053	0.0024	0.3066	0.0029
	$\pi = 0.8$	2.1409	0.5945	0.2037	0.0031	0.3058	0.0025
$\tau_1 = 1.5, c_1 = 0.5, \theta_1 = 1.5, \theta_2 = 2.5$							
80, 0.5, 2	$\pi = 0.6$	0.5359	0.0405	1.5699	0.2702	2.6132	0.8211
	$\pi = 0.7$	0.5265	0.0367	1.5778	0.2883	2.6577	0.7291
	$\pi = 0.8$	0.5287	0.0487	1.6202	0.3904	2.6398	0.7130
100, 0.5, 2	$\pi = 0.6$	0.5260	0.0321	1.5582	0.2207	2.6313	0.6288
	$\pi = 0.7$	0.5311	0.0336	1.5577	0.2627	2.5787	0.5555
	$\pi = 0.8$	0.5204	0.0275	1.5790	0.2669	2.6343	0.6503

Table 2: AE and MSE corresponding to $\mathcal{M} = 3$

n, T_1, T_2	π	Estimates							
		c_1		c_2		θ_1		θ_2	
		AE	MSE	AE	MSE	AE	MSE	AE	MSE
$\tau_1 = 1.0, \tau_2 = 1.5, c_1 = 2.0, c_2 = 1.5, \theta_1 = 0.2, \theta_2 = 0.3$									
100, 1.5, 9	$\pi = 0.6$	2.0901	0.4410	1.5586	0.4425	0.2047	0.0027	0.3094	0.0035
	$\pi = 0.7$	2.0716	0.5170	1.5262	0.4769	0.2098	0.0037	0.3090	0.0029
	$\pi = 0.8$	2.1082	0.8032	1.5871	0.6870	0.2157	0.0062	0.3057	0.0024
150, 1.5, 9	$\pi = 0.6$	2.0472	0.2905	1.5340	0.2978	0.2047	0.0016	0.3044	0.0021
	$\pi = 0.7$	2.0580	0.3520	1.5412	0.3148	0.2050	0.0022	0.3044	0.0019
	$\pi = 0.8$	2.1074	0.5004	1.6043	0.4539	0.2057	0.0033	0.3042	0.0015
$\tau_1 = 1.0, \tau_2 = 1.5, c_1 = 0.2, c_2 = 0.4, \theta_1 = 1.0, \theta_2 = 1.5$									
100, 1.5, 9	$\pi = 0.6$	0.2009	0.0042	0.3996	0.0178	1.0463	0.0479	1.5427	0.0646
	$\pi = 0.7$	0.2029	0.0046	0.4043	0.0193	1.0527	0.0642	1.5358	0.0572
	$\pi = 0.8$	0.2031	0.0054	0.4078	0.0259	1.0701	0.0993	1.5325	0.0468
150, 1.5, 9	$\pi = 0.6$	0.2007	0.0025	0.4031	0.0122	1.0297	0.0302	1.5054	0.0406
	$\pi = 0.7$	0.2032	0.0030	0.3968	0.0131	1.0342	0.0404	1.5307	0.0349
	$\pi = 0.8$	0.1986	0.0036	0.4000	0.0162	1.0556	0.0651	1.5139	0.0294

as missing values. Hence for any given $\mathcal{M}$, the likelihood of the observations becomes a function of the parameters $\boldsymbol{c}$, $\boldsymbol{\theta}$ and α, where the arrival times of the events from $N(t)$, being missing, are replaced with their expected values. Now for a given $\mathcal{M}$, we compute the maximum likelihood estimators of the unknown parameters, and then we finally choose $\mathcal{M}$ based on some information theoretic criteria, like Akike Information Criterion (AIC) or Bayes Information Criterion (BIC).

Suppose $\{N(t) : t \geq 0\}$ is a non-homogeneous Poisson process with intensity function $u(t) =$

Table 3: Average length and coverage probability corresponding to $\mathcal{M} = 2$

n, T_1, T_2	π	Estimates					
		c_1		θ_1		θ_2	
		AL	CP	AL	CP	AL	CP
$\tau_1 = 1.0, c_1 = 2.0, \theta_1 = 0.2, \theta_2 = 0.3$							
80, 1.5, 7	$\pi = 0.6$	2.3107	0.890	0.1573	0.890	0.2177	0.860
	$\pi = 0.7$	2.4757	0.898	0.1754	0.904	0.1982	0.884
	$\pi = 0.8$	2.8737	0.906	0.1976	0.896	0.1867	0.876
100, 1.5, 7	$\pi = 0.6$	1.9528	0.894	0.1408	0.882	0.1928	0.876
	$\pi = 0.7$	2.1153	0.894	0.1569	0.910	0.1753	0.898
	$\pi = 0.8$	2.4801	0.892	0.1773	0.902	0.1635	0.922
$\tau_1 = 1.5, c_1 = 0.5, \theta_1 = 1.5, \theta_2 = 2.5$							
80, 0.5, 2	$\pi = 0.6$	0.6745	0.880	1.7451	0.890	2.9682	0.880
	$\pi = 0.7$	0.6782	0.898	1.8488	0.882	2.9186	0.894
	$\pi = 0.8$	0.6597	0.912	1.9516	0.916	2.8318	0.892
100, 0.5, 2	$\pi = 0.6$	0.5661	0.886	1.5252	0.908	2.5777	0.910
	$\pi = 0.7$	0.5911	0.874	1.6258	0.870	2.5157	0.870
	$\pi = 0.8$	0.5663	0.890	1.7377	0.894	2.482	0.894

Table 4: Average length and coverage probability corresponding to $\mathcal{M} = 3$

n, T_1, T_2	π	Estimates							
		c_1		c_2		θ_1		θ_2	
		AL	CP	AL	CP	AL	CP	AL	CP
$\tau_1 = 1.0, \tau_2 = 1.5, c_1 = 2, c_2 = 1.5, \theta_1 = 0.2, \theta_2 = 0.3$									
100, 1.5, 9	$\pi = 0.6$	2.0931	0.902	2.1235	0.882	0.1647	0.896	0.1889	0.892
	$\pi = 0.7$	2.3791	0.894	2.2817	0.858	0.1952	0.886	0.1738	0.878
	$\pi = 0.8$	3.1842	0.888	2.8710	0.874	0.2524	0.876	0.1622	0.878
150, 1.5, 9	$\pi = 0.6$	1.6101	0.916	1.6728	0.866	0.1307	0.906	0.1486	0.880
	$\pi = 0.7$	1.8547	0.866	1.8349	0.876	0.1523	0.900	0.1379	0.892
	$\pi = 0.8$	2.2333	0.886	2.0707	0.874	0.1930	0.902	0.1258	0.904
$\tau_1 = 1.0, \tau_2 = 1.5, c_1 = 0.2, c_2 = 0.4, \theta_1 = 1.0, \theta_2 = 1.5$									
100, 1.5, 9	$\pi = 0.6$	0.1952	0.876	0.4127	0.890	0.6851	0.888	0.8180	0.864
	$\pi = 0.7$	0.2138	0.890	0.4508	0.854	0.7932	0.900	0.7539	0.896
	$\pi = 0.8$	0.2375	0.888	0.4986	0.864	1.0562	0.860	0.7066	0.888
150, 1.5, 9	$\pi = 0.6$	0.1624	0.888	0.3404	0.896	0.5410	0.884	0.6541	0.896
	$\pi = 0.7$	0.1712	0.890	0.3703	0.860	0.6391	0.880	0.5982	0.886
	$\pi = 0.8$	0.1894	0.900	0.3987	0.910	0.8079	0.882	0.5585	0.876

$\alpha t^{\alpha-1}$, $\alpha > 0$. Let $T_1, T_2, \ldots$ be the ordered arrival times of events from $N(t)$. Note that that the joint pdf of $T_1, T_2, \ldots, T_n$, given that $N(\tau) = n$, can be written as

$$f_{T_1,\ldots T_n|N(\tau)=n}(t_1, \ldots, t_n) = \begin{cases} \frac{n!\alpha^n}{\tau^{n\alpha}} t_1^{\alpha-1} \ldots t_n^{\alpha-1}, & 0 < t_1 < \cdots < t_n < \tau \\ 0 & \text{otherwise} \end{cases}.$$

Hence, for $j = 1, \dots, n$, we have

$$E\left(T_j | N(\tau) = n\right) = \frac{n!\alpha^n}{\tau^{n\alpha}} \int_0^\tau \int_0^{t_n} \cdots \int_0^{t_3} \int_0^{t_2} t_1^{\alpha-1} \dots t_{j-1}^{\alpha-1} t_j^{\alpha} t_{j+1}^{\alpha-1} \dots t_n^{\alpha-1} \, dt_1 dt_2 \dots dt_{n-1} dt_n$$
$$= \frac{n!\alpha^n}{\tau^{n\alpha}} \times \frac{\tau^{n\alpha+1}}{\alpha(2\alpha)\dots((j-1)\alpha)\,(j\alpha+1)\,((j+1)\alpha+1)\dots(n\alpha+1)}$$
$$= \frac{\tau}{\left(1+\frac{1}{j\alpha}\right)\dots\left(1+\frac{1}{n\alpha}\right)}.$$

Note that for homogeneous Poisson process, i.e., for $\alpha = 1$, $E\left(T_j | N(\tau) = n\right) = j\tau/(n+1)$. So in this case, the cut points are equidistant to each other.

The practical implementation of the above method is as follows. For any fixed $\mathcal{M}$, we compute the MLEs of α and $c_1, \dots, c_{\mathcal{M}-1}$ in two steps. First, for fixed α, we compute the MLEs of $\boldsymbol{c}$ and $\boldsymbol{\theta}$ using the methodology discussed earlier. Then using the profile likelihood method we compute the MLE of α. We vary the values of α and obtain different estimates of the parameters $\boldsymbol{c}$, $\boldsymbol{\theta}$ along with associated values of likelihood. Hence, the value of α, and consequently, the positions of the cut points, is chosen, which maximizes the likelihood as a function of α. Finally, we choose $\mathcal{M}$ based on AIC or BIC.

We now illustrate the proposed methodology using two simulated data sets corresponding to $\mathcal{M} = 2$ and $\mathcal{M} = 3$. We choose $\tau_1 = 15$, $c_1 = 2$ for $\mathcal{M} = 2$ and $\tau_1 = 5$, $\tau_2 = 15$, $c_1 = 2$, $c_2 = 3$ for $\mathcal{M} = 3$. For generating both data sets, we choose $T_1 = 10$, $T_2 = 25$, $\theta_1 = 0.02$, $\theta_2 = 0.04$ and the censoring proportion $\pi = 0.6$. The generated data sets are presented in Tables 5 and 6 corresponding to $\mathcal{M} = 2$ and $\mathcal{M} = 3$, respectively. In each data sets we fit the proposed models corresponding to $\mathcal{M} = 2$ and $\mathcal{M} = 3$. We report the maximum likelihood estimates of the parameters along with associated estimates of the cut points corresponding to that value of α which maximizes the observed likelihood function. Further, we obtain 90% bootstrap confidence intervals based on 1000 bootstrap samples. These results are reported in Tables 7 and 8 corresponding to $\mathcal{M} = 2$ and $\mathcal{M} = 3$, respectively. From both the tables, it is evident that the estimated values of the parameters are close to the true values. Further, the estimated cut points are also close to true values. It is worth noting that the value of the shape parameter α plays a crucial role in determining the distribution of cut-points. When $0 < \alpha < 1$, the intensity function $u(t)$ decreases with t, which results in cut-points being more densely concentrated near the origin. On the other hand, when $\alpha > 1$, the intensity function increases with t, so the cut-points are shifted toward later times. These distinctions are also evident in our data analysis, where the estimated values of α directly influence the placement of the cut-points and, consequently, the fitted shape of the hazard function. We also calculate AIC and BIC corresponding to different choices of $\mathcal{M}$ for the data sets considered. These results are reported in Table 9. It is observed that, in the case of both the data sets, the choice of $\mathcal{M} = 2$ is recommended based on the information criterion.

6. Extension to general *k*-step stage life testing

We now extend the proposed estimation methodology to the general k-step stage life testing with k stress levels denoted by stages $s_1, \dots, s_k$. Suppose n identical items are put on a life testing experiment at the stress level s_1 with pre-fixed stress-changing times $T_1, \dots, T_{k-1}$ (assume $T_0 = 0$). Let D_1 be the number of failures upto time T_1. Then, at time T_1, R_1^* the items are randomly removed and put at stress level s_2, where $R_1^* = \rho(D_1)$. Next at time T_2, R_2^* items are randomly withdrawn out of the surviving units at stress level s_1 these R_2^* items are put on stress

Table 5: Generated Data Set 1 ($M = 2$)

Stress level	Ordered observations								
s_1	00.9280	01.6351	02.3707	02.5222	03.7484	04.0334	05.2248	05.2500	05.8444
	05.9034	06.3881	06.7442	06.9310	08.7877	08.9542	09.2010	09.6936	09.7319
	10.2071	13.8650	15.7984	23.8000	24.2045	24.8741			
s_2	10.0178	10.2840	10.7906	11.4226	11.4987	11.7359	11.9272	12.0235	12.1481
	13.1969	13.4134	13.5388	13.6462	14.0821	14.9011	16.4377	18.5037	18.7145
	18.8673	19.6332	20.4992	22.7606	24.3096				

Table 6: Generated Data Set 2 ($M = 3$)

Stress level	Ordered observations								
s_1	00.0052	00.3464	00.5838	01.1657	01.1752	01.3162	01.3925	02.0568	03.5767
	03.7253	04.9738	05.3239	05.3743	06.1089	06.7935	08.9068	08.9121	08.9955
	09.1785	09.1922	09.3192	09.3364	09.8436	10.4567	10.8743	11.8596	12.7746
	14.1686	16.6238	18.1674	18.5138	20.4107				
s_2	10.1040	10.5691	11.1448	11.2879	11.3346	11.4742	11.8143	11.8680	12.0959
	12.2185	12.4028	12.8800	13.6625	14.7432	17.5247	18.1294	19.5607	20.3645
	21.7115	22.5395	23.3493						

Table 7: MLE and 90% CI of the parameters for $M = 2$

Parameters					
c_1		θ_1		θ_2	
MLE	90% CI	MLE	90% CI	MLE	90% CI
Data Set 1: $\hat{\alpha} = 1.32, \hat{\tau}_1 = 14.1525$					
2.0530	(1.2061, 4.0082)	0.0124	(0.0061, 0.0216)	0.0504	(0.0289, 0.0791)
Data Set 2: $\hat{\alpha} = 1.24, \hat{\tau}_1 = 12.9255$					
2.3667	(1.3551, 4.5217)	0.0163	(0.0086, 0.0286)	0.0512	(0.0294, 0.0811)

Table 8: MLE and 90% CI of the parameters for $M = 3$

Parameters							
c_1		c_2		θ_1		θ_2	
MLE	90% CI	MLE	90% CI	MLE	90% CI	MLE	90% CI
Data Set 1: $\hat{\alpha} = 1.33, \hat{\tau}_1 = 10.3191, \hat{\tau}_2 = 18.0779$							
1.9135	(0.8950, 4.9084)	0.9772	(0.5018, 1.9670)	0.0139	(0.0054, 0.0283)	0.0696	(0.0376, 0.1142)
Data Set 2: $\hat{\alpha} = 0.63, \hat{\tau}_1 = 5.0314, \hat{\tau}_2 = 13.0178$							
1.6189	(0.6967, 3.5090)	2.4656	(1.4175, 4.5140)	0.0185	(0.0096, 0.0315)	0.0499	(0.0296, 0.0785)

level s_3. Here $R_2^* = \rho(D_1 + D_2 + R_1^*)$ and D_2 is the number of failures observed in the time interval $(T_1, T_2]$. This process is continued in the same fashion until the stress-changing time T_{k-1} and the experiment is terminated at the pre-fixed time point T_k. Let $m = d_1 + \cdots + d_k$ be the total number

Table 9: AIC and BIC corresponding to different models

Data Set 1				Data Set 2			
$\mathcal{M}=2$		$\mathcal{M}=3$		$\mathcal{M}=2$		$\mathcal{M}=3$	
AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
-144.7623	-137.6163	-142.184	-132.6559	-193.7256	-186.5795	-164.2673	-154.7392

of failures observed in stress level s_1 and d_j^* be the total number of failures observed in stress level $s_{j+1},\ j=1,\ldots,k-1$. Suppose $\boldsymbol{y}=(y_{1:m:n},\ldots,y_{m:m:n})$ is the vector of ordered observations in stage s_1 and $\boldsymbol{z_j}=\left(z_{1:d_j^*:r_j^*},\ldots,z_{d_j^*:d_j^*:r_j^*}\right)$ is the vector of ordered observations at stage s_{j+1}, $j=1,\ldots,k-1$. We assume that the failure rate function $r(t)$ of an item that undergoes stress level S_j is given by

$$r(t)=\theta_j h_0(t),\ j=1,\ldots,k$$

where $h_0(t)$ is the piecewise constant hazard function as defined earlier.

Note that for $t\geq T_{j-1}$, the survival function $S_j(t)$ at stage s_j can be written as

$$\begin{aligned}S_j(t)&=\exp\left\{-\sum_{i=1}^{j-1}\int_{T_{i-1}}^{T_i}h(u)du-\int_{T_{j-1}}^{t}h(u)du\right\}\\&=\exp\left\{-\left(\theta_1-\theta_j\right)H_0(T_{j_1})-\theta_jH_0(t)\right\}.\end{aligned}$$

Hence the log-likelihood function, based on the observations $\boldsymbol{y},\boldsymbol{z}$, can be written as

$$\begin{aligned}\log\mathcal{L}(\mathbf{c},\theta|\mathbf{y},\mathbf{z})\propto\ &\log C+m\log\theta_1+\sum_{i=1}^{m}\log h_0\left(y_{i:m:n}\right)-\theta_1\sum_{i=1}^{m}H_0\left(y_{i:m:n}\right)+\sum_{i=1}^{k-1}d_i^*\log\theta_{i+1}\\&-\theta_1\left(n-m-\sum_{i=1}^{k-1}r_i^*\right)H_0(T_k)+\sum_{i=1}^{k-1}\sum_{j=1}^{d_i^*}\log h_0\left(z_{j:d_i^*r_i^*}\right)-\sum_{i=1}^{k-1}\theta_{i+1}\sum_{j=1}^{d_i^*}H_0\left(z_{j:d_i^*:r_i^*}\right)\\&-\sum_{i=1}^{k-1}r_i^*\left(\theta_1-\theta_{i+1}\right)H_0(T_i)-\sum_{i=1}^{k-1}\theta_{i+1}(r_i^*-d_i^*)H_0(T_k),\end{aligned}$$

where $C=\prod_{j=1}^{k}\binom{n-\sum_{i=1}^{j-1}d_i-\sum_{i=1}^{j-1}r_i^*}{d_j}d_j!$.

Now suppose $\boldsymbol{M}_1=\left(M_{1,1},\ldots,M_{1,n-m-\sum_{i=1}^{k-1}r_i^*}\right)$ is the vector of missing observations in stage s_1 and $\boldsymbol{M}_i=\left(M_{i,1},\ldots,M_{i,r_i^*-d_i^*}\right)$ is the vector of missing observations in stage s_i, $i=2,\ldots,k$. Arguing similarly as in Section 3, the expectation of the complete log-likelihood $\lambda(\boldsymbol{\theta},\boldsymbol{c}|\boldsymbol{X})$, given

$\boldsymbol{y}, \boldsymbol{z}$ and $\boldsymbol{\theta^{(h)}}, \boldsymbol{c^{(h)}}$ can be written as

$$
\begin{aligned}
Q\left(\boldsymbol{\theta}, \boldsymbol{c}|\boldsymbol{\theta}^{(h)}, \boldsymbol{c}^{(h)}\right) = & \left(n - \sum_{i=1}^{k-1} r_i^*\right) \log\theta_1 + \sum_{i=1}^{m} \log h_0(y_{i:m:n}) - \theta_1 \sum_{i=1}^{m} H_0(y_{i:m:n}) + \sum_{i=1}^{k-1} r_i^* \log\theta_{i+1} \\
& + \sum_{i=1}^{k-1}\sum_{j=1}^{d_i^*} \log h_0(z_{j:d_i^*:r_i^*}) - \sum_{i=1}^{k-1} \theta_{i+1} \sum_{j=1}^{d_i^*} H_0(z_{j:d_i^*:r_i^*}) - \sum_{i=1}^{k-1} r_i^* \left(\theta_1 - \theta_{i+1}\right) H_0(T_i) \\
& - \theta_1 \left(n - m - \sum_{i=1}^{k-1} r_i^*\right) \left\{H_0(T_k) + 1/\theta_1^{(h)}\right\} - \sum_{i=1}^{k-1} \theta_{i+1} \left(r_i^* - d_i^*\right) \left\{H_0(T_k) + 1/\theta_{i+1}^{(h)}\right\}.
\end{aligned}
$$

Let $\boldsymbol{\phi} = (\boldsymbol{\theta}, \boldsymbol{c})$ and suppose that we have the solutions $\boldsymbol{\phi^{(0)}}, \ldots \boldsymbol{\phi^{(h)}}$ from h iterations, where $\boldsymbol{\phi}^{(h)} = (\boldsymbol{\theta}^{(h)}, \boldsymbol{c}^{(h)})$. Now the iterative CM-steps at $(h+1)$-th iteration of the algorithm can be computed as follows:

(i) For $s = 1$, the solution of the constrained maximization problem can be obtained as $\boldsymbol{\phi}^{(h+1/2)} = \left(\boldsymbol{\theta}^{(h+1/2)}, \boldsymbol{c}^{(h+1/2)}\right)$, where $\boldsymbol{\theta}^{(h+1/2)} = \left(\theta_1^{(h)}, \theta_2^{(h)}\right)$ and

$$
c_q^{(h+1/2)} = \frac{A^{(h)}\left(\boldsymbol{y}, \boldsymbol{z}\right)}{B^{(h)}\left(\boldsymbol{y}, \boldsymbol{z}\right)}, \ q = 1, \ldots, \mathcal{M} - 1;
$$

where

$$
A^{(h)}\left(\boldsymbol{y}, \boldsymbol{z}\right) = \sum_{i=1}^{m} I_{[\tau_{q-1}, \tau_q)}(y_{i:m:n}) + \sum_{i=1}^{k-1}\sum_{j=1}^{d_i^*} I_{[\tau_{q-1}, \tau_q)}(z_{j:d_i^*:r_i^*})
$$

and

$$
\begin{aligned}
B^{(h)}\left(\boldsymbol{y}, \boldsymbol{z}\right) = & \theta_1^{(h)} \sum_{i=1}^{m} P(y_{i:m:n}, q) + \sum_{i=1}^{k-1} \theta_{i+1}^{(h)} \sum_{j=1}^{d_i^*} P(z_{j:d_i^*:r_i^*}, q) + \sum_{i=1}^{k-1} r_i^* \left(\theta_1^{(h)} - \theta_{i+1}^{(h)}\right) P(T_i, q) \\
& + (\tau_q - \tau_{q-1}) \left\{\theta_1^{(h)} \left(n - m - \sum_{i=1}^{k-1} r_i^*\right) + \sum_{i=1}^{k-1} \theta_{i+1}^{(h)} (r_i^* - d_i^*)\right\}.
\end{aligned}
$$

(ii) For $s = 2$, the solution to the constrained maximization problem can be obtained as $\boldsymbol{\phi}^{(h+1)} = \left(\boldsymbol{\theta}^{(h+1)}, \boldsymbol{c}^{(h+1)}\right)$, where $\boldsymbol{c}^{(h+1)} = \boldsymbol{c}^{(h+1/2)}$. The solution $\boldsymbol{\theta}^{(h+1)}$ can be obtained as

$$
\theta_1^{(h+1)} = \frac{n - \sum_{i=1}^{k-1} r_i^*}{\sum\limits_{i=1}^{m} H_0^{(h+1)}(y_{i:m:n}) + \sum\limits_{i=1}^{k-1} r_i^* H_0^{(h+1)}(T_i) + \left(n - m - \sum_{i=1}^{k-1} r_i^*\right) \left\{H_0^{(h+1)}(T_k) + \frac{1}{\theta_1^{(h)}}\right\}}
$$

and

$$
\theta_{i+1}^{(h+1)} = \frac{r_i^*}{\sum\limits_{j=1}^{d_i^*} H_0^{(h+1)}(z_{j:d_1^*:r_1^*}) - r_i^* H_0^{(h+1)}(T_i) + (r_i^* - d_i^*) \left\{H_0^{(h+1)}(T_K) + \frac{1}{\theta_{i+1}^{(h)}}\right\}}
$$

for $i = 1, \ldots, k-1$.

This completes the $(h+1)$-th iteration of the algorithm.

Remark 6.1. *Note that the existence and uniqueness of the solutions in the conditional maximization steps of the proposed algorithm are based on the fulfillment of the following conditions:*

(i) $\sum_{i=1}^{m} I_{[\tau_{q-1},\tau_q)}(y_{i:m:n}) + \sum_{i=1}^{k-1}\sum_{j=1}^{d_i^*} I_{[\tau_{q-1},\tau_q)}(z_{j:d_i^*:r_i^*}) > 0$, *i.e., each of the sub-intervals* $[\tau_{q-1},\tau_q)$ *contains at least one of the observations irrespective of stress-levels,*

(ii) $r_i^* > 0$ *for* $i = 1,\ldots,k-1$ *and* $n - \sum_{i=1}^{k-1} r_i^* > 0$.

The convergence of the proposed algorithm follows as in the case of $k = 2$.

7. Optimal stage life testing plan

In this section we discuss about the construction of the optimal stage life testing plan. In this case, the determination of the optimal plan is reduced to find out the values of $T_1,\ldots,T_{k-1}$ and $R_1,\ldots,R_{k-1}$ which will produce the optimal value of a certain scientific criterion. In the literature, several criteria had been discussed in the context of optimal life testing plan. One of the most frequently used criterion is the determinant of Fisher information matrix and the optimal plan maximizes the determinant of Fisher information matrix or equivalently minimizes the determinant of its inverse. This criterion is known as the D-optimality criterion in the literature. In this work, based on the D-optimality criterion we propose to find the optimal plan. As it has been mentioned before that even in simple progressive Type-II censoring case, it is not immediate what would be the optimum censoring scheme. In this situation in presence of stress factors, it is even more difficult to predict what would be the optimum censoring scheme. Due to this reason, we need a proper computational algorithm to determine the optimum censoring scheme.

In this work, determination of the optimal life testing plan is a mixture of both continuous and discrete optimization problem. For fixed n and pre-fixed stress-change points $T_1,\ldots,T_{k-1}$, the solution space of $R_1,\ldots,R_{k-1}$ has cardinality $\binom{n-1}{k-2}$ which is large enough even for moderate values of n and k. On the other hand, the solution space of $T_1,\ldots,T_{k-1}$ is infinite. Here, in computation we pre-fix the values of $T_1,\ldots,T_{k-1}$, along with n and k, and the optimization problem is reduced to a discrete optimization problem where the solution space consists of non-negative integers $R_1,\ldots,R_m$ with the restriction that $\sum_{i=1}^{m} R_i \leq n$. The solution space being quite large, the exact optimization is a numerically challenging problem. As a way out, we apply a meta-heuristic algorithm like variable neighborhood search (VNS) algorithm. Hansen and Mladenović (1999) introduced the VNS algorithm where at each step, an improved local optimum is found in the neighborhood of the previous step's local optimum. The algorithm is terminated if no further improvement is found. The VNS algorithm works on three strategies - the neighborhoods are connected, these connected neighborhoods may produce different local optima, and the global optimum is one of the local optima from those connected neighborhoods obtained so far. Bhattacharya et al. (2016) proposed a meta-heuristic algorithm to find the optimal plan under a progressive censoring scheme based on the variable neighborhood search method. Recently, Samanta et al. (2022) proposed a similar algorithm to obtain otimal plan for order-restricted estimation of Weibull parameters in the stage life testing experiment. Here we also apply a similar method.

In this context, the $R_1, R_2, \ldots, R_{k-1}$ are such that $\sum_{i=1}^{k-1} R_i \leq n$. We define $R_k = n - \sum_{i=1}^{k-1} R_i$ so that $\sum_{i=1}^{k} R_i = n$. We next define $R = (R_1, R_2, \ldots, R_k)$, and for an initial guess $R^0 = (R_1^0, R_2^0, \ldots, R_k^0)$, the i-th neighborhood can be constructed as

$$N_i(R^0) = \left\{ R : \sum_{i=1}^{k} R_i = n, \|R - R^0\| < i+1 \text{ and } |R_j^0 - R_j| \leq i \quad \text{for } j = 1,\ldots,k \right\},$$

where $\|R - R^0\|$ is the Euclidean distance between R^0 and R.

The pseudo code of the VNS algorithm is given below:

Pseudo code of the VNS Algorithm:

1. **Initial guess:** Choose an initial guess R^0 and set $i_{\max} = \max_{1 \le j \le k} R_j^0$.
2. **Iterations:** For $i = 1$ to $i_{\max}$, $i \leftarrow i + 1$, continue the process below.
3. **Local search:** Based on R^0, construct the i-th neighborhood $N_i(R^0)$. Apply local search to find out the local optimal, say R^{opt_i}.
4. **Decision to move:**
   - If $R^0 = R^{\text{opt}_i}$, set $i = i + 1$ and continue the local search.
   - If $R^0 \neq R^{\text{opt}_i}$, set the initial guess as R^{opt_i} and repeat local search.

Let $\phi = (\boldsymbol{c}, \boldsymbol{\theta})$. Then the Fisher information matrix can be obtained as

$$I(\phi) = -E\left[\frac{\partial^2 \log L}{\partial \phi_q \partial \phi_r}\right],$$

where

$$\frac{\partial^2 \log L}{\partial \phi_q^2} = -\frac{1}{c_q^2}\left[\sum_{i=1}^{M} I_{[\tau_{q-1}, \tau_q)}(Y_{i:M:n}) + \sum_{i=1}^{k-1}\sum_{j=1}^{D_i^*} I_{[\tau_{q-1}, \tau_q)}(Z_{j:D_i^*:r_i^*})\right] \text{ for } q = 1, \ldots, \mathcal{M} - 1,$$

$$\frac{\partial^2 \log L}{\partial \phi_{\mathcal{M}}^2} = -\frac{M}{\theta_1^2}, \; \frac{\partial^2 \log L}{\partial \phi_q^2} = -\frac{D_{q-\mathcal{M}+1}}{\phi_{q-\mathcal{M}+1}^2} \text{ for } q = \mathcal{M}+1, \ldots, \mathcal{M}+k-1,$$

$$\frac{\partial^2 \log L}{\partial \phi_q \partial \phi_r} = \frac{\partial^2 \log L}{\partial \phi_r \partial \phi_q} = 0 \text{ for all } q, r = 1, \ldots, \mathcal{M}-1 \text{ and } q \neq r,$$

$$\frac{\partial^2 \log L}{\partial \phi_q \partial \phi_r} = \frac{\partial^2 \log L}{\partial \phi_r \partial \phi_q} = 0 \text{ for all } q, r = \mathcal{M}, \ldots, \mathcal{M}+k-1 \text{ and } q \neq r,$$

$$\frac{\partial^2 \log L}{\partial \phi_q \partial \phi_{\mathcal{M}}} = -\sum_{i=1}^{M} P\left(Y_{i:M:n}, q\right) - \left(n - M - \sum_{j=1}^{k-1} R_j^*\right)(\tau_q - \tau_{q-1}) - \sum_{i=1}^{k-1} R_i^* \, P\left(T_i, q\right)$$

for $q \le \mathcal{M} - 1$, and

$$\frac{\partial^2 \log L}{\partial \phi_q \partial \phi_r} = R_{r-\mathcal{M}}^* P\left(T_{r-\mathcal{M}}, q\right) - \sum_{i=1}^{D_{r-\mathcal{M}}} P\left(Z_{i:D_{r-\mathcal{M}}^*:R_{r-\mathcal{M}}^*}, q\right) - (\tau_q - \tau_{q-1})\left(R_{r-\mathcal{M}}^* - D_{r-\mathcal{M}}^*\right)$$

for $q \le \mathcal{M} - 1$, $r \ge \mathcal{M} + 1$. The D-optimization is performed based on the expected Fisher information matrices and these expectations are obtained using Monte Carlo simulation based on 500 replications.

Note that under the proposed model, the MLEs will exist if $m > 0$, $d_i^* > 0$ for all i and each of the intervals $[\tau_q, \tau_{q-1})$ contains at least one observed failure times (irrespective of stress-levels). This is possible if $R_i > 0$ for $i = 1, \ldots, k$. Hence for the local search of the VNS algorithm, we restrict ourselves to the R's for which $R_i > 0$ for $i = 1, \ldots, k$.

We evaluate optimal plans for $k = 2$ and $k = 3$. The results are obtained for $\mathcal{M} = 2$ and $\mathcal{M} = 3$ and certain choices of the parameters. We choose $n = 30$ in each case. The results are reported in Tables 10 and 11 corresponding to $k = 2$ and $k = 3$, respectively. From the numerical results, it is evident that even if we start from some different initial guesses, it always converges to the same optimum value. It seems that the VNS algorithm works well in this case. Further, interestingly enough, it is also observed that in optimal plans, R_k is always 1, i.e., almost all the items are required withdrawal for optimal plans. This is justified since the observations of lifetimes in stress levels other than stage s_1 contain information of failure time distributions in stage s_1 whereas observations on stage s_1 contain no information of failure time distributions in different stages.

Table 10: Optimal life-testing plans obtained using VNS algorithm for $k = 2$

Parameters	Stress-change times	Initial guess	Optimal plan	Optimal value
$\mathcal{M} = 2$				
$\tau_1 = 9, c_1 = 0.5,$ $\theta_1 = 0.01, \theta_2 = 0.02$	$T_1 = 10, T_2 = 20$	15, 15	29, 1	2.9612×10^{-10}
		05, 25	29, 1	2.9612×10^{-10}
$\tau_1 = 25, c_1 = 1.5,$ $\theta_1 = 0.002, \theta_2 = 0.004$	$T_1 = 100, T_2 = 150$	15, 15	29, 1	1.4409×10^{-12}
		05, 25	29, 1	1.4409×10^{-12}
$\mathcal{M} = 3$				
$\tau_1 = 7, \tau_2 = 10c_1 = 0.5,$ $c_2 = 0.8, \theta_1 = 0.01, \theta_2 = 0.02$	$T_1 = 10, T_2 = 20$	15, 15	29, 1	3.5134×10^{-10}
		05, 25	29, 1	3.5134×10^{-10}
$\tau_1 = 25, \tau_2 = 50c_1 = 1.8,$ $c_2 = 1.5, \theta_1 = 0.002, \theta_2 = 0.004$	$T_1 = 100, T_2 = 150$	15, 15	29, 1	2.0079×10^{-12}
		05, 25	29, 1	2.0079×10^{-12}

Table 11: Optimal life-testing plans obtained using VNS algorithm for $k = 3$

Parameters	Stress-change times	Initial guess	Optimal plan	Optimal value
$\mathcal{M} = 2$				
$\tau_1 = 9, c_1 = 0.5, \theta_1 = 0.01,$	$T_1 = 10, T_2 = 20,$	10, 10, 10	14, 15, 1	7.7632×10^{-14}
$\theta_2 = 0.02, \theta_3 = 0.04$	$T_3 = 30$	05, 05, 20	14, 15, 1	7.7632×10^{-14}
$\tau_1 = 25, c_1 = 1.5, \theta_1 = 0.002,$	$T_1 = 100, T_2 = 150,$	10, 10, 10	13, 16, 1	1.7448×10^{-17}
$\theta_2 = 0.004, \theta_3 = 0.006$	$T_3 = 200$	05, 05, 20	13, 16, 1	1.7448×10^{-17}
$\mathcal{M} = 3$				
$\tau_1 = 7, \tau_2 = 10, c_1 = 0.5, c_2 = 0.8,$	$T_1 = 10, T_2 = 20,$	10, 10, 10	17, 12, 2001	9.4915×10^{-14}
$\theta_1 = 0.01, \theta_2 = 0.02, \theta_3 = 0.04$	$T_3 = 30$	05, 05, 20	17, 12, 2001	9.4915×10^{-14}
$\tau_1 = 25, \tau_2 = 50c_1 = 1.8, c_2 = 1.5,$	$T_1 = 100, T_2 = 150,$	10, 10, 10	16, 13, 1	3.0336×10^{-17}
$\theta_1 = 0.002, \theta_2 = 0.004, \theta_3 = 0.006$	$T_3 = 200$	05, 05, 20	16, 13, 1	3.0336×10^{-17}

8. Concluding Remarks

In this article, we have developed an estimation methodology for a flexible tampered failure rate model under stage life testing (SLT) experiments. The proposed approach relies on the assumption of piecewise constant baseline hazard functions and employs an Expectation Conditional Maximization (ECM) algorithm for efficient parameter estimation. This enables the method to handle complex likelihood structures that typically arise in SLT experiments. Furthermore, we have extended the methodology to incorporate situations with unknown cut points through a data-driven framework based on non-homogeneous Poisson processes. The versatility of the method has also been illustrated by its extension to general k-stage SLT setups.

Simulation studies confirm that the proposed estimators perform satisfactorily in terms of bias, mean squared error, and coverage probability, with performance improving as the sample size and censoring proportion increase. In addition, the developed framework also provides a computational strategy to determine optimal censoring schemes based on D-optimality criteria.

There are several promising directions for future research. An avenue is to explore adaptive or data-driven strategies for selecting the number and placement of cut points. Another is to extend the methodology to accommodate dependent failure mechanisms and more general censoring schemes.

References

Bagdonavicius, V., & Nikulin, M. (2001). *Accelerated life models: modeling and statistical analysis.* Chapman and Hall/CRC.

Balakrishnan, N. (2007). Progressive censoring methodology: an appraisal. *Test*, 16(2), 211–259.

Balakrishnan, N., & Aggarwala, R. (2000). *Progressive censoring: theory, methods, and applications.* Springer Science & Business Media.

Balakrishnan, N., Koutras, M. V., Milienos, F. S., & Pal, S. (2016). Piecewise linear approximations for cure rate models and associated inferential issues. *Methodology and Computing in Applied Probability*, 18(4), 937–966.

Balakrishnan, N., Kundu, D., Ng, K. T., & Kannan, N. (2007). Point and interval estimation for a simple step-stress model with Type-II censoring. *Journal of Quality Technology*, 39(1), 35–47.

Balakrishnan, N., & Ling, M. H. (2012). EM algorithm for one-shot device testing under the exponential distribution. *Computational Statistics & Data Analysis*, 56(3), 502–509.

Balakrishnan, N., & Xie, Q. (2007a). Exact inference for a simple step-stress model with Type-II hybrid censored data from the exponential distribution. *Journal of Statistical Planning and Inference*, 137(8), 2543–2563.

Balakrishnan, N., & Xie, Q. (2007b). Exact inference for a simple step-stress model with Type-I hybrid censored data from the exponential distribution. *Journal of Statistical Planning and Inference*, 137(11), 3268–3290.

Basak, P., Basak, I., & Balakrishnan, N. (2009). Estimation for the three-parameter lognormal distribution based on progressively censored data. *Computational Statistics & Data Analysis*, 53(10), 3580–3592.

Bhattacharya, R., Pradhan, B., & Dewanji, A. (2016). On optimum life-testing plans under Type-II progressive censoring scheme using variable neighborhood search algorithm. *Test*, 25(2), 309–330.

Bhattacharyya, G. K., & Soejoeti, Z. (1989). A tampered failure rate model for step-stress accelerated life test. *Communications in Statistics—Theory and Methods*, 18(5), 1627–1643.

Chen, D. G., & Lio, Y. L. (2010). Parameter estimations for generalized exponential distribution under progressive type-I interval censoring. *Computational Statistics & Data Analysis*, 54(6), 1581–1591.

Cohen, A. C. (1963). Progressively censored samples in life testing. *Technometrics*, 5(3), 327–339.

DeGroot, M. H., & Goel, P. K. (1979). Bayesian estimation and optimal designs in partially accelerated life testing. *Naval Research Logistics Quarterly*, 26(2), 223–235.

Dempster, A. (1977). Maximum likelihood estimation from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39, 1–38.

Efron, B. (1982). *The jackknife, the bootstrap and other resampling plans.* SIAM.

Goodman, M. S., Li, Y., & Tiwari, R. C. (2011). Detecting multiple change points in piecewise constant hazard functions. *Journal of Applied Statistics*, 38(11), 2523–2532.

Han, D., & Balakrishnan, N. (2010). Inference for a simple step-stress model with competing risks for failure from the exponential distribution under time constraint. *Computational Statistics & Data Analysis*, 54(9), 2066–2081.

Hansen, P., & Mladenović, N. (1999). An introduction to variable neighborhood search. In *Metaheuristics: Advances and trends in local search paradigms for optimization* (pp. 433–458). Boston, MA: Springer US.

Kundu, D. (2008). Bayesian inference and life testing plan for the Weibull distribution in presence of progressive censoring. *Technometrics*, 50(2), 144–154.

Kundu, D., & Ganguly, A. (2017). *Analysis of step-stress models: existing results and some recent developments.* Academic Press.

Kundu, D., & Joarder, A. (2006). Analysis of Type-II progressively hybrid censored data. *Computational Statistics & Data Analysis*, 50(10), 2509–2528.

Laumen, B., & Cramer, E. (2019). Stage life testing. *Naval Research Logistics (NRL)*, 66(8), 632–647.

Laumen, B., & Cramer, E. (2021). k-step stage life testing. *Statistica Neerlandica*, 75(2), 203–233.

Meeker, W. Q., Escobar, L. A., & Pascual, F. G. (2021). *Statistical methods for reliability data.* John Wiley & Sons.

Meeker, W. Q., & Hahn, G. J. (1985). *How to plan an accelerated life test: Some practical guidelines.* ASQC Statistics Division.

Meng, X. L., & Rubin, D. B. (1993). Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika*, 80(2), 267–278.

Murphy, A. (1996). A piecewise-constant hazard-rate model for the duration of unemployment in single-interview samples of the stock of unemployed. *Economics Letters*, 51(2), 177–183.

Neal, R. M., & Hinton, G. E. (1998). A view of the EM algorithm that justifies incremental, sparse, and other variants. In *Learning in Graphical Models* (pp. 355–368). Dordrecht: Springer Netherlands.

Nelson, W. B.(1980). Accelerated life testing-step-stress models and data analyses. *IEEE Transactions on Reliability*, 29(2), 103–108.

Nelson, W. B. (2009). *Accelerated testing: statistical models, test plans, and data analysis.* John Wiley & Sons.

Ng, H. K. T., Chan, P. S., & Balakrishnan, N. (2002). Estimation of parameters from progressively censored data using EM algorithm. *Computational Statistics & Data Analysis*, 39(4), 371–386.

Samanta, D., & Kundu, D. (2023). Bivariate semi-parametric model: Bayesian inference. *Methodology and Computing in Applied Probability*, 25(4), 87.

Samanta, D., & Kundu, D. (2024a). Inference of multi-sample stage life testing model under Weibull distribution. *Quality and Reliability Engineering International*, 40(8), 4462–4481.

Samanta, D., & Kundu, D. (2024b). On competing risk model under step-stress stage life testing. *Journal of Statistical Computation and Simulation*, 94(4), 699–721.

Samanta, D., & Kundu, D. (2025). Absolutely Continuous Semi-parametric Bivariate Distributions. *Sankhya B*, 1–29.

Samanta, D., Mondal, S., & Kundu, D. (2022). Optimal plan for ordered step-stress stage life testing. *Statistics*, 56(6), 1319–1344.

Sedyakin, N. M. (1966). On one physical principle in reliability theory. *Techn. Cybernetics*, 3, 80–87.