Computational Statistics (2022) 37:355–379
https://doi.org/10.1007/s00180-021-01124-0

ORIGINAL PAPER

Estimation of component reliability from superposed renewal processes by means of latent variables

Agatha Rodrigues[1] · **Pascal Kerschke**[2] · **Carlos Alberto de B. Pereira**[3,4] · **Heike Trautmann**[5] · **Carolin Wagner**[5] · **Bernd Hellingrath**[5] · **Adriano Polpo**[6]

Received: 2 July 2019 / Accepted: 23 June 2021 / Published online: 6 July 2021

Abstract

We present a new way to estimate the lifetime distribution of a reparable system consisted of similar (equal) components. We consider as a reparable system, a system where we can replace a failed component by a new one. Assuming that the lifetime distribution of all components (originals and replaced ones) are the same, the position of a single component can be represented as a renewal process. There is a considerable amount of works related to estimation methods for this kind of problem. However, the data has information only about the time of replacement. It was not recorded which component was replaced. That is, the replacement data are available in an aggregate form. Using both Bayesian and a maximum likelihood function approaches, we propose an estimation procedure for the lifetime distribution of components in a repairable system with aggregate data. Based on a latent variables method, our proposed method out-perform the commonly used estimators for this problem. The proposed procedure is generic and can be used with any lifetime probability model. Aside from point estimates, interval estimates are presented for both approaches. The performances of the proposed methods are illustrated through several simulated data, and their efficiency and applicability are shown based on the so-called cylinder problem. The computational implementation is available in the R package `srplv`.

✉ Agatha Rodrigues
agatha.rodrigues@ufes.br

1 Department of Statistics, Federal University of Espírito Santo, Vitória, ES, Brazil

2 Big Data Analytics in Transportation, Technical University Dresden, 01062 Dresden, Germany

3 Institute of Mathematics and Statistics, University of São Paulo, São Paulo, SP, Brazil

4 Institute of Mathematics, Federal University of Mato Grosso do Sul, Campo Grande, MS, Brazil

5 Information Systems and Statistics, University of Münster, 48149 Münster, Germany

6 Western Australia University, Perth, Australia

Keywords Aggregate data · Bayesian paradigm · Component lifetime · Maximum likelihood estimator · Parametric estimation · Repairable system

1 Introduction

A system is defined as a set of working components performing a given task. A repairable system is a system characterized by the replacement of a failed component by a new working component of the same kind, in each specific position. Assuming that the new component lifetime distribution is the same as the replaced one, a single component position can be represented by a renewal process (Rinne 2008). Under this assumption, some works presented methods to analyze the renewal process data (Crowder et al. 1994; Crow 1990; Meeker and Escobar 2014; Nelson 2003).

The objective is to estimate the lifetime distribution of components that form the system. It is important for many purposes, such as collecting information for future system design and maintenance planning for individual units (Zhang et al. 2017). However, there are situations in which the information about the exact position the failure occurred is not available (Miyakawa 1984; Liu et al. 2017; Rodrigues et al. 2018, 2019; Wang et al. 2015). As a reviewer of this manuscript pointed out, the reason that the position information was not in the database is that such databases were designed to provide financial information, and not engineering information. Such problems with warranty and maintenance databases are common.

For repairable system when the event time for each replacement is available, but we do not know which component underwent the replacement, data are available in an aggregate form. This aggregate data form a superposed renewal process (SRP) (Zhang et al. 2017). The scenario considered in this work is the following: a fleet of independent systems (sample) is observed. Within each system, there is a set of m identical components, and when a component fails, it is replaced by a functioning one in its position, which we will call as a socket. Although the number of failures r within the interval $[0, \tau]$, τ is the end-of-observation time, can be observed for a given system, this information is unknown for the single sockets.

The challenge is how to estimate the components' lifetime distribution in a SRP scenario. In Fig. 1, an example is presented with $m = 2$ sockets and the observed failures. The true failures occurred at socket 1 and socket 2 are shown in the first two lines. These underlying failure times are unobserved at the two positions. What it is really observed is the superposed failure times, which is shown in the third line. Once it is only observed the failure times without the information about the socket in which each failure occurs, it is unobserved the failure time of each component.

There are some studies that deal with SRP data. Peixoto (2009) propose assigning the event times to sockets randomly, and then using simulation to correct for bias in order to estimate the components' failure time distribution. Krivtsov et al. (2017) consider superposed recurrent data scenario, but their focus is of classification nature, that is, to classify a second claim as a repeat failure in a socket in which a component has already been replaced once, or as the first failure occurred in another socket. Dewanji et al. (2012) deal with the situation when the number of components is unknown, and they propose a nonparametric estimator was proposed to estimate this unknown num-

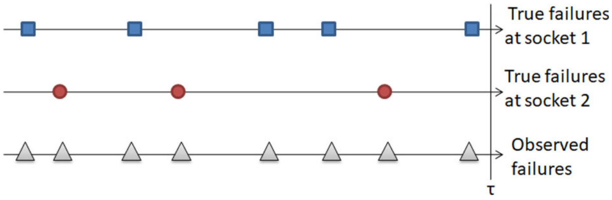


Fig. 1 An illustrated example for the case with $m = 2$ components and the observed failures. This figure is based on Figure 1 from Song and Xie (2018)

ber. Song and Xie (2018) develop a method to estimate superposed renewal processes by using the first few observed failure times without knowing the failure locations, considering all possible combinations to provide exact joint distribution in the likelihood function and for the rest observed failure times, they use some asymptotic distribution in the likelihood function.

Zhang et al. (2017) propose a procedure for estimating the component lifetime distribution from a collection of SRP by maximizing its likelihood function. The likelihood function is given by the sum of all possible data configurations, that is, all possible combinations in which the r failures might occur across the m sockets. However, the number of all possible data configurations increases exponentially with the number of failures, and for large numbers of m and r, the computation of the maximum likelihood is too expensive. Thus, depending on the numbers of failures and components for each system in the fleet, the computational time is very costly, and in some situations, it is not possible to compute by considering their method. In this way, as the authors discuss, the method proposed by them is only applicable for dealing with a fleet of SRPs where each SRP only has a relatively small number of failures.

This work aims to estimate the components' lifetime distribution of the part based on the superposed failure times without restrictions about the numbers of components and failures. Our two methods—a maximum likelihood and a Bayesian approach—consider latent variables during the estimation process. The contributions are as follows:

- Under the maximum likelihood approach, we expect that considering latent variables and estimating the parameters via the Expectation Maximization (EM) algorithm (Robert et al. 2010) solves the limitation of the approach by Zhang et al. (2017), i.e, not being able to compute the maximum likelihood estimator regardless of the number of failures and components. Besides, in situations in which the method of Zhang et al. (2017) is useful, we expect that both methods yield similar performances, once they propose maximizing the likelihood function.
- By proposing a Bayesian approach to solve the problem, we develop a useful method for incorporating expert knowledge and/or past experiences as a priori distribution, besides considering the statistical inference under the Bayesian paradigm.

Under the parametric approach, our proposed methods are generic, and any probability distribution on positive support can be considered for the components' lifetime distributions. Aside from point estimates, interval estimates are discussed for both approaches.

The remainder of this manuscript is organized as follows. In Sect. 2, we describe the data structure. Sections 3 and 4 present the maximum likelihood and Bayesian approaches in more detail. Both methods are evaluated by means of simulation studies, in which they are compared with the method proposed by Zhang et al. (2017), in scenarios this last is possible, and the corresponding results are given in Sect. 5. Section 6 shows the applicability of the methodology in the cylinder dataset and Sect. 7 concludes this work.

2 Data structure

Consider a series system with m components operating in m sockets (position of each component). When a component fails, it is recorded the failure time of the system, and the failed component is replaced by a new one in the same socket. However, there is no register of which socket the component was replaced. In the following paragraph, we will define quantities for a single socket and hence, for simplicity, we omit the socket indices.

Let Y_l be a random variable denoting the lifetime of the component in a specific socket, for $l = 1, 2, \ldots$. Note that, the second component (Y_2) will replace the first component (Y_1) when this one fails. The third component (Y_3) will replace the second component (Y_2), and so on. Also, assume that the components' lifetimes $(Y_1, Y_2, \ldots)$ are independent and identically distributed (i.i.d.). Let Z_k be a positive random variable that denotes the time of the k-th components lifetime in a specific socket. Then, $Z_k = \sum_{l=1}^{k} Y_l$, $k \geq 1$, and $\{Z_k\}$ is a renewal process (RP), that is, each socket in the system represents a RP.

Let T_k be the k-th failure time of the system, in which $T_1 = \min\{Y_{11}, Y_{21}, \ldots, Y_{m1}\}$ and Y_{j1} denotes the first component failure time in the j-th socket, $j = 1, \ldots, m$. Let $\mathcal{T} = (t_1, t_2, \ldots, t_r, \tau)$ denote the observed event history of a single SRP (superposed renewal process) with event times $t_1 < t_2 < \ldots < t_r$, and end-of-observation time τ with $\tau > t_r$. The data consist of n independents SRPs, corresponding to the n systems in the fleet.

The assumptions made here are:

1. the lifetime distribution is the same for all components, in all sockets and systems;
2. the failures are independent between all components, and all sockets;
3. all sockets within one system have the same end-of-observation time τ; and
4. the n systems in the fleet are independent.

3 Maximum likelihood approach

Under the assumption that the components' lifetimes are i.i.d., let $f(\cdot) = f(\cdot \mid \boldsymbol{\theta})$ and $R(\cdot) = R(\cdot \mid \boldsymbol{\theta})$ be the density and reliability functions of the component lifetime, where $\boldsymbol{\theta}$ is a vector, with dimension p, of unknown parameters.

Consider a sample of n systems. Let $\boldsymbol{t}_i = (t_{1i}, t_{2i}, \ldots, t_{r_i i})$ be the vector of observed r_i failure times for the i-th system and τ_i the end-observation time, with $i = 1, \ldots, n$, in which $\boldsymbol{\mathcal{T}}_i = (\boldsymbol{t}_i, \tau_i)$ is the observed data for the i-th system. Let $\boldsymbol{d}_i = (d_{1i}, d_{2i}, \ldots, d_{r_i i})$ the vector that indicates the position of the r_i failures, in which $d_{ki} = j$, if the k-th failure occurred in the j-th socket for the i-th system, for $j = 1, \ldots, m$, $k = 1, \ldots, r_i$ and $i = 1, \ldots, n$.

For instance, lets assume that $\boldsymbol{d}_i$ is observed, and consider a system i with $m = 16$ components, $r_i = 3$ failures, $d_{i1} = d_{3i} = 1$ and $d_{2i} = 13$, were observed. The likelihood contribution of this system is

$$f(t_{1i}) f(t_{3i} - t_{1i}) R(\tau_i - t_{3i}) f(t_{2i}) R(\tau_i - t_{2i}) [R(\tau_i)]^{m-2}. \tag{1}$$

Note that the likelihood contribution of system i given in (1) considers that $\boldsymbol{d}_i = (d_{1i}, d_{2i}, d_{3i})$ is known. In a aggregate scenario, the actual failure positions (sockets) $\boldsymbol{d}_i$ of system i are not observable. Hence, there are $V_i = m^{r_i} = 16^3 = 4{,}096$ different possible configurations of the likelihood contributions for this system, in which V_i is the number of possible data configurations of system i, with m sockets, and r_i failure times. The likelihood contribution of the i-th system is given by

$$L_i = \sum_{v=1}^{V_i} L_{iv},$$

in which L_{iv} is the likelihood contribution of the v-th configuration for system i. Considering that a fleet of n independent systems is observed, the likelihood function for $\boldsymbol{\theta}$ is

$$L(\boldsymbol{\theta} \mid \boldsymbol{\mathcal{T}}) = \prod_{i=1}^{n} \left[\sum_{v=1}^{V_i} L_{iv} \right], \tag{2}$$

where $\boldsymbol{\mathcal{T}} = (\boldsymbol{\mathcal{T}}_1, \ldots, \boldsymbol{\mathcal{T}}_n)$. Zhang et al. (2017) propose the maximization of the likelihood function given in (2).

In the aggregate scenario, $\boldsymbol{d}_i$ is a vector of latent variables. A suitable approach for estimating the parameter values, which maximize the likelihood function, is to consider an expectation-maximization (EM) algorithm, which is presented in the following subsection.

3.1 EM algorithm

The EM algorithm is an iterative method with Expectation (E), and Maximization (M) steps (Dempster et al. 1977). The E-step evaluates the expectation of the full log-likelihood function, and the M-step tries to find the parameter configuration, which maximizes the expectation found within the E-step.

The complete likelihood function [i.e., the likelihood function with augmented data (Robert et al. 2010)] of $\boldsymbol{\theta}$ is given by

$$L(\boldsymbol{\theta} \mid \mathcal{T}, \boldsymbol{d}) = \prod_{i=1}^{n} L_i(\boldsymbol{\theta} \mid \mathcal{T}_i, \boldsymbol{d}_i). \tag{3}$$

The form of $L_i(\boldsymbol{\theta} \mid \mathcal{T}_i, \boldsymbol{d}_i)$ depends on the number of failures r_i. For this reason, a general form is presented in the following.

Given $\boldsymbol{d}_i$, let Γ_i be the set of v_i component indexes that cause at least one failure for system i. Also, $v_i = 0$ when no failure is observed . Let x_{ilk} the k-th failure time caused by the l-th element of Γ_i, with $l = 1, \ldots, v_i$ and $k = 1, \ldots, n_l$. For instance, for system i with $r_i = 3$ failures observed, $d_{1i} = d_{3i} = 1$, and $d_{2i} = 13$, we have that $\Gamma_i = \{1, 13\}$, $v_i = 2$, $n_1 = 2$, $n_2 = 1$, $x_{i11} = t_{1i}$, $x_{i12} = t_{3i}$, and $x_{i21} = t_{2i}$. Thus, $\sum_{l=1}^{v_i} n_l = r_i$.

The likelihood contribution of the i-th system can be written as

$$L_i(\boldsymbol{\theta} \mid \mathcal{T}_i, \boldsymbol{d}_i) = \left\{ \prod_{l=1}^{v_i} \left[\prod_{k=1}^{n_l} f(x_{ilk} - x_{il(k-1)}) \right] R(\tau_i - x_{iln_l}) \right\}^{1-\mathrm{I}(v_i=0)} R(\tau_i)^{m-v_i},$$

where $x_{il0} = 0$, and indicator function $\mathrm{I}(A) = 1$, if A is true, and 0 otherwise.

Let $l_i(\boldsymbol{\theta} \mid \mathcal{T}_i, \boldsymbol{d}_i) = \log L_i(\boldsymbol{\theta} \mid \mathcal{T}_i, \boldsymbol{d}_i)$. The logarithm of the complete likelihood in (3) can be written as

$$\begin{aligned} l(\boldsymbol{\theta} \mid \mathcal{T}, \boldsymbol{d}) &= \sum_{i=1}^{n} l_i(\boldsymbol{\theta} \mid \mathcal{T}_i, \boldsymbol{d}_i) \\ &= \sum_{i=1}^{n} \left\{ \Big[1 - \mathrm{I}(v_i = 0)\Big] \left[\sum_{l=1}^{v_i} \sum_{k=1}^{n_l} \log f(x_{ilk} - x_{il(k-1)}) + \sum_{l=1}^{v_i} \log R(\tau_i - x_{iln_l}) \right] \right. \\ &\quad \left. + (m - v_i) \log R(\tau_i) \right\} \end{aligned} \tag{4}$$

Let $\boldsymbol{\theta}_r$ be the value assumed by $\boldsymbol{\theta}$ in the r-th iteration of the algorithm. The $(r+1)$-th E-step consists of calculating the expectation of (4), that is,

$$Q(\boldsymbol{\theta} \mid \boldsymbol{\theta}_r) = \mathrm{E}\big[l(\boldsymbol{\theta} \mid \mathcal{T}, \boldsymbol{d}) \mid \mathcal{T}; \boldsymbol{\theta}_r\big]. \tag{5}$$

Unfortunately, there exists no analytical expression of the expectation in (5). Instead, it can be approximated by Monte-Carlo simulations. Consider that L random samples $\boldsymbol{d}_i^{(1)}, \ldots, \boldsymbol{d}_i^{(L)}$ are simulated based on $f(\boldsymbol{d}_i \mid \mathcal{T})$, i.e., the density function of $\boldsymbol{d}$ conditional to $\mathcal{T}$, $i = 1, \ldots, n$ (for more details, see Sect. 3.1.1). Thus, the E-step results in calculating

$$Q_m(\boldsymbol{\theta} \mid \boldsymbol{\theta}_r) = \frac{1}{L} \sum_{l=1}^{L} l(\boldsymbol{\theta} \mid \mathcal{T}, \boldsymbol{d}^{(l)}) = \frac{1}{L} \sum_{l=1}^{L} \sum_{i=1}^{n} l_i\Big(\boldsymbol{\theta} \mid \mathcal{T}_i, \boldsymbol{d}_i^{(l)}\Big). \tag{6}$$

The M-step maximizes (6) with respect to $\boldsymbol{\theta}$ resulting in $\boldsymbol{\theta}_{r+1}$. The optimization method considered within this work is the Nelder–Mead algorithm (Nelder and Mead 1965). The E- and M-steps are alternated until the difference of estimates between two consecutive iteration values is less than 10^{-4}. The estimate $(\widehat{\boldsymbol{\theta}})$ of $\boldsymbol{\theta}$ is obtained when the convergence criterion is reached.

Let $g(\boldsymbol{\theta})$ be a function of $\boldsymbol{\theta}$. Due to the invariance property of the maximum likelihood estimator (MLE), the MLE of $g(\boldsymbol{\theta})$ is $g(\widehat{\boldsymbol{\theta}})$. For instance, if the Weibull distribution with parameters $\beta > 0$ (shape) and $\eta > 0$ (scale) is assumed for components' failure times, in which $\boldsymbol{\theta} = (\beta, \eta)$, the expected lifetime of the component is $\mathrm{E}(Y) = g(\boldsymbol{\theta}) = \eta\Gamma(1 + (1/\beta))$, and its MLE is by $g(\widehat{\boldsymbol{\theta}}) = \widehat{\eta}\Gamma(1 + (1/\widehat{\beta}))$, where $\widehat{\beta}$ and $\widehat{\eta}$ are the MLE of β and η, respectively (Casella and Berger 2002). Analogous, the MLE for the component reliability function is $\widehat{R}(y) = \exp\left[-(y/\widehat{\eta})^{\widehat{\beta}}\right]$, for $y > 0$.

3.1.1 Conditional distribution of d given $\mathcal{T}$

For a fixed i,

$$\begin{aligned}
f(\boldsymbol{d}_i \mid \boldsymbol{\mathcal{T}}_i) &= f(d_{1i}, d_{2i}, \ldots, d_{r_i i} \mid \boldsymbol{\mathcal{T}}_i)\\
&= f(d_{r_i i} \mid \boldsymbol{\mathcal{T}}_i, d_{(r_i-1)i}, d_{(r_i-2)i}, \ldots, d_{2i}, d_{1i}) f(d_{(r_i-1)i} \mid \boldsymbol{\mathcal{T}}_i, d_{(r_i-2)i}, \ldots, d_{1i})\\
&\quad \times \cdots \times f(d_{2i} \mid \boldsymbol{\mathcal{T}}_i, d_{1i}) f(d_{1i} \mid \boldsymbol{\mathcal{T}}_i). \qquad (7)
\end{aligned}$$

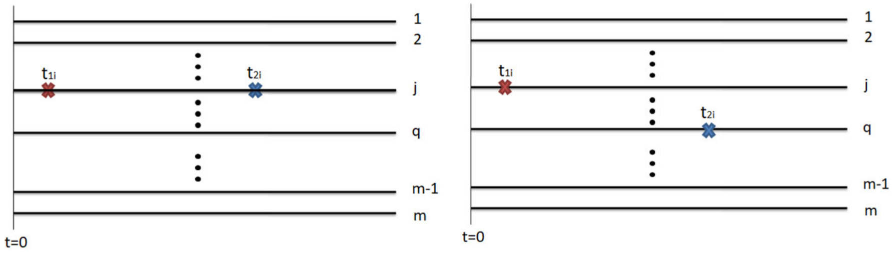
The conditional distribution of d given $\mathcal{T}$ is presented in Eq. (7) for any number of failures $r_i \geq 1$. In order to present the idea, consider a particular case with $r_i = 3$. Then, $\boldsymbol{\mathcal{T}}_i = (t_{1i}, t_{2i}, t_{3i}, \tau_i)$, and

$$\begin{aligned}
f(\boldsymbol{d}_i \mid \boldsymbol{\mathcal{T}}_i) &= f(d_{1i}, d_{2i}, d_{3i} \mid \boldsymbol{\mathcal{T}}_i)\\
&= f(d_{3i} \mid \boldsymbol{\mathcal{T}}_i, d_{2i}, d_{1i}) f(d_{2i} \mid \boldsymbol{\mathcal{T}}_i, d_{1i}) f(d_{1i} \mid \boldsymbol{\mathcal{T}}_i). \qquad (8)
\end{aligned}$$

The development of each term of the last side of the Eq. (8) is presented in the following.

Under i.i.d assumption, the distribution of $d_{1i} = j \mid \boldsymbol{\mathcal{T}}_i$ follows a Multinomial distribution, $Multin(1, \boldsymbol{p}_{1i})$, where $\boldsymbol{p}_{1i} = (p_{11i}, \ldots, p_{1mi})$ and $p_{1ji} = 1/m$, $j = 1, \ldots, m$. Note that, in this case, the multinomial distribution is, in fact, a discrete uniform distribution.

Before discussing the distribution of $d_{2i} \mid (\boldsymbol{\mathcal{T}}_i, d_{1i} = j)$, one comment is important to point: when we develop the distribution of d_{vi} conditional to the distribution of the indicator of the socket that each previous failure occurred, that is, conditional to $(d_{(v-1)i}, d_{(v-1)i}, \ldots, d_{2i}, d_{1i})$, with $v = 2, \ldots, m$, one only needs to worry about the number of sockets that the previous failures occurred and the index of the last failure that occurred in each socket that had failure, regardless the socket index, because we are assuming the lifetime distributions of the components at the sockets are i.i.d. So, the distribution of d_{2i} conditional to the socket that the first failure occurred, say at



(a) Situation 1 (b) Situation 2

Fig. 2 The two possible situations of how two failures can occur in the sockets

the socket j (any $j \in \{1, \ldots, m\}$), can be described as follows:

$$f(d_{2i} \mid \mathcal{T}_i, d_{1i} = j) \propto [f(t_{2i} - t_{1i})]^{\mathrm{I}(d_{2i}=j)} \prod_{l=1;l\neq j}^{m} [f(t_{2i})]^{\mathrm{I}(d_{2i}=l)},$$

that is, $d_{2i} \mid (\boldsymbol{t}_i, d_{1i} = j)$ follows $Multin(1, \boldsymbol{p}_{2i})$, in which $\boldsymbol{p}_{2i} = (p_{21i}, \ldots, p_{2mi})$, $p_{2ji} = f(t_{2i} - t_{1i})/C$ and $p_{2li} = f(t_{2i})/C$, $l = 1, \ldots, m$ and $l \neq j$, with $C = f(t_{2i} - t_{1i}) + (m-1)f(t_{2i})$.

Now, we need to specify the distribution of d_{3i} conditional to the values assumed by d_{2i} and d_{1i}. For that, one has to consider the following two situations: (1) the first failure occurs at a given socket, say at the socket j, and the second failure occurs at the same socket, that is, it also occurs at the socket j (Fig. 2a); and (2) the first failure occurs at a given socket, say at the socket j, and the second failure occurs at a different socket, say at the socket q (Fig. 2b), with $j \neq q$, $j = 1, \ldots, m$ and $q = 1, \ldots, m$.

So, depending on each situation, the conditional distribution of d_{3i} is given by:

- Situation 1: distribution of $d_{3i} \mid (\mathcal{T}_i, d_{1i} = j, d_{2i} = j)$:

$$f(d_{3i} \mid \mathcal{T}_i, d_{1i} = j, d_{2i} = j) \propto [f(t_{3i} - t_{2i})]^{\mathrm{I}(d_{3i}=j)} \prod_{l=1;l\neq j}^{m} [f(t_{3i})]^{\mathrm{I}(d_{3i}=l)},$$

  that is, $d_{3i} \mid (\mathcal{T}_i, d_{1i} = j, d_{2i} = j)$ follows $Multin(1, \boldsymbol{p}_{3i})$, in which $\boldsymbol{p}_{3i} = (p_{31i}, \ldots, p_{3mi})$, $p_{3ji} = f(t_{3i} - t_{2i})/C$ and $p_{3li} = f(t_{3i})/C$, $l = 1, \ldots, m$ and $l \neq j$, with $C = f(t_{3i} - t_{2i}) + (m-1)f(t_{3i})$.
- Situation 2: distribution of $d_{3i} \mid (\mathcal{T}_i, d_{1i} = j, d_{2i} = q)$, with $q \neq j$:

$$f(d_{3i} \mid \mathcal{T}_i, d_{1i} = j, d_{2i} = q) \propto [f(t_{3i} - t_{1i})]^{\mathrm{I}(d_{3i}=j)} [f(t_{3i} - t_{2i})]^{\mathrm{I}(d_{3i}=q)} \times \prod_{l=1;l\neq j,q}^{m} [f(t_{3i})]^{\mathrm{I}(d_{3i}=l)},$$

  that is, $d_{3i} \mid (\mathcal{T}_i, d_{1i} = j, d_{2i} = q)$ follows $Multin(1, \boldsymbol{p}_{3i})$, in which $\boldsymbol{p}_{3i} = (p_{31i}, \ldots, p_{3mi})$, $p_{3ji} = f(t_{3i} - t_{1i})/C$, $p_{3qi} = f(t_{3i} - t_{2i})/C$ and $p_{3li} =$

$f(t_{3i})/C, l = 1, \ldots, m$ and $l \neq j, q$, with $C = f(t_{3i} - t_{1i}) + f(t_{3i} - t_{2i}) + (m - 2) f(t_{3i})$.

The distribution in the Eq. (7) depends on the number of failures. In the supplementary material we present, as an example, the conditional distribution of d with $r_i = 5$.

3.2 Asymptotic distribution

The asymptotic distribution of the maximum likelihood estimator $\widehat{\boldsymbol{\theta}}$ can be approximated by a multivariate normal distribution with mean $\boldsymbol{\theta}$ and variance-covariance matrix $\boldsymbol{I_\theta}(\boldsymbol{\theta})^{-1}$, where $\boldsymbol{I_\theta}(\boldsymbol{\theta})$ is the observed information matrix for $\boldsymbol{\theta}$. As demonstrated by Louis (1982), $\boldsymbol{I_\theta}(\widehat{\boldsymbol{\theta}})$ is the sum of

$$I_1(\boldsymbol{\theta} \mid \widehat{\boldsymbol{\theta}}) = -\frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} Q(\boldsymbol{\theta} \mid \widehat{\boldsymbol{\theta}}) \quad \text{and} \quad I_2(\boldsymbol{\theta} \mid \widehat{\boldsymbol{\theta}}) = -\mathrm{Var}\left\{ \frac{\partial}{\partial \boldsymbol{\theta}} l(\boldsymbol{\theta} \mid \boldsymbol{\mathcal{T}}, \boldsymbol{d}) \Big| \boldsymbol{\mathcal{T}}; \widehat{\boldsymbol{\theta}} \right\}.$$

The matrix $I_1(\boldsymbol{\theta} \mid \widehat{\boldsymbol{\theta}})$ can be estimated by

$$-\frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} Q_m(\boldsymbol{\theta} \mid \widehat{\boldsymbol{\theta}}) = -\frac{1}{L} \sum_{i=1}^{n} \sum_{l=1}^{L} \frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} l_i \left(\boldsymbol{\theta} \mid \boldsymbol{\mathcal{T}}_i, \boldsymbol{d}_i^{(l)} \right) \bigg|_{\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}},$$

where $\boldsymbol{d}_i^{(l)}$, with $l = 1, \ldots, L$, being a random sample from the distribution of $f(\boldsymbol{d}_i \mid \boldsymbol{\mathcal{T}}_i)$ for the i-th system.

An estimate of $I_2(\boldsymbol{\theta} \mid \widehat{\boldsymbol{\theta}})$ results from the sum of

$$\sum_{i=1}^{n} \left\{ \frac{1}{L} \sum_{l=1}^{L} \frac{\partial}{\partial \boldsymbol{\theta}} l_i \left(\boldsymbol{\theta} \mid \boldsymbol{\mathcal{T}}_i, \boldsymbol{d}_i^{(l)} \right) \bigg|_{\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}} \right\} \left\{ \frac{1}{L} \sum_{l=1}^{L} \frac{\partial}{\partial \boldsymbol{\theta}} l_i \left(\boldsymbol{\theta} \mid \boldsymbol{\mathcal{T}}_i, \boldsymbol{d}_i^{(l)} \right) \bigg|_{\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}} \right\}^\top$$

and

$$-\frac{1}{L} \sum_{i=1}^{n} \sum_{l=1}^{L} \left\{ \frac{\partial}{\partial \boldsymbol{\theta}} l_i \left(\boldsymbol{\theta} \mid \boldsymbol{\mathcal{T}}_i, \boldsymbol{d}_i^{(l)} \right) \right\} \left\{ \frac{\partial}{\partial \boldsymbol{\theta}} l_i \left(\boldsymbol{\theta} \mid \boldsymbol{\mathcal{T}}_i, \boldsymbol{d}_i^{(l)} \right) \right\}^\top \bigg|_{\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}}.$$

Detailed derivation of $\boldsymbol{I_\theta}(\widehat{\boldsymbol{\theta}})^{-1}$ for the Weibull lifetime model, with parameters β (shape), and η (scale), is given in the "Appendix" section.

An asymptotic $\gamma\%$ confidence interval for $\boldsymbol{\theta}$ ($CI\gamma\%$) is given by

$$CI\gamma\% = \left(\widehat{\boldsymbol{\theta}} - z_{(1-\gamma/2)} \sqrt{(I_{11}, \ldots, I_{pp})}; \quad \widehat{\boldsymbol{\theta}} + z_{(1-\gamma/2)} \sqrt{(I_{11}, \ldots, I_{pp})} \right),$$

in which I_{jj} denotes the jth element of the main diagonal of $\boldsymbol{I_\theta}(\widehat{\boldsymbol{\theta}})^{-1}$.

Confidence intervals for functions of $\boldsymbol{\theta}$ can be obtained by the delta method (Casella and Berger 2002).

3.3 Model selection criteria

We have considered the the following model selection criteria, based on the maximized log-likelihood: AIC (Akaike Information Criterion), AICc (Corrected Akaike Information Criterion), BIC (Bayesian Information Criterion), HQIC (Hannan-Quinn Information Criterion) and CAIC (Consistent Akaike Information Criterion), which are computed, respectively, by $AIC = 2p - 2l$, $AICc = AIC + 2p(p+1)/(n-p-1)$, $BIC = p\log n - 2l$, $HQIC = 2p\log(\log n) - 2l$ and $CAIC = p(\log n + 1) - 2l$, where p is the number of parameters of the fitted model, n is the sample size and l is the maximized log-likelihood function value, obtained by evaluating (6) in the last iteration of EM algorithm estimates.

Given a set of candidate models, the preferred model is the one which provides the lowest criteria values.

4 Bayesian approach

In the Bayesian approach, the latent variable vector $\boldsymbol{d}$ is faced as parameter vector. Thus, the posterior distribution of $(\boldsymbol{\theta}, \boldsymbol{d})$ can be written as

$$\pi(\boldsymbol{\theta}, \boldsymbol{d} \mid \mathcal{T}) \quad \propto \quad \pi(\boldsymbol{\theta}, \boldsymbol{d}) L(\boldsymbol{\theta}, \boldsymbol{d} \mid \mathcal{T}), \tag{9}$$

where $L(\boldsymbol{\theta}, \boldsymbol{d} \mid \mathcal{T})$ has the same form as (3) in which $\boldsymbol{d}$ now is considered as parameter and $\pi(\boldsymbol{\theta}, \boldsymbol{d})$ is the prior distribution of $(\boldsymbol{\theta}, \boldsymbol{d})$.

In real-world settings, it is possible that the prior distributions can be influenced by expert knowledge and/or past experiences on the functioning of the components. In this work, no prior information about the functioning of the components is available, which is the reason for the choice of non-informative prior distributions, besides of the assumption that the parameters are independent a prior.

Given the posterior density in Eq. (9) does not have a closed form, statistical inferences about the parameters can rely on Markov-Chain Monte-Carlo (MCMC) simulations. Here, we consider the Metropolis within Gibbs algorithm (Tierney 1994) once it is possible to sample some of the parameters directly from the conditional distribution; however, this is not possible for other parameters. The algorithm works in the steps presented in Algorithm 1.

Discarding burn-in (i.e., the first generated values are discarded to eliminate the effect of the assigned initial values for parameters) and jump samples (i.e., gaps between the generated values in order to avoid correlation problems), a sample of size n_p from the joint posterior distribution of $(\boldsymbol{\theta}, \boldsymbol{d})$ is obtained. The sample from the posterior distribution can be expressed as $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_{n_p})$. Posterior quantities of $\boldsymbol{\theta}$ can be easily obtained (Robert et al. 2010). For instance, the posterior mean of $\boldsymbol{\theta}$ can be approximated by

$$\frac{1}{n_p}\sum_{k=1}^{n_p} \boldsymbol{\theta}_k.$$

Algorithm 1 The Metropolis within Gibbs algorithm.

1: Assign initial values $\boldsymbol{\theta}^{(0)}$ for $\boldsymbol{\theta}$ and set $b = 1$.
2: Draw $\boldsymbol{d}_i^{(b)}$ from $\pi(\boldsymbol{d}_i \mid \boldsymbol{\mathcal{T}}_i, \boldsymbol{\theta})$ from

$$\begin{aligned}\pi(\boldsymbol{d}_i \mid \boldsymbol{\mathcal{T}}_i, \boldsymbol{\theta}) &= \pi(d_{1i}, d_{2i}, \ldots, d_{r_i i} \mid \boldsymbol{\mathcal{T}}_i, \boldsymbol{\theta}) \\ &= \pi(d_{r_i i} \mid \boldsymbol{\mathcal{T}}_i, \boldsymbol{\theta}, d_{(r_i-1)i}, d_{(r_i-2)i}, \ldots, d_{2i}, d_{1i}) \times \\ &\times \pi(d_{(r_i-1)i} \mid \boldsymbol{\mathcal{T}}_i, \boldsymbol{\theta}, d_{(r_i-2)i}, \ldots, d_{2i}, d_{1i}) \ldots \pi(d_{2i} \mid \boldsymbol{\mathcal{T}}_i, \boldsymbol{\theta}, d_{1i}) \pi(d_{1i} \mid \boldsymbol{\mathcal{T}}_i, \boldsymbol{\theta}),\end{aligned}$$

in an analogous way presented in Sect. 3.1.1, for $i = 1, \ldots, n$, and $\boldsymbol{d}^{(b)} = (\boldsymbol{d}_1^{(b)}, \ldots, \boldsymbol{d}_n^{(b)})$.
3: Draw $\boldsymbol{\theta}^{(b)}$ from

$$\begin{aligned}\pi(\boldsymbol{\theta} \mid \boldsymbol{\mathcal{T}}, \boldsymbol{d}^{(b)}) \propto \pi(\boldsymbol{\theta}) \prod_{i=1}^{n} \Bigg\{ \Bigg[\prod_{l=1}^{v_i} \Bigg(\prod_{k=1}^{n_l} f(x_{ilk} - x_{il(k-1)}) \Bigg) \\ \times R(\tau_i - x_{iln_l}) \Bigg]^{1-\mathrm{I}(v_i=0)} R(\tau_i)^{m-v_i} \Bigg\},\end{aligned}$$

through Metropolis–Hastings algorithm (Robert et al. 2010).
4: Set $b = b + 1$ and repeat steps 2. and 3. until $b = B$, where B is the predefined number of simulated samples of $(\boldsymbol{\theta}, \boldsymbol{d})$.

The sample from the posterior distribution of $g(\boldsymbol{\theta})$ can be expressed as $(g(\boldsymbol{\theta}_1), g(\boldsymbol{\theta}_2), \ldots, g(\boldsymbol{\theta}_{n_p}))$ and posterior quantities of $g(\boldsymbol{\theta})$ can be obtained. For instance, the posterior mean of the reliability function can be approximated by

$$\frac{1}{n_p} \sum_{k=1}^{n_p} R(t \mid \boldsymbol{\theta}_k), \quad t > 0.$$

The proposed approach is generic and straightforward for any probability distribution. Thus, it may be of interest to consider a model selection criterion. Below a criterion based on the conditional predictive ordinates is presented.

4.1 Conditional predictive ordinate

A criterion for model selection that can be considered is based on the conditional predictive ordinates (CPO). For the i-th system, the conditional predictive ordinate (CPO) can be expressed as

$$\begin{aligned}CPO_i = f(\boldsymbol{\mathcal{T}}_i \mid \boldsymbol{\mathcal{T}}_{-i}) &= \sum_{\boldsymbol{d}} \int f(\boldsymbol{\mathcal{T}}_i \mid \boldsymbol{\theta}, \boldsymbol{d}) \pi(\boldsymbol{\theta}, \boldsymbol{d} \mid \boldsymbol{\mathcal{T}}_{-i}) \partial \boldsymbol{\theta} \\ &= \left\{ \sum_{\boldsymbol{d}} \int \frac{\pi(\boldsymbol{\theta}, \boldsymbol{d} \mid \boldsymbol{\mathcal{T}})}{f(\boldsymbol{\mathcal{T}}_i \mid \boldsymbol{\theta}, \boldsymbol{d})} \partial \boldsymbol{\theta} \right\}^{-1} \\ &\approx \left\{ \frac{1}{n_p} \sum_{k=1}^{n_p} \frac{1}{f(\boldsymbol{\mathcal{T}}_i \mid \boldsymbol{\theta}_k, \boldsymbol{d}_k)} \right\}^{-1},\end{aligned}$$

in which $\mathcal{T}_{-i} = (\mathcal{T}_1, \ldots, \mathcal{T}_{i-1}, \mathcal{T}_{i+1}, \ldots, \mathcal{T}_n)$ and $(\boldsymbol{\theta}_k, \boldsymbol{d}_k)$, for $k = 1, \ldots, n_p$, represent a sample from the posterior distribution of $(\boldsymbol{\theta}, \boldsymbol{d})$.

High values of CPO_i indicate that the model is capable of describing the i-th observation adequately (Gilks et al. 1995). The LPML (log pseudo marginal likelihood) measure is the sum of the logarithms of the CPO of all the observations, that is, $LPML = \sum_{i=1}^{n} \log\left(CPO_i\right)$ and the higher the $LPML$ value is, the better the model fit.

5 Model evaluation by means of a simulation study

This section presents the results from simulation studies to evaluate the performance of the estimation methods described above in regards to the estimation quality. In scenarios the method of Zhang et al. (2017) works, we compare its performance with those of the proposed methods.

Thus, the following estimation methods were fitted: Bayesian approach (BA), maximum likelihood estimator via EM algorithm (EM-ML) and maximum likelihood estimator obtained by Zhang et al. (2017) (Z-ML). The Z-ML estimates were obtained by means of the R-package `SRPML` (R Core Team 2020; Zhang et al. 2015). The proposed methods BA and EM-ML were fitted through the R package `srplv`. This package was built by authors in order to make the methodology available for applications.

The steps for generating the data of each simulated example, with m being the number of sockets and n the sample size, are presented in Algorithm 2. The mean (7) and variance (4) values of component failure time distribution are based on cylinder application data (Sect. 6).

Algorithm 2 Data generation.

1: **for** each system unit $i = 1, \ldots, n$ **do**
2: Draw τ_i from a Weibull distribution with mean m_c and variance 0.05.
3: Draw $Y_{11i}, Y_{21i}, \ldots, Y_{m1i}$ from a Weibull distribution with mean 7 and variance 4, where Y_{j1i} is the first component failure time in the j-th socket, for $j = 1, \ldots, m$.
4: Let $T_{1i} = \min\left\{Y_{11i}, Y_{21i}, \ldots, Y_{m1i}\right\}$.
5: **if** $T_{1i} \geq \tau_i$ **then**
6: stop simulation process and $r_i = 0$.
7: **else**
8: Let $Y_{l1i} = \min\left\{Y_{11i}, Y_{21i}, \ldots, Y_{m1i}\right\}$, then $t_{1i} = Y_{l1i}$.
9: Draw Y_{l2i} from Weibull distribution with mean 7 and variance 4 conditional to $Y_{l2i} > t_{1i}$, where Y_{l2i} is the second component failure time in the l-th socket, once the first failure occurred in the l-th socket.
10: Let $T_{2i} = \min\left\{Y_{11i}, Y_{21i}, \ldots, Y_{l2i}, \ldots, Y_{m1i}\right\}$.
11: **if** $T_{2i} \geq \tau_i$ **then**
12: stop simulation process and $r_i = 1$.
13: **else**
14: repeats steps 8 to 10 until $T_{r_i} < \tau_i < T_{(r_i+1)}$.
15: The dataset is $\mathcal{T}_i = \{t_{1i}, t_{2i}, \ldots, t_{r_i i}, \tau_i\}$, for $i = 1, \ldots, n$.

In this section, the Weibull distribution with parameters $\beta > 0$ (shape) and $\eta > 0$ (scale) is assumed for components' failure times, in wich $\boldsymbol{\theta} = (\beta, \eta)$. For BA, the priors of Weibull parameters are considered to be independent gamma distributed with mean 1 and variance 100. Besides, d_{li} follows $Multin(1, \boldsymbol{p}_{li})$, where $\boldsymbol{p}_{li} = (p_{l1i}, \ldots, p_{lmi})$ and $p_{lji} = 1/m$, with $j = 1, \ldots, m$.

To obtain posterior quantities, we used an MCMC procedure to generate a sample from the posterior distribution of the parameters. We generated 20,000 samples from the posterior distribution of each parameter. The first 10,000 of these samples were discarded as burn-in samples. A jump of size 10 was chosen to reduce correlation effects between the samples. As a result, the final sample size of the parameters generated from the posterior distribution was 1,000. The chains' convergence was monitored in all simulation scenarios for good convergence results to be obtained.

For the EM algorithm considered to obtain EM-ML estimates, we consider $L = 100$ and the discussion about the choice of this number based on the scenario considered in this work is presented in the supplementary material.

The mean absolute error (MAE) from each estimator to the true reliability of each method is considered as performance measure. $R(t)$ and $\widehat{R}(t)$ are the true reliability function and the estimate, respectively. Hence, the MAE is evaluated by $\frac{1}{l}\sum_{\ell=1}^{l} \mid \widehat{R}(g_\ell) - R(g_\ell) \mid$, where $\{g_1, \ldots, g_\ell, \ldots, g_l\}$ is a grid in the space of failure times.

First, we conducted two simulated examples, presented in the following. Second, scenarios with different sample sizes, number of sockets and censor mean time are considered.

5.1 Simulated examples

We conducted two simulated examples considering $n = 100$, $m = 16$ and $m_c = 4$ (Example 1) or $m_c = 8$ (Example 2), in which m_c represents the mean of censor distribution, considered in step 2 in Algorithm 2. It is worth noting that the expected number of failures with $m_c = 8$ is larger than with $m_c = 4$.

For the Bayesian approach, the Gelman–Rubin convergence diagnostic statistics (Gelman and Rubin 1992) for parameters β and η are 1.0011 and 1.0004, respectively, in Example 1 and they are 1.0002 and 1.0027 in Example 2. The measures are close to 1, which suggests that convergence chains have been reached.

For EM-ML, 8 and 17 EM iterations have been executed for Examples 1 and 2, respectively, and the corresponding values are listed in Table 1. For both examples, the initial values for (β, η) are (1, 1). After the first iteration it was (1.206, 32.335) for Example 1, after the second one it was (3.165, 8.766) and then reached the covergence region. For Example 2, it took about eight iterations to reach the covergence region. Figure 3 presents contour plots of the log-likelihood function, as well as the iteration values from the second to the eighth iteration for Example 1 and from third to 17-th iteration for Example 2. The convergence was obtained fast for both examples.

The Weibull parameter estimates obtained by BA, EM-ML and Z-ML are presented in Table 2. Note that the Z-ML estimation is not presented for Example 2, because those values could not be computed due to the high number of components and failures.

Table 1 EM algorithm iteration values of Weibull parameters for two simulated examples

Iterations	Example 1		Example 2	
	β	η	β	η
Initial value	1.000	1.000	1.000	1.000
1	1.206	32.335	0.400	31.479
2	3.165	8.766	0.627	17.346
3	3.716	7.793	0.796	13.848
4	3.726	7.780	0.979	11.866
5	3.727	7.778	1.220	10.445
6	3.726	7.780	1.557	9.401
7	3.727	7.778	2.007	8.694
8	3.727	7.778	2.550	8.254
9	–	–	3.050	8.025
10	–	–	3.379	7.924
11	–	–	3.539	7.884
12	–	–	3.603	7.869
13	–	–	3.630	7.863
14	–	–	3.634	7.862
15	–	–	3.638	7.861
16	–	–	3.639	7.861
17	–	–	3.639	7.861

The details about limitations of this method in situation of high numbers of failures and components are given in Zhang et al. (2017).

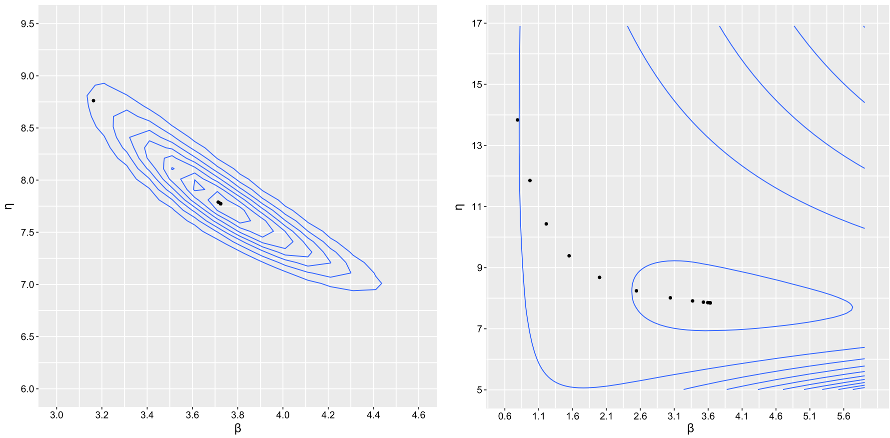
The estimates for the component reliability function obtained by BA, EM-ML, Z-ML, as well as the true reliability function, are presented in Fig. 4. Table 3 lists the MAE values, in which maximum likelihood approaches (EM-ML and Z-ML) present lower MAE values for Example 1, whereas BA and EM-ML present similar MAE values for Example 2.

5.2 Simulation studies in different scenarios

We conducted the simulations for all combinations of the following features: $n \in \{10, 50, 100, 200\}$, $m \in \{4, 8, 16, 32\}$, and $m_c \in \{4, 8\}$, resulting in 32 scenarios. For each scenario, 100 datasets were generated, and we compare the MAE from the estimators to the true distribution.

The boxplot graphs of 100 MAE values are presented in Fig. 5. The results can be split into two parts: 1) Z-ML is applicable and 2) Z-ML is not applicable. For the first one, in general, the methods present similar performance, as one can see in Fig. 5a, b in case of $m \in \{4, 16\}$. When $m_c = 8$ the BA method presents higher MAE means but the boxplot graph intersects with the boxplot graphs obtained by other methods.

Noticeably, Fig. 5b does not contain any boxplots for Z-ML in case of $m \in \{16, 32\}$. However, this is plausible as this method was not able to compute the respective



(a) Example 1 (b) Example 2

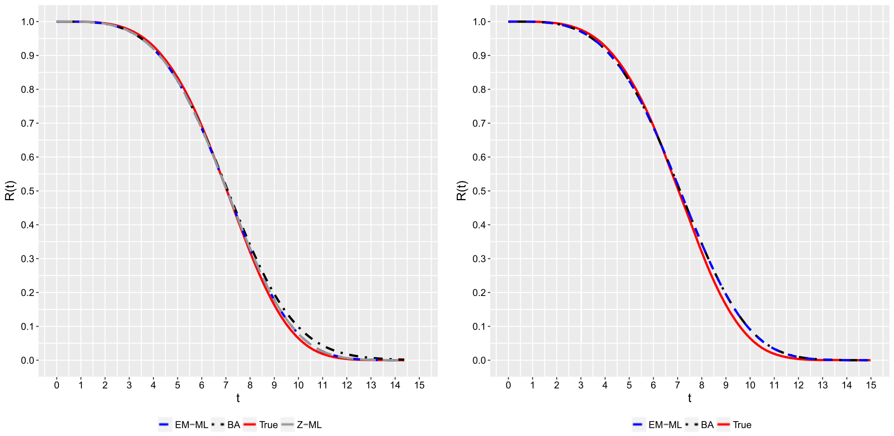
Fig. 3 Contour plots of the log-likelihood function and EM algorithm iteration values (dots) for Example 1 (with $m_c = 4$) and for Example 2 (with $m_c = 8$), in which m_c indicates the expected end-of-observation time

Table 2 Weibull model parameters (β, η) and expected components' time to failure ($\mathrm{E}(Y)$) estimation based on different estimation models: the Bayesian approach (BA), the EM maximum likelihood method (EM-ML) and the maximum likelihood approach from *Zhang et al.* (Z-ML) of simulated examples

Parameters	Example 1				Example 2			
	Mean	SD	HPD 95%		Mean	SD	HPD 95%	
BA								
β	3.696	0.321	3.095	4.357	3.638	0.118	3.410	3.858
η	7.890	0.522	6.906	8.912	7.861	0.074	7.733	8.007
$\mathrm{E}(Y)$	7.118	0.439	6.282	7.971	7.087	0.064	6.982	7.216
Parameters	Example 1				Example 2			
	MLE	SE	CI 95%		MLE	SE	CI 95%	
EM-ML								
β	3.728	0.377	2.989	4.466	3.641	0.089	3.467	3.815
η	7.777	0.585	6.631	8.924	7.860	0.053	7.757	7.964
$\mathrm{E}(Y)$	7.022	0.487	6.067	7.976	7.087	0.048	6.993	7.182
Z-ML								
β	3.729	0.323	3.096	4.361	–	–	–	–
η	7.776	0.488	6.820	8.732	–	–	–	–
$\mathrm{E}(Y)$	7.021	0.409	6.218	7.823	–	–	–	–

There are no Z-ML estimates in Example 2 because it could not be computed due to the high number of components and failures

SD means standard deviation; SE means standard error; HPD means highest posterior density; CI means confidence interval. The true parameters values are: $\beta = 3.924$, $\eta = 7.734$ and $\mathrm{E}(Y) = 7$



(a) Example 1 (b) Example 2

Fig. 4 Component reliability function estimation through the Bayesian approach (BA), the EM maximum likelihood method (EM-ML) and the maximum likelihood approach from *Zhang et al.* (Z-ML) for two scenarios of simulated examples, besides the generating curve (true). There is no Z-ML curve in Example 2 because it could not be computed due to the high number of components and failures

Table 3 MAE values obtained by the Bayesian approach (BA), the EM maximum likelihood method (EM-ML) and the maximum likelihood approach from *Zhang et al.* (Z-ML) of two simulated examples

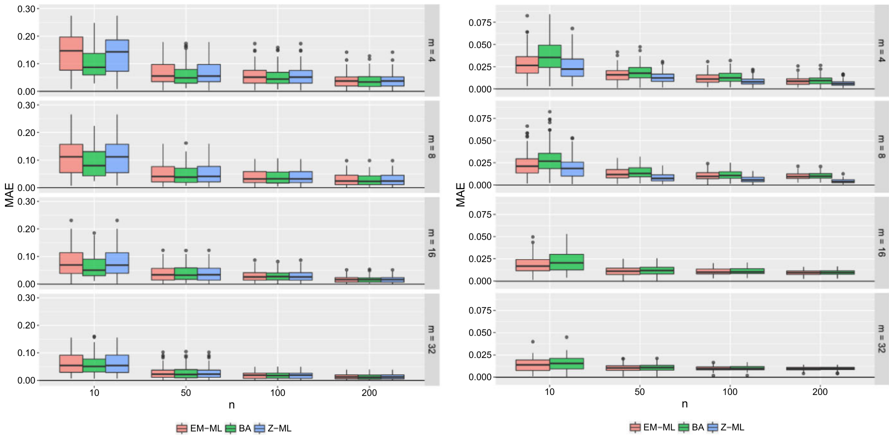
	BA	EM-ML	Z-ML
Example 1	0.0117	0.0057	0.0057
Example 2	0.0080	0.0079	–

There are no MAE values for Z-ML in Example 2 because they could not be computed due to the high number of components and failures

estimates due to the high number of failures and components. The computational time of each scenario was greater than four days and encountered errors in estimation.

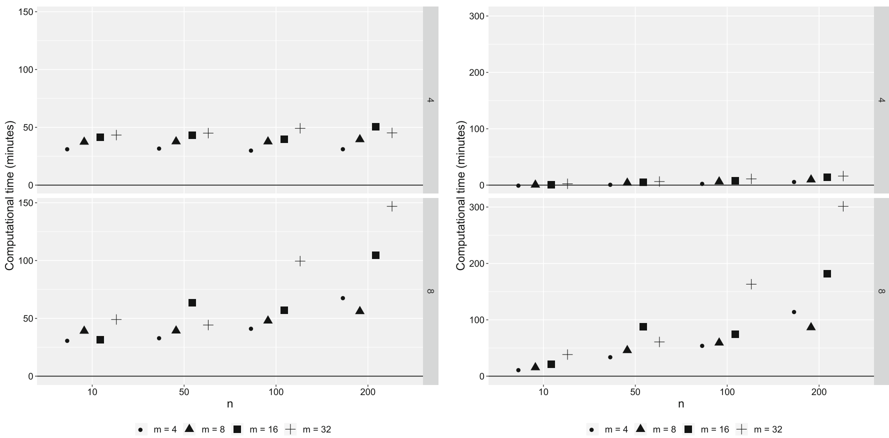
On the other hand, the computational times and availability of EM-ML and BA are not influenced that much (as the Z-ML method does) by the numbers of failures and components. In Fig. 6 are presented the average of the computational times to get the BA and EM-ML estimates over the 100 generated datasets. As one can see, for $m_c = 8$ and mainly when the sample size and the number of the components are high, the proposed methods take more computational time on average, but this is not impractical and it is possible to get estimates by the proposed methods. This way, the proposed methods are able to solve the problem for all cases regardless the scenario.

In short, in settings as those from Fig. 5b, Z-ML fails to compute the components' failure time distribution, whereas the two proposed methods find solutions. For the settings in which Z-ML finds solutions, the proposed methods also find solutions and present similar performance.



(a) $m_c = 4$ (b) $m_c = 8$

Fig. 5 Boxplot graphs of the 100 MAE values of the Bayesian approach (BA), the EM maximum likelihood method (EM-ML) and the maximum likelihood approach from *Zhang et al.* (Z-ML) in scenarios with different sample sizes (n) and number of components (m). There are no Z-ML MAE boxplots in case of $m \in \{16, 32\}$ and $m_c = 8$ because they could not be computed due to the high number of components and failures



(a) BA (b) EM-ML

Fig. 6 Average of the computational time of the Bayesian approach (BA) and the EM maximum likelihood method (EM-ML) in scenarios with different sample sizes (n), number of components (m) and for $m_c = 4$ and $m_c = 8$

6 Cylinder dataset analysis

In order to present the applicability of the proposed methodologies, a real cylinder dataset presented by Zhang et al. (2017) is also considered here. More details about this dataset can be found at Meeker and Escobar (2014) and Nelson and Doganaksoy (1989). A fleet of $n = 120$ diesel engines (systems) is observed. Each engine has 16

Table 4 Distributon of number of failures (r) of 120 systems from cylinder dataset

r	Number of systems	%
0	46	38.3
1	32	26.7
2	18	15.0
3	14	11.7
4	5	4.2
5	4	3.3
6	1	0.8
Total	120	100.0

Table 5 Selection criteria under frequentist approach obtained by the fitted models for cylinder dataset

Model	l	AIC	AICc	BIC	HQIC	CAIC
Weibull	− 677.81	1359.62	1359.72	1365.20	1361.89	1367.20
gamma	− 673.86	1351.73	1351.83	1357.31	1353.99	1359.31
lognormal	− 671.15	1346.30	1346.41	1351.88	1348.67	1353.88
log-logistic	− 677.00	1357.99	1358.10	1363.57	1360.26	1365.57
exponential	− 779.76	1561.53	1561.56	1564.31	1562.66	1565.31

identical cylinders working in series, that is, the first cylinder to fail causes the engine failure. When a cylinder fails, it is replaced by an identical functioning one in the socket (cylinder position), but the information about which socket each replacement comes from was not recorded. So, for each diesel engine, it is only observed the number of failures (r) and the failure time of each failure. Table 4 presents the distribution of the number of failures across all 120 systems.

We fitted models assuming the following distributions for components' failure times: Weibull, gamma, lognormal, log-logistic and exponential. All models were fit considering the R package `srplv`, which was built by authors and it is available for the community. Under the frequentist approach, the lognormal model presents the lowest value for all selection criteria (Table 5) and as a consequence, it is the selected model.

Under the Bayesian paradigm, for each model, we run the Metropolis within Gibbs sampler, discarding the first 20,000 as burn-in samples and using a jump of size 20 to avoid correlation problems, obtaining a sample size of 1,000. We evaluated the convergence of the chain by multiple runs of the algorithm from different starting values and the chains' convergence was monitored through graphical analysis, and good convergence results were obtained. Further, we considered the Gelman–Rubin convergence diagnostic statistics. The measures are close to 1 for all parameters in all fitted models, as shown in Table 6, which suggests that convergence chains have been reached.

The LPML values are presented in Table 6 and the lognormal model is the chosen one once it presents the largest LPML value.

Table 6 Gelman–Rubin statistics and LPML measures obtained by the fitted models for cylinder dataset

Model	Gelman–Rubin statistics	LPML
Weibull	1.0014–1.0024	– 687.05
gamma	1.0032–1.0048	– 680.55
lognormal	1.0020–1.0021	– 676.52
log-logistic	1.0030–1.0033	– 687.08
exponential	0.9997–0.9999	– 781.00

Table 7 Lognormal model parameters (μ_l, σ_l) and expected components' time to failure ($\mathrm{E}(Y)$) estimation based on the Bayesian approach (BA) and the EM Maximum Likelihood method (EM-ML) of cylinder dataset

Parameters	Posterior mean	Posterior SD	HPD 95%	
BA				
μ_l	2.2494	0.0597	2.1361	2.3677
σ_l	0.5464	0.0369	0.4749	0.6208
$\mathrm{E}(Y)$	11.0519	0.8887	9.5231	12.9137
Parameters	MLE	SE	CI 95%	
EM-ML				
μ_l	2.2443	0.0952	2.0577	2.4309
σ_l	0.5433	0.0554	0.4346	0.6520
$\mathrm{E}(Y)$	10.9345	1.3673	8.2546	13.6144

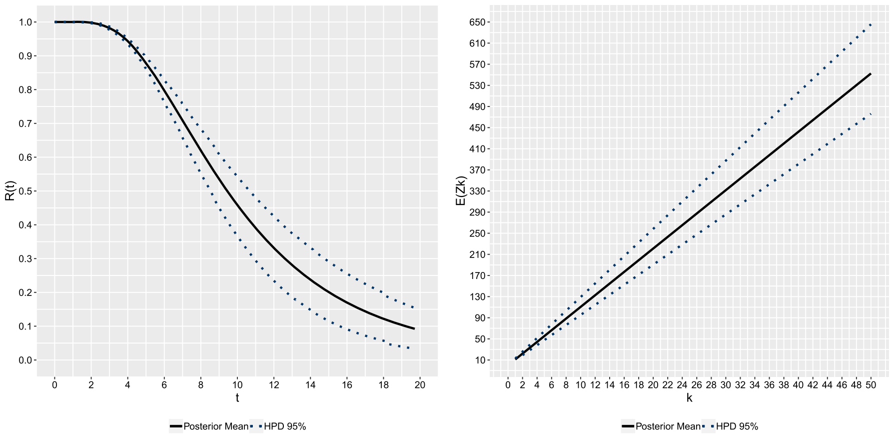
SD means standard deviation; SE means standard error; HPD means highest posterior density and CI means confidence interval

Table 7 lists the posterior mean obtained by BA and EM-ML estimates for the parameters of μ_l (mean of logarithm), σ_l (standard deviation of logarithm) and expected time of components' lifetime, $\mathrm{E}(Y) = \exp\left\{\mu_l + \sigma_l^2/2\right\}$. The expected times of the component lifetime obtained by BA and EM-ML are 11.05 and 10.93 years, respectively. In general, the BA and EM-ML estimates are close for all parameters, as expected.

The posterior mean and the 95% highest posterior density (HPD) point-wise band of the component reliability function are illustrated in Fig. 7a. Besides, the posterior mean and the 95% highest posterior density (HPD) point-wise band of $\mathrm{E}(Z_k)$, for $k = \{0, 1, \ldots, 49, 50\}$, are presented in Fig. 7b. The estimation for the reliability function obtained by EM-ML estimator is similar to the estimate obtained by the Bayesian approach.

7 Conclusion

A Bayesian model and a maximum likelihood estimator (MLE) were proposed in order to estimate identical components failure time distribution involved in a repairable series system with masked cause of failure. For both approaches, latent variables were con-



(a) Component reliability function (b) $E(Z_k)$ estimation

Fig. 7 Component reliability function and expected time of occurence of the k-th failure in the socket estimates through Bayesian approach of cylinder dataset

sidered in the estimation process through EM algorithm for MLE and Markov-Chain Monte-Carlo (MCMC) for the Bayesian approach. The proposed models are generic and straightforward for any probability distribution on positive support. In estimation processes, satisfactory results about the convergence of the MCMC's chains and EM algorithm were obtained, evaluated through graphical analysis and convergence performance measures.

Simulation studies were realized in scenarios with different sample sizes, number of components and distributions for censor lifetime. The mean absolute error (MAE) from each estimator to the true distribution was considered as performance measure. In situations of high numbers of failures and/or components, it was not possible to compute the maximum likelihood estimator proposed by Zhang et al. (2017) (Z-ML) through the package `SRPML`. In contrast to this well-established approach by Zhang et al. (2017), our proposed methods are not affected that much by the high numbers of failures and/or components. Instead they work perfectly even in these situations. Besides, in settings in which Z-ML finds solutions, the proposed methods also find a solution and achieve a similar performance. Thus, the huge advantage of our proposed methods is that they estimate the components' failure time distribution regardless of the number of failures and components.

The practical applicability was assessed in cylinder dataset, in which components' failure time quantities were estimated convincingly. The R package `srplv` was built by authors in order to make the methodology available for applications. The package is available in *github.com/agathasr/srplv*.

In this work, the assumption of independent and identically distributed (i.i.d.) components failure times has been made and found to be suitable for the cylinder dataset characteristics. However, this assumption might not be applicable to other scenarios. Thus, in future works, our proposed method can be extended to situations in which the assumption of independent and identically distributed failure times is violated.

Moreover, within future works we will also investigate the suitability of our approach for the assessment of system reliability rather than cylinder reliability, which has been the focus of this work.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00180-021-01124-0.

Acknowledgements This work was partially supported by the Brazilian agency CNPq: Grant 308776/2014-3. The agency had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. This study was financed in part by CAPES (Brazil) - Finance Code 001 and Federal University of Mato Grosso do Sul. Pascal Kerschke, Heike Trautmann, Bernd Hellingrath and Carolin Wagner acknowledge support by the European Research Center for Information Systems (ERCIS).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Appendix

We can write the logarithm of the complete likelihood function of i-th system if Weibull distribution with parameter β (shape) and η (scale) is assumed, as

$$
\begin{aligned}
l_i(\boldsymbol{\theta} \mid \boldsymbol{t}_i, \boldsymbol{d}_i) &= \Big[1 - \mathrm{I}(v_i = 0)\Big]\Bigg[\sum_{l=1}^{v_i}\sum_{k=1}^{n_l} \log f(x_{ilk} - x_{il(k-1)}) \\
&+ \sum_{l=1}^{v_i} \log R(\tau_i - x_{iln_l})\Bigg] + (m - v_i)\log R(\tau_i) \\
&= \Big[1 - \mathrm{I}(v_i = 0)\Big]\sum_{l=1}^{v_i}\sum_{k=1}^{n_l}\Bigg\{\log(\beta) - \log(\eta) + (\beta - 1)\Big[\log(x_{ilk} - x_{il(k-1)}) \\
&- \log(\eta)\Big] - \left(\frac{x_{ilk} - x_{il(k-1)}}{\eta}\right)^{\beta}\Bigg\} \\
&- \Big[1 - \mathrm{I}(v_i = 0)\Big]\sum_{l=1}^{v_i}\left(\frac{\tau_i - x_{iln_l}}{\eta}\right)^{\beta} - (m - v_i)\left(\frac{\tau_i}{\eta}\right)^{\beta} \\
&= \Big[1 - \mathrm{I}(v_i = 0)\Big]\Bigg\{r_i \log(\beta) - r_i \log(\eta) + (\beta - 1)\sum_{l=1}^{v_i}\sum_{k=1}^{n_l}\log(x_{ilk} - x_{il(k-1)}) \\
&- r_i(\beta - 1)\log(\eta) - \sum_{l=1}^{v_i}\Bigg[\sum_{k=1}^{n_l}\left(\frac{x_{ilk} - x_{il(k-1)}}{\eta}\right)^{\beta} + \left(\frac{\tau_i - x_{iln_l}}{\eta}\right)^{\beta}\Bigg]\Bigg\} \\
&- (m - v_i)\left(\frac{\tau_i}{\eta}\right)^{\beta}.
\end{aligned}
$$

The first derivatives 0f $l_i(\boldsymbol{\theta} \mid \boldsymbol{t}_i, \boldsymbol{d}_i)$ in relation to β and η, respectively, are

$$\frac{\mathrm{d}l_i(\boldsymbol{\theta} \mid \boldsymbol{t}_i, \boldsymbol{d}_i)}{\mathrm{d}\beta} = \Big[1 - \mathrm{I}(v_i = 0)\Big]\Bigg\{\frac{r_i}{\beta} + \sum_{l=1}^{v_i}\sum_{k=1}^{n_l}\log(x_{ilk} - x_{il(k-1)}) - r_i\log(\eta)$$
$$+\log(\eta)\left(\frac{1}{\eta}\right)^{\beta}\left[\sum_{l=1}^{v_i}\left(\sum_{k=1}^{n_l}(x_{ilk} - x_{il(k-1)})^{\beta} + (\tau_i - x_{iln_l})^{\beta}\right)\right]$$
$$-\left(\frac{1}{\eta}\right)^{\beta}\Bigg[\sum_{l=1}^{v_i}\sum_{k=1}^{n_l}\log(x_{ilk} - x_{il(k-1)})(x_{ilk} - x_{il(k-1)})^{\beta}$$
$$+\sum_{l=1}^{v_i}\log(\tau_i - x_{iln_l})(\tau_i - x_{iln_l})^{\beta}\Bigg]\Bigg\} + \left(\frac{1}{\eta}\right)^{\beta}(m - v_i)\tau_i^{\beta}[\log(\eta) - \log(\tau_i)],$$

and

$$\frac{\mathrm{d}l_i(\boldsymbol{\theta} \mid \boldsymbol{t}_i, \boldsymbol{d}_i)}{\mathrm{d}\eta} = \Big[1 - \mathrm{I}(v_i = 0)\Big]\Bigg\{-\frac{r_i}{\eta} - \frac{r_i(\beta - 1)}{\eta} + \beta\left(\frac{1}{\eta}\right)^{\beta+1}$$
$$\Bigg[\sum_{l=1}^{v_i}\sum_{k=1}^{n_l}(x_{ilk} - x_{il(k-1)})^{\beta} + \sum_{l=1}^{v_i}(\tau_i - x_{iln_l})^{\beta}\Bigg]\Bigg\} + \beta\left(\frac{1}{\eta}\right)^{\beta+1}(m - v_i)\tau_i^{\beta}.$$

The second derivatives are

$$\frac{\mathrm{d}^2 l_i(\boldsymbol{\theta} \mid \boldsymbol{t}_i, \boldsymbol{d}_i)}{\mathrm{d}\beta^2} = \Big[1 - \mathrm{I}(v_i = 0)\Big]\Bigg\{-\frac{r_i}{\beta^2} - [\log\eta]^2\left(\frac{1}{\eta}\right)^{\beta}$$
$$\times\left[\sum_{l=1}^{v_i}\left(\sum_{k=1}^{n_l}(x_{ilk} - x_{il(k-1)})^{\beta} + (\tau_i - x_{iln_l})^{\beta}\right)\right] + 2\log(\eta)\left(\frac{1}{\eta}\right)^{\beta}$$
$$\times\Bigg[\sum_{l=1}^{v_i}\sum_{k=1}^{n_l}\log(x_{ilk} - x_{il(k-1)})(x_{ilk} - x_{il(k-1)})^{\beta} + \sum_{l=1}^{v_i}\log(\tau_i - x_{iln_l})(\tau_i - x_{iln_l})^{\beta}\Bigg]$$
$$-\left(\frac{1}{\eta}\right)^{\beta}\Bigg[\sum_{l=1}^{v_i}\sum_{k=1}^{n_l}[\log(x_{ilk} - x_{il(k-1)})]^2(x_{ilk} - x_{il(k-1)})^{\beta}$$
$$+\sum_{l=1}^{v_i}[\log(\tau_i - x_{iln_l})]^2(\tau_i - x_{iln_l})^{\beta}\Bigg]\Bigg\}$$
$$+(m - v_i)\left(\frac{1}{\eta}\right)^{\beta}\tau_i^{\beta}\Big[-[\log(\tau_i)]^2 + 2\log(\tau_i)\log(\eta) - [\log(\eta)]^2\Big],$$

$$\frac{\mathrm{d}^2 l_i(\boldsymbol{\theta} \mid \boldsymbol{t}_i, \boldsymbol{d}_i)}{\mathrm{d}\beta\ \mathrm{d}\eta} = \Big[1 - \mathrm{I}(v_i = 0)\Big]\Bigg\{ -\frac{r_i}{\eta} + \Bigg[\left(\frac{1}{\eta}\right)^{\beta+1}(1 - \beta \log(\eta))\Bigg]$$
$$\times\Bigg[\sum_{l=1}^{v_i}\Bigg(\sum_{k=1}^{n_l}(x_{ilk} - x_{il(k-1)})^{\beta} + (\tau_i - x_{iln_l})^{\beta}\Bigg)\Bigg]$$
$$+\beta\left(\frac{1}{\eta}\right)^{\beta+1}\Bigg[\sum_{l=1}^{v_i}\sum_{k=1}^{n_l}\log(x_{ilk} - x_{il(k-1)})(x_{ilk} - x_{il(k-1)})^{\beta}$$
$$+\sum_{l=1}^{v_i}\log(\tau_i - x_{iln_l})(\tau_i - x_{iln_l})^{\beta}\Bigg]\Bigg\}$$
$$+(m - v_i)\left(\frac{1}{\eta}\right)^{\beta+1}\tau_i^{\beta}[1 - \beta\log(\eta) + \beta\log(\tau_i)],$$

and

$$\frac{\mathrm{d}^2 l_i(\boldsymbol{\theta} \mid \boldsymbol{t}_i, \boldsymbol{d}_i)}{\mathrm{d}\eta^2} = \Big[1 - \mathrm{I}(v_i = 0)\Big]\Bigg\{\frac{\beta r_i}{\eta^2} - \beta(\beta+1)\left(\frac{1}{\eta}\right)^{\beta+2}$$
$$\Bigg[\sum_{l=1}^{v_i}\Bigg(\sum_{k=1}^{n_l}(x_{ilk} - x_{il(k-1)})^{\beta} + (\tau_i - x_{iln_l})^{\beta}\Bigg)\Bigg]\Bigg\} - \beta(\beta+1)\left(\frac{1}{\eta}\right)^{\beta+2}(m - v_i)\tau_i^{\beta}.$$

Thus,

$$I = -\frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^{\top}}Q(\boldsymbol{\theta} \mid \widehat{\boldsymbol{\theta}}) = -\frac{1}{L}\sum_{i=1}^{n}\sum_{l=1}^{L}\frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^{\top}}l_i\Big(\boldsymbol{\theta} \mid \boldsymbol{t}_i, \boldsymbol{d}_i^{(l)}\Big)\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}$$
$$= \begin{bmatrix} -\frac{1}{L}\sum_{i=1}^{n}\sum_{l=1}^{L}\frac{\mathrm{d}^2 l_i(\boldsymbol{\theta}|\boldsymbol{t}_i,\boldsymbol{d}_i^{(l)})}{\mathrm{d}\eta^2}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} & -\frac{1}{L}\sum_{i=1}^{n}\sum_{l=1}^{L}\frac{\mathrm{d}^2 l_i(\boldsymbol{\theta}|\boldsymbol{t}_i,\boldsymbol{d}_i^{(l)})}{\mathrm{d}\eta\ \mathrm{d}\beta}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} \\ -\frac{1}{L}\sum_{i=1}^{n}\sum_{l=1}^{L}\frac{\mathrm{d}^2 l_i(\boldsymbol{\theta}|\boldsymbol{t}_i,\boldsymbol{d}_i^{(l)})}{\mathrm{d}\beta\ \mathrm{d}\eta}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} & -\frac{1}{L}\sum_{i=1}^{n}\sum_{l=1}^{L}\frac{\mathrm{d}^2 l_i(\boldsymbol{\theta}|\boldsymbol{t}_i,\boldsymbol{d}_i^{(l)})}{\mathrm{d}\beta^2}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} \end{bmatrix},$$

in which $\widehat{\boldsymbol{\theta}} = (\widehat{\eta}, \widehat{\beta})$. Besides,

$$II = \sum_{i=1}^{n}\Bigg\{\frac{1}{L}\sum_{l=1}^{L}\frac{\partial}{\partial\boldsymbol{\theta}}l_i\Big(\boldsymbol{\theta} \mid \boldsymbol{t}_i, \boldsymbol{d}_i^{(l)}\Big)\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}\Bigg\}\Bigg\{\frac{1}{L}\sum_{l=1}^{L}\frac{\partial}{\partial\boldsymbol{\theta}}l_i\Big(\boldsymbol{\theta} \mid \boldsymbol{t}_i, \boldsymbol{d}_i^{(l)}\Big)\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}\Bigg\}^{\top}$$
$$= \sum_{i=1}^{n}\Bigg\{\frac{1}{L}\sum_{l=1}^{L}\Bigg(\frac{\mathrm{d}l_i(\boldsymbol{\theta} \mid \boldsymbol{t}_i, \boldsymbol{d}_i^{(l)})}{\mathrm{d}\eta}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}, \frac{\mathrm{d}l_i(\boldsymbol{\theta} \mid \boldsymbol{t}_i, \boldsymbol{d}_i^{(l)})}{\mathrm{d}\beta}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}\Bigg)^{\top}\Bigg\}$$
$$\times\Bigg\{\frac{1}{L}\sum_{l=1}^{L}\Bigg(\frac{\mathrm{d}l_i(\boldsymbol{\theta} \mid \boldsymbol{t}_i, \boldsymbol{d}_i^{(l)})}{\mathrm{d}\eta}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}, \frac{\mathrm{d}l_i(\boldsymbol{\theta} \mid \boldsymbol{t}_i, \boldsymbol{d}_i^{(l)})}{\mathrm{d}\beta}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}\Bigg)^{\top}\Bigg\}^{\top}$$

and

$$
\begin{aligned}
III &= -\frac{1}{L}\sum_{i=1}^{n}\sum_{l=1}^{L}\left\{\frac{\partial}{\partial \boldsymbol{\theta}} l_i\left(\boldsymbol{\theta} \mid \boldsymbol{t}_i, \boldsymbol{d}_i^{(l)}\right)\right\}\left\{\frac{\partial}{\partial \boldsymbol{\theta}} l_i\left(\boldsymbol{\theta} \mid \boldsymbol{t}_i, \boldsymbol{d}_i^{(l)}\right)\right\}^{\top}\Bigg|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} \\
&= -\frac{1}{L}\sum_{i=1}^{n}\sum_{l=1}^{L}\left\{\left(\frac{\mathrm{d} l_i(\boldsymbol{\theta} \mid \boldsymbol{t}_i, \boldsymbol{d}_i^{(l)})}{\mathrm{d}\eta}, \frac{\mathrm{d} l_i(\boldsymbol{\theta} \mid \boldsymbol{t}_i, \boldsymbol{d}_i^{(l)})}{\mathrm{d}\beta}\right)^{\top}\right\} \\
&\quad \left\{\left(\frac{\mathrm{d} l_i(\boldsymbol{\theta} \mid \boldsymbol{t}_i, \boldsymbol{d}_i^{(l)})}{\mathrm{d}\eta}, \frac{\mathrm{d} l_i(\boldsymbol{\theta} \mid \boldsymbol{t}_i, \boldsymbol{d}_i^{(l)})}{\mathrm{d}\beta}\right)^{\top}\right\}^{\top}\Bigg|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}.
\end{aligned}
$$

The quantity $I_{\boldsymbol{\theta}}(\widehat{\boldsymbol{\theta}})$ can be estimated by $I + II + III$.

References

Casella G, Berger RL (2002) Statistical inference, vol 2. Duxbury, Pacific Grove
Crow LH (1990) Evaluating the reliability of repairable systems. In: Annual proceedings on reliability and maintainability symposium, pp 275–279. IEEE
Crowder MJ, Kimber A, Sweeting T, Smith R (1994) Statistical analysis of reliability data, vol 27. CRC Press, Boca Raton
Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. J Roy Stat Soc Ser B Methodol 39(1):1–22
Dewanji A, Kundu S, Nayak TK (2012) Nonparametric estimation of the number of components of a superposition of renewal processes. J Stat Plan Inference 142(9):2710–2718
Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. Stat Sci 7:457–472
Gilks WR, Richardson S, Spiegelhalter DJ (1995) Introducing Markov chain Monte Carlo. In: Markov chain Monte Carlo in practice, p 1
Krivtsov V, Frankstein M, Yevkin O (2017) On reliability of a multi-socket repairable system. Qual Reliab Eng Int 33(5):1011–1017
Liu B, Shi Y, Cai J, Bai X, Zhang C (2017) Nonparametric Bayesian analysis for masked data from hybrid systems in accelerated lifetime tests. IEEE Trans Reliab 66(3):662–676
Louis TA (1982) Finding the observed information matrix when using the EM algorithm. J Roy Stat Soc Ser B Methodol 44(2):226–233
Meeker WQ, Escobar LA (2014) Statistical methods for reliability data. Wiley, Hoboken
Miyakawa M (1984) Analysis of incomplete data in competing risks model. IEEE Trans Reliab 33(4):293–296
Nelder JA, Mead R (1965) A simplex method for function minimization. Comput J 7(4):308–313
Nelson W, Doganaksoy N (1989) A computer program for an estimate and confidence limits for the mean cumulative function for cost or number of repairs of repairable products. GE Research & Development Center, Niskayuna
Nelson WB (2003) Recurrent events data analysis for product repairs, disease recurrences, and other applications. SIAM, Seattle
Peixoto J (2009) Estimating lifetimes when several unidentified components are reported. In: Proceedings of the physical and engineering sciences section of the American statistical association, pp 5078–5092
R Core Team (2020) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/
Rinne H (2008) The Weibull distribution: a handbook. CRC Press, Boca Raton
Robert CP, Casella G, Casella G (2010) Introducing Monte Carlo methods with R, vol 18. Springer, Berlin
Rodrigues AS, Bhering F, Carlos ADB, Polpo A (2018) Bayesian estimation of component reliability in coherent systems. IEEE Access 6:18520–18535

Rodrigues AS, Pereira CADB, Polpo A (2019) Estimation of component reliability in coherent systems with masked data. IEEE Access 7:57476–57487
Song S, Xie M (2018) An integrated method for estimation with superimposed failure data. In: 2018 IEEE international conference on prognostics and health management (ICPHM), pp 1–5. IEEE
Tierney L (1994) Markov chains for exploring posterior distributions. Ann Stat 22:1701–1728
Wang R, Sha N, Gu B, Xu X (2015) Parameter inference in a hybrid system with masked data. IEEE Trans Reliab 64(2):636–644
Zhang W, Tian Y, Escobar L, Meeker W (2015) SRPML: estimating a parametric component lifetime distribution from SRP data. R package version 1
Zhang W, Tian Y, Escobar LA, Meeker WQ (2017) Estimating a parametric component lifetime distribution from a collection of superimposed renewal processes. Technometrics 59(2):202–214