IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received March 20, 2019, accepted April 11, 2019, date of publication April 29, 2019, date of current version May 14, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2913675

Estimation of Component Reliability in Coherent Systems With Masked Data

AGATHA S. RODRIGUES[1,2], CARLOS ALBERTO DE B. PEREIRA[1,3], AND ADRIANO POLPO[4], (Member, IEEE)

[1]Institute of Mathematics and Statistics, University of Sao Paulo, Sao Paulo 05508-060, Brazil
[2]Department of Obstetrics and Gynecology, Sao Paulo University Medical School, Sao Paulo 01246-903, Brazil
[3]Institute of Mathematics, Federal University of Mato Grosso do Sul, Campo Grande 79070-900, Brazil
[4]Department of Mathematics and Statistics, The University of Western Australia, Perth, WA 6009, Australia

Corresponding author: Agatha S. Rodrigues (agatha.srodrigues@gmail.com)

This work was supported in part by the Brazilian agency CNPq under Grant 308776/2014-3, in part by the CAPES, Brazil, under Grant 001, and in part by the Federal University of Mato Grosso do Sul.

ABSTRACT The reliability of a coherent system of components depends on the reliability of each component and the initial statistical work should be an estimation of the reliability of each component. This paper represents a challenging task because if the system fails, the failure time of a given component cannot be observed, that is, the phenomenon of censored data occurs. A solution for the reliability estimation of components exists when the system failure time and the status of each component are available at the time of system failure. However, it may be difficult to identify the status of the components at the moment of system failure. Such cases represent systems with masked causes of failure. Since parallel and series systems are the simplest systems, numerous solutions have been reported in the literature. To the best of our knowledge, this paper is the first to present the general case of coherent systems without the restriction of an identically distributed lifetime. The three-parameter Weibull Bayesian model is proposed. The Gibbs with the Metropolis–Hasting algorithm supports the statistical work of obtaining the posterior distribution quantities. With several simulations, the excellent performance of the model is evaluated. A real dataset of computer hard drives is analyzed to show the practical relevance of the proposed model.

INDEX TERMS Bayesian three-parameter Weibull model, coherent system, component reliability, masked data, metropolis within Gibbs algorithm.

ACRONYMS

ALT	accelerated life tests
BSNP	Bhattacharya-Samaniego nonparametric
CI	credibility interval
HPD	highest posterior density
MAE	mean absolute error
MCMC	Markov-Chain Monte Carlo
MLE	maximum likelihood estimator
PSS	parallel-series system
SPS	series-parallel system
SD	standard deviation
W3PM	Weibull 3-parameter model

NOTATION

β_j	shape parameter of 3-parameter Weibull.
$\boldsymbol{\delta}_j = (\delta_{1j}, \delta_{2j}, \delta_{3j})$	be an indicator vector of the censor, in which $\sum_{l=1}^{3} \delta_{lj} = 1$.
$\boldsymbol{d}_{ji} = (d_{1ji}, d_{2ji}, d_{3ji})$	latent variable vector, in which $\sum_{l=1}^{3} d_{lji} = 1$.
η_j	scale parameter of 3-parameter Weibull.
$f(\cdot \mid \boldsymbol{\theta}_j)$	density function of j-th component failure time distribution.
$F(\cdot \mid \boldsymbol{\theta}_j)$	distribution function of j-th component failure time distribution.
$h(\cdot)$	function that relates the system failure time to the components' functioning.
$\lambda_{1j}(t)$	probability that the j-th component is masked conditional to t and $\delta_{1j} = 1$.
$\lambda_{2j}(t)$	probability that the j-th component is masked conditional to t and $\delta_{2j} = 1$.
$\lambda_{3j}(t)$	probability that the j-th component is masked conditional to t and $\delta_{3j} = 1$.

The associate editor coordinating the review of this manuscript and approving it for publication was Xuerong Ye.

m	number of components in the system.
MAE	$\frac{1}{l}\sum_{\ell=1}^{l} \mid \widehat{R}(g_\ell) - R(g_\ell) \mid$.
μ_j	location parameter of 3-parameter Weibull.
n	number of systems in a sample.
n_p	number of posterior samples simulated.
p	proportion of the masked data system.
$R(\cdot \mid \boldsymbol{\theta}_j)$	reliability function of j-th component failure time distribution.
T	system failure time.
t_i	observed failure time of i-th system.
$\boldsymbol{\theta}_j$	parameter vector of Weibull 3-parameter $= (\beta_j, \eta_j, \mu_j)$ distribution for j-th component.
X_{ji}	j-th component failure time for i-th system.
υ_{ji}	indicator variable if the j-th component has a masked failure time for i-th system.
Υ_i	index set indicating the components candidates to cause the failure of the i-th system.

I. INTRODUCTION

An important step in statistical reliability studies investigating coherent systems is the estimation of the reliability of each system component. In general, the lifetime test can be conducted at the system level but not at the component level. Therefore, drawing statistical inferences regarding component reliability is a challenging task as follows: when a system fails, the failure time of a given component cannot be observed, i.e., censored data. Considering a random sample of a system with m components in which all n sample units (systems) are observed up to failure, each sampling unit produces a component failure time and a censored failure time for the remaining $m - 1$ components, although different types of censoring may occur. A specific component that has not failed at time t is either right-censored, in which case the component could continue to work after t, or is left-censored if it has failed before t. Depending on the system structure (the way the components are interconnected) and the component reliability, a high amount of component censored data is very common and is occasionally greater than 80%.

In some situations, the available information is the n system failure times and the status of each component at the system failure times (uncensored, right-censored or left-censored). Approaches for component reliability estimation in this situation have been proposed in the literature [1]–[9].

However, in certain situations, identifying the component whose failure leads to system failure (the component whose failure produced the system failure) or the status of the components at the moment of system failure is difficult. Such cases are known as masked cause of failure and are usually due to limited resources for failure diagnosis. For example, consider the reliability estimation of a series system with the following three components in the computer hard drive: electronic hard, head fly ability and disc magnetic [10]. The first component to fail causes the failure of the system (computer hard drive), and a small subset of components is identified as the possible cause of failure. In an attempt to repair the system as quickly as possible, the entire subset of components is replaced, and the component that leads to system failure cannot be identified.

The literature on the reliability of either parallel or series systems with masked cause failure is abundant, and different solutions have been presented. [11] studied reliability estimation for a series system with two components using a maximum likelihood estimator (MLE) in closed form and non-parametric estimates based on a Kaplan-Meier estimator. Reference [12] extended Miyakawa's results to a series system with three components. References [13]–[18] presented parametric models of a series system using MLE and Bayesian approaches. Reference [19] proposed an estimator of the reliability functions of the components in a parallel system. All these works assumed the symmetry assumption, i.e., the probability of a system to have masked cause of failure (masking probability) is the same regardless of which component causes the system failure. References [20], [21] presented an approach relaxing the symmetry assumption. Reference [22] considered the masking probability in the likelihood function construction, and in addition to depending on the cause of system failure [23] considered that the masking probability is a decreasing function of the system failure time.

Accelerated life tests (ALT) are usually used to quickly obtain information regarding the component reliability. In these tests, the components are subjected to high-stress levels to decrease the elapse time of failure occurrence. This resulting information is used to examine system failures under normal stress levels. To achieve this reversal, standard patterns of relationship between failures at different levels of stress are used. Using ALT, [24]–[26] proposed Bayesian estimators for masked series systems. For hybrid systems (systems with components in series and parallel), [27] considered MLE to draw inferences about component reliability.

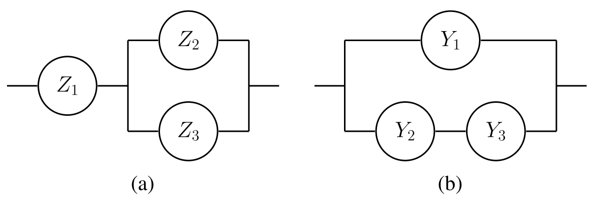


FIGURE 1. (a) Series-parallel systems (SPS) and (b) parallel-series systems (PSS) with three components.

For components involved in series-parallel systems (SPS) and in parallel-series systems (PSS) with three components (figures 1a and 1b), [28] proposed Bayesian nonparametric estimators for the reliability functions with masked data using ALT. These authors assumed that the components involved in SPS and PSS representations have mutually independent lifetimes and that the distribution of the components' lifetimes must have disjoint sets of jump points.

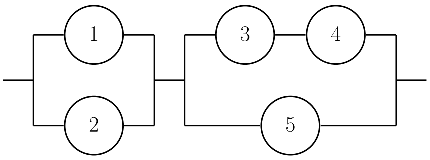


FIGURE 2. Complex system with 5 components.

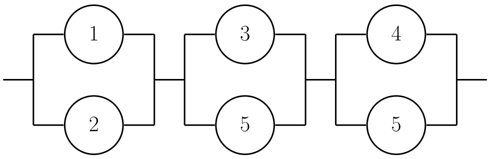


FIGURE 3. SPS representation of system in Figure 2.

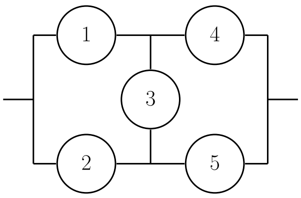


FIGURE 4. Bridge structure.

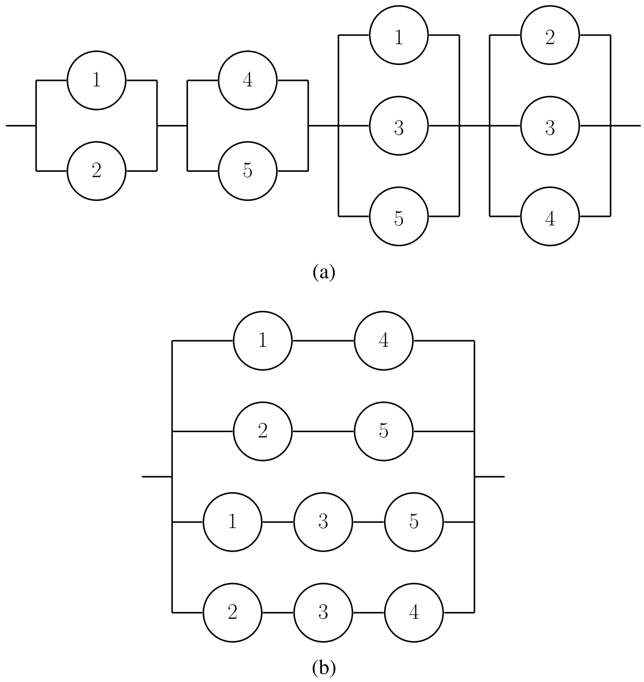


FIGURE 5. (a) SPS and (b) PSS representations of bridge system.

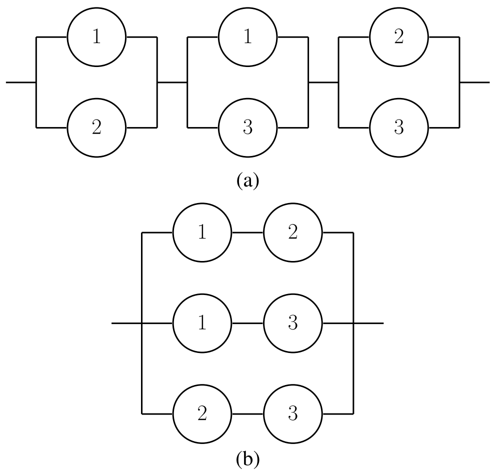


FIGURE 6. (a) SPS and (b) PSS representations of 2-out-of-3.

Their method can be applied to some components in a coherent system, once it is known that every coherent system can be written as SPS and PSS representations [29]. The authors discussed the estimation of component reliability in a coherent system in Figure 2 and estimated the reliability of components $j = 1, 2$ and 5 by representing the system as an SPS representation (Figure 1a). Let X_j be the lifetime of the j-th component ($j = 1, \ldots, 5$) in Figure 2 and Z_ℓ is the lifetime of the ℓ-th component in an SPS representation, $\ell = 1, 2, 3$. Considering component $j = 5$, the authors built a simplified system by considering $Z_1 = \max\{X_1, X_2\}$, $Z_2 = \min\{X_3, X_4\}$ and $Z_3 = X_5$. If the interest is the estimation of the reliability function of components $j = 3$ or $j = 4$, the complex system can be represented as Figure 3. If $Z_1 = \min\{\max\{X_1, X_2\}, \max\{X_4, X_5\}\}$ or $Z_1 = \min\{\max\{X_1, X_2\}, \max\{X_3, X_5\}\}$, $Z_2 = X_3$ or $Z_2 = X_4$ and $Z_3 = X_5$; thus, the coherent system can be simplified as an SPS representation as in Figure 1a. However, their assumption that the components are mutually independent in the representation is violated because of the presence of component $j = 5$. Therefore, the method proposed by [28] cannot be considered for the reliability estimation of components $j = 3$ and $j = 4$.

Other complex coherent systems present the same problem as follows: some components may appear in two or more places in SPS or PSS representations. Figure 4 is a bridge system described in the literature [29] that Figure 5 illustrates SPS and PSS representations. Notably, each of the five components appears twice for both representations. Another interesting structure is the k-out-of-m system (which only works if at least k of the m components work). For instance, Figure 6 presents a simple 2-out-of-3 case of SPS and PSS representations. Notably, each of the three components also appears twice in both representations. These situations violate the assumption proposed by [28], and their estimator is not suitable for the reliability function of components involved in these complex coherent systems.

The nonparametric estimator of the reliability of components involved in coherent systems proposed by [30] can be considered in a scenario with masked data since the only necessary information is the system failure time and system structure; thus, knowledge regarding the cause of failure is not necessary, which is suitable for masked data situations. This scenario occurred because the authors assumed a restrictive assumption in which the component lifetimes are s-independent and identically distributed; therefore, there is only one estimator for all different components, which can be a restrictive and not applicable/realistic assumption.

Other works can be cited by estimating the reliability function of components involved in coherent systems and also restricted to the identical components assumption. Reference [31] proposed nonparametric inference for components in a scenario that system structure is known and may vary between the samples; [32] considered Bayesian inference via MCMC for components distributions and [33] discussed statistical inference of the lifetime distribution of components based on observing the system lifetimes when the system structure is known.

To the best of our knowledge, no work reported in the literature has considered the reliability estimation of components involved in any coherent system with masked data in which identically distributed failure times are not imposed. Therefore, a Bayesian three-parameter Weibull model of component reliability with masked data is proposed. The presented model is general and can be considered for any coherent system, the symmetry assumption is not necessary and the accelerated life tests (ALT) may also be considered. An important highlight of the proposed approach is that the identically distributed components' failure times assumption is not required. The statistical work performed to obtain the quantities of the posterior distribution is supported by the Metropolis within Gibbs algorithm.

We consider a sample of n identical coherent systems with m components. In the observed sample, the system failure times are available for all n units. Let Υ be the index set indicating the possible components that produce the failure of the system, i.e., the components with masked failure times; Υ is a subset of $\{1, \ldots, m\}$. For some sample units, system autopsy is possible and the diagnose the cause of the failure is realized, that is, Υ is unitary, and the statuses of all the components in these cases are observed.

In this work, the structure of the system need not be known, but if it is, it brings additional information to the estimation process. As an example, consider $\Upsilon = \{1, 2, 3\}$ and so, the components 1, 2 and 3 present masked failure time. In a situation that system structure is known to be bridge structure (Figure 4), for example, the failure of component 3 could not lead the system failure and this information can be incorporated into the model.

The performance of the component reliability estimator obtained from the proposed model is compared to the nonparametric estimator considered by [30] in scenarios with different proportions of masked data, different system structures and different distributions of component lifetimes. We also consider a real dataset to present the applicability of the proposed model. The dataset consists of 172 computer hard drives that were monitored over a period of 4 years, and their failure times were observed. However, for some hard drives (38%), the cause of hard drive failure was not identified.

This paper is organized as follows. The proposed model and estimation method are described in Section II. In Section III, we present simulated examples, and simulation studies are presented in Section IV. In Section V, the applicability of the proposed model to computer hard drives problem is presented. Finally, some final remarks and additional comments are provided in Section VI.

II. WEIBULL MODEL AND ESTIMATION PROCEDURE

Consider a system with m components and let index j represent the j-th component. The random variable X_j denotes the failure time of the j-th component, $j = 1, \ldots, m$, and we assume that $X_1, X_2, \ldots, X_m$ are mutually independent. Let T be a random variable that represents the system failure time, and t is an observation of T. Associated with each component j, let $\boldsymbol{\delta}_j = (\delta_{1j}, \delta_{2j}, \delta_{3j})$ be an indicator vector of the censor. The observation of X_j can be as follows: if $X_j = t$, the failure time of X_j is not censored ($\delta_{1j} = 1$); if $X_j > t$, the failure time is right-censored ($\delta_{2j} = 1$); and if $X_j \leq t$, the failure time is left-censored ($\delta_{3j} = 1$), in which $\sum_{l=1}^{3} \delta_{lj} = 1$. Additionally, the j-th component may belong to the masked set.

Let $t_1, \ldots, t_n$ be a sample of system failure times of size n, and Υ_i is the set of masked components in the i-th sample, $i = 1, \ldots, n$. Additionally, $\upsilon_{ji} = 1$ if the j-th component has a masked failure time ($j \in \Upsilon_i$); otherwise, $\upsilon_{ji} = 0$ ($j \notin \Upsilon_i$), $j = 1, \ldots, m$. The observation of j-th component can be one of the following:

$$\begin{aligned} \text{uncensored; not masked:} &\quad \delta_{1ji} = 1 \text{ and } \upsilon_{ji} = 0;\\ \text{right-censored; not masked:} &\quad \delta_{2ji} = 1 \text{ and } \upsilon_{ji} = 0;\\ \text{left-censored; not masked:} &\quad \delta_{3ji} = 1 \text{ and } \upsilon_{ji} = 0; \text{ or}\\ \text{masked:} &\quad \upsilon_{ji} = 1. \end{aligned}$$

If a component has a masked failure time, the component could have led to the system failure (uncensored) or it is right-censored or left-censored observation. Consider

$$\begin{aligned} \lambda_{1j}(t) &= \Pr(\upsilon_j = 1 \mid t, \delta_{1j} = 1),\\ \lambda_{2j}(t) &= \Pr(\upsilon_j = 1 \mid t, \delta_{2j} = 1),\\ \lambda_{3j}(t) &= \Pr(\upsilon_j = 1 \mid t, \delta_{3j} = 1), \end{aligned}$$

where $\lambda_{1j}(t)$ is the conditional probability that the j-th component is masked given the failure time of the system t and the censor type $\delta_{1j} = 1$. $\lambda_{2j}(t)$ and $\lambda_{3j}(t)$ are analogous to $\lambda_{1j}(t)$. Here, we consider $\lambda_{1j}(t) = \lambda_{1j}$, $\lambda_{2j}(t) = \lambda_{2j}$, and $\lambda_{3j}(t) = \lambda_{3j}$, i.e., the probability that a component is masked does not depend on the failure time t.

For each component, we can observe the triple $\{t_i, \boldsymbol{\delta}_{ji}, \upsilon_{ji} : i = 1, \ldots, n\}$. Our interest is the estimation of the distribution function, F, of the j-th component. We consider a parametric family model of F with parameter $\boldsymbol{\theta}_j$; then, the estimation of parameter $\boldsymbol{\theta}_j$ induces the distribution function F. The available information from the data is one of the following types:

1) $\Pr(X_{ji} \in (t_i, t_i], \upsilon_{ji} = 0 \mid \boldsymbol{\theta}_j) = f(t_i \mid \boldsymbol{\theta}_j)(1 - \lambda_{1j})$, if the i-th observation is uncensored and not masked;
2) $\Pr(X_{ji} \in (t_i, \infty), \upsilon_{ji} = 0 \mid \boldsymbol{\theta}_j) = [1 - F(t_i \mid \boldsymbol{\theta}_j)](1 - \lambda_{2j})$, if the i-th observation is right-censored and not masked;

3) $\Pr(X_{ji} \in (0, t_i], \upsilon_{ji} = 0 \mid \boldsymbol{\theta}_j) = F(t_i \mid \boldsymbol{\theta}_j)(1-\lambda_{3j})$, if the i-th observation is left-censored and not masked;
4) $\Pr(X_{ji} \in (t_i, t_i], \upsilon_{ji} = 1 \mid \boldsymbol{\theta}_j) = f(t_i \mid \boldsymbol{\theta}_j)\lambda_{1j}$, if the i-th observation is uncensored and masked;
5) $\Pr(X_{ji} \in (t_i, \infty), \upsilon_{ji} = 1 \mid \boldsymbol{\theta}_j) = [1 - F(t_i \mid \boldsymbol{\theta}_j)]\lambda_{2j}$, if the i-th observation is right-censored and masked; and
6) $\Pr(X_{ji} \in (0, t_i], \upsilon_{ji} = 1 \mid \boldsymbol{\theta}_j) = F(t_i \mid \boldsymbol{\theta}_j)\lambda_{3j}$, if the i-th observation is left-censored and masked.

However, we do not have information about cases 4 to 6 since the data are masked, and we do not know whether that component was censored. Using an augmented data procedure (latent variable), consider $d_{1ji} = 1$ if the masked observation is not censored; otherwise, consider $d_{1ji} = 0$. Consider $d_{2ji} = 1$ if the masked observation is right-censored; otherwise, consider $d_{2ji} = 0$. Consider $d_{3ji} = 1$ if the masked observation is left-censored; otherwise, consider $d_{3ji} = 0$. In addition, $\boldsymbol{d}_{ji} = (d_{1ji}, d_{2ji}, d_{3ji})$ and $\sum_{l=1}^{3} d_{lji} = 1$.

Let $R(t_i \mid \boldsymbol{\theta}_j) = 1 - F(t_i \mid \boldsymbol{\theta}_j)$ be the reliability function. Under the s-independent assumption of the components lifetimes, the likelihood function of the j-th component can be written as a part of the non-masked data and a part of the masked data (augmented data) as follows:

$$\begin{aligned} L(\boldsymbol{\theta}_j, \lambda_{1j}, \lambda_2 j, \lambda_3 j, \boldsymbol{d}_j \mid \boldsymbol{t}, \boldsymbol{\delta}_j, \boldsymbol{\upsilon}_j) &= \prod_{i:\, \upsilon_{ji}=0} \Big\{ \big[f(t_i \mid \boldsymbol{\theta}_j)\,(1-\lambda_{1j})\big]^{\delta_{1ji}} \\ &\quad \times \big[R(t_i \mid \boldsymbol{\theta}_j)\,(1-\lambda_{2j})\big]^{\delta_{2ji}} \\ &\quad \times \big[F(t_i \mid \boldsymbol{\theta}_j)\,(1-\lambda_{3j})\big]^{\delta_{3ji}} \Big\} \\ &\quad \prod_{i:\, \upsilon_{ji}=1} \Big\{ \big[f(t_i \mid \boldsymbol{\theta}_j)\,\lambda_{1j}\big]^{d_{1ji}} \big[R(t_i \mid \boldsymbol{\theta}_j)\,\lambda_{2j}\big]^{d_{2ji}} \\ &\quad \times \big[F(t_i \mid \boldsymbol{\theta}_j)\,\lambda_{3j}\big]^{d_{3ji}} \Big\}, \end{aligned} \tag{1}$$

where $I(A) = 1$ if A is true and 0 otherwise, $\boldsymbol{t} = \{t_1, \ldots, t_n\}$, $\boldsymbol{\upsilon}_j = \{\upsilon_{j1}, \ldots, \upsilon_{jn}\}$, $\boldsymbol{d}_j = (\boldsymbol{d}_{ji} : i \in \{\upsilon_{ji} = 1\})$ and $\boldsymbol{\delta}_j = (\boldsymbol{\delta}_{ji} : i \in \{\upsilon_{ji} = 0\})$.

It is worth noting that in our approach, it is not necessary to know the structure of the system. In situations there are more information from data, we can incorporate in the model. For example, if one knows the system works in series and the j-th component presents $\upsilon_{ji} = 1$ for i-th sample unit, the component j cannot be left-censored failure time and thus, $\lambda_{3j} = 0$. The model changes according to the data and this approach avoids identifiability problems, as discussed by [34].

The likelihood function in (1) is generic and straightforward for any probability distribution. The considered distribution is a three-parameter Weibull distribution. Due to its characteristics, the Weibull distribution is a great candidate for modeling the component lifetimes. One of these characteristics is that by changing parameter values the distribution takes a variety of shapes and it has important distributions as special cases, besides allowing modeling increasing, decreasing or constant hazard rates [35].

The Weibull reliability function is as follows:

$$R(t \mid \boldsymbol{\theta}_j) = \exp\left[-\left(\frac{t-\mu_j}{\eta_j}\right)^{\beta_j}\right],$$

for $t > 0$, where $\boldsymbol{\theta}_j = (\beta_j, \eta_j, \mu_j)$ and $\beta_j > 0$ (shape), $\eta_j > 0$ (scale) and $0 < \mu_j < t$ (location).

The Weibull distribution with two parameters ($\mu_j = 0$) is the most celebrated case in the literature. However, the location parameter that represents the baseline lifetime has an important meaning in reliability and survival analyses. In reliability analyses, the tested component may not be new. For example, in medicine, a patient may have the disease before the onset medical appointment. Failure to consider the initial time can result in the underestimation of the other parameters. Clearly, in testing a new component, μ_j may be 0.

The estimation work is performed under a Bayesian perspective of inference; thus, the priori distribution of $(\boldsymbol{\theta}_j, \lambda_{1j}, \lambda_{2j}, \lambda_{3j}, \boldsymbol{d}_j)$ needs to be defined. All parameters are considered independent a prior, assuming a gamma distribution with mean 1 and variance 1000 for β_j, η_j, μ_j and a uniform distribution over (0, 1) for λ_{1j}, λ_{2j} and λ_{3j}. In addition, $\boldsymbol{d}_{ji} \sim \text{Multinomial}(1; p_{1ji}, p_{2ji}, p_{3ji})$, in which $p_{lji} = 1/3$, for $l = 1, 2, 3$.

We do not have prior information about the component's operation, and the noninformative prior is considered. However, it is possible to express the prior information about the component functioning through expert opinion and/or past experience.

The posterior density of $(\boldsymbol{\theta}_j, \lambda_{1j}, \lambda_{2j}, \lambda_{3j}, \boldsymbol{d}_j)$ is as follows:

$$\begin{aligned} &\pi(\boldsymbol{\theta}_j, \lambda_{1j}, \lambda_{2j}, \lambda_{3j}, \boldsymbol{d}_j \mid \boldsymbol{t}, \boldsymbol{\delta}_j, \boldsymbol{\upsilon}_j) \propto \\ &\quad \pi(\boldsymbol{\theta}_j, \lambda_{1j}, \lambda_{2j}, \lambda_{3j}, \boldsymbol{d}_j) L(\boldsymbol{\theta}_j, \lambda_{1j}, \lambda_{2j}, \lambda_{3j}, \boldsymbol{d}_j \mid \boldsymbol{t}, \boldsymbol{\delta}_j, \boldsymbol{\upsilon}_j), \end{aligned} \tag{2}$$

where $\pi(\boldsymbol{\theta}_j, \lambda_{1j}, \lambda_{2j}, \lambda_{3j}, \boldsymbol{d}_j)$ is the prior distribution of $(\boldsymbol{\theta}_j, \lambda_{1j}, \lambda_{2j}, \lambda_{3j}, \boldsymbol{d}_j)$.

The posterior density in Equation (2) does not have a closed form. An alternative is to rely on Markov-Chain Monte-Carlo (MCMC) simulations. We considered the Gibbs with Metropolis-Hasting steps algorithm. This algorithm is suitable for this situation because it is possible to directly sample some parameters from the conditional distribution; however, this sampling is not possible for other parameters [36]. The algorithm works in the following steps:

1) *Attribute initial values* $\boldsymbol{\theta}_j^{(0)}$, $\lambda_{1j}^{(0)}$, $\lambda_{2j}^{(0)}$ *and* $\lambda_{3j}^{(0)}$ *for* $\boldsymbol{\theta}_j = (\beta_j, \eta_j, \mu_j)$, λ_{1j}, λ_{2j} *and* λ_{3j}, *respectively, and set* $b = 1$;
2) *For* $i \in \{\boldsymbol{\upsilon}_{ji} = 1\}$, *draw* $\boldsymbol{d}_{ji}^{(b)}$ *from*

$$\boldsymbol{d}_{ji} \mid \boldsymbol{t}, \boldsymbol{\delta}_j, \boldsymbol{\upsilon}_j, \boldsymbol{\theta}_j, \lambda_{1j}, \lambda_{2j}, \lambda_{3j} \sim \text{Multinomial}(1; p_{1ji}, p_{2ji}, p_{3ji}),$$

where

$$\begin{aligned} p_{1ji} &= \lambda_{1j} f(t_i|\boldsymbol{\theta}_j)/C, \\ p_{2ji} &= \lambda_{2j} R(t_i|\boldsymbol{\theta}_j)/C, \\ p_{3ji} &= \lambda_{3j} F(t_i|\boldsymbol{\theta}_j)/C, \text{ and} \\ C &= \lambda_{1j} f(t_i|\boldsymbol{\theta}_j) + \lambda_{2j} R(t_i|\boldsymbol{\theta}_j) + \lambda_{3j} F(t_i|\boldsymbol{\theta}_j); \end{aligned}$$

3) *Using the Metropolis-Hastings algorithm [37], draw* $\boldsymbol{\theta}_j^{(b)}$ *from*

$$\begin{aligned}\pi(\boldsymbol{\theta}_j \mid \boldsymbol{t}, \boldsymbol{\delta}_j, \boldsymbol{\upsilon}_j, \boldsymbol{d}_j, \lambda_{1j}, \lambda_{2j}, \lambda_{3j}) \propto \\ \pi(\boldsymbol{\theta}_j) \prod_{i:\ \upsilon_{ji}=0} \Big\{ \big[f(t_i \mid \boldsymbol{\theta}_j)(1-\lambda_{1j})\big]^{\delta_{1ji}} \\ \times \big[R(t_i \mid \boldsymbol{\theta}_j)(1-\lambda_{2j})\big]^{\delta_{2ji}} \\ \times \big[F(t_i \mid \boldsymbol{\theta}_j)(1-\lambda_{3j})\big]^{\delta_{3ji}} \Big\} \\ \times \prod_{i:\ \upsilon_{ji}=1} \Big\{ \big[f(t_i \mid \boldsymbol{\theta}_j)\lambda_{1j}\big]^{d_{1ji}} \\ \times \big[R(t_i \mid \boldsymbol{\theta}_j)\lambda_{2j}\big]^{d_{2ji}} \\ \times \big[F(t_i \mid \boldsymbol{\theta}_j)\lambda_{3j}\big]^{d_{3ji}} \Big\}.\end{aligned}$$

4) *Draw* λ_{1j} *from*

$$\lambda_{1j} \mid \boldsymbol{t}, \boldsymbol{\delta}_j, \boldsymbol{\upsilon}_j, \boldsymbol{d}_j, \boldsymbol{\theta}_j, \lambda_{2j}, \lambda_{3j} \sim \text{Beta}\left(\sum_{\{i:\upsilon_{ji}=1\}} d_{1ji} + 1, n_f + 1 \right),$$

where n_f *is the number of systems in which component j leads to* ***system failure***.

5) *Draw* λ_{2j} *from*

$$\lambda_{2j} \mid \boldsymbol{t}, \boldsymbol{\delta}_j, \boldsymbol{\upsilon}_j, \boldsymbol{d}_j, \boldsymbol{\theta}_j, \lambda_{1j}, \lambda_{3j} \sim \text{Beta}\left(\sum_{\{i:\upsilon_{ji}=1\}} d_{2ji} + 1, n_r + 1 \right),$$

where n_r *is the number of systems in which component j is observed to be* ***right-censored***.

6) *Draw* λ_{3j} *from*

$$\lambda_{3j} \mid \boldsymbol{t}, \boldsymbol{\delta}_j, \boldsymbol{\upsilon}_j, \boldsymbol{d}_j, \boldsymbol{\theta}_j, \lambda_{1j}, \lambda_{2j} \sim \text{Beta}\left(\sum_{\{i:\upsilon_{ji}=1\}} d_{3ji} + 1, n_l + 1 \right),$$

where n_l *is the number of systems in which component j is observed to be* ***left-censored***.

7) *Let* $b = b + 1$ *and repeat steps* 2) *to* 7) *until* $b = B$*, where B is the pre-set number of simulated samples of* $(\boldsymbol{\theta}_j, \lambda_{1j}, \lambda_{2j}, \lambda_{3j}, \boldsymbol{d}_j)$.

Discarding the burn-in (first generated values discarded to eliminate the effect of the assigned initial values for the parameters) and jump samples (spacing among the generated values to avoid a correlation between the simulated samples), a sample of size n_p is obtained from the joint posterior distribution of $(\boldsymbol{\theta}_j, \lambda_{1j}, \lambda_{2j}, \lambda_{3j}, \boldsymbol{d}_j)$. For the j-th component, the sample from the posterior can be expressed as $(\boldsymbol{\theta}_j^{(1)}, \boldsymbol{\theta}_j^{(2)}, \ldots, \boldsymbol{\theta}_j^{(n_p)})$, $(\lambda_{1j}^{(1)}, \lambda_{1j}^{(2)}, \ldots, \lambda_{1j}^{(n_p)})$, $(\lambda_{2j}^{(1)}, \lambda_{2j}^{(2)}, \ldots, \lambda_{2j}^{(n_p)})$, $(\lambda_{3j}^{(1)}, \lambda_{3j}^{(2)}, \ldots, \lambda_{3j}^{(n_p)})$ and $(\boldsymbol{d}_j^{(1)}, \boldsymbol{d}_j^{(2)}, \ldots, \boldsymbol{d}_j^{(n_p)})$. Consequently, the posterior quantities of reliability function $R(t \mid \boldsymbol{\theta}_j)$ can be easily obtained [37]. For example, the posterior mean of the reliability function can be approximated as follows:

$$E[R(t \mid \boldsymbol{\theta}_j) \mid Data] = \frac{1}{n_p} \sum_{k=1}^{n_p} R\Big(t \mid \boldsymbol{\theta}_j^{(k)}\Big), \qquad \text{for each } t > 0.$$

The proposed method is general and it can be considered in situations that j-th component in Υ can be susceptible to failure at t, before or after t. In situations one has more information from data, for example: the interest component in Υ cannot fail after t, the restriction $\lambda_{2j} = 0$ is considered. As a consequence, in step 2 of the MCMC algorithm, $\boldsymbol{d}_{ji}$ conditional to $(\boldsymbol{t}, \boldsymbol{\delta}_j, \boldsymbol{\upsilon}_j, \boldsymbol{\theta}_j, \lambda_{1j}, \lambda_{3j})$ follows a Multinomial$(1; p_{1ji}, p_{3ji})$, in which $p_{1ji} + p_{3ji} = 1$. Besides, step 5 is eliminated.

Note that the algorithm is general and is suitably adjusted according to the restrictions done by information from data.

A. SYMMETRIC MASKING PROBABILITIES

The symmetric masking probabilities assumption suggests that the masking probabilities are the same regardless of the component that causes the system failure, i.e., $\lambda_{1j} = \lambda_{2j} = \lambda_{3j}$ is plausible in some masked data system situations.

Under this assumption, we have the following:

$$\boldsymbol{d}_{ji} \mid \boldsymbol{t}, \boldsymbol{\delta}_j, \boldsymbol{\upsilon}_j, \boldsymbol{\theta}_j, \lambda_{1j}, \lambda_{2j}, \lambda_{3j} \sim \text{Multinomial}(1; p_{1ji}, p_{2ji}, p_{3ji}),$$

where $p_{1ji} = f(t_i|\boldsymbol{\theta}_j)/C$, $p_{2ji} = R(t_i|\boldsymbol{\theta}_j)/C$, $p_{3ji} = F(t_i|\boldsymbol{\theta}_j)/C$, and $C = 1 + f(t_i|\boldsymbol{\theta}_j)$. Thus, the estimation process does not depend on masking probabilities λ_{lj}, $l = 1, 2, 3$, and the previously presented algorithm can be considered in eliminating steps 4 to 6.

B. INCORPORATION OF COVARIATES

In a sample of n systems, the units are not exposed to exactly the same temperature and pressure conditions, for example, and depending on their values, the reliability of the components can increase or decrease. Thus, considering these different conditions in the reliability estimation of each component is important and can be achieved by incorporating covariates into the model.

In general, evaluating the performance of the components in a system under normal conditions can be time-consuming and costly. Therefore, another importance emerges from the incorporation of covariates as follows: accelerated life tests in which covariates are called stress variables. In accelerated life tests (ALT), the components are subjected to stress levels to reduce their time to failure, and inferences are obtained about their behavior under normal operating conditions.

The analysis of stress-response relationships and extrapolation to usual operating conditions can be achieved through regression models of data from accelerated tests called accelerated life models. In ALT models, there is multiplicative effect with the reliability time, i.e., $R(t) = R_0(\varphi t)$, where φ is the acceleration factor, and $R_0(t)$ is the baseline reliability function. Thus, if $\varphi > 1$, $R(t)$ behaves as $R_0(t)$ "in the future''; if $\varphi < 1$, $R(t)$ behaves as $R_0(t)$ ''in the past''; and if $\varphi = 1$, $R(t) = R_0(t)$.

Some parametric models have the properties of accelerated life tests. The three-parameter Weibull distribution is an ALT model in which $R(t \mid \boldsymbol{\theta}_j) = R_0((t - \mu_j)\varphi_j \mid \beta_j)$, and $R_0(\cdot \mid \beta_j)$ is the reliability function of a Weibull distribution with scale 1 and shape β_j. Thus, the covariates can be included in the scale parameter through a log link function, i.e., $\eta_j = \exp(\boldsymbol{w}_j^\top \boldsymbol{\gamma}_j)$, where $\boldsymbol{\gamma}_j$ is a vector $(k \times 1)$ of the regression coefficients, and $\boldsymbol{w}_j$ is a vector of the covariables of the j-th component.

III. SIMULATED SYSTEM DATASETS

We consider two simulated examples of the complex system structure presented in figures 2 and 6. As previously mentioned, there is no solution in the literature for the reliability estimation of all components involved in these complex systems with masked cause of failure.

In this section, the structure of the systems is observed and, as consequence, Υ_i consists of components that no longer work in the system failure, i.e., there are no right-censored components in Υ_i that lead to $\lambda_{2j} = 0$. In addition, $j \in \Upsilon_i$ only if j belongs to the minimal cut that caused the i-th system fail. A cut set is a set of components that by failing causes the system to fail. A cut set is considered minimal if it cannot be reduced without losing its status as a cut set. For example, in the system presented in Figure 2, there are three minimal cut set, i.e., $\{1, 2\}$, $\{3, 5\}$ and $\{4, 5\}$.

In a situation in which this system fails and components 1, 2 and 3 do not work at the moment of system failure, only components 1 and 2 belong to set Υ because component 3 does not belong to the minimal cut set that caused the system failure, and in fact, component 3 is observed to have a left-censored failure time.

To generate data in each simulated example with m being the number of components and n being the sample size, the following steps are considered:

For each system unit i, where $i = 1, \ldots, n$:

1) Draw X_{ji} from a given distribution for $j = 1, \ldots, m$;
2) Let $T_i = h(X_{1i}, \ldots, X_{mi})$, where T_i is the system failure time, and $h(\cdot)$ is the function that relates the system failure time to the components' functioning depending on the system structure;
3) Draw $c_i \sim$ Bernoulli (p), where p is the proportion of the masked data system. If $c_i = 1$, the system i has a masked failure cause and obtain Υ_i, which is the set of index components that could lead to system failure; and
4) For each j component, where t_i is the system failure time:
   - If $j \in \Upsilon_i$: let $\upsilon_{ji} = 1$;
   - Otherwise, let $\upsilon_{ji} = 0$ and observe δ_{ji}, in which:
     - If $X_{ji} = t_i, \quad \delta_{ji} = 1$;
     - If $X_{ji} > t_i, \quad \delta_{ji} = 2$ and
     - If $X_{ji} < t_i, \quad \delta_{ji} = 3$.

The dataset for the j-th component is $\{(t_1, \delta_{j1}, \upsilon_{j1}), (t_2, \delta_{j2}, \upsilon_{j2}), \ldots, (t_n, \delta_{jn}, \upsilon_{jn})\}$, where δ_{ji} is empty set if $\upsilon_{ji} = 1$.

The simulated systems have the following characteristics:

- System structure 1 (Figure 6): $m = 3$, X_1 is generated from a Weibull distribution with a mean of 15 and a variance of 8, X_2 is generated from a gamma distribution with a mean of 18 and a variance of 12, X_3 is generated from a lognormal distribution with a mean of 20 and a variance of 10, and the system failure time is
$$\begin{aligned} T &= h(X_1, X_2, X_3) \\ &= \max(\min(X_1, X_2), \min(X_1, X_3), \min(X_2, X_3)). \end{aligned}$$
In addition, $n = 300$, and the proportion of the masked system is $p = 0.4$.
- System structure 2 (Figure 2): $m = 5$, X_1 is generated from a Weibull distribution with a mean of 12 and a variance of 15, X_2 is generated from a gamma distribution with a mean of 11 and a variance of 11, X_3 is generated from a three-parameter Weibull distribution with a mean of 12 and a variance of 9, X_4 is generated from a lognormal distribution with a mean of 12 and a variance of 7 and X_5 is generated from a three-parameter Weibull distribution with a mean of 11 and a variance of 14. In this structure, the system failure time is
$$\begin{aligned} T &= h(X_1, \ldots, X_5) \\ &= \min(\max(X_1, X_2), \max(\min(X_3, X_4), X_5)). \end{aligned}$$
In this case, $n = 100$, and the proportion of the masked data systems is $p = 0.3$.

To obtain the posterior quantities, we used an MCMC procedure to generate a sample from the posterior distribution of the parameters. We generated 30,000 values for each parameter and disregarded the first 10,000 iterations to eliminate the effect of the initial values and of the spacing size 20 to avoid correlation problems; finally, we obtained a sample size of $n_p = 1{,}000$. The chains convergence was monitored to obtain good convergence results, as acceptance rates between 20% and 35%.

A. SYSTEM 1

The posterior quantities of $R(t \mid \boldsymbol{\theta}_j)$ for some values of t are shown in Table 1. In the following discussion in this paper, the posterior mean of the reliability function is considered the performing posterior measure obtained by the proposed model.

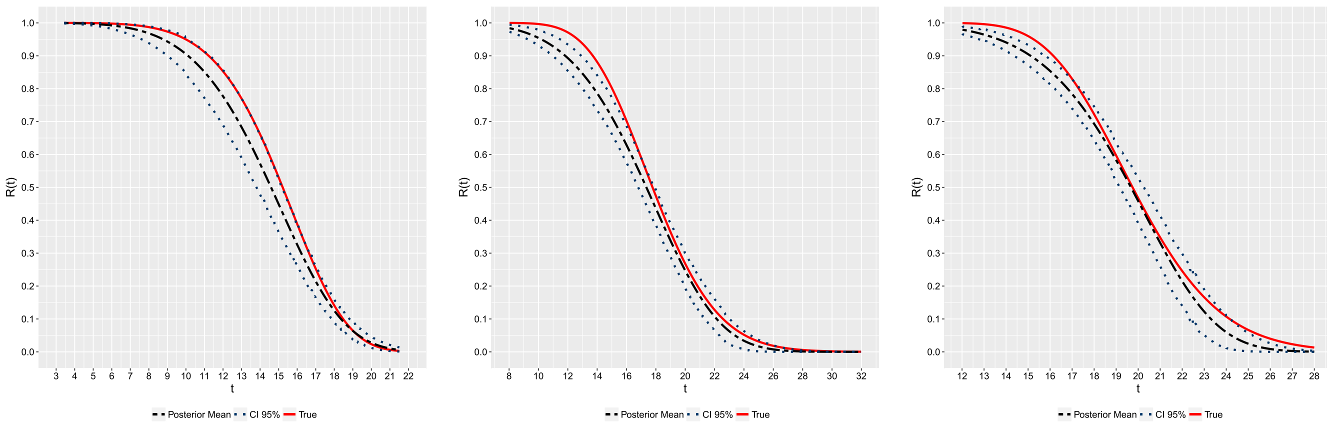
The true curves and posterior mean of the reliability function are shown in Figure 7, which also presents the 95% HPD point-wise band (CI 95%) obtained by the proposed model. For component 1, the true curve is contained in the all HPD point-wise band. By considering components 2 and 3 and t values in which the posterior mean is more distant from the true curve, the upper limit of the HPD band is very close to the true curve.

B. SYSTEM 2

The true reliability functions, posterior means and 95% HPD point-wise bands (CI 95%) are shown in Figure 8. In general,

TABLE 1. Posterior measures of components reliability functions involved in system structure 1 for some values of t.

t	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum	Standard deviation	HPD 95%
Component 1								
3.5	0.995	0.999	1.000	0.999	1.000	1.000	0.001	(0.998; 1.000)
10.0	0.788	0.889	0.908	0.905	0.927	0.966	0.029	(0.842; 0.951)
20.0	0.007	0.021	0.026	0.028	0.033	0.059	0.009	(0.012; 0.046)
Component 2								
8.5	0.938	0.975	0.980	0.979	0.984	0.994	0.007	(0.965; 0.992)
15.0	0.602	0.692	0.712	0.712	0.732	0.793	0.030	(0.660; 0.772)
26.0	0.001	0.003	0.006	0.007	0.009	0.056	0.006	(0.001; 0.020)
Component 3								
12.5	0.936	0.968	0.974	0.973	0.978	0.989	0.008	(0.958; 0.986)
20.0	0.347	0.437	0.460	0.458	0.480	0.555	0.034	(0.392; 0.528)
25.0	0.001	0.012	0.021	0.024	0.033	0.119	0.016	(0.002; 0.056)



(a) (b) (c)

FIGURE 7. True reliability functions and estimated curves by proposed model for the components 1 to 3 involved in system structure 1. (a) Component 1. (b) Component 2. (c) Component 3.

the true curves are contained in the HPD point-wise bands, and the proposed model presents an almost perfect reliability function estimation for component 3.

IV. MODEL EVALUATION WITH SIMULATION STUDIES

To evaluate the performance of the proposed model, our approach is compared to the approach presented by [30] in scenarios with different sample sizes, system structures and proportions of masked data. The posterior mean of the reliability function is considered the performing posterior measure obtained by the proposed model and is denoted by W3PM (Weibull 3-Parameter Model). Reference [30] has been considered the best approach for masked data in complex systems. This method estimates the reliability of components involved in any coherent system from the simplest to the most complex. The only necessary types of information for the computation of the estimates are the system structure and the observed system failure times; thus, knowing the cause of failure is not necessary, which is suitable for masked data situations. The authors assumed a restrictive assumption in which the component lifetimes are s-independent and identically distributed; therefore, all components have the same reliability. For simplification, we refer to this estimator as BSNP (Bhattacharya-Samaniego Nonparametric Estimator).

The following two types of system structures are used: 2-out-of-3 (Figure 6) and the bridge system (Figure 4). Four sample sizes ($n = 50,\ 100,\ 300$, and 1000) and the following three proportions of masked data are considered: $p = 0.2,\ 0.4$ and 0.7. For each scenario (combination of sample size and proportion of masked data), 1000 samples were generated, and the distributions are described as follows:

- 2-out-of-3 system: The same generation as that in example III-A is used.
- Bridge system: $m = 5$, X_1 is generated from a Weibull distribution with a mean of 4 and a variance of 15, X_2 is generated from a modified Weibull distribution [38] with a mean 5.6 and a variance of 14.9, X_3 is generated from a lognormal distribution with a mean of 6 and a variance of 7, X_4 is generated from a gamma distribution with a mean of 5 and a variance of 8 and X_5 is generated from a three-parameter Weibull distribution with a mean of 4 and a variance of 8. In this structure, the system lifetime is expressed as follows:

$$\begin{aligned} T &= h(X_1, \ldots, X_5) \\ &= \max(\min(X_1, X_4), \min(X_2, X_5), \\ &\quad \min(X_1, X_3, X_5), \min(X_2, X_3, X_4)). \end{aligned}$$

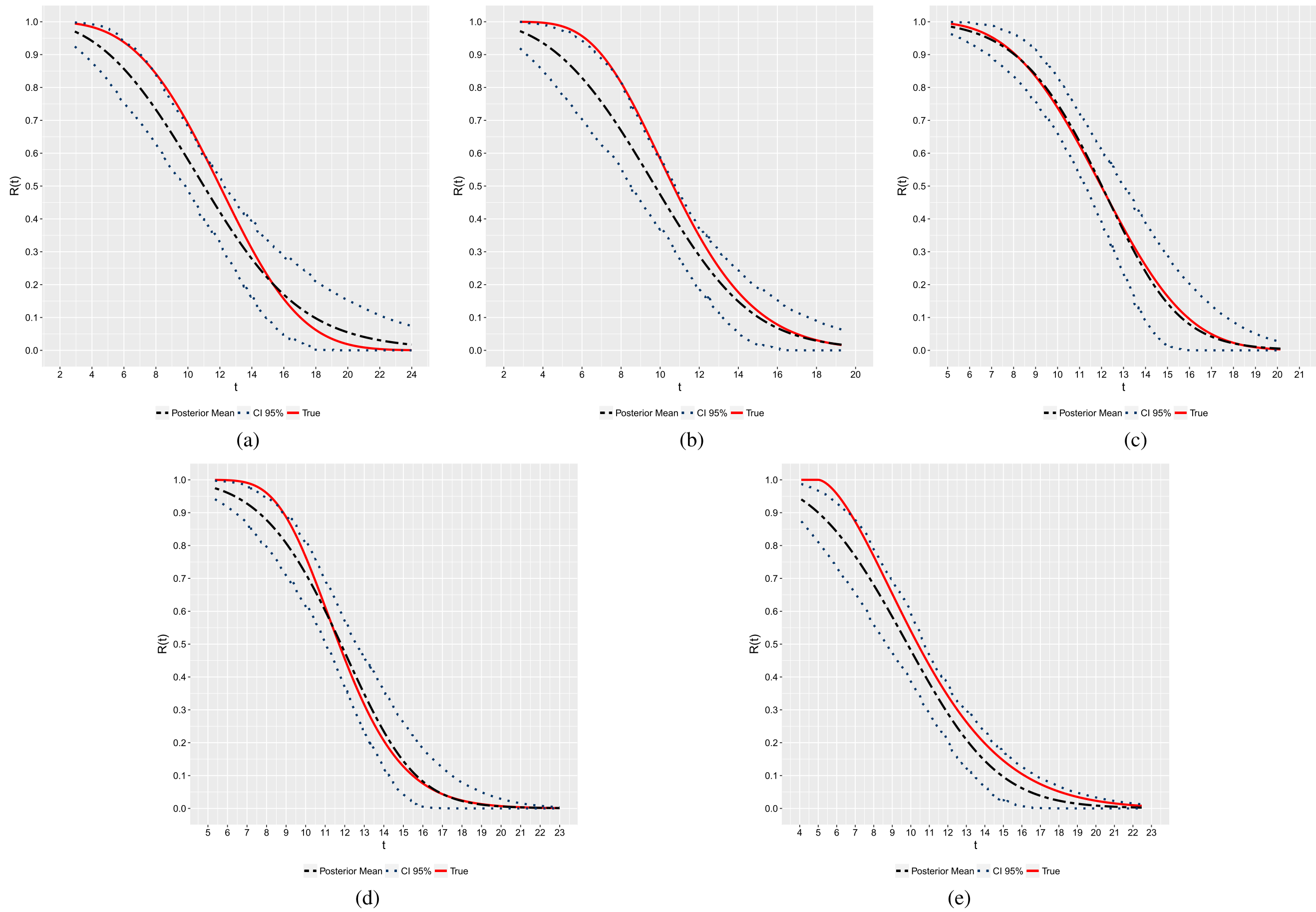


FIGURE 8. True reliability functions and estimated curves by the proposed model for the components 1 to 5 involved in system structure 2. (a) Component 1. (b) Component 2. (c) Component 3. (d) Component 4. (e) Component 5.

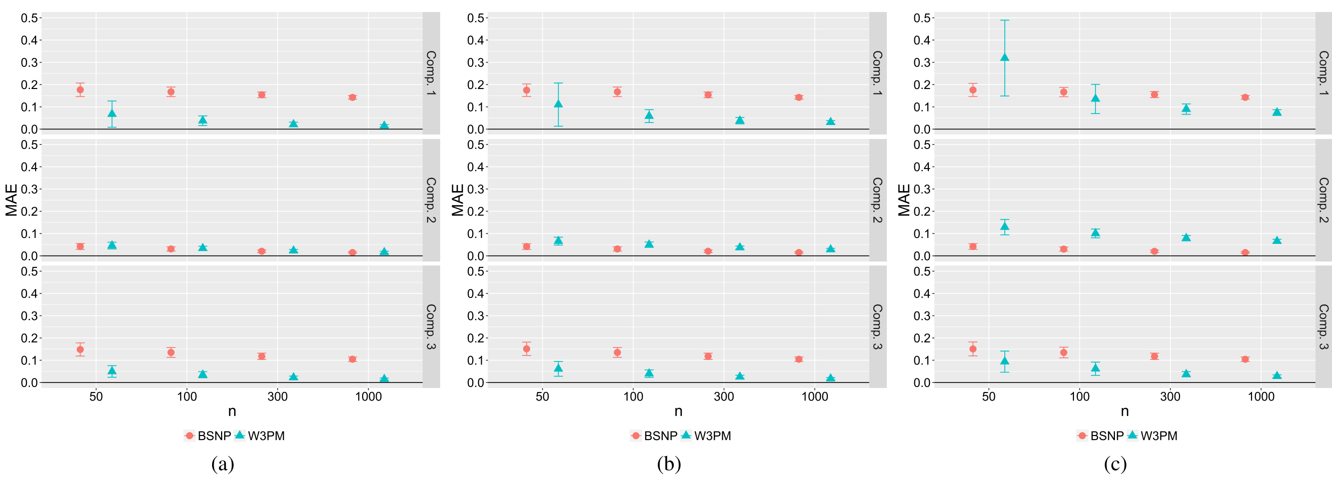
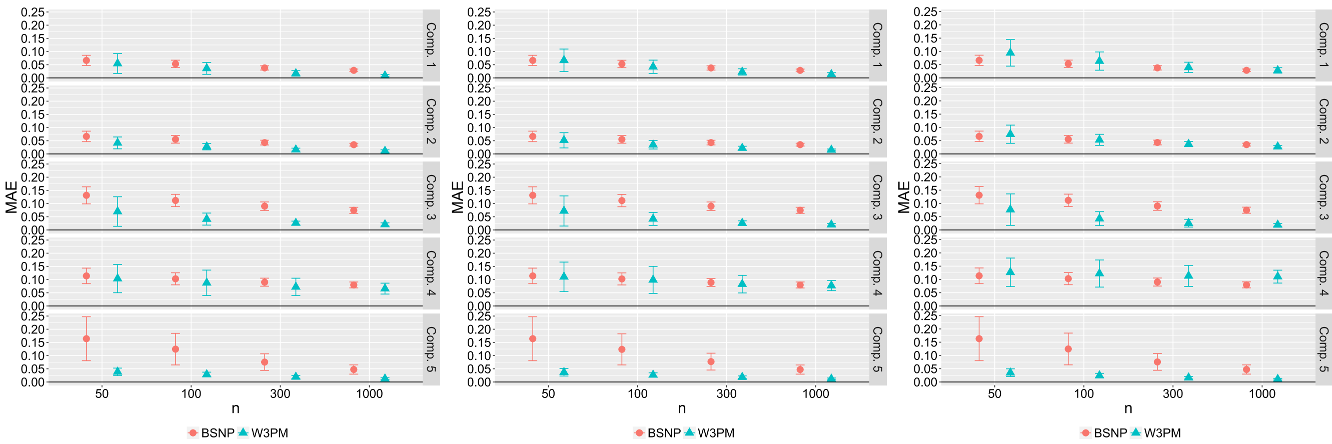


FIGURE 9. Mean (symbol) and standard deviation (bars) of MAE obtained by W3PM and BSNP for components 1 to 3 of 2-out-of-3 structure considering $p = 0.2,\ 0.4,\ 0.7$ and $n = 50,\ 100,\ 300,\ 1000$. (a) $p = 0.2$. (b) $p = 0.4$. (c) $p = 0.7$.

The mean absolute error (MAE) from the estimators to the true distribution is considered the comparison measure. $R(t)$ and $\widehat{R}(t)$ represent 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. The means and standard deviations of 1000 MAE values obtained by W3PM and BSNP are presented in figures 9 and 10 for the 2-out-of-3 and bridge systems, respectively. In general, W3PM presents lower MAE values mean. The exception is component 2 of



(a) (b) (c)

FIGURE 10. Mean (symbol) and standard deviation (bars) of MAE obtained by W3PM and BSNP for components 1 to 5 of the bridge structure considering $p = 0.2,\ 0.4,\ 0.7$ and $n = 50,\ 100,\ 300,\ 1000$. (a) $p = 0.2$. (b) $p = 0.4$. (c) $p = 0.7$.

TABLE 2. Distribution of $n = 172$ systems of hard drives dataset among causes of failure.

	Failure by $j = 1$	Failure by $j = 2$	Failure by $j = 3$	$\Upsilon = \{1, 3\}$	$\Upsilon = \{1, 2, 3\}$
n (%)	35 (20.35)	19 (11.05)	52 (30.23)	32 (18.60)	34 (19.77)

the 2-out-of-3 system, in which the BSNP presents better performance. However, the difference between the two methods decreases as the sample size increases mainly because the performance of the proposed estimator improves as n increases.

V. APPLICATION

In this section, a real dataset is considered to present the applicability of the proposed model to reliability estimations of components involved in coherent systems. The dataset is available in [10] and consists of 172 observed failure times of computer hard drives monitored over a period of 4 years. There were three possible causes of failure as follows: electronic hard (component $j = 1$), head fly ability (component $j = 2$) and head/disc magnetics (component $j = 3$). However, for some hard drives (38%), the cause of hard drive failure was not identified. For these masked data systems, $\Upsilon = \{1, 3\}$ or $\Upsilon = \{1, 2, 3\}$, i.e., there is no possible masked set $\Upsilon = \{1, 2\}$ or $\Upsilon = \{2, 3\}$. Notably, in our proposed approach, the configuration of set Υ is not important, and the critical information for the estimation of the reliability of the j-th component is whether j belongs to Υ. More details about the detection of failure causes are presented in [39] and [40].

As shown in Table 2, component 1 is observed to cause the failure of 20.35% of the systems, 11.05% of the systems had observed failures because of component 2 and component 3 led to 30.23% of the system failures. In addition, component 1 or component 3 caused the failure of 18.60% of the systems, and the remaining 19.77% of system failures were due to any of the three components.

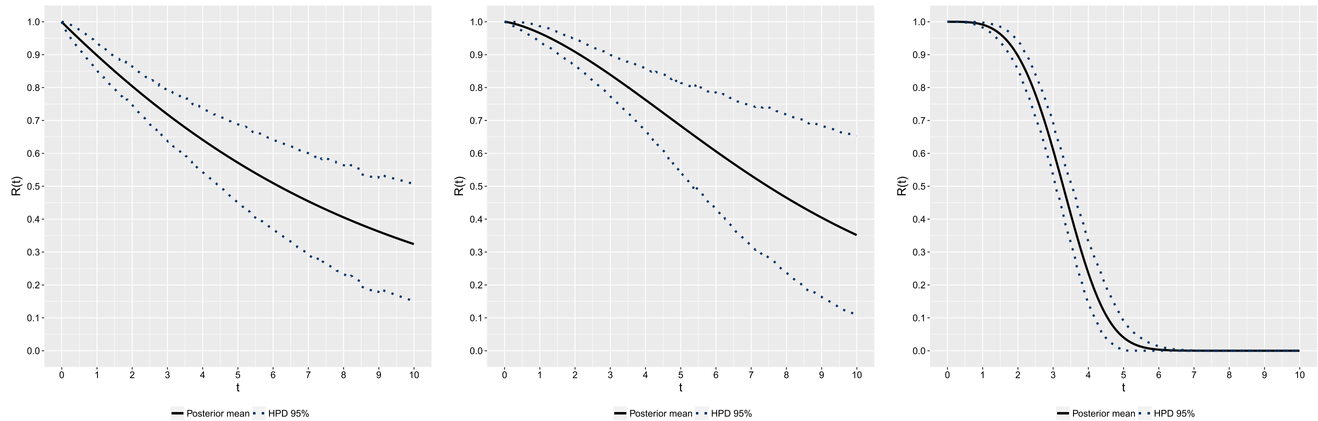
Since the components in Υ are right-censored or lead to system failure, $\lambda_{3j} = 0$ for $j = 1, 2, 3$, and no hard drive is subject to left-censored failure times.

TABLE 3. Parameters posterior quantities for components 1 to 3 involved in hard drives dataset under symmetric assumption and under relaxing this assumption.

	Component 1			
	Symmetric Assumption		Symmetric Assumption Relaxation	
	Posterior mean	Posterior SD	Posterior mean	Posterior SD
β	1.031	0.161	0.949	0.148
η	9.656	3.007	12.530	4.348
μ	3.33E-38	6.11E-37	2.22E-47	2.95E-46
	Component 2			
	Symmetric Assumption		Symmetric Assumption Relaxation	
	Posterior mean	Posterior SD	Posterior mean	Posterior SD
β	1.531	0.309	1.490	0.314
η	10.490	4.455	9.595	3.480
μ	1.29E-38	2.35E-37	2.17E-42	2.29E-41
	Component 3			
	Symmetric Assumption		Symmetric Assumption Relaxation	
	Posterior mean	Posterior SD	Posterior mean	Posterior SD
β	3.728	0.389	3.339	0.401
η	3.629	0.132	3.965	0.198
μ	2.43E-08	1.46E-07	8.08E-13	3.12E-12

The proposed model is fit under the symmetric assumption and under relaxing this assumption. To obtain posterior quantities related to the posterior distribution of $(\boldsymbol{\theta}_j, \lambda_{1j}, \lambda_{2j}, \boldsymbol{d}_j)$, with $\boldsymbol{\theta}_j = (\beta_j, \eta_j, \mu_j)$ and for $j = 1, 2, 3$, from (2) through MCMC simulations, we discarded the first 10,000 iterations as burn-in samples and used a jump size of 30 to avoid correlation problems; finally, we obtained a sample size of $n_p = 1{,}000$. The chain convergence was monitored to obtain good convergence results.

The proposed model parameter estimates are presented in Table 3. The two versions of the proposed model provided close estimates for $\boldsymbol{\theta}_j = (\beta_j, \eta_j, \mu_j)$ for all j. Notably, the posterior mean of μ_j is close to zero for all j, indicating that the beginning of the computer lifetimes coincides with the beginning of the experiment, which is logical given that this



(a) (b) (c)

FIGURE 11. Estimated reliability curves by the proposed model of three components involved in computer hard drive dataset. (a) Electronic hard ($j = 1$). (b) Head fly ability ($j = 2$). (c) Head/disc magnetics ($j = 3$).

is a controlled experiment and that the hard drives have not been previously tested.

In Figure 11, we present the estimated curves of electronic hard, head fly ability and head/disc magnetics by BSNP and the proposed model. The proposed model under the symmetric assumption and under relaxing this assumption obtain overlapping curves; thus, only one curve is presented.

VI. FINAL REMARKS

A Bayesian three-parameter Weibull model was proposed for component reliability. The assumption that the component lifetimes are identically distributed is not imposed. The presented model is considered general because it can be used for any coherent system; the symmetry assumption is not necessary, and its application in accelerated life tests can be considered. We worked with the Weibull model; however, extending this work to other distributions or even to the pure likelihood approach is simple.

The proposed model was compared to the nonparametric estimator proposed by [30] (BSNP), which can be considered for components involved in any system for which the only necessary information is the system failure time and the structure. However, these authors assumed a restrictive assumption in which the component lifetimes are *s*-independent and identically distributed. Therefore, there is only one estimator for all the different components in the system. The simulation study showed that our proposed estimator had excellent performance and was superior over BSNP. The advantage of our model is more evident as the sample size increases.

The practical relevance and applicability of our model were demonstrated using a real dataset of computer hard drives with three components in series. Thus, our estimator of the component reliability function demonstrated profound performance in situations in which the lifetime distribution is not the same for all components in a coherent system, with different proportions of masked systems. In estimation processes, satisfactory results of convergence were obtained, and the posterior quantities of the reliability functions were easily obtained.

ACKNOWLEDGMENT

CNPq had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

[1] A. S. Rodrigues, F. Bhering, C. A. de B. Pereira, and A. Polpo, "Bayesian estimation of component reliability in coherent systems," *IEEE Access*, vol. 6, pp. 18520–18535, 2018.
[2] V. H. Salinas-Torres, C. A. de B. Pereira, and R. C. Tiwari, "Convergence of Dirichlet measures arising in context of Bayesian analysis of competing risks models," *J. Multivariate Anal.*, vol. 62, no. 1, pp. 24–35, 1997.
[3] V. Salinas-Torres, C. A. de B. Pereira, and R. Tiwari, "Bayesian nonparametric estimation in a series system or a competing-risks model," *J. Nonparam. Statist.*, vol. 14, no. 4, pp. 449–458, 2002.
[4] A. Polpo and C. A. de B. Pereira, "Reliability nonparametric Bayesian estimation in parallel systems," *IEEE Trans. Rel.*, vol. 58, no. 2, pp. 364–373, Jun. 2009.
[5] A. Polpo, M. Coque-Jr, and C. A. de B. Pereira, "Statistical analysis for Weibull distributions in presence of right and left censoring," in *Proc. 8th Int. Conf. Rel., Maintainability Saf.*, Jul. 2009, pp. 219–223.
[6] A. Polpo, D. Sinha, P. R. Simoni, and C. A. de B. Pereira, "On Bayesian estimation of a survival curve: Comparative study and examples," in *Proc. AIP Conf.*, Secaucus, NJ USA: Springer, vol. 1490, 2012, pp. 268–277.
[7] A. S. Rodrigues, T. C. M. Dias, M. S. Lauretto, and A. Polpo, "Reliability analysis in series systems: An empirical comparison between Bayesian and classical estimators," in *Proc. 31st AIP Conf.*, 2012, vol. 1443, no. 1, pp. 214–221.
[8] A. Polpo, D. Sinha, and C. A. de B. Pereira, "Nonparametric Bayesian estimation of reliabilities in a class of coherent systems," *IEEE Trans. Rel.*, vol. 62, no. 2, pp. 455–465, Jun. 2013.
[9] F. L. Bhering, C. A. de B. Pereira, and A. Polpo, "Reliability estimators for the components of series and parallel systems: The Weibull model," *Appl. Math.*, vol. 5, no. 11, pp. 1633–1640, 2014.
[10] B. J. Flehinger, B. Reiser, and E. Yashchin, "Parametric modeling for survival with competing risks and masked failure causes," *Lifetime Data Anal.*, vol. 8, no. 2, pp. 177–203, 2002.
[11] M. Miyakawa, "Analysis of incomplete data in competing risks model," *IEEE Trans. Rel.*, vol. R-33, no. 4, pp. 293–296, Oct. 1984.
[12] J. S. Usher and T. J. Hodgson, "Maximum likelihood analysis of component reliability using masked system life-test data," *IEEE Trans. Rel.*, vol. R-37, no. 5, pp. 550–555, Dec. 1988.

[13] B. Reiser, I. Guttman, D. K. J. Lin, F. M. Guess, and J. S. Usher, "Bayesian inference for masked system lifetime data," *J. Roy. Stat. Soc.*, vol. 44, no. 1, pp. 79–90, 1995.
[14] F. M. Guess, J. S. Usher, and T. J. Hodgson, "Estimating system and component reliabilities under partial information on cause of failure," *J. Stat. Planning Inference*, vol. 29, pp. 75–85, Sep./Oct. 1991.
[15] D. K. J. Lin, J. S. Usher, and F. M. Guess, "Bayes estimation of component-reliability from masked system-life data," *IEEE Trans. Rel.*, vol. 45, no. 2, pp. 233–237, Jun. 1996.
[16] A. M. Sarhan, "Reliability estimations of components from masked system life data," *Rel. Eng. Syst. Saf.*, vol. 74, no. 1, pp. 107–113, 2001.
[17] A. M. Sarhan, "Estimation of system components reliabilities using masked data," *Appl. Math. Comput.*, vol. 136, no. 1, pp. 79–92, 2003.
[18] J. S. Usher, "Weibull component reliability-prediction in the presence of masked data," *IEEE Trans. Rel.*, vol. 45, no. 2, pp. 229–232, Jun. 1996.
[19] A. M. Sarhan and A. H. El Bassiouny, "Estimation of components reliability in a parallel system using masked system life data," *Appl. Math. Comput.*, vol. 138, pp. 61–75, Jun. 2003.
[20] C. Mukhopadhyay, "Maximum likelihood analysis of masked series system lifetime data," *J. Stat. Planning Inference*, vol. 136, no. 3, pp. 803–838, 2006.
[21] I. Guttman, D. K. J. Lin, B. Reiser, and J. S. Usher, "Dependent masking and system life data analysis: Bayesian inference for two-component systems," *Lifetime Data Anal.*, vol. 1, no. 1, pp. 87–100, 1995.
[22] D. K. J. Lin and F. M. Guess, "System life data analysis with dependent partial knowledge on the exact cause of system failure," *Microelectron. Rel.*, vol. 34, no. 3, pp. 535–544, 1994.
[23] L. Kuo and T. Y. Yang, "Bayesian reliability modeling for masked system lifetime data," *Statist. Probab. Lett.*, vol. 17, pp. 229–241, Apr. 2000.
[24] Y. Tan, C. Zhang, and X. Chen, "Bayesian analysis of incomplete data from accelerated life testing with competing failure modes," in *Proc. 8th Int. Conf. Rel., Maintainability Safety*, vol. 58, 2009, pp. 1268–1272.
[25] T.-H. Fan and T.-M. Hsu, "Accelerated life tests of a series system with masked interval data under exponential lifetime distributions," *IEEE Trans. Rel.*, vol. 61, no. 3, pp. 798–808, Sep. 2012.
[26] T.-H. Fan and T.-M. Hsu, "Constant stress accelerated life test on a multiple-component series system under Weibull lifetime distributions," *Commun. Statist.-Theory Methods*, vol. 43, pp. 2370–2383, May 2014.
[27] R. Wang, N. Sha, B. Gu, and X. Xu, "Parameter inference in a hybrid system with masked data," *IEEE Trans. Rel.*, vol. 64, no. 2, pp. 636–644, Jun. 2015.
[28] B. Liu, Y. Shi, J. Cai, X. Bai, and C. Zhang, "Nonparametric Bayesian analysis for masked data from hybrid systems in accelerated lifetime tests," *IEEE Trans. Rel.*, vol. 66, no. 3, pp. 662–676, Sep. 2017.
[29] R. Barlow and F. Proschan, *Statistical Theory of Reliability and Life Testing*. New York, NY, USA: Winston, 1981.
[30] D. Bhattacharya and F. J. Samaniego, "Estimating component characteristics from system failure-time data," *Naval Res. Logistics*, vol. 57, no. 4, pp. 380–389, 2010.
[31] W. Volterman, N. Balakrishnan, K. F. Davies, and H. K. T. Ng, "Exact nonparametric meta-analysis of lifetime data from systems with known signatures," *IEEE Trans. Rel.*, vol. 65, no. 2, pp. 796–801, Jun. 2016.
[32] L. J. M. Aslett, "MCMC for inference on phase-type and masked system lifetime models," Ph.D. dissertation, Trinity College Dublin, Dublin, Ireland, 2012.
[33] H. K. T. Ng, J. Navarro, and N. Balakrishnan, "Parametric inference from system lifetime data under a proportional hazard rate model," *Metrika*, vol. 75, no. 3, pp. 367–388, 2012.
[34] Y. Liang and D. Sun, "Identifiability of masking probabilities in competing risks models with emphasis on Weibull models," *Commun. Statist.-Theory Methods*, vol. 45, no. 7, pp. 2143–2157, 2016.
[35] H. Rinne, *The Weibull Distribution*. London, U.K.: Chapman & Hall, 2008.
[36] L. Tierney, "Markov chains for exploring posterior distributions," *Ann. Statist.*, vol. 22, pp. 1701–1762, Dec. 1994.
[37] C. Robert and G. Casella, *Introducing Monte Carlo Methods With R*. New York, NY, USA: Springer, 2010.
[38] C. D. Lai, M. Xie, and D. N. P. Murthy, "A modified Weibull distribution," *IEEE Trans. Rel.*, vol. 52, no. 1, pp. 33–37, Mar. 2003.
[39] B. J. Flehinger, B. Reiser, and E. Yashchin, "Statistical analysis for masked data," in *Handbook of Statistics*, vol. 20. 2001, ch. 18, pp. 499–522.
[40] R. V. Craiu and B. Reiser, "Inference for the dependent competing risks model with masked causes of failure," *Lifetime Data Anal*, vol. 12, no. 1, pp. 21–33, 2006.

AGATHA S. RODRIGUES received the B.Sc. degree in statistics from the Federal University of Sao Carlos, Brazil, in 2010, and the M.Sc. and Ph.D. degrees in statistics from the University of Sao Paulo, Brazil, in 2013 and 2018, respectively. She has authored or coauthored more than 10 papers. Her current research interests include biostatistics, reliability, and survival analysis.

CARLOS ALBERTO DE B. PEREIRA received the B.Sc. degree in statistics from the National School of Statistical Science, Brazil, in 1968, the M.Sc. degree in statistics from the University of Sao Paulo (USP), Brazil, in 1971, and the Ph.D. degree in statistics from Florida State University, USA, in 1980. He was the Director of the Institute of Mathematics and Statistics, USP, from 1994 to 1998, the Head of the Statistical Department for three times, and the Director of the Bioinformatics Scientific Center, USP, from 2006 to 2009. He is currently a Senior Professor with the Department of Statistics, USP, and a Visiting Professor with the Federal University of Mato Grosso do Sul. He has authored or coauthored more than 200 papers and seven books. He was the President of the Brazilian Statistical Society, from 1998 to 1990. He was an elected member of the International Statistical Institute. He has edited two special editions of Entropy, from 2018 to 2019.

ADRIANO POLPO received the B.Sc. degree in statistics from the State University of Campinas, Brazil, in 2001, and the Ph.D. degree in statistics from the University of Sao Paulo, Brazil, in 2005. He held a postdoctoral position at Florida State University, USA, from 2011 to 2012. He served as the President of ISBrA, the Brazilian Society of Bayesian statistics, from 2013 to 2014. He is currently an Associate Professor with The University of Western Australia, Perth, WA, Australia. He has authored or coauthored more than 30 papers and two books. His current research interest includes reliability and survival analysis, in both theoretical and practical sense.