Semi-parametric Bivariate Singular Proportional Reversed Hazard Model

Sanjay Kumar[1], Debasis Kundu[2] & Sharmishtha Mitra[3]

Department of Mathematics & Statistics, Indian Institute of Technology Kanpur,

Kanpur, Pin 208016, India

Abstract

Modeling bivariate lifetime data with ties and complex hazard structures remains a challenging problem, particularly when classical parametric models fail to capture non-monotone failure behaviors. In this paper, we develop a flexible semi-parametric framework for bivariate distributions that incorporates a singular component to accommodate ties and allows for highly adaptable marginal hazard functions through a piecewise constant specification. The proposed approach does not assume any specific parametric form for the baseline distribution, enabling the modeling of increasing, decreasing, bathtub-shaped, and other non-standard hazard patterns. Dependence between variables is characterized through an induced copula structure, facilitating the study of dependence measures. Parameter estimation is performed using an efficient EM algorithm that involves lower-dimensional optimization and is computationally straightforward. A simulation study demonstrates the accuracy and stability of the proposed estimation procedure. Methods for selecting the number and locations of cut points are also discussed. The practical utility of the proposed framework is illustrated through an application to real data.

Keywords: Marshall-Olkin distribution, Semi-parametric singular distribution, Proportional reversed hazard class, Copula function, Maximum likelihood estimators, Piecewise constant hazard function.

1 Introduction

Bivariate singular distributions play an important role in the analysis of bivariate data sets with ties arising in various fields, including reliability engineering, survival analysis, medical sciences, and social sciences, where assessing the association between two variables is of interest. The class of such distributions was first introduced by Marshall and Olkin [27] and has since become one of the most widely used frameworks for modeling tied observations in bivariate data. The main feature of the Marshall-Olkin distributions is that it has a singular component and the marginals

✉[1] kumar.sanjaymk@gmail.com, ✉[2] kundu@iitk.ac.in, ✉[3] smitra@iitk.ac.in

can be equal with a positive probability. An extensive amount of work on Marshall-Olkin type bivariate singular distributions have been done and several extensions have been proposed in the literature. Some of the well known distributions are Marshall-Olkin bivariate generalized exponential (MOBGE) distribution, see Kundu and Gupta [23]; Marshall-Olkin bivariate inverse Weibull (MOBIW) distribution, see Kundu and Gupta [22]; bivariate proportional reversed hazard model, see Kundu and Gupta [24]; bivariate exponentiated modified Weibull extension distribution, see Gohary et al. [9]; bivariate modified Weibull distrubtion, see Kumar et al. [20, 18, 19] and Kumar [16, 17] etc. In all these cases, authors use parametric baseline distribution, but these are very limited to analyze real bivariate data sets because each has specific hazard rate. Due to this limitation and motivated by the MOBGE distribution, we propose a very general and more flexible bivariate semi-parametric singular distribution whose marginals are belong to proportional reversed class. In this case, we do not assume parametric form of the baselise distribution, instead use a piecewise constant hazard function for baseline distribution. This is a semi-parametric extension of Marshall-Olkin distribution and capable of modeling a large class of bivariate data sets with ties whose marginals have bathtub-shaped, inverted bathtub-shaped, increasing, decreasing, constant and non-monotone hazard functions.

The hazard rate has long been established in the statistical literature, whereas the concept of the reversed hazard rate is comparatively more recent. Notably, the reversed hazard function has found extensive applications in forensic studies and related fields; see, for instance, Block et al. [8] and the references therein. Recently, the proportional reversed hazard model (PRHM) has received considerable attention in the statistical literature. Several PRHMs with various baseline distribution functions $F_0(t)$ have been proposed and studied extensively. For example, the generalized exponential family, see Gupta and Kundu [13, 14, 15], exponentiated Weibull family, see Mudholkar et al. [30, 31, 29], exponentiated Rayleigh of Surles and Padgett [40], generalized linear failure rate model of Sarhan and Kundu [39]. In all these cases, it has been observed that PRHMs can be employed quite effectively for the analysis of lifetime data.

The semi-parametric nature allows flexibility by blending parametric components with non-parametric elements, making the model more adaptable to real-world data. This is particularly useful when the underlying distribution cannot be fully captured by a purely parametric form. Many real-world phenomena involve singular dependencies (e.g., perfect or near-perfect correlation in certain scenarios). By focusing on singular distributions, this study addresses an important aspect often overlooked in standard models. Bivariate distributions are fundamental in multivariate analysis, and developing a more comprehensive class of such distributions enhances our ability to model and understand joint behavior in paired data. The combination of semi-parametric methods with the Marshall-Olkin framework in a bivariate singular setting suggests a novel approach that could expand the theoretical understanding and practical applicability of such models in fields like finance, engineering, and epidemiology.

The primary objective of this article is to introduce a highly general and flexible class of bivariate semi-parametric singular distributions, whose marginal distributions belong to the proportional reversed hazard family. Owing to the wide variety of hazard rate shapes accommodated by this framework, the proposed class offers substantial modeling flexibility. Notably, the MOBGE distribution arises as a special case within this family.

In contrast to fully parametric approaches, no specific functional form is imposed on the baseline distribution. The only assumption considered is that the baseline distribution is absolutely continuous, with a piecewise constant baseline hazard function. The proposed model also admits meaningful physical interpretations in reliability and lifetime analysis contexts.

Various structural and analytical properties of the proposed class are investigated, including the marginal distributions, their shapes, and the joint probability density function. Furthermore, the model possesses a convenient copula representation, which facilitates the derivation of several dependence measures and the study of dependence properties.

We further developed the classical inference procedure based on the proposed bivariate distribution. In developing the estimation procedure, we assume that the baseline distribution has piecewise constant hazard function with a number of cut points. The maximum likelihood estimators (MLEs) of unknown parameters cannot be obtained in closed form. We implement the expectation-maximization (EM) algorithm to compute the MLEs and it reduces the dimension of the optimization problem. The proposed EM algorithm is easy to implement and demonstrates satisfactory numerical performance. We develop a procedure for estimating the cut points and their number in a piecewise constant hazard function. A bivariate goodness-of-fit test is also addressed. Simulation experiments have been performed to see how the proposed new model and implementation of the EM algorithm for the inference procedure work in practice. The EM algorithm perfomed quite well and the results are satisfactory. A bivariate real data set with ties have been analyzed for to show how the proposed model and method work in practice.

The remainder of the article is organized as follows. Section 2 presents the preliminaries. Section 3 introduces the semi-parametric distribution, while Section 4 discusses the associated copula. Maximum likelihood estimation is developed in Section 5, and the estimation of cut points is addressed in Section 6. A bivariate goodness-of-fit test is presented in Section 7. Extensive simulation studies are carried out in Section 8, followed by a real-data analysis in Section 9. Finally, Section 10 concludes the article.

2 Preliminary

2.1 Proportional Reversed Hazard Class

Let us recall the reversed hazard rate (RHR) for a non-negative continuous random variable T with probability density function (PDF) $f(t)$ and cumulative distribution function (CDF) $F(t)$. The RHR, introduced by Barlow et al. [6], is defined as

$$r(t) = \frac{f(t)}{F(t)}, \quad t \geq 0. \tag{1}$$

Suppose $F_0(t)$ is an absolutely continuous distribution function defined on the positive real axis, referred to as the baseline distribution function, and let $S_0(t) = 1 - F_0(t)$ denote the corresponding SF. A class of distribution functions $\{F_{PRHC}(t; \alpha) : \alpha > 0\}$ is said to be a proportional reversed hazard class (PRHC) if its CDF $F_{PRHC}(t; \alpha)$ can be expressed as

$$F_{PRHC}(t; \alpha) = \big[F_0(t)\big]^{\alpha}; \quad t > 0,$$

with the corresponding PDF given by

$$f_{PRHC}(t; \alpha) = \alpha f_0(t) \big[F_0(t)\big]^{\alpha - 1}; \quad t > 0,$$

where $f_0(t)$ denotes the baseline PDF. Clearly, $F_{PRHC}(t;\alpha)$ is also an absolutely continuous distribution function. Throughout this article, we assume that $F_0(t) = 0$ for $t \le 0$, although all the results can be extended for general $F_0(\cdot)$ also. The results remain valid even if $F_0(t)$ has bounded support. From now it will be denoted by PRHC(F_0, α). Note that if $f_0(t)(f_{PRHC}(t;\alpha))$ and $r_0(t)(r_{PRHC}(t;\alpha))$ denote the PDF and the reversed hazard function (RHF) associated with $F_0(t)(F_{PRHC}(t;\alpha))$, then

$$r_{PRHC}(t;\alpha) = \frac{f_{PRHC}(t;\alpha)}{F_{PRHC}(t;\alpha)} = \alpha\frac{f_0(t)}{F_0(t)} = \alpha r_0(t).$$

The PRHC was originally introduced by Lehmann [26] in the context of hypothesis testing. Subsequently, Gupta et al. [11] investigated several important properties of various PRHCs. For some notable applications of PRHC in breast cancer studies, refer to Tsodikov et al. [41]. Several structural properties of PRHC are straightforward to observe. For instance, the distribution of the random variable T is always skewed, even when the baseline distribution function $F_0(t)$ is symmetric. Specifically, the distribution is positively skewed when $\alpha > 1$ and negatively skewed when $\alpha < 1$. A recent review of the PRHC can be found in Gupta et al. [12].

It can be easily seen that the generalized exponential family, see Gupta and Kundu [13, 14, 15], exponentiated Weibull family, see Mudholkar et al. [30, 31, 29], exponentiated Rayleigh of Surles and Padgett [40], generalized linear failure rate model of Sarhan and Kundu [39], are members of the PRHC. An extensive amount of work has been done in developing several properties and establishing the inference procedures of several univariate PRHC of distributions, see for example Lehmann [26] and Block et al [8] in this respect.

2.2 Piecewise Constant Hazard Function

We assume here a piecewise constant approximation (PCA) of the baseline hazard function $h_0(t)$, i.e. it is constructed based on: (i). the number of positive constant (M) and (ii). the associated cut (or change) time points $0 = \tau_0 < \tau_1 < \ldots < \tau_{M-1} < \tau_M = \infty$. At the beginning both M and the cut points are assumed to be known. Later we will discuss how to estimate these constants. Further, it is assumed that $h_0(t)$ has the following form

$$h_0(t) = \sum_{j=1}^{M} c_j I_{[\tau_{j-1},\tau_j)}(t), \tag{2}$$

where c_j's are unknown positive constants, and

$$I_{[\tau_{j-1},\tau_j)}(t) = \begin{cases} 1, & t \in [\tau_{j-1},\tau_j) \\ 0, & \text{otherwise} \end{cases}$$

is the indicator function. It is clear that under this set up, $c_1, \ldots, c_M$ are identifiable only up to a multiplicative constant. Hence, without loss of generality it is assumed that $c_M = 1$. Note that when $M = 1$, $h_0(t)$ becomes an exponential hazard with scale parameter one. The distribution function associated with a piecewise constant hazard (PCH) function, as defined in (2), represents a highly flexible class of lifetime distributions. This flexibility allows the hazard function to exhibit a wide range of shapes, including increasing, decreasing, bathtub-shaped, univerted bathtub-shaped, constant and unimodal forms. Importantly, there is no need to assume the shape of the hazard function a priori; instead, it is determined in a data-driven manner, allowing the observed data to guide the estimation of the distribution. While this approach is not fully non-parametric, thus avoiding the need for very large sample sizes, it also does not rely entirely on rigid parametric assumptions. Rather, it occupies a middle ground as a semi-parametric method, enabling effective estimation of both hazard and survival functions from reasonably sized data sets. The use of piecewise constant hazard models in the analysis of lifetime data is well-established in the statistical literature; see, for instance, Murphy [32], Goodman et al. [10], Balakrishnan et al. [5], and the references therein.

Now let us compute $F_0(t)$ and $f_0(t)$ which will be needed later. Note that, the cumulative baseline hazard function $H_0(t)$ can be written as

$$H_0(t) = \sum_{j=1}^{M} c_j(\min\{t, \tau_j\} - \tau_{j-1}) I_{[\tau_{j-1}, \tau_M)}(t).$$

Equivalenty for $1 \leq m \leq M$, we have

$$H_0(t) = \sum_{j=1}^{m-1} c_j(\tau_j - \tau_{j-1}) + c_m(t - \tau_{m-1}) \qquad \text{if} \quad \tau_{m-1} \leq t < \tau_m.$$

Once $H_0(t)$ is obtained, $F_0(t)$ and $f_0(t)$ can be obtained from the following formulas

$$F_0(t) = 1 - e^{-H_0(t)} \quad \text{and} \quad f_0(t) = h_0(t)e^{-H_0(t)}, \quad \text{where} \quad H_0(t) = \int_0^t h_0(u)du.$$

The semi-parametric PRHC of distribution is given as

$$F_{PRHC}(t; \alpha) = [1 - e^{-H_0(t)}]^{\alpha}.$$

For $M = 1$, F_{PRHC} is the CDF of generalized exponential distribution with scale parameter one and shape parameter α, denoted as $GE(1, \alpha)$.

3 Semi-parametric Bivariate Proportional Reversed Hazard Class

Kundu and Gupta [23] recently introduced the MOBGE distribution, which has generated significant interest in the study of bivariate lifetime models. Further, Kundu and Gupta [24] studied a class of bivariate models with proportional reversed hazard marginals based on parametric baseline distributions. Motivated by these works and by the univariate proportional reversed hazard class (PRHC) of distributions, we propose a more general and flexible bivariate proportional reversed hazard class of distributions that does not require specifying a parametric baseline distribution. This generalized framework is referred to as the semi-parametric bivariate proportional reversed hazard (BPRH) model.

The BPRH model is formally defined as follows. Let $V_1 \sim \text{PRHC}(F_0, \alpha_1)$, $V_2 \sim \text{PRHC}(F_0, \alpha_2)$, $V_3 \sim \text{PRHC}(F_0, \alpha_3)$, where V_1, V_2 and V_3 are independent RVs. Define

$$X_1 = \max\{V_1, V_3\} \quad \text{and} \quad X_2 = \max\{V_2, V_3\}$$

then (X_1, X_2) is said to have BPRH distribution with the baseline CDF $F_0(\cdot)$ and with parameters α_1, α_2 and α_3. From now on it will be denoted by $BPRH(F_0, \alpha_1, \alpha_2, \alpha_3)$. Before deriving the properties let us first provide some physical interpretation of the BPRH model.

Stress Model: Originally introduced by Kundu and Gupta [23], this model considers a two-component system where each component experiences its own stress, denoted by V_1 and V_2, along with a common stress V_3 that affects both components simultaneously. The effective stresses acting on the components are therefore given by $X_1 = \max V_1, V_3$ and $X_2 = \max V_2, V_3$.
Maintenance Model: Consider a system with two components that undergo both individual and common maintenance. Let V_1 and V_2 represent the lifetime improvements due to individual maintenance, while V_3 denotes the additional improvement from a shared maintenance effort. The resulting lifetimes of the components are then $X_1 = \max V_1, V_3$ and $X_2 = \max V_2, V_3$ (see Kundu and Gupta [23]).
The generation from the BPRH of distributions is straightforward. One can adopt the following simple generation technique to generate random samples from BPRH.

Step 1 Generate $V_1 \sim \text{PRHC}(F_0, \alpha_1)$, $V_2 \sim \text{PRHC}(F_0, \alpha_2)$, and $V_3 \sim \text{PRHC}(F_0, \alpha_3)$ independently.

Step 2 Set $X_1 = \max\{V_1, V_3\}$ and $X_2 = \max\{V_2, V_3\}$.

3.1 Some Distributional Properties

The joint CDF of (X_1, X_2) can be obtained in the following theorem, see Kundu and Gupta [21]

Theorem 3.1. *If $(X_1, X_2) \sim BPRH(F_0, \alpha_1, \alpha_2, \alpha_3)$ then the joint CDF of (X_1, X_2) for $z = \min\{x_1, x_2\}$ can be written as*

$$\begin{aligned}
F_{X_1,X_2}(x_1, x_2) &= P(X_1 \leq x_1, X_2 \leq x_2) = P(V_1 \leq x_1, V_2 \leq x_2, V_3 \leq z) \\
&= [F_0(x_1)]^{\alpha_1}[F_0(x_2)]^{\alpha_2}[F_0(z)]^{\alpha_3} \\
&= \begin{cases} [F_0(x_1)]^{\alpha_1+\alpha_3}[F_0(x_2)]^{\alpha_2} & \text{if} \;\; x_1 < x_2 \\ [F_0(x_1)]^{\alpha_1}[F_0(x_2)]^{\alpha_2+\alpha_3} & \text{if} \;\; x_1 > x_2 \\ [F_0(x)]^{\alpha_1+\alpha_2+\alpha_3} & \text{if} \;\; x_1 = x_2 = x. \end{cases}
\end{aligned}$$

Proof. It is trivial. The proof is avoided. □

The associated joint PDF of (X_1, X_2) can be obtained as follows

Theorem 3.2. *If $(X_1, X_2) \sim BPRH(F_0, \alpha_1, \alpha_2, \alpha_3)$ then the joint PDF of (X_1, X_2) is*

$$f_{X_1,X_2}(x_1, x_2) = \begin{cases} (\alpha_1+\alpha_3)\alpha_2 f_0(x_1)f_0(x_2)[F_0(x_1)]^{\alpha_1+\alpha_3-1}[F_0(x_2)]^{\alpha_2-1} & \text{if} \;\; x_1 < x_2 \\ \alpha_1(\alpha_2+\alpha_3)f_0(x_1)f_0(x_2)[F_0(x_1)]^{\alpha_1-1}[F_0(x_2)]^{\alpha_2+\alpha_3-1} & \text{if} \;\; x_1 > x_2 \\ \alpha_3 f_0(x)[F_0(x)]^{\alpha_1+\alpha_2+\alpha_3-1} & \text{if} \;\; x_1 = x_2 = x. \end{cases}$$

Proof. The proof can be obtained along the same line as the proof of Theorem 2.2 of Kundu and Gupta [23]. The details are avoided. □

The CDF and PDF of $(X_1, X_2) \sim BPRH(F_0, \alpha_1, \alpha_2, \alpha_3)$ also can be written as

$$F_{X_1,X_2}(x_1, x_2) = \begin{cases} F_{PRHC}(x_1; \alpha_1 + \alpha_3)F_{PRHC}(x_2; \alpha_2) & \text{if } \ x_1 < x_2 \\ F_{PRHC}(x_1; \alpha_1)F_{PRHC}(x_1; \alpha_2 + \alpha_3) & \text{if } \ x_1 > x_2 \\ F_{PRHC}(x; \alpha_1 + \alpha_2 + \alpha_3) & \text{if } \ x_1 = x_2 = x \end{cases}$$

and

$$f_{X_1,X_2}(x_1, x_2) = \begin{cases} f_{PRHC}(x_1; \alpha_1 + \alpha_3)f_{PRHC}(x_2; \alpha_2) & \text{if } \ x_1 < x_2 \\ f_{PRHC}(x_1; \alpha_1)f_{PRHC}(x_2; \alpha_2 + \alpha_3) & \text{if } \ x_1 > x_2 \\ \dfrac{\alpha_3}{\alpha_1 + \alpha_2 + \alpha_3} f_{PRHC}(x; \alpha_1 + \alpha_2 + \alpha_3) & \text{if } \ x_1 = x_2 = x \end{cases}$$

respectively. The BPRH distribution has both an absolutely continuous part and a singular part, similar to Marshall and Olkin's bivariate exponential model. It should be mentioned here that the joint PDF does not exist with respect two dimensional Lebesgue measure as it is not an absolutely continuous distribution function. Here, the joint PDF is with respect to two dimensional Lebesgue measure on the set $\{(x_1, x_2) : x_1 \neq x_2\}$ and with respect to one dimensional Lebesgue measure on the set $\{(x_1, x_2) : x_1 = x_2\}$, similar to MOBE distribution, see for example Bemis et al. [7]. The following theorem establishes the unique decomposition of the joint CDF.

Theorem 3.3. *If $(X_1, X_2) \sim BPRH(F_0, \alpha_1, \alpha_2, \alpha_3)$, then the joint CDF of (X_1, X_2) can be written as a mixture of an absolutely continuous part and a singular part as follows;*

$$F_{X_1,X_2}(x_1, x_2) = \frac{\alpha_1 + \alpha_2}{\alpha_1 + \alpha_2 + \alpha_3} F_{ac}(x_1, x_2) + \frac{\alpha_3}{\alpha_1 + \alpha_2 + \alpha_3} F_{si}(x_1, x_2)$$

where $F_{ac}(.,.)$ is the absolutely continuous part and $F_{si}(.,.)$ is the singular part. It is to note here that for $z = \max\{x_1, x_2\}$,

$$F_{si}(x_1, x_2) = [F_0(z)]^{\alpha_1+\alpha_2+\alpha_3}$$

and $F_{ac}(x_1, x_2)$ can be obtained by subtraction as

$$F_{ac}(x_1, x_2) = \frac{\alpha_1 + \alpha_2 + \alpha_3}{\alpha_1 + \alpha_2}[F_0(x_1)]^{\alpha_1}[F_0(x_2)]^{\alpha_2}[F_0(z)]^{\alpha_3} - \frac{\alpha_3}{\alpha_1 + \alpha_2}[F_0(z)]^{\alpha_1+\alpha_2+\alpha_3}.$$

Proof. The proof is trivial so we omitted it. □

Corollary 3.3.1. *The corresponding joint PDF (X_1, X_2) also can be written as a mixture of absolutely continuous and singular parts as*

$$f_{X_1,X_2}(x_1, x_2) = \frac{\alpha_1 + \alpha_2}{\alpha_1 + \alpha_2 + \alpha_3} f_{ac}(x_1, x_2) + \frac{\alpha_3}{\alpha_1 + \alpha_2 + \alpha_3} f_{si}(x_1, x_2)$$

where where $f_{ac}(.,.)$ is the absolutely continuous part and $f_{si}(.,.)$ is the singular part. It is to note here that for $z = \max\{x_1, x_2\}$,

$$f_{ac}(x_1, x_2) = \frac{\alpha_1 + \alpha_2 + \alpha_3}{\alpha_1 + \alpha_2} \begin{cases} (\alpha_1 + \alpha_3)\alpha_2 f_0(x_1) f_0(x_2) [F_0(x_1)]^{\alpha_1+\alpha_3-1} [F_0(x_2)]^{\alpha_2-1} & \text{if } x_1 < x_2 \\ \alpha_1(\alpha_2 + \alpha_3) f_0(x_1) f_0(x_2) [F_0(x_1)]^{\alpha_1-1} [F_0(x_2)]^{\alpha_2+\alpha_3-1} & \text{if } x_1 > x_2 \end{cases}$$

and

$$f_{si}(x) = f_{PRHC}(x; \alpha_1 + \alpha_2 + \alpha_3) = (\alpha_1 + \alpha_2 + \alpha_3) f_0(x) [F_0(x)]^{\alpha_1+\alpha_2+\alpha_3-1}.$$

Proof. The proof is trivial so we omitted it. □

The following theorem provides the marginal and conditional results of the BPRH distribution.

Theorem 3.4. *If $(X_1, X_2) \sim BPRH(F_0, \alpha_1, \alpha_2, \alpha_3)$, then*

(1) $X_1 \sim PRHC(F_0, \alpha_1 + \alpha_3)$ *and* $X_2 \sim PRHC(F_0, \alpha_2 + \alpha_3)$.

(2) $\max\{X_1, X_2\} \sim PRHC(F_0, \alpha_1 + \alpha_2 + \alpha_3)$.

(3) $X_1 | \{X_1 > X_2\} \sim PRHC(F_0, \alpha_1 + \alpha_2 + \alpha_3)$.

(4) $X_2 | \{X_1 < X_2\} \sim PRHC(F_0, \alpha_1 + \alpha_2 + \alpha_3)$.

Proof. The Proof is not difficult to obtain. Hence, the details are avoided. □

The above results will help us to find the initial estimates of parameters in our proposed EM algorithm to compute the maximum likelihood estimators. It can easily see that

$$P(X_1 < X_2) = \frac{\alpha_1}{\alpha_1 + \alpha_3 + \alpha_3}, \; P(X_1 > X_2) = \frac{\alpha_2}{\alpha_1 + \alpha_3 + \alpha_3}, \; P(X_1 = X_2) = \frac{\alpha_3}{\alpha_1 + \alpha_3 + \alpha_3}.$$

Note that the joint SF can be written by using the following relation

$$S_{X_1,X_2}(x_1, x_2) = 1 - F_{PRHC}(x_1; \alpha_1 + \alpha_3) - F_{PRHC}(x_2; \alpha_2 + \alpha_3) + F_{X_1,X_2}(x_1, x_2). \tag{3}$$

The correlation coefficient of X_1 and X_2 varies from zero to one and the correlation tend to one as $\alpha_3 \to \infty$. If $\alpha_3 = 0$, then X_1 and X_2 are independent.

3.2 Bivariate Reversed Hazard Rates

We now turn to the study of the bivariate reversed failure rate associated with the proposed BPRH model. It is important to note that multiple formulations of the bivariate reversed hazard rate exist in the literature. For instance, the definitions provided by Block et al. [8] and Sankaran and Gleeja [38] are developed under the assumption of absolute continuity of the joint distribution. However, these formulations are not suitable in the present context, since the BPRH model does not belong to the class of absolutely continuous distributions.

Furthermore, the approach of Sankaran and Gleeja [38] does not necessarily ensure a unique characterization of the joint distribution function, which limits its practical usefulness. To overcome these limitations, we adopt the definition proposed by Roy [37], where the bivariate reversed hazard rate is formulated as a vector, providing a more appropriate and consistent framework for the BPRH model. Hence, we have used the definition of Roy [37], as defined bivariate RHR as a vector

$$r(x_1, x_2) = \big(r_1(x_1, x_2), r_2(x_1, x_2)\big),$$

where,

$$\begin{aligned} r_i(x_1, x_2) &= \lim_{\Delta x_i \to 0} P(x_i - \Delta x_i < X_i \le x_i \big| X_1 \le x_1, X_2 \le x_2)/\Delta x_i \\ &= \partial \log F(x_1, x_2)/\partial x_i, \qquad i = 1, 2. \end{aligned}$$

For $i = 1, 2$, $r_1(x_1, x_2)\Delta x_i$ is the probability of failure of the first component in the interval $(x_1 - \Delta x_1, x_1]$ given that it has failed before x_1 and the second has failed before x_2. The interpretation for $r_2(x_1, x_2)$ is similar. From Roy [37], it follows that $r_i(x_1, x_2)$ determine $F(x_1, x_2)$ uniquely by the relationships

$$F(x_1, x_2) = \exp\left[-\int_{x_1}^{\infty} r_1(u, \infty)\, du - \int_{x_2}^{\infty} r_2(x_1, u)\, du\right], \tag{4}$$

or

$$F(x_1, x_2) = \exp\left[-\int_{x_1}^{\infty} r_1(u, x_2)\, du - \int_{x_2}^{\infty} r_2(\infty, u)\, du\right], \tag{5}$$

where $r_1(x_1, \infty) = r_1(x_1)$ and $r_2(\infty, x_2) = r_2(x_2)$ are the marginal RHR of X_1 and X_2, respectively. Furthermore, from (4) and (5), it follows that X_1 and X_2 are independent if and only if

$$r_i(x_1, x_2) = r_i(x_i), \quad i = 1, 2.$$

Theorem 3.5. *If $(X_1, X_2) \sim BPRH(F_0, \alpha_1, \alpha_2, \alpha_3)$, then the reverse hazard gradient of the BPRH model are defined as*

$$r_1(x_1, x_2) = \begin{cases} (\alpha_1 + \alpha_3) r_0(x_1) & \text{if } x_1 < x_2 \\ \alpha_1 r_0(x_1) & \text{if } x_1 > x_2 \end{cases}$$

$$r_2(x_1, x_2) = \begin{cases} \alpha_2 r_0(x_2) & \text{if } x_1 < x_2 \\ (\alpha_2 + \alpha_3) r_0(x_2) & \text{if } x_1 > x_2. \end{cases}$$

Note that as $\alpha_3 \to 0$, we have $r_i(x_1, x_2) = r_i(x_i)$ for $i = 1, 2$, implying that X_1 and X_2 become independent.

Another definition of the reversed hazard rate (RHR), which plays a vital role in the analysis of dependent data, is given by

$$r^*(t_1, t_2) = \left(r_1^*(t_1, t_2),\, r_2^*(t_1, t_2)\right),$$

where

$$r_i^*(t_1, t_2) = \lim_{\Delta t_i \to 0} \frac{\mathbb{P}\left(t_i - \Delta t_i \le T_i \le t_i \,|\, T_i \le t_i,\, T_j = t_j\right)}{\Delta t_i} = \frac{f(t_i \mid T_j = t_j)}{F(t_i \mid T_j = t_j)}, \quad i = 1, 2, \quad i \ne j, \tag{2.6}$$

with $f(t_i \mid T_j = t_j)$ denoting the conditional density function of T_i given $T_j = t_j$ and $F(t_i \mid T_j = t_j)$ denoting the conditional distribution function of T_i given $T_j = t_j$.

Thus, definition (2.6) is nothing but the univariate RHR of the conditional variable T_i given $T_j = t_j$. Since conditional distributions, in general, do not uniquely determine the joint density, (2.6) does not uniquely provide $F(t_1, t_2)$.

However, it can be shown that T_1 and T_2 are independent if and only if

$$r_i^*(t_1, t_2) = r_i(t_i), \quad i = 1, 2.$$

Furthermore, if T_1 and T_2 are independent, then

$$r_i(t_1, t_2) = r_i^*(t_1, t_2).$$

Therefore, the ratio

$$\frac{r_i^*(t_1, t_2)}{r_i(t_1, t_2)}$$

can be considered as a measure of association between T_1 and T_2, which is analogous to the well-known cross ratio function of Oakes Oakes [35].

Theorem 3.6. *If $(X_1, X_2) \sim BPRH(F_0, \alpha_1, \alpha_2, \alpha_3)$, then we have*

$$r_1^*(x_1, x_2) = \begin{cases} (\alpha_1 + \alpha_3) r_0(x_1) & \text{if } x_1 < x_2 \\ \alpha_1 r_0(x_1) & \text{if } x_1 > x_2 \end{cases}$$

$$r_2^*(x_1, x_2) = \begin{cases} \alpha_2 r_0(x_2) & \text{if } x_1 < x_2 \\ (\alpha_2 + \alpha_3) r_0(x_2) & \text{if } x_1 > x_2. \end{cases}$$

3.3 Different properties

In probability theory and statistics, a stochastic order provides a formal framework for comparing random variables in terms of their magnitude. Depending on the probabilistic criterion employed, several types of stochastic orders can be defined. Among the most important and widely used are the usual stochastic order, the failure rate order, and the likelihood ratio order.

Let (X_1, X_2) be a bivariate random vector with joint CDF $F_{X_1,X_2}(x_1, x_2)$ and marginal CDFs $F_{X_1}(x_1)$ and $F_{X_2}(x_2)$. The pair (X_1, X_2) is said to be *positively quadrant dependent (PQD)* if, for all (x_1, x_2),

$$F_{X_1,X_2}(x_1, x_2) \;\geq\; F_{X_1}(x_1) \cdot F_{X_2}(x_2).$$

If $(X_1, X_2) \sim BPRH(F_0, \alpha_1, \alpha_2, \alpha_3)$, then (X_1, X_2) exhibits the PQD. It indicates a positive dependence structure, but not necessarily a linear one. Notably, the Marshall–Olkin copula, often used in reliability and survival analysis, also satisfies the PQD property, which is generally weaker than positive association or correlation.

Total Positivity of Order Two (TP$_2$): For a bivariate random vector (X, Y) with joint probability density function (pdf) $f_{X,Y}(x, y)$, the pair (X, Y) is said to be *TP$_2$* if, for all $x_1 < x_2$ and $y_1 < y_2$,

$$f_{X,Y}(x_1, y_1)\, f_{X,Y}(x_2, y_2) \;\geq\; f_{X,Y}(x_1, y_2)\, f_{X,Y}(x_2, y_1).$$

Equivalently, the joint pdf is *log-supermodular*, that is,

$$\frac{\partial^2}{\partial x\,\partial y}\log f_{X,Y}(x,y)\ \geq\ 0,$$

whenever the second-order derivative exists. This property indicates a strong positive dependence structure, which is stronger than positive quadrant dependence (PQD), and is often applied in reliability theory, survival analysis, and stochastic ordering.

Theorem 3.7. *If $(X_1, X_2) \sim BPRH(F_0, \alpha_1, \alpha_2, \alpha_3)$, then (X_1, X_2) has the total positivity of order two (TP_2) property.*

Proof. Note that (X_1, X_2) has the TP_2 property if and only if for any $\boldsymbol{x} = (u_1, v_1)$, $\boldsymbol{y} = (u_2, v_2)$ whenever, $0 < u_1 < u_2$, $0 < v_1 < v_2$, then

$$f_{X_1,X_2}(u_1, v_1) f_{X_1,X_2}(u_2, v_2) \geq f_{X_1,X_2}(u_2, v_1) f_{X_1,X_2}(u_1, v_2)$$

Now the above inequality can be shown by considering all possible different twenty four cases as follows;
Case 1: $v_1 < v_2 < u_1 < u_2$. In this case, it is clear that

$$f_{X_1,X_2}(u_1, v_1) f_{X_1,Y_2}(u_2, v_2) - f_{X_1,X_2}(u_2, v_1) f_{X_1,X_2}(u_1, v_2) = 0.$$

Case 2: $v_1 < u_1 < v_2 < u_2$. In this case it can be easily observed by simple calculation that if

$$\left(\frac{F_0(u_1)}{F_0(v_2)}\right)^{\theta_3} \geq \frac{\theta_2(\theta_1 + \theta_3)}{\theta_1(\theta_2 + \theta_3)},$$

then (X_1, X_2) has the total positivity of order two (TP_2) property. Similarly, for other cases also we can find the conditions such that (X_1, X_2) has the total positivity of order two (TP_2) property. □

4 Associated Copula

Now we would like to provide the CDF copula associated with the BPRH model. Based on the marginal distributions, namely $PRHC(F_0, \alpha_1 + \alpha_3)$ and $PRHC(F_0, \alpha_2 + \alpha_3)$, respectively. The BPRH model has the following CDF copula for $0 \leq u, v \leq 1$;

$$C(u,v) = F[F_1^{-1}(u), F_2^{-1}(v)] = \begin{cases} u v^{\frac{\alpha_2}{\alpha_2+\alpha_3}} & \text{if } u < v^{\frac{\alpha_1+\alpha_3}{\alpha_2+\alpha_3}} \\ u^{\frac{\alpha_1}{\alpha_1+\alpha_3}} v & \text{if } u \geq v^{\frac{\alpha_1+\alpha_3}{\alpha_2+\alpha_3}}. \end{cases}$$

If we take $\lambda = \frac{\alpha_3}{\alpha_1+\alpha_3}$ and $\delta = \frac{\alpha_3}{\alpha_2+\alpha_3}$, then the copula can be equivalently expressed as

$$C(u,v) = \begin{cases} u v^{1-\delta} & \text{if } u^{\lambda} < v^{\delta} \\ u^{1-\lambda} v & \text{if } u^{\lambda} \geq v^{\delta}. \end{cases}$$

Further, in the symmetric case, i.e., when $\alpha_1 = \alpha_2$, if we denote $\gamma = \frac{\alpha_3}{\alpha_1+\alpha_3} = \frac{\alpha_3}{\alpha_2+\alpha_3}$,

$$C(u,v) = \begin{cases} uv^{1-\gamma} & \text{if } u < v \\ u^{1-\gamma}v & \text{if } u \geq v. \end{cases}$$

This copula corresponds to the well-known Marshall–Olkin type copula, see Nelson [34]. Consequently, the measures of dependence, namely Kendall's τ and Spearman's ρ, are given by

$$\tau = \frac{\alpha_1\alpha_2}{\alpha_1 - \alpha_1\alpha_2 + \alpha_2} \quad \text{and} \quad \rho = \frac{3\alpha_1\alpha_2}{2\alpha_1 - \alpha_1\alpha_2 + 2\alpha_2}.$$

5 Maximum Likelihood Estimators

In this section, we focus on the computation of the maximum likelihood estimates (MLEs) of the unknown parameters of the BPRH distribution within a semi-parametric framework, where no specific parametric form is assumed on the baseline distribution. Previous studies, such as those by Kundu and Gupta [23, 22, 25], have derived MLEs under fully parametric assumptions. In contrast, we adopt a more flexible approach by modeling the baseline hazard function $h_0(t)$ as a piecewise constant hazard (PCH) function, as described in Subsection 2.2, enabling the model to capture a broad range of baseline behaviors.

5.1 Log-likelihood function

Consider a random sample of size n, denoted by $\mathcal{D} = \{(x_{11}, x_{21}), (x_{12}, x_{22}), \ldots, (x_{1n}, x_{2n})\}$, drawn from a BPRH$(F_0, \alpha_1, \alpha_2, \alpha_3)$ distribution. Within the semi-parametric framework described in Subsection 2.2, the aim is to estimate a total of $M + 2$ unknown parameters, namely α_1, α_2, α_3, and $c_1, c_2, \ldots, c_{M-1}$. At this stage, the number of intervals M and the associated cut points $\tau_0 < \tau_1 < \cdots < \tau_M$ are assumed to be known. In Section 6, a more general setting is considered, where both M and the locations of the cut points are treated as unknown and are estimated from the data. For notational convenience, partition the index set $\{1, 2, \ldots, n\}$ into the following three subsets:

$$I_1 = \{i : x_{1i} < x_{2i}\}, \quad I_2 = \{i : x_{1i} > x_{2i}\}, \quad I_3 = \{i : x_{1i} = x_{2i} = x_i\},$$

with cardinalities $n_j = |I_j|$, for $j = 1, 2, 3$. To obtain the MLEs of the unknown parameters, we first derive the contributions to the log-likelihood function corresponding to an observation (x_1, x_2) under the different cases.

Case 1: When $x_1 < x_2$

Suppose $\tau_{m-1} \leq x_1 < \tau_m,\ \tau_{k-1} \leq x_2 < \tau_k$ for $1 \leq m, k \leq M$, we have

$$
\begin{aligned}
& \ln(\alpha_1+\alpha_3) + \ln\alpha_2 + \ln f_0(x_1) + \ln f_0(x_2) + (\alpha_1+\alpha_3-1)\ln F_0(x_1) + (\alpha_2-1)\ln F_0(x_2) \\
= \; & \ln(\alpha_1+\alpha_3) + \ln\alpha_2 + \ln h_0(x_1) + \ln h_0(x_2) - H_0(x_1) - H_0(x_2) + (\alpha_1+\alpha_3-1)\ln\left[1-e^{-H_0(x_1)}\right] \\
& +(\alpha_2-1)\ln\left[1-e^{-H_0(x_2)}\right] \\
= \; & \ln(\alpha_1+\alpha_3) + \ln\alpha_2 + \ln c_m + \ln c_k - \left\{\sum_{j=1}^{m-1} c_j(\tau_j-\tau_{j-1}) + c_m(x_1-\tau_{m-1})\right\} \\
& -\left\{\sum_{j=1}^{k-1} c_j(\tau_j-\tau_{j-1}) + c_k(x_2-\tau_{k-1})\right\} \\
& +(\alpha_1+\alpha_3-1)\ln\left[1-\exp\left\{-\sum_{j=1}^{m-1} c_j(\tau_j-\tau_{j-1}) - c_m(x_1-\tau_{m-1})\right\}\right] \\
& +(\alpha_2-1)\ln\left[1-\exp\left\{-\sum_{j=1}^{k-1} c_j(\tau_j-\tau_{j-1}) - c_k(x_2-\tau_{k-1})\right\}\right]. \qquad (6)
\end{aligned}
$$

Case 2: When $x_2 < x_1$
Suppose $\tau_{m-1} \leq x_1 < \tau_m,\ \tau_{k-1} \leq x_2 < \tau_k$ for $1 \leq m, k \leq M$, we have

$$
\begin{aligned}
& \ln\alpha_1 + \ln(\alpha_2+\alpha_3) + \ln f_0(x_1) + \ln f_0(x_2) + (\alpha_1-1)\ln F_0(x_1) + (\alpha_2+\alpha_3-1)\ln F_0(x_2) \\
= \; & \ln\alpha_1 + \ln(\alpha_2+\alpha_3) + \ln h_0(x_1) + \ln h_0(x_2) - H_0(x_1) - H_0(x_2) + (\alpha_1-1)\ln\left[1-e^{-H_0(x_1)}\right] \\
& +(\alpha_2+\alpha_3-1)\ln\left[1-e^{-H_0(x_2)}\right] \\
= \; & \ln\alpha_1 + \ln(\alpha_2+\alpha_3) + \ln c_m + \ln c_k - \left\{\sum_{j=1}^{m-1} c_j(\tau_j-\tau_{j-1}) + c_m(x_1-\tau_{m-1})\right\} \\
& -\left\{\sum_{j=1}^{k-1} c_j(\tau_j-\tau_{j-1}) + c_k(x_2-\tau_{k-1})\right\} \\
& +(\alpha_1-1)\ln\left[1-\exp\left\{-\sum_{j=1}^{m-1} c_j(\tau_j-\tau_{j-1}) - c_m(x_1-\tau_{m-1})\right\}\right] \\
& +(\alpha_2+\alpha_3-1)\ln\left[1-\exp\left\{-\sum_{j=1}^{k-1} c_j(\tau_j-\tau_{j-1}) - c_k(x_2-\tau_{k-1})\right\}\right]. \qquad (7)
\end{aligned}
$$

Case 3: When $x_2 = x_1 = x$

Suppose $\tau_{m-1} \leq x < \tau_m$ for $1 \leq m \leq M$, we have

$$
\begin{aligned}
& \ln\alpha_3 - \ln(\alpha_1+\alpha_2+\alpha_3) + \ln f_0(x) + (\alpha_1+\alpha_2+\alpha_3-1)\ln F_0(x) \\
= \; & \ln\alpha_3 - \ln(\alpha_1+\alpha_2+\alpha_3) + \ln h_0(x) - H_0(x) + (\alpha_1+\alpha_2+\alpha_3-1)\ln\left[1-e^{-H_0(x)}\right] \\
= \; & \ln\alpha_3 - \ln(\alpha_1+\alpha_2+\alpha_3) + \ln c_m - \left\{\sum_{j=1}^{m-1} c_j(\tau_j-\tau_{j-1}) + c_m(x-\tau_{m-1})\right\} \\
& +(\alpha_1+\alpha_2+\alpha_3-1)\ln\left[1-\exp\left\{\sum_{j=1}^{m-1} c_j(\tau_j-\tau_{j-1}) + c_m(x-\tau_{m-1})\right\}\right]. \qquad (8)
\end{aligned}
$$

Based on the above three cases, the log-likelihood function corresponding to the observed sample $\mathcal{D}$ can be constructed in a straightforward manner. To simplify the notation, we re-express the observations x_{1i}, x_{2i}, and x_i as follows. For $i \in I_1 \cup I_2$, any observation x_{1i} is denoted by x_{1uvw}, where the index $u = j$ if $i \in I_j$, $j = 1,2$. The index v specifies the interval such that $\tau_{v-1} < x_{1i} \leq \tau_v$, for $v = 1,\ldots,M$, and $w \in \{1,\ldots,n_{1uv}\}$. Here, n_{1uv} represents the number of observations x_{1i} falling in the interval $(\tau_{v-1},\tau_v]$ for indices $i \in I_u$. Similarly, for $i \in I_1 \cup I_2$, each x_{2i} is denoted by x_{2uvw}, where $u = j$ if $i \in I_j$, $j = 1,2$. The index v satisfies $\tau_{v-1} < x_{2i} \leq \tau_v$, for $v = 1,\ldots,M$, and $w \in \{1,\ldots,n_{2uv}\}$, where n_{2uv} denotes the number of such observations in the corresponding interval for $i \in I_u$. For $i \in I_3$, the common observations are represented by x_{vw}, where $\tau_{v-1} < x_i \leq \tau_v$, for $v = 1,\ldots,M$, and $w \in \{1,\ldots,n_{3v}\}$. Here, n_{3v} is the number of observations in I_3 that lie within the v-th interval. Let $\boldsymbol{\theta} = (\alpha_1,\alpha_2,\alpha_3)$ and $\boldsymbol{c} = (c_1,c_2,\ldots,c_{M-1})$. Further, define

$$
n_{\bullet v} = n_{11v}+n_{12v}+n_{21v}+n_{22v}+n_{3v}, \quad n_{1\bullet v} = n_{11v}+n_{12v}, \quad n_{2\bullet v} = n_{21v}+n_{22v}.
$$

Then, the associated log-likelihood function is given by

$$
\begin{aligned}
l(\boldsymbol{\theta},\boldsymbol{c}|\mathcal{D}) = \; & n_1\ln(\alpha_1+\alpha_3) + n_1\ln\alpha_2 + n_2\ln\alpha_1 + n_2\ln(\alpha_2+\alpha_3) + n_3\ln\alpha_3 + \sum_{i\in I_1\cup I_2}\ln f_0(x_{1i}) \\
& + \sum_{i\in I_1\cup I_2}\ln f_0(x_{2i}) + \sum_{i\in I_3}\ln f_0(x_{1i}) + (\alpha_1+\alpha_3-1)\sum_{i\in I_1}\ln F_0(x_{1i}) + (\alpha_2-1)\sum_{i\in I_1}\ln F_0(x_{2i}) \\
& + (\alpha_1-1)\sum_{i\in I_2}\ln F_0(x_{1i}) + (\alpha_2+\alpha_3-1)\sum_{i\in I_2}\ln F_0(x_{2i}) + (\alpha_1+\alpha_2+\alpha_3-1)\sum_{i\in I_3}\ln F_0(x_i),
\end{aligned}
$$

$$
\begin{aligned}
l(\boldsymbol{\theta},\boldsymbol{c}|\mathcal{D}) = \; & n_1\ln(\alpha_1+\alpha_3) + n_1\ln\alpha_2 + n_2\ln\alpha_1 + n_2\ln(\alpha_2+\alpha_3) + n_3\ln\alpha_3 + \sum_{v=1}^{M} n_{\bullet v}\ln c_v \\
& - A_1(\boldsymbol{c}|\mathcal{D}) - A_2(\boldsymbol{c}|\mathcal{D}) - A_3(\boldsymbol{c}|\mathcal{D}) - \big[B_{11}(\boldsymbol{c}|\mathcal{D}) + B_{12}(\boldsymbol{c}|\mathcal{D})\big] - \big[B_{21}(\boldsymbol{c}|\mathcal{D}) + B_{22}(\boldsymbol{c}|\mathcal{D})\big] \\
& - B_3(\boldsymbol{c}|\mathcal{D}) + \alpha_1\big[B_{11}(\boldsymbol{c}|\mathcal{D}) + B_{12}(\boldsymbol{c}|\mathcal{D}) + B_3(\boldsymbol{c}|\mathcal{D})\big] + \alpha_2\big[B_{11}(\boldsymbol{c}|\mathcal{D}) + B_{12}(\boldsymbol{c}|\mathcal{D}) + B_3(\boldsymbol{c}|\mathcal{D})\big] \\
& + \alpha_3\big[B_{11}(\boldsymbol{c}|\mathcal{D}) + B_{22}(\boldsymbol{c}|\mathcal{D}) + B_3(\boldsymbol{c}|\mathcal{D})\big], \qquad (9)
\end{aligned}
$$

where

$$A_1(\boldsymbol{c}|\mathcal{D}) = \sum_{i\in I_1\cup I_2} H_0(x_{1i}) = \left\{\sum_{j=1}^{M-1}\sum_{v=1+j}^{M} c_j(n_{11v}+n_{12v})(\tau_j-\tau_{j-1}) + \sum_{v=1}^{M}\sum_{u=1}^{2}\sum_{w=1}^{n_{1uv}} c_j(x_{1uvw}-\tau_{j-1})\right\}$$

$$A_2(\boldsymbol{c}|\mathcal{D}) = \sum_{i\in I_1\cup I_2} H_0(x_{2i}) = \left\{\sum_{j=1}^{M-1}\sum_{v=1+j}^{M} c_j(n_{21v}+n_{22v})(\tau_j-\tau_{j-1}) + \sum_{v=1}^{M}\sum_{u=1}^{2}\sum_{w=1}^{n_{2uv}} c_j(x_{2uvw}-\tau_{j-1})\right\}$$

$$A_3(\boldsymbol{c}|\mathcal{D}) = \sum_{i\in I_3} H_0(x_i) = \left\{\sum_{j=1}^{M-1}\sum_{v=1+j}^{M} c_j n_{3v}(\tau_j-\tau_{j-1}) + \sum_{v=1}^{M}\sum_{w=1} c_j(x_{vw}-\tau_{j-1})\right\},$$

$$\begin{aligned} B_{11}(\boldsymbol{c}|\mathcal{D}) &= \sum_{i\in I_1} \ln\left[1-e^{-H_0(x_{1i})}\right] = \sum_{v=1}^{M}\sum_{w=1}^{n_{11v}} \ln\left[1-e^{-H_0(x_{11vw})}\right] \\ &= \sum_{w=1}^{n_{111}} \ln\left[1-e^{-c_1(x_{111w}-\tau_0)}\right] + \sum_{v=2}^{M}\sum_{w=1}^{n_{11v}} \ln\Big[1-\exp\Big\{-\sum_{j=1}^{v-1} c_j(\tau_j-\tau_{j-1}) - c_v(x_{11vw}-\tau_{v-1}\Big\}\Big] \end{aligned}$$

$$\begin{aligned} B_{12}(\boldsymbol{c}|\mathcal{D}) &= \sum_{i\in I_2} \ln\left[1-e^{-H_0(x_{1i})}\right] = \sum_{v=1}^{M}\sum_{w=1}^{n_{12v}} \ln\left[1-e^{-H_0(x_{12vw})}\right] \\ &= \sum_{w=1}^{n_{121}} \ln\left[1-e^{-c_1(x_{121w}-\tau_0)}\right] + \sum_{v=2}^{M}\sum_{w=1}^{n_{12v}} \ln\Big[1-\exp\Big\{-\sum_{j=1}^{v-1} c_j(\tau_j-\tau_{j-1}) - c_v(x_{12vw}-\tau_{v-1}\Big\}\Big] \end{aligned}$$

$$\begin{aligned} B_{21}(\boldsymbol{c}|\mathcal{D}) &= \sum_{i\in I_1} \ln\left[1-e^{-H_0(x_{2i})}\right] = \sum_{v=1}^{M}\sum_{w=1}^{n_{21v}} \ln\left[1-e^{-H_0(x_{21vw})}\right] \\ &= \sum_{w=1}^{n_{211}} \ln\left[1-e^{-c_1(x_{211w}-\tau_0)}\right] + \sum_{v=2}^{M}\sum_{w=1}^{n_{21v}} \ln\Big[1-\exp\Big\{-\sum_{j=1}^{v-1} c_j(\tau_j-\tau_{j-1}) - c_v(x_{21vw}-\tau_{v-1}\Big\}\Big] \end{aligned}$$

$$\begin{aligned} B_{22}(\boldsymbol{c}|\mathcal{D}) &= \sum_{i\in I_2} \ln\left[1-e^{-H_0(x_{2i})}\right] = \sum_{v=1}^{M}\sum_{w=1}^{n_{22v}} \ln\left[1-e^{-H_0(x_{22vw})}\right] \\ &= \sum_{w=1}^{n_{221}} \ln\left[1-e^{-c_1(x_{221w}-\tau_0)}\right] + \sum_{v=2}^{M}\sum_{w=1}^{n_{22v}} \ln\Big[1-\exp\Big\{-\sum_{j=1}^{v-1} c_j(\tau_j-\tau_{j-1}) - c_v(x_{22vw}-\tau_{v-1}\Big\}\Big] \end{aligned}$$

$$\begin{aligned} B_{3}(\boldsymbol{c}|\mathcal{D}) &= \sum_{i\in I_3} \ln\left[1-e^{-H_0(x_{i})}\right] = \sum_{v=1}^{M}\sum_{w=1}^{n_{3v}} \ln\left[1-e^{-H_0(x_{vw})}\right] \\ &= \sum_{w=1}^{n_{31}} \ln\left[1-e^{-c_1(x_{1w}-\tau_0)}\right] + \sum_{v=2}^{M}\sum_{w=1}^{n_{3v}} \ln\Big[1-\exp\Big\{-\sum_{j=1}^{v-1} c_j(\tau_j-\tau_{j-1}) - c_v(x_{vw}-\tau_{v-1}\Big\}\Big]. \end{aligned}$$

It is evident that the MLEs of the unknown parameters cannot be derived in closed-form expressions. Instead, they must be obtained by solving an $(M+2)$-dimensional optimization problem, where the computational complexity increases as M grows. To address this challenge, we reformulate the estimation as a missing data problem and propose the use of an efficient EM algorithm. This approach significantly reduces the computational burden, as only an $(M-1)$-dimensional optimization needs to be solved at each E-step of the EM procedure.

5.2 EM Algorithm

To treat the estimation problem as missing data problem, we first define a pair of random variables (Δ_1, Δ_2) associated with each (X_1, X_2) as follows:

$$(\Delta_1, \Delta_2) = \begin{cases} (1,2), & \text{if } X_1 = V_1,\ X_2 = V_2 \\ (1,3), & \text{if } X_1 = V_1,\ X_2 = V_3 \\ (3,2), & \text{if } X_1 = V_3,\ X_2 = V_2 \\ (3,3), & \text{if } X_1 = V_3,\ X_2 = V_3. \end{cases}$$

If $X_1 < X_2$, then (Δ_1, Δ_2) is either (1, 2) or (3, 2) with positive probability and hence in the observed data Δ_1 is missing. Similarly, if $X_2 < X_1$, then (Δ_1, Δ_2) is either (1, 2) or (1, 3) with positive probability and hence in the observed data Δ_2 is missing. Further, if $X_1 = X_2$, then (Δ_1, Δ_2) is (3, 3) with probability one.

In our problem, the entire set $I_1 \cup I_2$ consists of the missing observations. Note that in I_1, Δ_1 is missing while Δ_2 is known. Similarly, note that in I_2, Δ_1 is known while Δ_2 is missing. The observations in Table 1 will be useful in constructing E-step of the EM algorithm.

Table 1: All possible cases of V_1, V_2, V_3 corresponding probabilities and (Δ_1, Δ_2)

Different cases	Probability	(Δ_1, Δ_2)	X_1 & X_2	Set
$V_3 < V_1 < V_2$	$\frac{\alpha_1\alpha_2}{(\alpha_1+\alpha_3)(\alpha_1+\alpha_2+\alpha_3)}$	(1, 2)	$X_1 < X_2$	I_1
$V_1 < V_3 < V_2$	$\frac{\alpha_3\alpha_2}{(\alpha_1+\alpha_3)(\alpha_1+\alpha_2+\alpha_3)}$	(3, 2)	$X_1 < X_2$	I_1
$V_3 < V_2 < V_1$	$\frac{\alpha_2\alpha_1}{(\alpha_2+\alpha_3)(\alpha_1+\alpha_2+\alpha_3)}$	(1, 2)	$X_1 > X_2$	I_2
$V_2 < V_3 < V_1$	$\frac{\alpha_3\alpha_1}{(\alpha_2+\alpha_3)(\alpha_1+\alpha_2+\alpha_3)}$	(1, 3)	$X_1 > X_2$	I_2
$V_2 < V_1 < V_3$	$\frac{\alpha_1\alpha_3}{(\alpha_1+\alpha_2)(\alpha_1+\alpha_2+\alpha_3)}$	(3, 3)	$X_1 = X_2$	I_3
$V_1 < V_2 < V_3$	$\frac{\alpha_2\alpha_3}{(\alpha_1+\alpha_2)(\alpha_1+\alpha_2+\alpha_3)}$	(3, 3)	$X_1 = X_2$	I_3

Thus if $(x_1, x_2) \in I_1$, then the observed data can be formed by fractioning as (x_1, x_2, u_1) and (x_1, x_2, u_2), where the conditional probabilities u_1 and u_2 can be obtained as

$$\begin{aligned} u_1 &= P(\Delta_1 = 1, \Delta_2 = 2 | X_1 < X_2) = \frac{\alpha_1}{\alpha_1 + \alpha_3}, \\ u_2 &= P(\Delta_1 = 3, \Delta_2 = 2 | X_1 < X_2) = \frac{\alpha_3}{\alpha_1 + \alpha_3}. \end{aligned} \tag{10}$$

Thus if $(x_1, x_2) \in I_2$, then the observed data can be formed by fractioning as (x_1, x_2, v_1) and (x_1, x_2, v_2), where the conditional probabilities v_1 and v_2 can be obtained as

$$\begin{aligned} v_1 &= P(\Delta_1 = 1, \Delta_2 = 2 | X_1 > X_2) = \frac{\alpha_2}{\alpha_2 + \alpha_3}, \\ v_2 &= P(\Delta_1 = 1, \Delta_2 = 3 | X_1 > X_2) = \frac{\alpha_3}{\alpha_2 + \alpha_3}. \end{aligned} \tag{11}$$

Hence, based on the observations $\{(x_{1i}, x_{2i}), i = 1, 2, \ldots, n\}$, the pseudo log-likelihood function is obtained as follows;

$$\begin{aligned} l_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{c}) = & \; u_1 \big[n_1 \ln \alpha_1 + \sum_{i \in I_1} \ln f_0(x_{1i}) + (\alpha_1 + \alpha_3 - 1) \sum_{i \in I_1} \ln F_0(x_{1i}) \big] \\ & + u_2 \big[n_1 \ln \alpha_3 + \sum_{i \in I_1} \ln f_0(x_{1i}) + (\alpha_1 + \alpha_3 - 1) \sum_{i \in I_1} \ln F_0(x_{1i}) \big] \\ & + n_1 \ln \alpha_2 + \sum_{i \in I_1} \ln f_0(x_{2i}) + (\alpha_2 - 1) \sum_{i \in I_1} \ln F_0(x_{2i}) \\ & + n_2 \ln \alpha_1 + \sum_{i \in I_2} \ln f_0(x_{1i}) + (\alpha_1 - 1) \sum_{i \in I_2} \ln F_0(x_{1i}) \\ & + v_1 \big[n_2 \ln \alpha_2 + \sum_{i \in I_2} \ln f_0(x_{2i}) + (\alpha_2 + \alpha_3 - 1) \sum_{i \in I_2} \ln F_0(x_{2i}) \big] \\ & + v_2 \big[n_2 \ln \alpha_3 + \sum_{i \in I_2} \ln f_0(x_{2i}) + (\alpha_2 + \alpha_3 - 1) \sum_{i \in I_2} \ln F_0(x_{2i}) \big] \\ & + n_3 \ln \alpha_3 + \sum_{i \in I_3} \ln f_0(x_i) + (\alpha_1 + \alpha_2 + \alpha_3 - 1) \sum_{i \in I_3} \ln F_0(x_i) \end{aligned}$$

$$\begin{aligned} = & \; (n_1 u_1 + n_2) \ln \alpha_1 + (n_1 + n_2 v_1) \ln \alpha_2 + (n_1 u_2 + n_2 v_2 + n_3) \ln \alpha_3 \\ & + \sum_{i \in I_1 \cup I_2} \big[\ln h_0(x_{1i}) + \ln h_0(x_{2i}) \big] + \sum_{i \in I_3} \ln h_0(x_i) - \sum_{i \in I_1 \cup I_2} \big[H_0(x_{1i}) + H_0(x_{2i}) \big] \\ & - \sum_{i \in I_3} H_0(x_i) - \sum_{i \in I_1 \cup I_2} \ln[1 - e^{-H_0(x_{1i})}] - \sum_{i \in I_1 \cup I_2} \ln[1 - e^{-H_0(x_{2i})}] - \sum_{i \in I_3} \ln[1 - e^{-H_0(x_i)}] \\ & + \alpha_1 \sum_{i \ \in I} \ln[1 - e^{-H_0(x_{1i})}] + \alpha_2 \sum_{i \ \in I} \ln[1 - e^{-H_0(x_{2i})}] \\ & + \alpha_3 \Bigg(\sum_{i \in I_1} \ln[1 - e^{-H_0(x_{1i})}] + \sum_{i \in I_2} \ln[1 - e^{-H_0(x_{2i})}] + \sum_{i \in I_3} \ln[1 - e^{-H_0(x_i)}] \Bigg) \end{aligned}$$

$$\begin{aligned} = & \; (n_1 u_1 + n_2) \ln \alpha_1 + (n_1 + n_2 v_1) \ln \alpha_2 + (n_1 u_2 + n_2 v_2 + n_3) \ln \alpha_3 + \sum_{v=1}^{M} n_{\bullet v} \ln c_v \\ & - A_1(\boldsymbol{c}|\mathcal{D}) - A_2(\boldsymbol{c}|\mathcal{D}) - A_3(\boldsymbol{c}|\mathcal{D}) - B_1(\boldsymbol{c}|\mathcal{D}) - B_2(\boldsymbol{c}|\mathcal{D}) - B_3(\boldsymbol{c}|\mathcal{D}) \\ & + \alpha_1 \big[B_{11}(\boldsymbol{c}|\mathcal{D}) + B_{12}(\boldsymbol{c}|\mathcal{D}) + B_3(\boldsymbol{c}|\mathcal{D}) \big] + \alpha_2 \big[B_{21}(\boldsymbol{c}|\mathcal{D}) + B_{22}(\boldsymbol{c}|\mathcal{D}) + B_3(\boldsymbol{c}|\mathcal{D}) \big] \\ & + \alpha_3 \big[B_{11}(\boldsymbol{c}|\mathcal{D}) - B_{22}(\boldsymbol{c}|\mathcal{D}) + B_3(\boldsymbol{c}|\mathcal{D}) \big]. \end{aligned} \tag{12}$$

where $B_j(\boldsymbol{c}|\mathcal{D}) = B_j(\boldsymbol{c}|\mathcal{D}) + B_3(\boldsymbol{c}|\mathcal{D})$, $j = 1, 2$. For fixed $\boldsymbol{c}$, the MLE's of the unknown parameters are obtained by maximizing the (12) with respect to $\alpha_1, \alpha_2, \alpha_3$, which are given as;

$$\begin{aligned}
\hat{\alpha}_1(\boldsymbol{c}) &= -\frac{n_1 u_1 + n_2}{B_1(\boldsymbol{c}|\mathcal{D}) + B_3(\boldsymbol{c}|\mathcal{D})}, \quad \hat{\alpha}_2(\boldsymbol{c}) = -\frac{n_1 + n_2 v_1}{B_2(\boldsymbol{c}|\mathcal{D}) + B_3(\boldsymbol{c}|\mathcal{D})}, \\
\hat{\alpha}_3(\boldsymbol{c}) &= -\frac{n_1 u_2 + n_2 v_2 + n_3}{B_{11}(\boldsymbol{c}|\mathcal{D}) + B_{22}(\boldsymbol{c}|\mathcal{D}) + B_3(\boldsymbol{c}|\mathcal{D})}.
\end{aligned} \tag{13}$$

The MLE of $\boldsymbol{c}$ can be obtained by maximizing

$$\begin{aligned}
g(\boldsymbol{c}|\hat{\boldsymbol{\theta}}) &= \sum_{v=1}^{M} n_{\bullet v} \ln c_v - A_1(\boldsymbol{c}|\mathcal{D}) - A_2(\boldsymbol{c}|\mathcal{D}) - A_3(\boldsymbol{c}|\mathcal{D}) - B_1(\boldsymbol{c}|\mathcal{D}) - B_2(\boldsymbol{c}|\mathcal{D}) - B_3(\boldsymbol{c}|\mathcal{D}) \\
&\quad -(n_1 u_1 + n_2) \ln \big[- \{B_1(\boldsymbol{c}|\mathcal{D}) + B_3(\boldsymbol{c}|\mathcal{D})\}\big] \\
&\quad -(n_1 + n_2 v_1) \ln \big[- \{B_2(\boldsymbol{c}|\mathcal{D}) + B_3(\boldsymbol{c}|\mathcal{D})\}\big] \\
&\quad -(n_1 u_2 + n_2 v_2 + n_3) \ln \big[- \{B_{11}(\boldsymbol{c}|\mathcal{D}) + B_{22}(\boldsymbol{c}|\mathcal{D}) + B_3(\boldsymbol{c}|\mathcal{D})\}\big]
\end{aligned} \tag{14}$$

with respect to $\boldsymbol{c}$. But closed form expression of the MLE of $\boldsymbol{c}$ does not exist. So that it becomes $(M-1)$-dimensional optimization problem. Now we discuss how to find the $(k+1)^{th}$ iteration from the k^{th} iteration of the EM algorithm. We assume that at the k^{th} iteration the estimate are denoted as $(\theta^{(k)}, \boldsymbol{c}^{(k)})$, $k = 0, 1, 2, \ldots$
Step 1 (E-Step)

- Compute $u_1^{(k)}, u_2^{(k)}, v_1^{(k)}, v_2^{(k)}$ using (10) and (11).

Step 2 (M-Step)

- Find the MLE $\boldsymbol{c}^{(k+1)}$ by maximizing (14) using some standard optimization technique, e.g. Nelder-Mead algorithm see [33].
- Once $\boldsymbol{c}^{(k+1)}$ is obtained, we compute $\alpha_1^{(k+1)} = \alpha_1^{(k+1)}(\boldsymbol{c}^{(k+1)})$, $\alpha_2^{(k+1)} = \alpha_2^{(k+1)}(\boldsymbol{c}^{(k+1)})$, $\alpha_3^{(k+1)} = \alpha_3^{(k+1)}(\boldsymbol{c}^{(k+1)})$ using (13).

5.3 Some special cases

In practical applications, the procedure performs well for $M = 1, 2$, and 3. In these cases, the corresponding maximization reduces to low-dimensional optimization problems (of dimension 0, 1, or 2), which can be efficiently handled through the EM algorithm to obtain the MLEs. In this subsection, we give explicit expressions $A_1(\boldsymbol{c}), A_2(\boldsymbol{c}), A_3(\boldsymbol{c}), B_1(\boldsymbol{c}), B_2(\boldsymbol{c}), B_3(\boldsymbol{c}), B_{11}(\boldsymbol{c}), B_{22}(\boldsymbol{c})$ and $g(\boldsymbol{c}|\hat{\theta})$, for some special cases namely M = 1, 2, 3 which are useful in practice to analyzed real data sets.
***Case* 1 :** M = 1.
In this case clearly, $A_1(\boldsymbol{c}), A_2(\boldsymbol{c}), A_3(\boldsymbol{c}), B_1(\boldsymbol{c}), B_2(\boldsymbol{c}), B_3(\boldsymbol{c}), B_{11}(\boldsymbol{c})$ and $B_{22}(\boldsymbol{c})$ do not depend on $\boldsymbol{c}$, and let us denote them as $A_1, A_2, A_3, B_1, B_2, B_3, B_{11}$ and B_{22} respectively. Here they are as follows;

$$A_1 = \sum_{i \in I_1 \cup I_2} x_{1i}, \quad A_2 = \sum_{i \in I_1 \cup I_2} x_{2i}, \quad A_3 = \sum_{i \in I_3} x_i$$

$$B_1 = \sum_{i \in I_{12}} \ln(1 - e^{-x_{1i}}), \quad B_2 = \sum_{i \in I_{12}} \ln(1 - e^{-x_{2i}}), \quad B_3 = \sum_{i \in I_3} \ln(1 - e^{-x_i}),$$

$$B_{11} = \sum_{i \in I_1} \ln(1 - e^{-x_{1i}}), \quad B_{22} = \sum_{i \in I_2} \ln(1 - e^{-x_{2i}}).$$

In this case the BPRH model becomes bivariate generalized exponential (BVGE) model and it is denoted as $BVGE(\alpha_1, \alpha_2, \alpha_3)$, for more details see Kundu and Gupta [23]. The pseudo log likelihood function is

$$\begin{aligned} l_{pseudo}(\theta) = {} & (n_1 u_1 + n_2) \ln \alpha_1 + (n_1 + n_2 v_1) \ln \alpha_2 + (n_1 u_2 + n_2 v_2 + n_3) \ln \alpha_3 - \sum_{i \in I_1 \cup I_2} x_{1i} \\ & - \sum_{i \in I_1 \cup I_2} x_{2i} - \sum_{i \in I_3} x_i - \sum_{i \in I_1 \cup I_2} \ln\left(1 - e^{-x_{1i}}\right) - \sum_{i \in I_1 \cup I_2} \ln\left(1 - e^{-x_{2i}}\right) \\ & - \sum_{i \in I_3} \ln\left(1 - e^{-x_i}\right) + \alpha_1 \sum_{i \ \in I} \ln\left(1 - e^{-x_{1i}}\right) + \alpha_2 \sum_{i \ \in I} \ln\left(1 - e^{-x_{2i}}\right) \\ & + \alpha_3 \left[\sum_{i \in I_1} \ln\left(1 - e^{-x_{1i}}\right) + \sum_{i \in I_2} \ln\left(1 - e^{-x_{2i}}\right) + \sum_{i \in I_3} \ln\left(1 - e^{-x_i}\right) \right] \end{aligned}$$

The MLEs are obtained as

$$\hat{\alpha_1} = -\frac{n_1 u_1 + n_2}{\sum_{i=1}^{n} \ln\left(1 - e^{-x_{1i}}\right)}, \quad \hat{\alpha_2} = -\frac{n_1 + n_2 v_1}{\sum_{i=1}^{n} \ln\left(1 - e^{-x_{2i}}\right)},$$

$$\hat{\alpha_3} = -\frac{n_1 u_2 + n_2 v_2 + n_3}{\sum_{i \in I_1} \ln\left(1 - e^{-x_{1i}}\right) + \sum_{i \in I_2} \ln\left(1 - e^{-x_{2i}}\right) + \sum_{i \in I_3} \ln\left(1 - e^{-x_i}\right)}$$

Case 2 : M = 2

In this case we have $c_1 = c$, $c_2 = 1$ and $\tau_0 = 0$, $\tau_1 = \tau$, $\tau_2 = \infty$. Then we have,

$$n_{11} = \#\{x_{1i} \in I_{12} : x_{1i} \leq \tau, 1 \leq i \leq n\}, \quad n_{12} = \#\{x_{1i} \in I_{12} : x_{1i} > \tau, 1 \leq i \leq n\},$$

$$n_{21} = \#\{x_{2i} \in I_{12} : x_{2i} \leq \tau, 1 \leq i \leq n\}, \quad n_{22} = \#\{x_{2i} \in I_{12} : x_{ji} > \tau, 1 \leq i \leq n\},$$

$$n_{31} = \#\{x_i \in I_3 : x_i \leq \tau, 1 \leq i \leq n\}, \quad n_{32} = \#\{x_i \in I_3 : x_i > \tau, 1 \leq i \leq n\}$$

$$A_1(c) = c\tau n_{12} + c \sum_{i \in I_{12} : x_{1i} < \tau} x_{1i} + \sum_{i \in I_{12} : x_{1i} > \tau} (x_{1i} - \tau)$$

$$A_2(c) = c\tau n_{22} + c \sum_{i \in I_{12} : x_{2i} < \tau} x_{2i} + \sum_{i \in I_{12} : x_{2i} > \tau} (x_{2i} - \tau)$$

$$A_3(c) = c\tau n_{32} + c \sum_{i \in I_3 : x_i < \tau} x_i + \sum_{i \in I_3 : x_i > \tau} (x_i - \tau)$$

$$B_1(c) = \sum_{i \in I_{12} : x_{1i} < \tau} \ln(1 - e^{-c x_{1i}}) + \sum_{i \in I_{12} : x_{1i} > \tau} \ln\left[1 - \exp\{-c\tau - (x_{1i} - \tau)\}\right]$$

$$B_2(c) = \sum_{i \in I_{12}: x_{2i} < \tau} \ln(1 - e^{-cx_{2i}}) + \sum_{i \in I_{12}: x_{2i} > \tau} \ln\left[1 - \exp\{-c\tau - (x_{2i} - \tau)\}\right]$$

$$B_3(c) = \sum_{i \in I_3: x_i < \tau} \ln(1 - e^{-cx_i}) + \sum_{i \in I_3: x_i > \tau} \ln\left[1 - \exp\{-c\tau - (x_i - \tau)\}\right]$$

$$B_{11}(c) = \sum_{i \in I_1: x_{1i} < \tau} \ln(1 - e^{-cx_{1i}}) + \sum_{i \in I_1: x_{1i} > \tau} \ln\left[1 - \exp\{-c\tau - (x_{1i} - \tau)\}\right]$$

$$B_{22}(c) = \sum_{i \in I_2: x_{2i} < \tau} \ln(1 - e^{-cx_{2i}}) + \sum_{i \in I_2: x_{2i} > \tau} \ln\left[1 - \exp\{-c\tau - (x_{2i} - \tau)\}\right]$$

$$\sum_{v=1}^{M} n_{\bullet v} \ln c_v = (n_{11} + n_{21} + n_{31}) \ln c$$

***Case* 3 :** M = 3
In this case we have,

$$m_{11} = \#\{x_{1i} \in I_{12} : x_{1i} \le \tau_1\}, \quad m_{12} = \#\{x_{1i} \in I_{12} : \tau_1 < x_{1i} \le \tau_2\}, \quad m_{13} = \#\{x_{1i} \in I_{12} : x_{1i} > \tau_2\}$$

$$m_{21} = \#\{x_{2i} \in I_{12} : x_{2i} \le \tau_1\}, \quad m_{22} = \#\{x_{2i} \in I_{12} : \tau_1 < x_{2i} \le \tau_2\}, \quad m_{23} = \#\{x_{2i} \in I_{12} : x_{2i} > \tau_2\}$$

$$m_{31} = \#\{x_i \in I_3 : x_i \le \tau_1\}, \quad m_{32} = \#\{x_i \in I_3 : \tau_1 < x_i \le \tau_2\}, \quad m_{33} = \#\{x_i \in I_3 : x_i > \tau_2\}$$

$$A_1(c_1, c_2) = c_1\tau_1(m_{12}+m_{13})+c_2 m_{13}(\tau_2-\tau_1)+c_1 \sum_{i \in I_{12}: x_{1i} < \tau_1} x_{1i} + c_2 \sum_{i \in I_{12}: \tau_1 < x_{1i} \le \tau_2} (x_{1i}-\tau_1) + \sum_{i \in I_{12}: x_{1i} > \tau_2} (x_{1i}-\tau_2)$$

$$A_2(c_1, c_2) = c_1\tau_1(m_{22}+m_{23})+c_2 m_{23}(\tau_2-\tau_1)+c_1 \sum_{i \in I_{12}: x_{2i} < \tau_1} x_{2i} + c_2 \sum_{i \in I_{12}: \tau_1 < x_{2i} \le \tau_2} (x_{2i}-\tau_1) + \sum_{i \in I_{12}: x_{2i} > \tau_2} (x_{2i}-\tau_2)$$

$$A_3(c_1, c_2) = c_1\tau_1(m_{32}+m_{33})+c_2 m_{33}(\tau_2-\tau_1)+c_1 \sum_{i \in I_3: x_i < \tau_1} x_i + c_2 \sum_{i \in I_3: \tau_1 < x_i \le \tau_2} (x_i-\tau_1) + \sum_{i \in I_3: x_i > \tau_2} (x_i-\tau_2)$$

$$\begin{aligned} B_1(c_1, c_2) = &\sum_{i \in I_{12}: x_{1i} < \tau_1} \ln(1 - e^{-c_1 x_{1i}}) + \sum_{i \in I_{12}: \tau_1 < x_{1i} \le \tau_2} \ln\left[1 - \exp\{-c_1\tau_1 - c_2(x_{1i} - \tau_1)\}\right] \\ &+ \sum_{i \in I_{12}: x_{1i} > \tau_2} \ln\left[1 - \exp\{-c_1\tau_1 - c_2(\tau_2 - \tau_1) - (x_{1i} - \tau_2)\}\right] \end{aligned}$$

$$\begin{aligned} B_2(c_1, c_2) = &\sum_{i \in I_{12}: x_{2i} < \tau_1} \ln(1 - e^{-c_1 x_{2i}}) + \sum_{i \in I_{12}: \tau_1 < x_{2i} \le \tau_2} \ln\left[1 - \exp\{-c_1\tau_1 - c_2(x_{2i} - \tau_1)\}\right] \\ &+ \sum_{i \in I_{12}: x_{1i} > \tau_2} \ln\left[1 - \exp\{-c_1\tau_1 - c_2(\tau_2 - \tau_1) - (x_{1i} - \tau_2)\}\right] \end{aligned}$$

$$\begin{aligned} B_3(c_1, c_2) = &\sum_{i \in I_3: x_i < \tau_1} \ln(1 - e^{-c_1 x_i}) + \sum_{i \in I_3: \tau_1 < x_i \le \tau_2} \ln\left[1 - \exp\{-c_1\tau_1 - c_2(x_i - \tau_1)\}\right] \\ &+ \sum_{i \in I_3: x_i > \tau_2} \ln\left[1 - \exp\{-c_1\tau_1 - c_2(\tau_2 - \tau_1) - (x_i - \tau_2)\}\right] \end{aligned}$$

$$\sum_{v=1}^{M} n_{\bullet v} \ln c_v = (m_{11} + m_{21} + m_{31}) \ln c_1 + (m_{12} + m_{22} + m_{32}) \ln c_2.$$

$$\begin{aligned} B_{11}(c_1, c_2) = & \sum_{i \in I_1 : x_{1i} < \tau_1} \ln(1 - e^{-c_1 x_{1i}}) + \sum_{i \in I_1 : \tau_1 < x_{1i} \le \tau_2} \ln\left[1 - \exp\{-c_1\tau_1 - c_2(x_{1i} - \tau_1)\}\right] \\ & + \sum_{i \in I_1 : x_{1i} > \tau_2} \ln\left[1 - \exp\{-c_1\tau_1 - c_2(\tau_2 - \tau_1) - (x_{1i} - \tau_2)\}\right] \end{aligned}$$

$$\begin{aligned} B_{22}(c_1, c_2) = & \sum_{i \in I_2 : x_{2i} < \tau_1} \ln(1 - e^{-c_1 x_{2i}}) + \sum_{i \in I_2 : \tau_1 < x_{2i} \le \tau_2} \ln\left[1 - \exp\{-c_1\tau_1 - c_2(x_{2i} - \tau_1)\}\right] \\ & + \sum_{i \in I_2 : x_{2i} > \tau_2} \ln\left[1 - \exp\{-c_1\tau_1 - c_2(\tau_2 - \tau_1) - (x_{2i} - \tau_2)\}\right]. \end{aligned}$$

6 Estimation of Cut Points

In much of the existing literature, it is typically assumed that both the number of cut points and their locations are known a priori. A similar assumption was adopted in Section 5, where the number of cut points M and their positions $0 = \tau_0 < \tau_1 < \cdots < \tau_{M-1} < \tau_M = \infty$ were treated as fixed and known. However, in practical applications, such information is rarely available. To address this limitation, we consider a more realistic framework in which both the number of cut points and their locations are treated as unknown and must be estimated from the data.

We focus on the estimation of the parameters of the piecewise constant hazard (PCH) model when both M and the change-points $\tau_1, \tau_2, \ldots, \tau_{M-1}$ are unknown. The primary interest lies in capturing the shape of the hazard function over the interval $[0, \tau]$, where τ denotes the largest observed time. The cut points $\{\tau_1, \tau_2, \ldots, \tau_{M-1}\}$ are assumed to arise as event times from a non-homogeneous Poisson process $\{N(t) : t \ge 0\}$ with intensity function $\lambda(t) = \delta t^{\delta - 1}$, where $\delta > 0$. Since these event times are not directly observed, they are treated as latent variables.

For a fixed value of M, the likelihood function is constructed in terms of the unknown parameters δ, $c_1, c_2, \ldots, c_M$, and $\alpha_1, \alpha_2, \alpha_3$, where the unobserved event times are replaced by their corresponding expected values. The maximum likelihood estimates of the parameters are then obtained for each candidate value of M. Finally, model selection criteria such as the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC) are employed to determine the optimal number of cut points.

For the subsequent development, it is essential to obtain the joint probability density function of the event times associated with a non-homogeneous Poisson process. Let $\{N(t) : t \ge 0\}$ be a non-homogeneous Poisson process with intensity function $\lambda(t) = \delta t^{\delta - 1}$, where $\delta > 0$. Denote the ordered event times by $T_1, T_2, \ldots$.

Then, conditional on the event that $N(\tau) = k$, the joint probability density function of $(T_1, T_2, \ldots, T_k)$ can be expressed as follows:

$$f_{T_1, \ldots, T_k | N(\tau) = k}(t_1, \ldots, t_k) = \frac{k!\, \delta^k}{\tau^{k\delta}} \prod_{i=1}^{k} t_i^{\delta - 1}, \quad 0 < t_1 < t_2 < \cdots < t_k < \tau, \tag{15}$$

and is zero otherwise. For $j = 1, 2, \dots, k$, with $k = M - 1$, the conditional expectation of T_j is

$$\begin{aligned}
E(T_j \mid N(\tau) = k) &= \frac{k!\,\delta^k}{\tau^{k\delta}} \int_0^{\tau} \int_0^{t_k} \cdots \int_0^{t_2} \left(\prod_{i=1}^{j-1} t_i^{\delta-1} \right) t_j^{\delta} \left(\prod_{i=j+1}^{k} t_i^{\delta-1} \right) dt_1 \cdots dt_k \\
&= \frac{k!\,\delta^k}{\tau^{k\delta}} \cdot \frac{\tau^{k\delta+1}}{\delta(2\delta)\cdots((j-1)\delta)\,(j\delta+1)\,((j+1)\delta+1)\cdots(k\delta+1)} \\
&= \frac{k!}{(j-1)!} \frac{\delta^{\,k-j+1}\,\tau}{(j\delta+1)((j+1)\delta+1)\cdots(k\delta+1)} \\
&= \tau \prod_{l=j}^{k} \frac{1}{1+\frac{1}{l\delta}}.
\end{aligned}$$

In the special case of a homogeneous Poisson process, i.e., when $\delta = 1$, we have

$$E(T_j \mid N(\tau) = n) = \frac{j\tau}{n+1},$$

which corresponds to equally spaced cut-points. Thus, the equidistant partition is a special case of the proposed method. In particular when $M = 2$, we have $k = 1$ and $E(T_1 \mid N(\tau) = 1) = \tau/(1+\frac{1}{\delta})$. For $M = 3$, we have $k = 2$ with $E(T_1|N(\tau) = 2) = \frac{\tau}{(1+1/\delta)(1+1/2\delta)}$ and $E(T_2|N(\tau) = 2) = \tau\frac{1}{1+1/2\delta}$.

For practical implementation, the estimation procedure is carried out as follows. For a given value of M, we first obtain the MLEs of δ and the parameter set $(c_1, \dots, c_M, \alpha_1, \alpha_2, \alpha_3)$, along with the corresponding maximized log-likelihood value. The choice of M is then guided by model selection criteria such as the AIC or the BIC.

The estimation is performed in two stages:

(i) For a fixed value of δ, compute the MLEs of $(c_1, \dots, c_M, \alpha_1, \alpha_2, \alpha_3)$. Subsequently, estimate δ by maximizing the corresponding profile likelihood function.

(ii) Determine the optimal number of cut points M by minimizing an information criterion such as AIC or BIC.

6.1 Simulation

In this subsection, we study the effect of the number of cut points on the estimation of δ. A simulation study based on 1000 replications is conducted by taking $\tau = 1$ for different sample sizes n and different numbers of cut points M. For $M = 2$ and $M = 3$, the average estimates of δ and the corresponding mean squared errors (MSEs) are reported in Tables 2 and 3, respectively. The results indicate that, in both cases, the average bias and MSE decrease as the sample size n increases. Moreover, for a fixed sample size, both the average bias and MSE decrease as the number of cut points M increases.

Table 2: The AE and MSE for δ when $M = 2$.

n		$\delta = 0.5$	$\delta = 1$	$\delta = 1.5$	$\delta = 2$
20	AE	0.5217	1.0570	1.5887	2.1054
	MSE	0.0173	0.0657	0.1524	0.2517
30	AE	0.5158	1.0314	1.5557	2.0994
	MSE	0.0103	0.0382	0.0970	0.1567
50	AE	0.5113	1.0202	1.5263	2.0548
	MSE	0.0060	0.0207	0.0515	0.0936
75	AE	0.5078	1.0126	1.5271	2.0408
	MSE	0.0034	0.0137	0.0301	0.0590
100	AE	0.5077	1.0004	1.5128	2.0171
	MSE	0.0029	0.0093	0.0240	0.0406
250	AE	0.5023	1.0017	1.5058	2.0163
	MSE	0.0010	0.0041	0.0085	0.0176

Table 3: The AE and MSE for δ when $M = 3$.

n		$\delta = 0.5$	$\delta = 1$	$\delta = 1.5$	$\delta = 2$
20	AE	0.5174	1.0220	1.5299	2.0581
	MSE	0.0072	0.0293	0.0654	0.1195
30	AE	0.5047	1.0079	1.5377	2.0395
	MSE	0.0044	0.0177	0.0426	0.0726
50	AE	0.5043	1.0139	1.5096	2.0228
	MSE	0.0028	0.0106	0.0237	0.0406
75	AE	0.5021	1.0069	1.5165	2.0169
	MSE	0.0017	0.0065	0.0149	0.0289
100	AE	0.5020	1.0045	1.5099	2.0116
	MSE	0.0013	0.0050	0.0111	0.0209
250	AE	0.5003	1.0038	1.5030	2.0038
	MSE	0.0005	0.0020	0.0045	0.0079

7 Bivariate goodness-of-fit

In this section, using the same idea of Alam et al. [2, 3] we consider a bivariate GoF based on SEBs. For more details see, Anderson [4], Pyke [36]. The SEBs yield spacings which are distributed as sample spacings in the univariate case. Let $\mathcal{D} = \{(x_{11}, x_{21}), ..., (x_{1n}, x_{2n})\}$ be n observations from a population with unkown CDF $F(\cdot)$. We want to test that the given sample is drawn from a family of $BPRH(F_0, \alpha_1, \alpha_2, \alpha_3)$, i.e.,

$$H_0 : F = F_0 \ \forall \ (x_1, x_2) \quad versus \quad H_0 : F \neq F_0 \ \text{ for at least one } (x_1, x_2),$$

where $F_0 = F_0(x_1, x_2; \theta)$ is the CDF of the BPRH distribution. Since the parameters $\theta = (\alpha_1, \alpha_2, \alpha_3, \beta, \lambda, c_1, c_2, ..., c_{M-1})$ are unknown, the null hypothesis H_0 is composite and the test is known as composite GoF test.

Let us consider the alternating components of the bivariare observations in the sample, denoted as $z_{(1),1}, z_{(2),2}, z_{(3),1}, z_{(4),2}, ...$, where $z_{(r),j}$ denotes the smallest value among the j^{th} components, $j = 1, 2$, in the sample, excluding the observations associated with $z_{(1),1}, z_{(2),2}, z_{(3),1}, ..., z_{(r-1),j'}$, where r=1,2,...,n and

$$j' = \begin{cases} j+1 & \text{for } j = 1 \\ j-1 & \text{for } j = 2. \end{cases}$$

That is, $z_{(1),1}$ denotes the smallest value among the first components, $z_{(2),2}$ denotes the smallest value among the second components, discarding the observation associated with $z_{(1),1}$, etc. The partial sums of coverages are given by

$$v_1 = 1 - S_{X_1}(z_{(1),1}), \qquad v_2 = 1 - S_{(X_1,X_2)}(z_{(1),1}, z_{(2),2})$$

and for $1 \leq t \leq N$, $N = \frac{n-1}{2}$ or $N = \frac{n-2}{2}$ according as n is odd or even

$$v_{2t+1} = 1 - S_{(X_1,X_2)}(z_{(2t+1),1}, z_{(2t),2}) \qquad v_{2t+2} = 1 - S_{(X_1,X_2)}(z_{(2t+1),1}, z_{(2t+2),2})$$

where, $S_{(X_1,X_2)}(x_1, x_2)$ is the joint SF (3) of the BPRH model and $S_{X_1}(x_1)$ is the marginal SF of X_1. The KS test statistic (two-sided) is

$$D_n = \max\left\{0, \max_{1\leq i\leq n}\left(\frac{i}{n} - v_i\right), \max_{1\leq i\leq n}\left(v_i - \frac{i-1}{n}\right)\right\}. \tag{16}$$

In our GoF test problem the null hypothesis is composite so the asymptotic distribution of D_n is obtained by the parametric bootstrap method. Thus for real data case the critical value and p-value of the GoF test are obtained by the parametric bootstrap method.

8 Simulation Study

In this section, we present an extensive simulation experiments to evaluate the performance of the proposed EM algorithm for the BPRH model. The experiments are conducted for various values of M, different parameter sets, and varying sample sizes n. In each scenario, we computed the maximum likelihood estimators (MLEs) using the EM algorithm with the same stopping criterion. The algorithm was terminated when the ratio $\left|[l(k) - l(k-1)]/l(k-1)\right|$ become less than 10^{-8}, where $l(k)$ denotes the log-likelihood value at the k^{th} iteration. The EM algorithm is consistent with the choices of initial guesses of the unknown parameters, see Kumar et al. [20, 18]. For $M = 1$, there is no need for an optimization algorithm. However, for $M > 1$, each iteration of the EM algorithm involves solving an $(M-1)$ dimensional optimization problem. For this purpose, we use Nelder-Mead method, see [33], available in the R under the *optim()* function, to find the MLEs of the parameters $\alpha_1, \alpha_2, \alpha_3, c_1, c_2, ..., c_{M-2}$ and c_{M-1}. In each case, we have reported the average biases (ABs), the associated mean squared errors (MSEs), the average length of confidence intervals (ALCIs), the coverage probabilities (CPs) based on 95% confidence interval (CI), and the average number of iterations (ANIs). All the results are based on 1000 replications for different sample sizes $n = 25, 50, 75, 100, 250, 500$.

Simulation results are presented for several model configurations and parameter settings. For the case $M = 1$, two sets of true parameter values, $(\alpha_1, \alpha_2, \alpha_3) = (0.3, 0.4, 0.5)$ and $(1.5, 1.5, 1.5)$, are considered, and the corresponding results are reported in Table 4. For $M = 2$, simulations are conducted under the parameter combinations $(\alpha_1, \alpha_2, \alpha_3, c, \tau) = (0.4, 0.6, 0.5, 0.5, 0.5)$ and $(1.0, 1.0, 1.0, 2.0, 0.3)$, with the results summarized in Table 5. For $M = 3$, two sets of true parameter values, $(\alpha_1, \alpha_2, \alpha_3, c_1, c_2, \tau_1, \tau_2) = (0.4, 0.6, 0.5, 0.5, 0.8, 0.1, 0.5)$ and $(1.5, 1.4, 1.0, 0.4, 0.8, 1.5, 2.0)$, are examined, and the corresponding results are provided in Table 6. Additional simulation results for $M = 3$ under alternative parameter combinations are reported in Table 7. Other parameter settings were also examined but are not reported for brevity; similar patterns were observed in all cases.

Several observations can be drawn from the simulation results. In all scenarios, the ABs, MSEs, and ALCIs decrease monotonically as the sample size n increases. At the same time, the CPs of all estimators converge to the nominal confidence level, supporting the adequacy of the likelihood-based inference. These results provide empirical evidence that the EM-based maximum likelihood estimators are consistent and asymptotically efficient. Although the estimators exhibit positive bias for small sample sizes, this bias diminishes as n increases. Consequently, for moderate to large samples, confidence intervals constructed using asymptotic normality yield satisfactory performance. For fixed sample size n, comparing Tables 5 and 6 we can infer that if α_1, α_2 and α_3 are fixed and the number of cut points increases, the biases and MSEs of the corresponding MLEs of α_1, α_2 and α_3 increase. It is quite clear that the proposed EM algorithm for the BPRH model is working quite well and it can be used in practice quite conveniently.

9 Real Data Analysis

In this section, we analyze a real data set to illustrate the practical advantages of the proposed semi-parametric model. We consider the Union of European Football Associations (UEFA) Champions League football data, which consist of 37 matches in which the home team scored at least one goal and at least one goal was scored directly from a kick (including penalty kicks, free kicks, foul kicks, or other direct kicks) by either team. The data, reported in Table 8, were originally analyzed by Meintanis [28] using the MOBE model. Subsequently, several authors have revisited this data set using the conventional MLE for different bivariate lifetime models; for example, Kundu and Gupta [23] applied the MLE to estimate the parameters of the Marshall–Olkin bivariate Weibull distribution and see Kumar et al. [20, 19].

In the present study, we fit the data using the proposed flexible semi-parametric BPRH framework and focus on inference based on the MLE. Here, X_1 denotes the time (in minutes) at which the first goal scored directly from a kick occurs in a match, while X_2 represents the time (in minutes) when the home team scores its first goal of any type. The data set exhibits all three possible orderings: $n_1 = 6$ observations with $X_1 < X_2$, $n_2 = 17$ observations with $X_1 > X_2$ and $n_3 = 14$ observations with $X_1 = X_2$, satisfying $n_1 + n_2 + n_3 = 37$. A scatter plot of the data is shown in Figure 1.

To explore the hazard functions of the marginal distributions, we present the scaled total time on test (TTT) plot, as proposed by Aarset [1], which provides insights into the shape of a distribution's hazard function. Aarset [1] demonstrated that the scaled TTT transform is convex (concave) if the hazard rate is decreasing (increasing). For a bathtub-shaped (unimodal) hazard rate, the scaled TTT transform is first convex (concave) and then concave (convex). In this case,

Table 4: The average bias (AB), mean square error (MSE), average length of confidence interval (ALCI), coverage probability (CP) and average number of iteration (ANI) of α_1, α_2 and α_3 for BPRH model with $M = 1$.

True.par $\rightarrow$		$\alpha_1 = 0.3$, $\alpha_2 = 0.4$, $\alpha_3 = 0.5$				$\alpha_1 = 1.5$, $\alpha_2 = 1.5$, $\alpha_3 = 1.5$			
$n \downarrow$		α_1	α_2	α_3	ANI	α_1	α_2	α_3	ANI
25	AB	0.0118	0.0115	0.0301	16.21	0.0475	0.0385	0.0737	16.91
	MSE	0.0148	0.0211	0.0222		0.2516	0.2600	0.2143	
	ALCI	0.3957	0.4830	0.4702		1.7071	1.7007	1.4898	
	CP	0.90	0.90	0.92		0.915	0.909	0.905	
50	AB	0.0098	0.0077	0.0120	15.74	0.0286	0.0287	0.0347	15.91
	MSE	0.0074	0.0093	0.0094		0.1172	0.1292	0.1046	
	ALCI	0.2778	0.3380	0.3237		1.1952	1.1943	1.0360	
	CP	0.899	0.921	0.916		0.917	0.903	0.908	
75	AB	0.0063	0.0070	0.0055	15.61	0.0185	0.0207	0.0127	15.61
	MSE	0.0045	0.0062	0.0062		0.0766	0.0768	0.0675	
	ALCI	0.2245	0.2752	0.2615		0.9681	0.9692	0.8366	
	CP	0.91	0.914	0.91		0.912	0.928	0.901	
100	AB	0.0043	0.0055	0.0041	15.56	0.0102	0.0155	0.0213	15.32
	MSE	0.0031	0.0045	0.0044		0.0569	0.0574	0.0498	
	ALCI	0.1937	0.2378	0.2259		0.8371	0.8394	0.7278	
	CP	0.91	0.93	0.92		0.924	0.92	0.908	
250	AB	0.0005	0.0015	0.0017	15.49	0.0073	0.0079	0.0093	14.71
	MSE	0.0012	0.0016	0.0017		0.0220	0.0233	0.0186	
	ALCI	0.1216	0.1492	0.1422		0.5283	0.5284	0.4577	
	CP	0.92	0.94	0.91		0.927	0.907	0.905	
500	AB	0.0009	0.0014	0.0007	15.47	0.0020	0.0034	0.0074	14.21
	MSE	0.0007	0.0008	0.0009		0.0116	0.0100	0.0101	
	ALCI	0.0860	0.1055	0.1004		0.3725	0.3729	0.3233	
	CP	0.91	0.94	0.90		0.917	0.934	0.901	

the scaled TTT transform for the UEFA data is shown in Figure 2. The results clearly indicate that the marginals X_1, X_2 and $\max\{X_1, X_2\}$ exhibit increasing hazard rates. Therefore, the BPRH model is appropriate for analyzing this data set.

Before analyzing the UEFA data using the BPRH model, we scaled all the data points by dividing them by 100 such that all the data points lies in the interval $[0, 1]$. After scaling the data, all observations lie in the interval $[0, 1]$. The largest observed scaled value is 0.85. We set $\tau = 1$ as the known upper bound of the support after scaling, rather than the maximum observed value. All the models are fitted using the same transformed data. We fitted the proposed model with $M = 1, 2, 3$ to the data. In all cases, we employed the EM algorithm as described in Subsection 5.2. The EM algorithm is initialized by fitting the marginal distributions of X_1, X_2, and $\min\{X_1, X_2\}$, and using the resulting parameter estimates as the initial values. The algorithm was terminated when the ratio $\big|[l(k) - l(k-1)]/l(k-1)\big|$ become less than 10^{-8}. Furthermore, the EM algorithm has been shown to be consistent with respect to the choice of initial parameter values for this class of models; see Kumar et al. [20].

For the BPRH model with $M > 1$, there are $(M-1)$ cut points, denoted as $\boldsymbol{\tau} = (\tau_1, \tau_2, ..., \tau_{M-1})$, which need to be optimally chosen based on the data to determine the MLEs of the $(M+2)$ unknown parameters. Since the MLEs $\hat{\alpha}_1(\boldsymbol{\tau})$, $\hat{\alpha}_2(\boldsymbol{\tau})$, $\hat{\alpha}_3(\boldsymbol{\tau})$, $\hat{c}_1(\boldsymbol{\tau}), \hat{c}_2(\boldsymbol{\tau}), ..., \hat{c}_{M-1}(\boldsymbol{\tau})$ are the function

Table 5: The average bias (AB), mean square error (MSE), average length of confidence interval (ALCI), coverage probability (CP) and average number of iteration (ANI) of α_1, α_2, α_3 and c for BPRH model with $M = 2$.

True.par $\rightarrow$		$\alpha_1 = 0.4$, $\alpha_2 = 0.6$, $\alpha_3 = 0.5$, $c = \tau = 0.5$					$\alpha_1 = \alpha_2 = \alpha_3 = 1.0$, $c = 2$, $\tau = 0.3$				
$n \downarrow$		α_1	α_2	α_3	c	ANI	α_1	α_2	α_3	c	ANI
25	AB	0.0568	0.0947	0.0666	0.1051	9.53	0.3136	0.2969	0.3325	0.3520	9.39
	MSE	0.0511	0.1101	0.0695	0.1417		1.7149	1.3008	2.0470	1.9556	
	ALCI	0.7132	1.0496	0.8016	1.3128		2.8213	2.7275	2.7888	4.3000	
	CP	0.92	0.94	0.93	0.91		0.89	0.89	0.90	0.90	
50	AB	0.0201	0.0221	0.0258	0.0353	8.81	0.0944	0.0901	0.0961	0.1351	8.68
	MSE	0.0166	0.0306	0.0227	0.0522		0.1715	0.1849	0.1706	0.5740	
	ALCI	0.4411	0.6272	0.4914	0.8356		1.4206	1.4152	1.3334	2.8383	
	CP	0.92	0.92	0.93	0.91		0.94	0.93	0.93	0.94	
75	AB	0.0159	0.0234	0.0100	0.0223	8.55	0.0638	0.0810	0.0730	0.1120	8.50
	MSE	0.0118	0.0185	0.0124	0.0315		0.1104	0.1182	0.1051	0.3972	
	ALCI	0.3538	0.5065	0.3855	0.6706		1.1124	1.1295	1.0504	2.3065	
	CP	0.91	0.94	0.91	0.93		0.93	0.94	0.94	0.93	
100	AB	0.0092	0.0183	0.0171	0.0215	8.28	0.0549	0.0525	0.0410	0.0866	8.40
	MSE	0.0070	0.0135	0.0099	0.0227		0.0718	0.0705	0.0674	0.2681	
	ALCI	0.3018	0.4356	0.3368	0.5815		0.9397	0.9362	0.8692	1.9688	
	CP	0.94	0.94	0.93	0.94		0.94	0.94	0.93	0.95	
150	AB	0.0073	0.0064	0.0092	0.0158	8.07	0.0305	0.0361	0.0207	0.0395	8.20
	MSE	0.0047	0.0089	0.0063	0.0152		0.0416	0.0472	0.0369	0.1739	
	ALCI	0.2442	0.3473	0.2694	0.4708		0.7399	0.7443	0.6848	1.5797	
	CP	0.93	0.94	0.93	0.93		0.94	0.94	0.93	0.94	
250	AB	0.0064	0.0066	0.0073	0.0132	7.81	0.0218	0.0232	0.0182	0.0324	8.18
	MSE	0.0027	0.0056	0.0035	0.0093		0.0239	0.0252	0.0206	0.0994	
	ALCI	0.1880	0.2685	0.2071	0.3631		0.5656	0.5671	0.5260	1.2215	
	CP	0.94	0.94	0.94	0.95		0.947	0.942	0.934	0.95	
500	AB	0.0014	0.0018	0.0014	0.0020	7.55	0.0111	0.0099	0.0105	0.0225	8.22
	MSE	0.0014	0.0025	0.0018	0.0041		0.0120	0.0116	0.0107	0.0506	
	ALCI	0.1308	0.1871	0.1439	0.2521		0.3934	0.3933	0.3661	0.8578	
	CP	0.92	0.94	0.92	0.95		0.94	0.94	0.93	0.95	

of $\boldsymbol{\tau}$, and hence so does the log-likelihood (LL) function. For a fixed M, the optimal $\boldsymbol{\tau}$ is obtained by minimizing the LL value $ll(\hat{\alpha}_1(\boldsymbol{\tau}), \hat{\alpha}_2(\boldsymbol{\tau}), \hat{\alpha}_3(\boldsymbol{\tau}), \hat{c}_1(\boldsymbol{\tau}), \hat{c}_2(\boldsymbol{\tau}), ..., \hat{c}_{M-1}(\boldsymbol{\tau}))$ using a fine grid search over $\boldsymbol{\tau}$, i.e.

$$\boldsymbol{\tau}^* = \min_{\tau_1, \tau_2, \ldots, \tau_{M-1}} ll(\hat{\alpha}_1(\boldsymbol{\tau}), \hat{\alpha}_2(\boldsymbol{\tau}), \hat{\alpha}_3(\boldsymbol{\tau}), \hat{c}_1(\boldsymbol{\tau}), \hat{c}_2(\boldsymbol{\tau}), ..., \hat{c}_{M-1}(\boldsymbol{\tau})).$$

Along with the AIC as dicussed in Section 6. When $M = 1$, the BPRH model reduces to the $MOBGE(\alpha_1, \alpha_2, \alpha_3)$ model. For the model with $M = 1$, the EM algorithm stopped after $k = 11$ iterations, and the MLEs, the associated SEs and 95% CIs are given Table 9. The corresponding log-likelihood (LL) value is -41.9476 and AIC value is 89.8951. Since it has been observed by several authors see for example Balakrishnan et al. [5] that in practice M = 2 or M = 3 works quite well, we restrict ourselves in this case for M = 2 and M = 3 only. For M = 2, the maximum likelihood estimates of the model parameters along with log-likelihood and AIC for different choice of δ, with grid size 0.02, are provided in Table 10.

Table 6: The average bias (AB), mean square error (MSE), average length of confidence interval (ALCI) and coverage probability (CP) of α_1, α_2, α_3, c_1 and c_2 for BPRH model with $M = 3$.

True.par →		$\alpha_1 = 0.4$, $\alpha_2 = 0.6$, $\alpha_3 = 0.5$ $c_1 = 0.5$, $c_2 = 0.8$, $\tau_1 = 0.1$, $\tau_2 = 0.5$					$\alpha_1 = 1.5$, $\alpha_2 = 1.4$, $\alpha_3 = 1.0$ $c_1 = 0.4$, $c_2 = 0.8$, $\tau_1 = 1.5$, $\tau_2 = 2.0$				
$n \downarrow$		α_1	α_2	α_3	c_1	c_2	α_1	α_2	α_3	c_1	c_2
25	AB	0.0323	0.0636	0.0347	0.1753	0.0370	0.2687	0.2512	0.1916	0.0445	0.0102
	MSE	0.0399	0.0766	0.0553	0.3149	0.1446	0.9321	0.8542	0.6923	0.0580	0.0893
	ALCI	0.8705	1.3037	1.0405	3.6113	1.8069	4.0682	3.8040	2.7982	0.9832	1.1781
	CP	0.89	0.92	0.89	0.85	0.90	0.91	0.91	0.89	0.91	0.91
50	AB	0.0243	0.0384	0.0275	0.1174	0.0398	0.2115	0.2117	0.1381	0.0435	0.0178
	MSE	0.0199	0.0386	0.0283	0.2273	0.0844	0.5228	0.4941	0.2633	0.0302	0.0417
	ALCI	0.5814	0.8468	0.6809	2.3131	1.2527	2.6785	2.5152	1.7617	0.6902	0.8123
	CP	0.92	0.94	0.91	0.86	0.94	0.96	0.95	0.93	0.95	0.95
75	AB	0.0194	0.0395	0.0248	0.1003	0.0361	0.1693	0.1530	0.1117	0.0322	0.0117
	MSE	0.0151	0.0283	0.0207	0.1729	0.0597	0.3828	0.3234	0.1941	0.0221	0.0285
	ALCI	0.4566	0.6762	0.5386	1.7993	1.0149	2.1044	1.9441	1.3836	0.5558	0.6581
	CP	0.93	0.95	0.92	0.88	0.95	0.93	0.94	0.92	0.94	0.94
100	AB	0.0193	0.0255	0.0219	0.0937	0.0283	0.1310	0.1124	0.0685	0.0230	0.0102
	MSE	0.0107	0.0204	0.0150	0.1460	0.0491	0.2467	0.2171	0.1120	0.0155	0.0207
	ALCI	0.3902	0.5650	0.4551	1.5301	0.8666	1.7428	1.6116	1.1291	0.4734	0.5668
	CP	0.95	0.95	0.94	0.90	0.94	0.95	0.94	0.94	0.94	0.95
150	AB	0.0116	0.0182	0.0168	0.0676	0.0145	0.0644	0.0619	0.0409	0.0103	0.0075
	MSE	0.0070	0.0146	0.0099	0.0974	0.0329	0.1408	0.1199	0.0685	0.0102	0.0146
	ALCI	0.3079	0.4483	0.3603	1.1912	0.6971	1.3358	1.2446	0.8777	0.3768	0.4603
	CP	0.94	0.94	0.94	0.92	0.94	0.93	0.95	0.92	0.94	0.94
250	AB	0.0101	0.0113	0.0137	0.0481	0.0117	0.0412	0.0411	0.0278	0.0101	0.0013
	MSE	0.0047	0.0087	0.0064	0.0607	0.0190	0.0722	0.0603	0.0336	0.0054	0.0084
	ALCI	0.2360	0.3399	0.2749	0.8968	0.5398	1.0067	0.9372	0.6614	0.2902	0.3529
	CP	0.93	0.94	0.95	0.92	0.94	0.95	0.96	0.94	0.95	0.94
500	AB	0.0030	0.0016	0.0028	0.0103	0.0027	0.0225	0.0254	0.0137	0.0052	0.0034
	MSE	0.0020	0.0039	0.0027	0.0262	0.0100	0.0375	0.0330	0.0170	0.0030	0.0040
	ALCI	0.1618	0.2333	0.1875	0.5974	0.3795	0.6994	0.6525	0.4592	0.2041	0.2497
	CP	0.922	0.94	0.926	0.915	0.937	0.94	0.94	0.92	0.94	0.96

For $M = 2$, the optimal cut point τ_1 is determined to be $\hat{\tau}_1 = 0.6610$, as shown in Table 10. From the table values of LL and model selection criteria such as AIC (see Kumar et al. [20]), it is evident that the optimal value is $\hat{\tau}_1 = 0.6610$. For $M = 2$, with optimal $\hat{\tau}_1 = 0.6610$, the MLEs of parameters, the corresponding SEs and 95% CIs are reported in Table 11. For $M = 3$, the optimal cut points τ_1 and τ_2 are obtained as $\hat{\tau}_1 = 0.4689$ and $\hat{\tau}_2 = 0.7619$, as shown in Table 12. The MLEs, associated SEs and 95% CI for $M = 3$ with optimal $\hat{\tau}_1 = 0.4689$ and $\hat{\tau}_2 = 0.7619$ are reported in Table 13.

We now select the best model among the three models, $M = 1$, $M = 2$ and $M = 3$, using standard model selection criteria such as AIC. Also compare the proposed BPRH model with some existing parametric models, namely the $MOBGE(\alpha_1, \alpha_2, \alpha_3)$, see Kundu and Gupta [23] and $MOBIW(\alpha_1, \alpha_2, \alpha_3, \beta)$, see Kundu and Gupta [22], models. The estimates MOBGE model parameters are reported in Table 9. For the MOBIW model, the MLEs are $\hat{\alpha}_1 = 0.1117$, $\hat{\alpha}_2 = 0.0242$, $\hat{\alpha}_3 = 0.1716$, and $\hat{\beta} = 1.0522$, with a log-likelihood value is LL $= -171.3142$. Based on the results provided in Table 14, it is observed that the best fitting model is BPRH with $M = 3$

Table 7: The average bias (AB), mean square error (MSE), average length of confidence interval (ALCI) and coverage probability (CP) of $\alpha_1, \alpha_2, \alpha_3$ and c for BPRH model with $M = 3$.

True.par $\rightarrow$		$\alpha_1 = 0.6,\ \alpha_2 = 0.4,\ \alpha_3 = 0.5$ $c_1 = 0.8,\ c_2 = 0.5,\ \tau_1 = 0.2,\ \tau_2 = 0.8$					$\alpha_1 = 1.0,\ \alpha_2 = 1.0,\ \alpha_3 = 1.0$ $c_1 = 0.5,\ c_2 = 1.5,\ \tau_1 = 1.0,\ \tau_2 = 2.0$				
$n \downarrow$		α_1	α_2	α_3	c_1	c_2	α_1	α_2	α_3	c_1	c_2
25	AB	0.1402	0.0825	0.1163	0.3376	0.0437	0.3163	0.3048	0.2975	0.1198	0.0412
	MSE	0.2026	0.0996	0.1581	1.0379	0.0636	0.7845	0.7510	0.8640	0.1443	0.1238
	ALCI	1.3993	0.9288	1.1051	3.2796	0.9206	3.0289	2.9874	2.8936	1.3711	1.3299
	CP	0.93	0.90	0.92	0.89	0.90	0.91	0.91	0.91	0.90	0.91
50	AB	0.0552	0.0370	0.0399	0.1470	0.0129	0.1514	0.1429	0.1403	0.0563	0.0280
	MSE	0.0606	0.0303	0.0468	0.3325	0.0267	0.2684	0.2781	0.2657	0.0619	0.0579
	ALCI	0.8032	0.5428	0.6191	2.0303	0.6221	1.7631	1.7557	1.6566	0.9164	0.9189
	CP	0.94	0.92	0.92	0.91	0.93	0.93	0.94	0.94	0.93	0.94
75	AB	0.0384	0.0256	0.0271	0.1058	0.0157	0.1229	0.0890	0.1108	0.0492	0.0148
	MSE	0.0335	0.0162	0.0214	0.2017	0.0187	0.1812	0.1451	0.1615	0.0416	0.0379
	ALCI	0.6281	0.4248	0.4821	1.6157	0.5105	1.3667	1.3212	1.2686	0.7363	0.7376
	CP	0.93	0.93	0.94	0.92	0.94	0.96	0.94	0.95	0.94	0.95
100	AB	0.0244	0.0208	0.0206	0.0650	0.0116	0.0813	0.0853	0.0743	0.0369	0.0138
	MSE	0.0213	0.0109	0.0145	0.1341	0.0135	0.1179	0.1094	0.1016	0.0303	0.0268
	ALCI	0.5300	0.3627	0.4111	1.3730	0.4428	1.1230	1.1252	1.0467	0.6288	0.6366
	CP	0.94	0.93	0.93	0.93	0.94	0.93	0.95	0.93	0.93	0.94
150	AB	0.0204	0.0141	0.0201	0.0674	0.0078	0.0647	0.0678	0.0555	0.0294	0.0118
	MSE	0.0152	0.0075	0.0097	0.0982	0.0088	0.0622	0.0676	0.0610	0.0181	0.0180
	ALCI	0.4251	0.2892	0.3303	1.1166	0.3572	0.8901	0.8937	0.8280	0.5083	0.5167
	CP	0.94	0.92	0.93	0.95	0.94	0.96	0.96	0.95	0.96	0.95
250	AB	0.0106	0.0093	0.0099	0.0369	0.0034	0.0352	0.0347	0.0262	0.0144	0.0067
	MSE	0.0078	0.0043	0.0049	0.0530	0.0055	0.0320	0.0332	0.0275	0.0102	0.0112
	ALCI	0.3220	0.2198	0.2493	0.8467	0.2756	0.6626	0.6625	0.6152	0.3866	0.3997
	CP	0.94	0.92	0.92	0.93	0.94	0.96	0.96	0.95	0.94	0.94
500	AB	0.0050	0.0030	0.0045	0.0160	0.0013	0.0135	0.0074	0.0129	0.0064	-0.0019
	MSE	0.0034	0.0019	0.0021	0.0230	0.0025	0.0158	0.0146	0.0125	0.0049	0.0049
	ALCI	0.2240	0.1522	0.1730	0.5887	0.1942	0.4554	0.4528	0.4247	0.2703	0.2815
	CP	0.95	0.92	0.95	0.95	0.94	0.94	0.94	0.95	0.95	0.96

and $\hat{\tau}_1 = 0.4689$ and $\hat{\tau}_2 = 0.7619$, because it has smallest AIC. The MLEs, their associated SEs, and 95% CIs for the optimal τ's are presented in Table 13.

A natural question arises: does the model with $M = 3$ and $\hat{\tau}_1 = 0.4689$ and $\hat{\tau}_2 = 0.7619$ appropriately fit the UEFA data? For this purpose, first we consider the bivariate GoF test based on the statistically equivalence blocks (SEB), see Kumar et al. [20]. The nul hypothesis is $H_0 : F = F_{H_0}$, where F_{H_0} is the CDF of the $BPRH(\hat{c}_1, \hat{c}_2, \hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3)$ model with $M = 3$ and $\hat{\tau}_1$, $\hat{\tau}_2$. The p-value obtained by the parametric bootstrap approach with 1000 replications. Thus, the KS test statistic and the corresponding p-value are 0.0745 and 0.6633 respectively. The result indicates that the model with $M = 3$ and $\hat{\tau}_1 = 0.4689$ and $\hat{\tau}_2 = 0.7619$ is fit the UEFA data quite well.

Table 8: UEFA Champion's League data.

2005-2006	X_1	X_2	2004-2005	X_1	X_2
Lyon-Real Madrid	26	20	Internazionale-Bremen	34	34
Milan-Fenerbahce	63	18	Real Madrid-Roma	53	39
Chelsea-Anderlecht	19	19	Man. United-Fenerbache	54	7
Club Brugge-Juventus	66	85	Bayern-Ajax	51	28
Fenerbache-PSV	40	40	Moscow-PSG	76	64
Internazionale-Rangers	49	49	Barcelona-Shakhtar	64	15
Panathinaikos-Bremen	8	8	Leverkusen-Roma	26	48
Ajax-Arsenal	69	71	Arsenal- Panathinaikos	16	16
Man. United-Benfica	39	39	Dynamo Kyiv-Real Madrid	44	13
Real Madrid–Rosenborg	82	48	Man. United-Sparta	25	14
Villarreal-Benfica	72	72	Bayern-M. TelAviv	55	11
Juventus-Bayern	66	62	Bremen-Internazionale	49	49
Club Brugge-Rapid	25	9	Anderlecht-Valencia	24	24
Olympiacos-Lyon	41	3	Panathinaikos-PVS	44	30
Internazionale-Porto	16	75	Arsenal-Rosenborg	42	3
Schalke-PSV	18	18	Liverpool-Olympiacos	27	47
Barcelona-Bremen	22	14	M. TelAviv-Juventus	28	28
Milan-Schalke	42	42	Bremen-Panathinaikos	2	2
Rapid-Juventus	36	52			

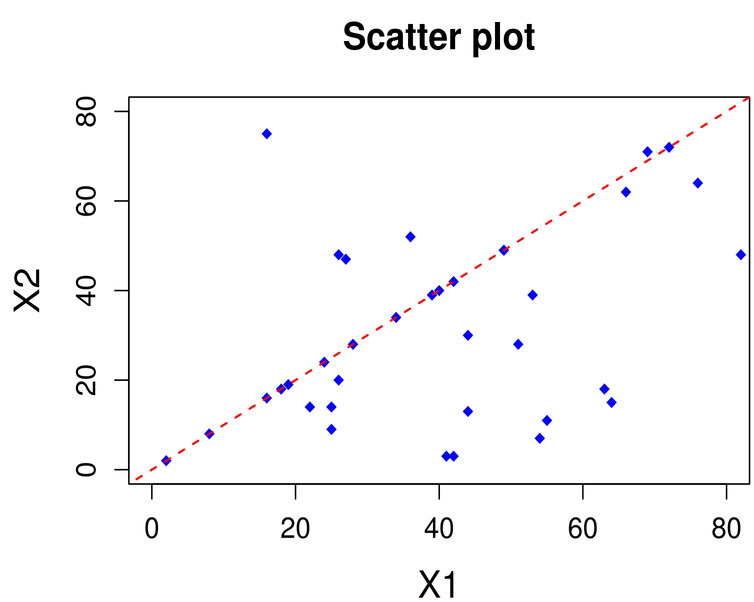
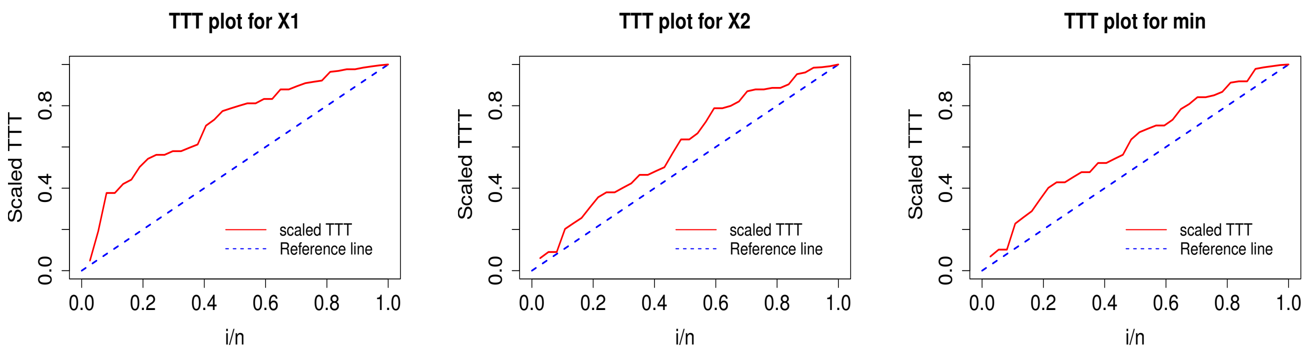


Figure 1: The scatter plot for UEFA data.



(a) (b) (c)

Figure 2: The TTT plots for UEFA data.

Table 9: The MLEs, SEs and 95% CIs for the model $M = 1$.

Par.	MLE	SE	95% CI
α_1	0.4236	0.0951	(0.2371, 0.6100)
α_2	0.1800	0.0551	(0.0720, 0.2881)
α_3	0.4763	0.0877	(0.3045, 0.6481)

10 Conclusion

In this paper, we introduced a new class of bivariate semi-parametric distributions with a singular component, where the marginals belong to the proportional reversed hazard class (PRHC). Rather than assuming a specific form for the baseline distribution, we considered a piecewise constant hazard function, enhancing the flexibility of the proposed distribution. We developed an efficient EM algorithm for estimating the unknown parameters, which proved to be highly convenient. Additionally, we developed properties of the proposed distribution. We discussed a method to find the estimates of the cut points for the piecewise constant hazard function. We implement a bivariate GoF test. While this paper primarily focuses on the classical inference of the unknown parameters, exploring Bayesian inference in this context would be an interesting approach for future research. Further work is needed along that direction.

Acknowledgements

The authors sincerely thank the Editor and the anonymous Reviewers for their valuable comments and constructive suggestions, which have significantly improved the quality and presentation of this manuscript.

Table 10: The MLEs along with log-likelihood and AIC for different choice of δ for $M = 2$.

δ	τ_1	α_1	α_2	α_3	c	LL	AIC
1.45	0.5918	1.1023	0.3834	0.9727	3.1153	-33.4063	74.8126
1.47	0.5951	1.0976	0.3821	0.9695	3.1016	-33.4932	74.9864
1.49	0.5984	1.0930	0.3809	0.9664	3.0881	-33.5782	75.1565
1.51	0.6016	1.0885	0.3796	0.9633	3.0749	-33.6614	75.3229
1.53	0.6047	1.0840	0.3784	0.9603	3.0620	-33.7429	75.4857
1.55	0.6078	1.0797	0.3772	0.9574	3.0493	-33.8226	75.6452
1.57	0.6109	1.0755	0.3761	0.9545	3.0368	-33.9006	75.8013
1.59	0.6139	1.0713	0.3749	0.9517	3.0247	-33.9771	75.9541
1.61	0.6169	1.0673	0.3738	0.9490	3.0127	-34.0519	76.1039
1.63	0.6198	1.0633	0.3727	0.9463	3.0010	-34.1253	76.2505
1.65	0.6226	1.1031	0.3832	0.9713	3.1038	-33.0780	74.1559
1.67	0.6255	1.0995	0.3822	0.9689	3.0933	-33.1466	74.2931
1.69	0.6283	1.0960	0.3812	0.9664	3.0830	-33.2139	74.4277
1.71	0.6310	1.1351	0.3913	0.9907	3.1823	-32.1375	72.2749
1.73	0.6337	1.1319	0.3905	0.9886	3.1732	-32.1997	72.3993
1.75	0.6364	1.1288	0.3896	0.9865	3.1642	-32.2607	72.5215
1.77	0.6390	1.1257	0.3888	0.9844	3.1554	-32.3207	72.6414
1.79	0.6416	1.2053	0.4092	1.0334	3.3546	-30.0147	68.0295
1.81	0.6441	1.2028	0.4085	1.0317	3.3476	-30.0664	68.1329
1.83	0.6466	1.2003	0.4078	1.0301	3.3406	-30.1173	68.2345
1.85	0.6491	1.1978	0.4072	1.0285	3.3338	-30.1673	68.3345
1.87	0.6516	1.1954	0.4065	1.0268	3.3270	-30.2164	68.4329
1.89	0.6540	1.1931	0.4059	1.0253	3.3204	-30.2648	68.5296
1.91	0.6564	1.1907	0.4053	1.0237	3.3139	-30.3124	68.6248
1.93	0.6587	1.1884	0.4047	1.0222	3.3074	-30.3592	68.7184
1.95	0.6610	1.2657	0.4241	1.0689	3.4961	-27.9539	63.9078
1.97	0.6633	1.2639	0.4237	1.0677	3.4911	-27.9926	63.9851
1.99	0.6656	1.2620	0.4232	1.0665	3.4863	-28.0307	64.0613
2.01	0.6678	1.2602	0.4228	1.0654	3.4815	-28.0682	64.1364
2.03	0.6700	1.2585	0.4223	1.0642	3.4768	-28.1052	64.2103
2.05	0.6721	1.2567	0.4219	1.0631	3.4722	-28.1416	64.2832
2.07	0.6743	1.2550	0.4215	1.0620	3.4676	-28.1775	64.3549
2.09	0.6764	1.2533	0.4210	1.0609	3.4631	-28.2128	64.4257
2.11	0.6785	1.2517	0.4206	1.0598	3.4586	-28.2477	64.4954
2.13	0.6805	1.2500	0.4202	1.0588	3.4543	-28.2821	64.5641
2.15	0.6825	1.2484	0.4198	1.0577	3.4499	-28.3159	64.6319
2.17	0.6845	1.2468	0.4194	1.0567	3.4457	-28.3493	64.6987
2.19	0.6865	1.2452	0.4190	1.0557	3.4415	-28.3823	64.7645
2.21	0.6885	1.2437	0.4186	1.0547	3.4373	-28.4147	64.8294

Table 11: The MLEs, SEs and 95% CIs for the model $M = 2$ with optimal $\hat{\tau}_1 = 0.6610$.

Par.	MLE	SE	95% CI
α_1	1.2657	0.3652	(0.5500, 1.9814)
α_2	0.4241	0.1415	(0.1468, 0.7015)
α_3	1.0689	0.2469	(0.5849, 1.5529)
c	3.4961	0.5739	(2.3713, 4.6209)

Conflict of Interest and Funding Statements

The author does not have any conflict of interest. This manuscript has not been funded by any funding agent.

Data Availability

The data set analyzed during this study are available in the article on page 26 taken from [https://doi.org/10.1016/j.jspi.2007.04.013].

References

[1] Aarset, M. V. (1987). How to identify a bathtub hazard rate. *IEEE Transactions on Reliability*, 36(1):106–108.

[2] Alam, K., Abernathy, R., and Williams, C. L. (1993). Multivariate goodness-of-fit tests based on statistically equivalent blocks. *Communications in Statistics-Theory and Methods*, 22(6):1515–1533.

[3] Alam, K. and Williams, C. L. (1995). A multivariate goodness-of-fit test for stochastically ordered distributions. *Biometrical Journal*, 37(8):945–956.

[4] Anderson, T. W. (1966). Some nonparametric multivariate procedures based on statistically equivalent blocks. *Proccedings of the International Symposium on Multivariate Analysis, Academic Press, New York*, pages 5–27.

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

[6] Barlow, R. E., Marshall, A. W., and Proschan, F. (1963). *Mathematical Theory of Reliability.* John Wiley & Sons, New York.

[7] Bemis, B. M., Bain, L. J., and Higgins, J. J. (1972). Estimation and hypothesis testing for the parameters of a bivariate exponential distribution. *Journal of the American Statistical Association*, 67(340):927–929.

Table 12: MLEs along with log-likelihood and AIC for different choice of δ for $M = 3$.

δ	τ_1	τ_2	α_1	α_2	α_3	c_1	c_2	LL	AIC
1.10	0.3601	0.6875	1.0031	0.3515	0.8765	2.6474	4.2094	-27.2433	64.4866
1.12	0.3652	0.6914	0.9953	0.3488	0.8675	2.6029	4.4711	-25.6090	61.2180
1.14	0.3703	0.6951	0.9741	0.3429	0.8524	2.5374	4.5572	-25.4063	60.8125
1.16	0.3753	0.6988	0.9535	0.3372	0.8377	2.4740	4.6468	-25.1848	60.3696
1.18	0.3802	0.7024	0.9331	0.3314	0.8232	2.4115	4.7387	-24.9449	59.8899
1.20	0.3850	0.7059	0.9131	0.3258	0.8091	2.3509	4.8332	-24.6869	59.3737
1.22	0.3898	0.7093	0.8934	0.3203	0.7952	2.2918	4.9315	-24.4108	58.8216
1.24	0.3945	0.7126	0.9831	0.3458	0.8598	2.5770	4.9082	-23.9679	57.9357
1.26	0.3991	0.7159	0.9647	0.3408	0.8473	2.5230	5.0107	-23.7250	57.4499
1.28	0.4037	0.7191	0.9937	0.3492	0.8694	2.6219	4.9326	-24.1488	58.2976
1.30	0.4082	0.7222	0.9867	0.3471	0.8626	2.5882	5.2621	-22.2719	54.5438
1.32	0.4127	0.7253	1.0157	0.3554	0.8843	2.6850	5.1815	-22.7042	55.4084
1.34	0.4170	0.7283	0.9996	0.3510	0.8736	2.6388	5.2929	-22.4626	54.9252
1.36	0.4214	0.7312	1.0735	0.3716	0.9268	2.8729	4.9880	-23.4970	56.9940
1.38	0.4256	0.7340	1.0589	0.3677	0.9170	2.8305	5.0912	-23.3171	56.6343
1.40	0.4298	0.7368	1.0449	0.3638	0.9075	2.7898	5.1995	-23.1250	56.2500
1.42	0.4340	0.7396	1.0308	0.3600	0.8981	2.7490	5.3110	-22.9207	55.8413
1.44	0.4381	0.7423	1.0169	0.3562	0.8889	2.7091	5.4251	-22.7041	55.4082
1.46	0.4421	0.7449	1.0905	0.3765	0.9408	2.9365	5.0572	-23.7594	57.5188
1.48	0.4461	0.7475	1.0787	0.3732	0.9327	2.9020	5.1595	-23.6081	57.2162
1.50	0.4500	0.7500	1.0670	0.3699	0.9247	2.8680	5.2645	-23.4473	56.8945
1.52	0.4539	0.7525	1.0633	0.3684	0.9203	2.8454	5.6575	-21.5578	53.1156
1.54	0.4577	0.7549	1.0519	0.3652	0.9125	2.8125	5.7809	-21.3449	52.6898
1.56	0.4615	0.7573	1.0407	0.3620	0.9050	2.7805	5.9082	-21.1219	52.2438
1.58	0.4652	0.7596	1.0296	0.3588	0.8975	2.7486	6.0378	-20.8887	51.7774
1.60	0.4689	0.7619	1.0254	0.3571	0.8930	2.7268	6.4856	-18.7899	47.5797
1.62	0.4725	0.7642	1.0569	0.3658	0.9151	2.8239	6.3383	-19.3670	48.7340
1.64	0.4761	0.7664	1.0466	0.3628	0.9082	2.7952	6.4802	-19.1182	48.2365
1.66	0.4796	0.7685	1.0364	0.3600	0.9015	2.7674	6.6267	-18.8601	47.7201
1.68	0.4831	0.7706	1.1099	0.3803	0.9522	2.9868	6.1256	-20.2397	50.4793
1.70	0.4865	0.7727	1.1012	0.3781	0.9467	2.9638	6.2532	-20.0520	50.1040
1.72	0.4899	0.7748	1.0925	0.3758	0.9412	2.9409	6.3817	-19.8574	49.7148
1.74	0.4933	0.7768	1.1644	0.3954	0.9900	3.1495	5.8201	-21.0855	52.1709
1.76	0.4966	0.7788	1.1566	0.3934	0.9851	3.1283	5.9350	-20.9445	51.8890
1.78	0.4999	0.7807	1.1487	0.3914	0.9802	3.1073	6.0520	-20.7976	51.5951
1.80	0.5031	0.7826	1.1407	0.3893	0.9752	3.0859	6.1729	-20.6446	51.2893
1.82	0.5063	0.7845	1.1331	0.3873	0.9704	3.0655	6.2996	-20.4857	50.9713
1.84	0.5094	0.7863	1.1254	0.3853	0.9656	3.0450	6.4285	-20.3206	50.6413
1.86	0.5126	0.7881	1.1577	0.3941	0.9874	3.1380	6.1618	-20.8823	51.7647
1.88	0.5156	0.7899	1.1505	0.3922	0.9829	3.1188	6.2829	-20.7419	51.4839
1.90	0.5187	0.7917	1.1439	0.3905	0.9788	3.1011	6.4077	-20.5963	51.1925
1.92	0.5217	0.7934	1.1762	0.3993	1.0003	3.1931	6.1102	-21.1520	52.3040

Table 13: MLEs, SEs and 95% CIs for the model $M = 3$ with optimal $\hat{\tau}_1 = 0.4689$ and $\hat{\tau}_2 = 0.7619$.

Par.	MLE	SE	95% CI
α_1	1.0254	0.3132	(0.4116, 1.6391)
α_2	0.3571	0.1232	(0.1157, 0.5985)
α_3	0.8930	0.2214	(0.4589, 1.3270)
c_1	2.7268	0.6118	(1.5278, 3.9259)
c_2	6.4856	1.4131	(3.7161, 9.2552)

Table 14: Comparing the BPRH model for $M = 1, 2, 3$ with MOBIW model.

Model	τ_1	τ_2	LL	AIC
MOBIW			-171.3142	350.6284
MOBGE (M = 1)			-41.9476	89.8951
BPRH (M = 2)	0.6610		-27.9539	63.9078
BPRH (M = 3)	0.4689	0.7619	-18.7899	47.5797

[8] Block, H. W., Savits, T. H., and Singh, H. (1998). The reversed hazard rate function. *Probability in the Engineering and Informational Sciences*, 12(1):69–90.

[9] El-Gohary, A., El-Bassiouny, A., and El-Morshedy, M. (2015). Bivariate exponentiated modified Weibull extension distribution. *arXiv preprint arXiv:1501.03528*.

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

[11] Gupta, R. C., Gupta, P. L., and Gupta, R. D. (1998). Modeling failure time data by Lehman alternatives. *Communications in Statistics-Theory and Methods*, 27(4):887–904.

[12] Gupta, R. C. and Gupta, R. D. (2007). Proportional reversed hazard rate model and its applications. *Journal of Statistical Planning and Inference*, 137(11):3525–3536.

[13] Gupta, R. D. and Kundu, D. (1999). Theory & methods: Generalized exponential distributions. *Australian & New Zealand Journal of Statistics*, 41(2):173–188.

[14] Gupta, R. D. and Kundu, D. (2001). Generalized exponential distribution: different method of estimations. *Journal of Statistical Computation and Simulation*, 69(4):315–337.

[15] Gupta, R. D. and Kundu, D. (2007). Generalized exponential distribution: Existing results and some recent developments. *Journal of Statistical Planning and Inference*, 137(11):3537–3547.

[16] Kumar, S. (2025). Some properties of bivariate modified Weibull distribution with competing risk application. *Journal of Statistics Applications & Probability Letters*, 12(3):125–148.

[17] Kumar, S. (2026). Absolutely continuous bivariate modified Weibull distribution: Properties and estimation. *Journal of Statistical Theory and Applications*, 25(1):12.

[18] Kumar, S., Kundu, D., and Mitra, S. (2025). Robust and efficient estimation for the Block-Basu bivariate exponential distribution. *Statistics and Computing*, 35(5):150.

[19] Kumar, S., Kundu, D., and Mitra, S. (2026a). Divergence-based robust inference for the Marshall-Olkin bivariate exponential distribution. *Journal of Statistical Computation and Simulation*, 96(10):2454–2481.

[20] Kumar, S., Kundu, D., and Mitra, S. (2026b). Estimation of five parameter bivariate modified Weibull singular distribution. *Communications in Statistics-Simulation and Computation*, 55(4):1267–1290.

[21] Kundu, D. and Gupta, A. K. (2013). Bayes estimation for the Marshall–Olkin bivariate Weibull distribution. *Computational Statistics & Data Analysis*, 57(1):271–281.

[22] Kundu, D. and Gupta, A. K. (2017). On bivariate inverse Weibull distribution. *Communications in Statistics—Theory and Methods*, 46(15):7370–7387.

[23] Kundu, D. and Gupta, R. D. (2009). Bivariate generalized exponential distribution. *Journal of Multivariate Analysis*, 100(4):581–593.

[24] Kundu, D. and Gupta, R. D. (2010a). A class of bivariate models with proportional reversed hazard marginals. *Sankhya B*, 72(2):236–253.

[25] Kundu, D. and Gupta, R. D. (2010b). Modified Sarhan–Balakrishnan singular bivariate distribution. *Journal of Statistical Planning and Inference*, 140(2):526–538.

[26] Lehmann, E. L. (1953). The power of rank test. *Annals of Mathematical Statistics, vol. 24, 23-42.*

[27] Marshall, A. W. and Olkin, I. (1967). A multivariate exponential distribution. *Journal of the American Statistical Association*, 62(317):30–44.

[28] Meintanis, S. G. (2007). Test of fit for Marshall–Olkin distributions with applications. *Journal of Statistical Planning and Inference*, 137(12):3954–3963.

[29] Mudholkar, G. S. and Hutson, A. D. (1996). The exponentiated Weibull family: some properties and a flood data application. *Communications in Statistics–Theory and Methods*, 25(12):3059–3083.

[30] Mudholkar, G. S. and Srivastava, D. K. (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE Transactions on Reliability*, 42(2):299–302.

[31] Mudholkar, G. S., Srivastava, D. K., and Freimer, M. (1995). The exponentiated Weibull family: A reanalysis of the bus-motor-failure data. *Technometrics*, 37(4):436–445.

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

[33] Nelder, J. A. and Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, 7(4):308–313.

[34] Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer Series in Statistics. Springer, New York, 2 edition.

[35] Oakes, D. (1989). Bivariate survival models induced by frailties. *Journal of the American Statistical Association*, 84(406):487–493.

[36] Pyke, R. (1965). Spacings. *Journal of the Royal Statistical Society: Series B (Methodological)*, 27(3):395–436.

[37] Roy, D. (2002). A characterization of model approach for generating bivariate life distributions using reversed hazard rates. *Journal of the Japan Statistical Society*, 32(2):239–245.

[38] Sankaran, P. and Gleeja, V. L. (2006). On bivariate reversed hazard rates. *Journal of the Japan Statistical Society*, 36(2):213–224.

[39] Sarhan, A. M. and Kundu, D. (2009). Generalized linear failure rate distribution. *Communications in Statistics-Theory and Methods*, 38(5):642–660.

[40] Surles, J. and Padgett, W. (2001). Inference for reliability and stress-strength for a scaled burr type x distribution. *Lifetime Data Analysis*, 7:187–200.

[41] Tsodikov, A. D., Asselain, B., and Yakovlev, A. Y. (1997). A distribution of tumor size at detection: An application to breast cancer data. *Biometrics*, pages 1495–1502.