

# The Bivariate Kumaraswamy Weibull regression model: a complete classical and Bayesian analysis

Juliana B. Fachini-Gomes<sup>a</sup>, Edwin M. M. Ortega<sup>b</sup>, Gauss M. Cordeiro<sup>c</sup>,  
Adriano K. Suzuki<sup>1,d</sup>

<sup>a</sup>Department of Statistics, University of Brasilia, Brazil;

<sup>b</sup>Department of Exact Sciences, University of São Paulo, Brazil;

<sup>c</sup>Department of Statistics, Federal University of Pernambuco, Brazil;

<sup>d</sup>Department of Applied Mathematics and Statistics, University of São Paulo, Brazil

---

## Abstract

Bivariate distributions play a fundamental role in survival and reliability studies. We consider a regression model for bivariate survival times under right-censored based on the bivariate Kumaraswamy Weibull (Cordeiro *et al.*, *Journal of the Franklin Institute*, **347**, 1399–1429, 2010) distribution to model the dependence of bivariate survival data. We describe some structural properties of the marginal distributions. The method of maximum likelihood and a Bayesian procedure are adopted to estimate the model parameters. We use diagnostic measures based on the local influence and Bayesian case influence diagnostics to detect influential observations in the new model. We also show that the estimates in the bivariate Kumaraswamy Weibull regression model are robust to deal with the presence of outliers in the data. In addition, we use some measures of goodness-of-fit to evaluate the bivariate Kumaraswamy Weibull regression model. The methodology is illustrated by means of a real lifetime data set for kidney patients.

**Keywords:** Bayesian inference, bivariate failure time, censored data, diagnostics, survival analysis

---

## 1. Introduction

Statistical applications for the time of occurrence of an event of interest generally use the exponential, Weibull, gamma, log-normal and log-logistic distributions. However, there has been growing interest in new distributions to model skewness, kurtosis and different types of hazard rates. Among them, we cite the exponentiated Weibull (Mudholkar *et al.*, 1995) and beta modified Weibull (Silva *et al.*, 2010) distributions. More recently, Cordeiro *et al.* (2010) introduced the Kumaraswamy Weibull (KwW) distribution from the generator defined by Cordeiro and de Castro (2011). By considering more than one response variable in the experiment, Cordeiro *et al.* (2010) proposed the bivariate Kumaraswamy Weibull (BKwW) distribution based on the construction of the bivariate Weibull (Hougaard, 1986) distribution.

In practice, regressor variables associated with the response variable of each observation are always presented; however, statistical analysis always gives consideration to models that account for all the information existing in the observations. Due to these factors, this work extends the well-known KwW distribution to include covariates by means of the scale parameter leading to the BKwW regression model. The inferential part is conducted using the asymptotic distribution of the maximum

---

<sup>1</sup> Corresponding author: Department of Applied Mathematics and Statistics, University of São Paulo, Avenida Trabalhador São-carlense, 400 - Centro CEP: 13566-590, São Carlos-SP, Brazil. E-mail: [suzuki@icmc.usp.br](mailto:suzuki@icmc.usp.br)

likelihood estimators (MLEs) subject to restrictions on parameters. To implement this method, we use the adjusted barrier function (Lange, 1999). Situations with small samples may present difficult results to justify. We explore the use of a Bayesian method as an alternative to classic analysis. Markov chain Monte Carlo (MCMC) methods are used to develop a Bayesian analysis for the regression model.

After fitting the data, it is important to check model assumptions and conduct a robustness study to detect influential or extreme observations that can cause distortions in the results of the analysis. Influence diagnostics is an important step in the analysis of a data set, since it provides an indication of bad model fit or influential observations. Cook (1986) proposed a diagnostic approach named local influence to assess the effect of small perturbations in the model and/or data on parameter estimates. Several authors have applied the local influence method in more general regression models than the normal regression model. Some authors have also investigated the assessment of local influence in survival analysis models. For instance, Ortega *et al.* (2013) proposed the log-beta Weibull regression model with application to predict recurrence of prostate cancer, Hashimoto *et al.* (2013) adapted local influence methods to log-generalized gamma regression model for interval-censored data, Ortega *et al.* (2015) considered the problem of assessing local influence in a power series beta Weibull regression model to predict breast carcinoma and da Cruz *et al.* (2016) explored global and local influence methods to the log-odd log-logistic Weibull regression model with censored data. For the BKwW regression model, we propose a similar method to detect influential subjects by considering the global and local influence and Bayesian case influence.

The article is organized as follows. In Section 2, we present the BKwW regression model and some properties of the marginal distributions. In Section 3, we examine the performance of the likelihood function by computing the maximum likelihood, while the estimated equations are considered under parameter constraints and derive several diagnostic measures by considering the normal curvatures of local influence under various perturbation schemes. In Section 4, we consider a Bayesian approach and influence diagnostics for the BKwW regression model. In Section 5, we conduct various simulation studies to evaluate the behavior of the estimators in the BKwW regression model. In Section 6, we present a reanalysis of the dataset from patients of a renal insufficiency study reported by McGilchrist and Aisbett (1991). Finally, Section 7 provides some conclusions remark.

## 2. A bivariate KwW regression model

In practice, the majority of studies involve covariates related to survival times. Regression models can be formulated in various ways. In survival analysis, the class of parametric regression models and Cox regression model are well-known. However, we adopt a reparameterization of the BKwW (Cordeiro *et al.*, 2010) distribution to define a new regression model. The bivariate Kumaraswamy (BKw) cumulative distribution is defined by

$$F(t_1, t_2) = 1 - [1 - G(t_1, t_2)]^a]^b, \quad t_1, t_2 > 0, \quad (2.1)$$

where  $a > 0$  and  $b > 0$  are additional shape parameters, and  $G(t_1, t_2)$  is an arbitrary joint cumulative distribution function (cdf). We also consider the bivariate Weibull (Hougaard, 1986) distribution due to its evident applicability and popularity in the literature. Its cdf  $G(t_1, t_2)$  is given by

$$G(t_1, t_2) = \exp \left\{ - \left[ (\lambda_1 t_1)^{\frac{c_1}{\alpha}} + (\lambda_2 t_2)^{\frac{c_2}{\alpha}} \right]^\alpha \right\} - \exp [ - (\lambda_1 t_1)^{c_1} ] - \exp [ - (\lambda_2 t_2)^{c_2} ] + 1. \quad (2.2)$$

Let  $T_1$  and  $T_2$  be two non-negative random variables denoting the survival times of two components of a system. We define a bivariate random variable  $T = (T_1, T_2)^T$  having a BKwW distribution.

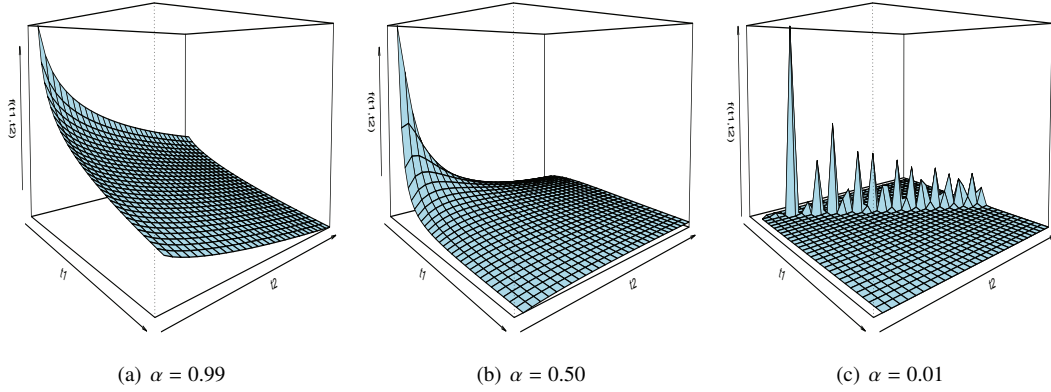


Figure 1: Joint probability density function for some parameter values.

Its cumulative distribution (for  $t_k > 0, k = 1, 2$ ) follows from equations (2.1) and (2.2) as

$$F(t_1, t_2) = 1 - \left\{ 1 - \left[ \exp \left\{ - \left[ (\lambda_1 t_1)^{\frac{c_1}{\alpha}} + (\lambda_2 t_2)^{\frac{c_2}{\alpha}} \right]^\alpha \right\} - \exp [ - (\lambda_1 t_1)^{c_1} ] - \exp [ - (\lambda_2 t_2)^{c_2} ] + 1 \right]^\alpha \right\}^b,$$

where  $a > 0, b > 0$ , and  $c_k > 0$  are shape parameters,  $\lambda_k > 0$  is a scale parameter and  $0 < \alpha \leq 1$  is an association parameter between  $T_1$  and  $T_2$ . We propose the reparameterization  $\lambda_k = \exp(-\mu_k)$  and  $c_k = \sigma_k^{-1}$ , for which  $-\infty < \mu_k < \infty$  and  $0 < \sigma_k < \infty$ . We introduce a vector of regressor variables  $\mathbf{x} = (x_0, x_1, \dots, x_p)^T$  and the linear structure  $\mu = \mathbf{x}^T \boldsymbol{\beta}$ , where  $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_p)^T$  is the unknown parameter vector associated with the covariates. Therefore, the reparameterization can be expressed as:  $\lambda_k = \exp(-\mathbf{x}^T \boldsymbol{\beta}_k) = \exp[-(\beta_{0k} x_0 + \beta_{1k} x_1 + \dots + \beta_{pk} x_p)]$  and  $c_k = 1/\sigma_k$ , for  $k = 1, 2$ . The BKwW regression model is defined by

$$F(t_1, t_2 | \mathbf{x}) = 1 - \left\{ 1 - \left\{ \exp \left[ - \left\{ \left[ \exp(-\mathbf{x}^T \boldsymbol{\beta}_1) t_1 \right]^{\frac{1}{\sigma_1 \alpha}} + \left[ \exp(-\mathbf{x}^T \boldsymbol{\beta}_2) t_2 \right]^{\frac{1}{\sigma_2 \alpha}} \right\}^\alpha \right] - \exp \left\{ - \left[ \exp(-\mathbf{x}^T \boldsymbol{\beta}_1) t_1 \right]^{\frac{1}{\sigma_1}} \right\} - \exp \left\{ - \left[ \exp(-\mathbf{x}^T \boldsymbol{\beta}_2) t_2 \right]^{\frac{1}{\sigma_2}} \right\} + 1 \right\}^\alpha \right\}^b, \quad (2.3)$$

with probability density function (pdf) given by

$$f(t_1, t_2 | \mathbf{x}) = \frac{abG^{a-2}(t_1, t_2 | \mathbf{x}) [A(t_1, t_2 | \mathbf{x}) + B(t_1, t_2 | \mathbf{x}) + C(t_1, t_2 | \mathbf{x})]}{[1 - G^a(t_1, t_2 | \mathbf{x})]^{1-b}}, \quad (2.4)$$

where the functions  $A(\cdot), B(\cdot), C(\cdot)$ , and  $G(\cdot)$  are given in the Appendix.

To illustrate some possible shapes of the pdf, we set the parameter values  $\mu_1 = 4.5, \sigma_1 = 1.5, \mu_2 = 5, \sigma_2 = 1.5, a = 1.5, b = 0.9$ , and take  $\alpha = 0.99$  (when the lifetimes,  $t_1$  and  $t_2$ , are independent),  $\alpha = 0.50$  (when there is some dependency between the lifetimes  $t_1$  and  $t_2$ ), and  $\alpha = 0.01$  (when the lifetimes,  $t_1$  and  $t_2$ , are highly dependent), respectively. Figure 1 displays the plots of the pdf.

The corresponding marginal pdfs  $f(t_k | \mathbf{x})$  and marginal cdfs  $F(t_k | \mathbf{x})$  are given by

$$f(t_k | \mathbf{x}) = \frac{ab}{\sigma_k} \left[ \exp(-\mathbf{x}^T \boldsymbol{\beta}_k) \right]^{\frac{1}{\sigma_k}} t_k^{\frac{1}{\sigma_k} - 1} \exp \left\{ - \left[ \exp(-\mathbf{x}^T \boldsymbol{\beta}_k) t_k \right]^{\frac{1}{\sigma_k}} \right\} \times \left\{ 1 - \exp \left\{ - \left[ \exp(-\mathbf{x}^T \boldsymbol{\beta}_k) t_k \right]^{\frac{1}{\sigma_k}} \right\} \right\}^{a-1} \left\{ 1 - \left[ 1 - \exp \left\{ - \left[ \exp(-\mathbf{x}^T \boldsymbol{\beta}_k) t_k \right]^{\frac{1}{\sigma_k}} \right\} \right]^a \right\}^{b-1} \quad (2.5)$$

and

$$F(t_k|\mathbf{x}) = 1 - \left\{ 1 - \left[ 1 - \exp \left\{ - \left[ \exp \left( -\mathbf{x}^T \boldsymbol{\beta}_k \right) t_k \right]^{\frac{1}{\sigma_k}} \right\} \right]^a \right\}^b, \tag{2.6}$$

respectively, for  $k = 1, 2$ .

If  $|z| < 1$  and  $b$  is real non-integer, the power series

$$(1 - z)^{b-1} = \sum_{i=0}^{\infty} (-1)^i \binom{b-1}{i} z^i \tag{2.7}$$

holds, where the binomial coefficient is defined for any real number. Let  $T_k$  be a random variable with pdf (2.5) for  $k = 1, 2$ . We consider here only the general case: the parameters  $a$  and  $b$  are real non-integers. By applying (2.7) to (2.5) and expanding the binomial term, the marginal density function of  $T_k$  reduces to

$$f(t_k|\mathbf{x}) = \sum_{r=0}^{\infty} w_r g_{\lambda_{r,k}, c_k}(t_k|\mathbf{x}), \tag{2.8}$$

where  $g_{\lambda_{r,k}, c_k}(x)$  denotes the Weibull density function with shape parameter  $c_k = \sigma_k^{-1}$  and scale parameter  $\lambda_{r,k} = (r + 1)^{\sigma_k} \exp(-\mathbf{x}^T \boldsymbol{\beta}_k)$  and the coefficients  $w_r$  are given by

$$w_r = \frac{ab}{r+1} \sum_{j=0}^{\infty} (-1)^{j+r} \binom{b-1}{j} \binom{(j+1)a-1}{r}.$$

Equation (2.8) reveals that the marginal density function of  $T_k$  is an infinite linear combination of Weibull densities. We can easily check using computer software that  $\sum_{r=0}^{\infty} w_r = 1$  as expected. Thus, the ordinary, inverse and factorial moments and generating function of  $T_k$  can be obtained from an infinite weighted linear combination of those quantities for Weibull distributions.

The quantile function of  $T_k$  is readily obtained by inverting (2.6) as

$$t_k = \left( \frac{1}{\lambda_k} \right) \left[ -\log \left\{ 1 - \left[ 1 - (1 - u)^{\frac{1}{b}} \right]^{\frac{1}{a}} \right\} \right]^{\sigma_k}, \tag{2.9}$$

where  $\lambda_k = \exp(-\mathbf{x}^T \boldsymbol{\beta}_k)$  and  $u$  is uniform on the unit interval  $(0, 1)$ .

The survival function associated with (2.3) can be expressed as

$$S(t_1, t_2|\mathbf{x}) = 1 - F(t_1|\mathbf{x}) - F(t_2|\mathbf{x}) + F(t_1, t_2|\mathbf{x}),$$

where the functions  $F(t_1, t_2|\mathbf{x})$  and  $F(t_1|\mathbf{x})$ ,  $F(t_2|\mathbf{x})$  are defined in (2.3) and (2.6), respectively. Consequently, the joint survival function for the model (2.3) reduces to

$$\begin{aligned} S(t_1, t_2|\mathbf{x}) &= \left[ 1 - \left[ 1 - \exp \left\{ - \left[ \exp \left( -\mathbf{x}^T \boldsymbol{\beta}_1 \right) t_1 \right]^{\frac{1}{\sigma_1}} \right\} \right]^a \right]^b + \left[ 1 - \left[ 1 - \exp \left\{ - \left[ \exp \left( -\mathbf{x}^T \boldsymbol{\beta}_2 \right) t_2 \right]^{\frac{1}{\sigma_2}} \right\} \right]^a \right]^b \\ &\quad - \left[ 1 - \left[ \exp \left\{ - \left\{ \left[ \exp \left( -\mathbf{x}^T \boldsymbol{\beta}_1 \right) t_1 \right]^{\frac{1}{\sigma_1 a}} + \left[ \exp \left( -\mathbf{x}^T \boldsymbol{\beta}_2 \right) t_2 \right]^{\frac{1}{\sigma_2 a}} \right\} \right]^a \right] \right]^b \\ &\quad - \exp \left\{ - \left[ \exp \left( -\mathbf{x}^T \boldsymbol{\beta}_1 \right) t_1 \right]^{\frac{1}{\sigma_1}} \right\} - \exp \left\{ - \left[ \exp \left( -\mathbf{x}^T \boldsymbol{\beta}_2 \right) t_2 \right]^{\frac{1}{\sigma_2}} \right\} + 1 \right\}^b. \end{aligned} \tag{2.10}$$

The marginal survival function of  $T_k$  is given by

$$S(t_k|\mathbf{x}) = \left\{ 1 - \left[ 1 - \exp \left\{ - \left[ \exp \left( -\mathbf{x}^T \boldsymbol{\beta}_k \right) t_k \right]^{\frac{1}{\sigma_k}} \right\} \right]^a \right\}^b. \tag{2.11}$$

### 3. Classical inference and diagnostics analysis

Let  $(t_{1k}, \delta_{1k}, \mathbf{x}_1), \dots, (t_{nk}, \delta_{nk}, \mathbf{x}_n)$  be an observed sample of  $n$  independent observations, where  $t_{ik}$  represents the failure-time or the censoring-time,  $\delta_{ik}$  is a censoring indicator and  $\mathbf{x}_i = (\mathbf{x}_{i1}, \dots, \mathbf{x}_{ip})^T$  is the vector of explanatory variables associated with the  $i^{th}$  individual, where  $k = 1, 2$  and  $i = 1, 2, \dots, n$ . The likelihood function for bivariate data was considered by Lawless (2003) and He and Lawless (2005). By taking explanatory variables  $\mathbf{x}_i$  and the joint survival function (2.10), the log-likelihood function for the BKwW regression model follows by summing the contributions from each one of the  $n$  individuals

$$l(\boldsymbol{\Psi}) = \sum_{i=1}^n \left\{ \delta_{i1} \delta_{i2} \log(f(t_{i1}, t_{i2}|\mathbf{x}_i)) + \delta_{i1}(1 - \delta_{i2}) \log \left[ \frac{-\partial S(t_{i1}, t_{i2}|\mathbf{x}_i)}{\partial t_{i1}} \right] + (1 - \delta_{i1}) \delta_{i2} \log \left[ \frac{-\partial S(t_{i1}, t_{i2}|\mathbf{x}_i)}{\partial t_{i2}} \right] + (1 - \delta_{i1})(1 - \delta_{i2}) \log(S(t_{i1}, t_{i2}|\mathbf{x}_i)) \right\}. \tag{3.1}$$

The density function  $f(t_{i1}, t_{i2}|\mathbf{x}_i)$  is defined in (2.4), where  $\boldsymbol{\Psi} = (a, b, \alpha, \boldsymbol{\beta}_k^T, \sigma_k)^T$  is an unknown parameter vector and  $\boldsymbol{\beta}_k^T = (\beta_{0k}, \beta_{1k}, \dots, \beta_{pk})$  for  $k = 1, 2$ . The dimension of  $\boldsymbol{\Psi}$  is  $(2p + 7)$ .

The log-likelihood function (3.1) has the following restrictions on the parameters:  $a > 0$ ,  $b > 0$ ,  $0 < \alpha \leq 1$ , and  $\sigma_k > 0$  for  $k = 1, 2$  and  $i = 1, \dots, n$ . So, to continue the estimation process it is necessary to rewrite the log-likelihood function including restrictions on the parametric space. We then consider the general problem of maximizing (3.1) subject to the linear constraints  $\boldsymbol{o}_j^T \boldsymbol{\Psi} - c_j^* \geq 0$ , where  $\boldsymbol{o}_j, j = 1, 2, \dots, q$  are  $(2p + 7) \times 1$  vectors and  $c_j^*$  are scalars, both known and fixed numbers. In this study,  $c_j^*$  has value zero for the cases:  $a > 0$ ,  $b > 0$ ,  $\alpha > 0$ , and  $\sigma_k > 0$  and value one for  $\alpha \leq 1$ . To solve this problem, we use the adaptive barrier method (Lange, 1999). Thus, the log-likelihood function subject to linear constraints on the parameters reduces to

$$l_R(\boldsymbol{\Psi}, \vartheta) = l(\boldsymbol{\Psi}) + \vartheta \sum_{j=1}^q (\boldsymbol{o}_j^T \boldsymbol{\Psi} - c_j^*), \tag{3.2}$$

where  $\vartheta > 0$  is the multiplier barrier term,  $\boldsymbol{o}_j^T \boldsymbol{\Psi} - c_j^* \geq 0$  is the linear inequality constraint for  $j = 1, \dots, q$ ,  $\boldsymbol{\Psi} = (a, b, \alpha, \boldsymbol{\beta}_k^T, \sigma_k)^T$ , and  $\boldsymbol{\beta}_k^T = (\beta_{0k}, \beta_{1k}, \dots, \beta_{pk})$ .

The MLEs under constraints on the parameters in  $\boldsymbol{\Psi}$  can be calculated numerically by maximizing (3.2). We can adopt the statistical software R to compute the estimate  $\hat{\boldsymbol{\Psi}}$ . However, it is usually more convenient to use the `constrOptim` and function to maximize this function numerically. Initial values for  $\hat{\boldsymbol{\Psi}}$  are taken from the fit of the bivariate Weibull regression model corresponding to  $a = b = 1$ .

Under standard regularity conditions, the asymptotic distribution of  $(\hat{\boldsymbol{\Psi}} - \boldsymbol{\Psi})$  is multivariate normal  $N_{(2p+7)}(0, I(\boldsymbol{\Psi})^{-1})$ , where  $I(\boldsymbol{\Psi})$  is the expected information matrix for the  $(2p + 7) \times 1$  vector of parameters. The asymptotic covariance matrix  $I(\boldsymbol{\Psi})$  of  $\hat{\boldsymbol{\Psi}}^{-1}$  can be approximated by the  $(2p + 7) \times (2p + 7)$  inverse of the observed information matrix  $\check{L}_R(\boldsymbol{\Psi}, \vartheta) = -\{\partial^2 l_R(\boldsymbol{\Psi}, \vartheta) / (\partial \boldsymbol{\Psi} \partial \boldsymbol{\Psi}^T)\}$  evaluated at  $\boldsymbol{\Psi} = \hat{\boldsymbol{\Psi}}$ .

The approximate multivariate normal  $N_{(2p+7)}(0, -\check{L}_R(\boldsymbol{\Psi}, \vartheta)^{-1})$  distribution can be used to construct confidence intervals for the parameters in  $\boldsymbol{\Psi}$  in the usual way.

Besides estimation, hypothesis tests is another key issue. Let  $\Psi_1$  and  $\Psi_2$  be proper disjoint subsets of  $\Psi$ . Suppose that we want to test  $H_0 : \Psi_1 = \Psi_{10}$  versus  $H_1 : \Psi_1 \neq \Psi_{10}$ , where  $\Psi_2$  is a nuisance parameter vector. Let  $\hat{\Psi}_0$  be the MLEs under  $H_0$ , and define the likelihood ratio (LR) statistic  $w = 2\{\ell(\hat{\Psi}) - \ell(\hat{\Psi}_0)\}$ . Under  $H_0$  and standard regularity conditions,  $w$  converges to a chi-square distribution with  $\dim(\Psi_1)$  degrees of freedom.

### 3.1. Diagnostics analysis: global influence

The first tool to perform sensitivity analysis, is by means of global influence starting from case deletion (Cook, 1977), which is a common approach to study the effect of dropping the  $i^{th}$  case from the data. Case deletion for model (2.3) can be expressed as

$$F(t_{(i)1}, t_{(i)2} | \mathbf{x}) = 1 - \left[ 1 - \left\{ \exp \left[ - \left\{ \exp(-\mathbf{x}_{(i)}^T \boldsymbol{\beta}_1) t_{(i)1} \right]^{\frac{1}{\sigma_1^\alpha}} + \left[ \exp(-\mathbf{x}_{(i)}^T \boldsymbol{\beta}_2) t_{(i)2} \right]^{\frac{1}{\sigma_2^\alpha}} \right\}^\alpha \right] - \exp \left\{ - \left[ \exp(-\mathbf{x}_{(i)}^T \boldsymbol{\beta}_1) t_{(i)1} \right]^{\frac{1}{\sigma_1}} \right\} - \exp \left\{ - \left[ \exp(-\mathbf{x}_{(i)}^T \boldsymbol{\beta}_2) t_{(i)2} \right]^{\frac{1}{\sigma_2}} + 1 \right\}^\alpha \right]^{a^b}. \quad (3.3)$$

Accordingly, a quantity with subscript “ $(i)$ ” means the original quantity with the  $i^{th}$  observation deleted. For model (3.3), the log-likelihood function is denoted by  $l_{R(i)}(\Psi, \vartheta)$ .

Let  $\hat{\Psi}_{(i)} = (\hat{a}_{(i)}, \hat{b}_{(i)}, \hat{\alpha}_{(i)}, \hat{\boldsymbol{\beta}}_{k(i)}^T, \hat{\sigma}_{k(i)})^T$  be the MLEs under constraints on the parameters in  $\Psi$  obtained by maximizing  $l_{R(i)}(\Psi, \vartheta)$ . To assess the influence of the  $i^{th}$  observation on the MLEs under constraints on the parameters in  $\Psi$ , the idea is to compare the difference between  $\hat{\Psi}_{(i)}$  and  $\hat{\Psi}$ . If deletion of an observation seriously influences the estimates, more attention should be directed to that observation. Hence, if  $\hat{\Psi}_{(i)}$  is far from  $\hat{\Psi}$ , then the  $i^{th}$  case is regarded as an influential observation. A first global influence measure of the  $i^{th}$  observation is the standardized norm of  $\hat{\Psi}_{(i)} - \hat{\Psi}$  (generalized Cook distance; GD) given by

$$GD_i(\Psi) = (\hat{\Psi}_{(i)} - \hat{\Psi})^T [\ddot{L}_R(\Psi, \vartheta)] (\hat{\Psi}_{(i)} - \hat{\Psi}).$$

So, we can assess the values of  $GD_i(a)$ ,  $GD_i(b)$ ,  $GD_i(\alpha)$ ,  $GD_i(\boldsymbol{\beta}_k)$ , and  $GD_i(\sigma_k)$  to estimate the impact of the  $i^{th}$  observation on the estimates of  $a$ ,  $b$ ,  $\alpha$ ,  $\boldsymbol{\beta}_k$ , and  $\sigma_k$ , respectively. Another popular measure of the difference between  $\hat{\Psi}_{(i)}$  and  $\hat{\Psi}$  is the likelihood displacement (LD)

$$LD_i(\Psi) = 2 \left[ l_R(\hat{\Psi}, \vartheta) - l_{R(i)}(\hat{\Psi}_{(i)}, \vartheta) \right].$$

### 3.2. Diagnostics analysis: local influence

Since regression models are sensitive to the underlying model assumptions, performing sensitivity analysis in general is strongly advisable. Another approach is suggested by Cook (1986), where instead of removing observations, weights are given to them. Local influence calculation can be conducted for model (2.3). If likelihood displacement  $LD(\omega) = 2\{l(\hat{\Psi}) - l(\hat{\Psi}_\omega)\}$  is used, where  $\hat{\Psi}_\omega$  denotes the MLE under the perturbed model, the normal curvature for  $\Psi$  at the direction  $\mathbf{d}$ ,  $\|\mathbf{d}\| = 1$ , is given by  $C_d(\Psi) = 2[\mathbf{d}^T \Delta^T [\ddot{L}(\Psi)]^{-1} \Delta \mathbf{d}]$ , where  $\Delta$  is a  $(2p + 7) \times n$  matrix that depends on the perturbation scheme, and whose elements are given by  $\Delta_{vi} = \partial^2 l(\Psi | \omega) / \partial \Psi_v \partial \omega_i$ ,  $i = 1, \dots, n$  and  $v = 1, 2, \dots, (2p + 7)$  evaluated at  $\hat{\Psi}$  and  $\omega_0$ , where  $\omega_0$  is the no perturbation vector (Cook, 1986). For the bivariate regression model, the elements of  $\ddot{L}(\Psi)$  can be evaluated numerically. We can also calculate normal curvatures  $C_d(a)$ ,  $C_d(b)$ ,  $C_d(\alpha)$ ,  $C_d(\boldsymbol{\beta}_k)$ , and  $C_d(\sigma_k)$  to perform various index plots,

for instance, the index plot of  $\mathbf{d}_{\max}$ , the eigenvector corresponding to  $C_{\mathbf{d}_{\max}}$ , the largest eigenvalue of the matrix  $\mathbf{B} = -\mathbf{\Delta}^T [\ddot{\mathbf{L}}(\boldsymbol{\Psi})]^{-1} \mathbf{\Delta}$ . The index plots of  $C_{\mathbf{d}_i}(a)$ ,  $C_{\mathbf{d}_i}(b)$ ,  $C_{\mathbf{d}_i}(\alpha)$ ,  $C_{\mathbf{d}_i}(\boldsymbol{\beta}_k)$ , and  $C_{\mathbf{d}_i}(\sigma_k)$ , called total local influence, where  $\mathbf{d}_i$  denotes an  $n \times 1$  vector of zeros with one at the  $i^{\text{th}}$  position. Thus, the curvature at direction  $\mathbf{d}_i$  takes the form  $C_i = 2|\Delta_i^T [\ddot{\mathbf{L}}(\boldsymbol{\Psi})]^{-1} \Delta_i|$ , where  $\Delta_i^T$  denotes the  $i^{\text{th}}$  row of  $\mathbf{\Delta}$ . It is usual to show those cases such that  $C_i \geq 2\bar{C}$ , where  $\bar{C} = (1/n) \sum_{i=1}^n C_i$ .

Consider the vector of weights  $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^T$ . Under three perturbation schemes (case-weight perturbation, response perturbation and explanatory variable perturbation), we can easily derive from the log-likelihood (3.2) the matrix

$$\mathbf{\Delta} = (\Delta_{vi})_{[(2p+7) \times n]} = \left( \frac{\partial^2 l(\boldsymbol{\Psi}|\boldsymbol{\omega})}{\partial \Psi_v \partial \omega_i} \right)_{[(2p+7) \times n]},$$

where  $v = 1, \dots, (2p + 7)$  and  $i = 1, \dots, n$ .

#### 4. Bayesian inference and influence diagnostics

In this section, we consider the Bayesian method as an alternative analysis to incorporate previous knowledge of the parameters through informative prior density functions.

Let  $(t_{1k}, \delta_{1k}, \mathbf{x}_1), \dots, (t_{nk}, \delta_{nk}, \mathbf{x}_n)$  be an observed sample of  $n$  independent observations, where  $t_{ik}$  represents the failure-time or the censoring-time,  $\delta_{ik}$  is a censoring indicator and  $\mathbf{x}_i = (\mathbf{x}_{i1}, \dots, \mathbf{x}_{ip})^T$  is the vector of explanatory variables associated with the  $i^{\text{th}}$  individual, where  $k = 1, 2$  and  $i = 1, \dots, n$ . The log-likelihood function  $l(\boldsymbol{\Psi})$  for the model parameters  $\boldsymbol{\Psi} = (a, b, \alpha, \boldsymbol{\beta}_k^T, \sigma_1, \sigma_2)^T$  of the BKwW regression model is given by equation (3.1).

We consider that  $a, b, \alpha, \boldsymbol{\beta}_k, \sigma_1, \sigma_2$  have independent priors,

$$\pi(\boldsymbol{\Psi}) = \pi(\alpha)\pi(\sigma_1)\pi(\sigma_2)\pi(a)\pi(b) \prod_{k=1}^2 \pi(\boldsymbol{\beta}_k), \tag{4.1}$$

where

$$\pi(\boldsymbol{\beta}_k) = \pi(\beta_{0k}, \beta_{1k}, \dots, \beta_{pk}) = \prod_{j^*=0}^p \pi(\beta_{j^*k}).$$

Further, we assume the following prior distributions  $\beta_{j^*k} \sim N(0, 10^2)$  for  $k = 1, 2$  and  $j^* = 0, \dots, p$ ,  $\log(\alpha/(1 + \alpha)) \sim N(0, 10^2)$ ,  $\log(\sigma_1) \sim N(0, 10^2)$ ,  $\log(\sigma_2) \sim N(0, 10^2)$ ,  $\log(a) \sim N(0, 10^2)$ , and  $\log(b) \sim N(0, 10^2)$ . The normal distribution with mean  $\mu$  and variance  $\tau^2$  is denoted by  $N(\mu, \tau^2)$ . All the hyper-parameters have been specified to express non-informative priors.

By combining the likelihood function (exponential of the log-likelihood (3.1)) and the prior distribution (4.1), we obtain the joint posterior distribution, which is analytically intractable. Then, we have based our inference on the MCMC simulation methods. By changing variables  $\boldsymbol{\xi} = (\log[\alpha/(1 + \alpha)], \log(\sigma_1), \log(\sigma_2), \boldsymbol{\beta}_k, \log(a), \log(b))$ , the parameter space is transformed into  $\mathcal{R}^{(2p+7)}$  (necessary for the work with Gaussian densities).

To implement the Metropolis-Hastings algorithm, we proceed as follows.

- (1) Start with any point  $\boldsymbol{\xi}_{(0)}$  and stage indicator  $\nu = 0$ ;
- (2) Generate a point  $\boldsymbol{\xi}'$  according to the transitional kernel  $Q(\boldsymbol{\xi}', \boldsymbol{\xi}_\nu) = N_{p+2}(\boldsymbol{\xi}_\nu, \tilde{\boldsymbol{\Sigma}})$ , where  $\tilde{\boldsymbol{\Sigma}}$  is the covariance matrix of  $\boldsymbol{\xi}$ , which is the same in any stage;

- (3) Update  $\xi_{(v)}$  to  $\xi_{(v+01)} = \xi'$  with probability  $p_v^* = \min\{1, \pi(\xi'|\mathcal{D})/\pi(\xi_{(v)}|\mathcal{D})\}$ , or keep  $\Psi_{(v)}$  with probability  $1 - p_v^*$ ;
- (4) Repeat Steps (2) and (3) by increasing the stage indicator until the process has reached a stationary distribution.

All computations are performed in R software (R Development Core Team, 2016). In all the work, after 20,000 sample burn-in, we use every tenth sample from the 60,000 MCMC posterior samples to reduce the autocorrelations and yield better convergence results, thus obtaining an effective sample of size 6,000 upon which the posterior is based on. We monitor the convergence of the Metropolis-Hasting algorithm using the method proposed by Geweke (1992) as well as trace plots.

#### 4.1. Model comparison criteria

A variety of methodologies can be applied for comparing several competing models for a given dataset and selecting those which provide the best fits to the data. We consider some of the Bayesian model selection criteria, namely, the deviance information criterion (DIC) proposed by Spiegelhalter *et al.* (2002), the expected Akaike information criterion (EAIC) by Brooks (2002) and the expected Bayesian (or Schwarz) information criterion (EBIC) by Carlin and Louis (2001). Let  $\Psi^{(1)}, \dots, \Psi^{(Q)}$  be a sample of size  $Q$  of  $\pi(\Psi|\mathcal{D})$  after the burn-in, where  $\mathcal{D}$  denote the full data. They are based on the posterior mean of the deviance, which can be approximated by  $\bar{d} = \sum_{q^*=1}^Q d(\Psi_{q^*})/Q$ , where  $d(\Psi) = -2 \sum_{i=1}^n \log[f(t_{1i}, t_{2i}|\Psi)]$ . The DIC criterion can be estimated using the MCMC output by  $\widehat{\text{DIC}} = \bar{d} + \hat{\rho}_d = 2\bar{d} - \hat{d}$ , where  $\rho_d$  is the effective number of parameters defined as  $E\{d(\Psi)\} - d\{E(\Psi)\}$ , and  $d\{E(\Psi)\}$  is the deviance evaluated at the posterior mean.

Similarly, the EAIC and EBIC criteria can be estimated by  $\widehat{\text{EAIC}} = \bar{d} + 2\#(\Psi)$  and  $\widehat{\text{EBIC}} = \bar{d} + \#(\Psi) \log(n)$ , where  $\#(\Psi)$  is the number of model parameters. Comparing alternative models, the preferred model is the one with the smallest criteria values.

Other criteria such as LPML is derived from conditional predictive ordinate (CPO) statistics. Let  $\mathcal{D}^{(-i)}$  denote the data with the deleted  $i^{\text{th}}$  observation. We denote the posterior density of  $\Psi$  given  $\mathcal{D}^{(-i)}$  by  $\pi(\Psi|\mathcal{D}^{(-i)})$ ,  $i = 1, \dots, n$ . For the  $i^{\text{th}}$  observation, CPO<sub>*i*</sub> is given by

$$\text{CPO}_i = \int_{\Psi} f(t_{1i}, t_{2i}|\Psi) \pi(\Psi|\mathcal{D}^{(-i)}) d\Psi = \left\{ \int_{\Psi} \frac{\pi(\Psi|\mathcal{D})}{f(t_{1i}, t_{2i}|\Psi)} d\Psi \right\}^{-1}. \quad (4.2)$$

The CPO<sub>*i*</sub> can be interpreted as the height of the marginal density of the time for an event at  $t_i$ . Therefore, high CPO<sub>*i*</sub> implies a better fit of the model. No closed-form of CPO<sub>*i*</sub> is available for the proposed model. However, a Monte Carlo estimate of CPO<sub>*i*</sub> can be obtained by using a single MCMC sample from the posterior distribution  $\pi(\Psi|\mathcal{D})$ . A Monte Carlo approximation of CPO<sub>*i*</sub> (Ibrahim *et al.*, 2001) is given by

$$\widehat{\text{CPO}}_i = \left\{ \frac{1}{Q} \sum_{q^*=1}^Q \frac{1}{f(t_{1i}, t_{2i}|\Psi^{(q^*)})} \right\}^{-1}.$$

For model comparisons, we use the log pseudo marginal likelihood (LPML) defined by  $\text{LPML} = \sum_{i=1}^n \log(\widehat{\text{CPO}}_i)$ . The higher the LPML value, the better the fit of the model.



### 4.2. Bayesian case influence diagnostics

It is well known that regression models may be sensitive to underlying model assumptions. Therefore, a sensitivity analysis is strongly advisable. Cook (1986) uses this idea to motivate his assessment of influence analysis and suggests that more confidence should be put in a model relatively stable under small modifications. In order to investigate if some of the observations are influential for the Bayesian analysis, we consider the Bayesian case-deletion influence diagnostic measures for the joint posterior distribution based on the  $\psi$ -divergence (Peng and Dey, 1995)

Let  $D_\psi(P, P_{(-i)})$  denote the  $\psi$ -divergence between  $P$  and  $P_{(-i)}$ , in which  $P$  denotes the posterior distribution of  $\Psi$  for the full data, and  $P_{(-i)}$  denotes the posterior distribution of  $\Psi$  without the  $i^{th}$  case. Therefore,

$$D_\psi(P, P_{(-i)}) = \int \psi \left( \frac{\pi(\Psi|\mathcal{D}^{(-i)})}{\pi(\Psi|\mathcal{D})} \right) \pi(\Psi|\mathcal{D}) d\Psi, \tag{4.3}$$

where  $\psi$  is a convex function with  $\psi(1) = 0$ . Several choices concerning the  $\psi$  are given by Dey and Birimiwal (1994). For example,  $\psi(z^*) = -\log(z^*)$  defines the Kullback-Leibler (K-L) divergence,  $\psi(z^*) = 0.5|z^* - 1|$  defines the  $L_1$  norm (or variational distance) and  $\psi(z^*) = (z^* - 1) \log(z^*)$  gives the  $J$ -distance (or the symmetric version of the K-L divergence).

Let  $\Psi^{(1)}, \dots, \Psi^{(Q)}$  be a sample of size  $Q$  from  $\pi(\Psi|\mathcal{D})$ . Then,  $D_\psi(P, P_{(-i)})$  can be calculated by

$$\widehat{D}_\psi(P, P_{(-i)}) = \frac{1}{Q} \sum_{q^*=1}^Q \psi \left( \frac{\widehat{CPO}_i}{f(t_{1i}, t_{2i}|\Psi^{(q^*)})} \right), \tag{4.4}$$

where  $\widehat{CPO}_i = \{(1/Q) \sum_{q^*=1}^Q 1/f(t_{1i}, t_{2i}|\Psi^{(q^*)})\}^{-1}$  is the numerical approximation of the conditional predictive ordinate statistic of the  $i^{th}$  observation (Ibrahim, 2001).

Note that  $D_\psi(P, P_{(-i)})$  can be interpreted as the  $\psi$ -divergence of the effect of deleting the  $i^{th}$  case from the full data on the joint posterior distribution of  $\Psi$ . As pointed out by Peng and Dey (1995), it may be difficult for a practitioner to judge the cutoff point of the divergence measure so as to determine whether a small subset of observations is influential or not. By using a biased coin procedure (Peng and Dey, 1995), which has probability value  $\phi$ , the  $\psi$ -divergence between the biased and unbiased coins is given by

$$D_\psi(f_0, f_1) = \int \psi \left( \frac{f_0(y)}{f_1(y)} \right) f_1(y) dy, \tag{4.5}$$

where  $f_0(y) = \phi^y(1 - \phi)^{1-y}$  and  $f_1(y) = 0.5, y = 0, 1$ . If  $D_\psi(f_0, f_1) = d_\psi(\phi)$ , it can be easily checked that  $d_\psi$  satisfies the following equation

$$d_\psi(\phi) = \frac{\psi(2\phi) + \psi(2(1 - \phi))}{2}. \tag{4.6}$$

It is not difficult to see for the divergence measures considered that  $d_\psi$  increases as  $\phi$  moves away from 0.5. In addition,  $d_\psi(\phi)$  is symmetric about  $\phi = 0.5$  and  $d_\psi$  achieves its minimum value at  $\phi = 0.5$ . For this point,  $d_\psi(0.5) = 0$  and  $f_0 = f_1$ . Therefore, if we consider  $\phi > 0.90$  (or  $\phi \leq 0.10$ ) as a strong bias in a coin, then  $d_{K-L}(0.90) = 0.51, d_J(0.90) = 0.88$ , and  $d_{L_1}(0.90) = 0.4$ . In addition, an observation which  $d_J > 0.88$  can also be considered influential if we use the  $J$ -distance. Similarly, using the K-L divergence and the  $L_1$  norm, we can consider an influential observation when  $d_{K-L} > 0.51$  and  $d_{L_1} > 0.4$ , respectively.

## 5. Simulation study: maximum likelihood and Bayesian estimation

A simulation study is conducted to evaluate the parameter estimates for the proposed model. The results are obtained from 1,000 Monte Carlo simulations using the R software. In each replication, a random sample of size  $n$  is drawn from the BKwW regression model and the parameters are estimated by maximum likelihood and Bayesian estimation.

The samples denoted by  $(t_{11}, t_{21}), \dots, (t_{1n}, t_{2n})$  are generated according to the following steps:

- Step 0: Set up the sample size  $n$  and start with stage indicator  $i = 1$ .
- Step 1: Generate the covariate  $x_{1i}$  from a Bernoulli distribution with parameter 0.5 and the censorship time  $C_{1i}$  from a uniform distribution  $U(0, \tau_1)$ , where  $\tau_1$  controls the percentage of censored observations.
- Step 2: Use the quantile function given in (2.9) to obtain

$$T_{1i} = \left( \frac{1}{\lambda_{1i}} \right) \left[ -\log \left\{ 1 - \left[ 1 - (1 - u_{1i})^{\frac{1}{b}} \right]^{\frac{1}{a}} \right\} \right]^{\sigma_1},$$

where  $u_{1i} \sim U(0, 1)$  and  $\lambda_{1i} = \exp(-\beta_{01} - \beta_{11}x_{1i})$ .

- Step 3: Compare  $T_{1i}$  with  $C_{1i}$  in order to determine the indicator of censorship  $\delta_{1i}$  and the observed value given by  $t_{1i} = \min(T_{1i}, C_{1i})$ .
- Step 4: Generate the covariate  $x_{2i}$  from a Bernoulli distribution with parameter 0.5 and the censorship time  $C_{2i}$  from a uniform distribution  $U(0, \tau_2)$ , where  $\tau_2$  controls the percentage of censored observations.
- Step 5: Next,  $T_{2i}$  is generated using a random variable  $\eta_i \sim U(0, 1)$  and the solution of the nonlinear equation,  $\exp(-[{-\log(1 - (1 - (1 - u_{1i})^{1/b})^{1/a})}^{1/\alpha} + (-\log(1 - u_{2i}))^{1/\alpha}]^\alpha - 1 + u_{2i} + (1 - (1 - u_{1i})^{1/b})^{1/a} - \eta_i u_{1i} = 0$ , by considering  $T_{2i} = (1/\lambda_{2i})[-\log\{1 - [1 - (1 - u_{2i})^{1/b}]^{1/a}\}]^{\sigma_2}$ , where  $\lambda_{2i} = \exp(-\beta_{02} - \beta_{21}x_{2i})$ .
- Step 6: Compare  $T_{2i}$  with  $C_{2i}$  in order to determine the indicator of censorship  $\delta_{2i}$  and the observed value given by  $t_{2i} = \min(T_{2i}, C_{2i})$ .
- Step 7: Do  $i = i + 1$ . If  $i = n$  stop, else return to Step 1.

The simulation study is performed for  $n = 50, 100$ , and  $150$ . We consider the following values for the parameters of the model:  $\alpha = 0.75, \sigma_1 = 1.0, \beta_{01} = 4.0, \beta_{11} = 0.75, \sigma_2 = 1.25, \beta_{02} = 3.25, \beta_{12} = 1.75, a = 1.0$ , and  $b = 0.75$ . For all sample sizes, the percentage of censored observations in the times 1 and 2 are approximately 30% and 20%, respectively.

Tables 1 and 2 provide the averages (Mean), biases and the mean square errors (MSEs) of the MLEs and Bayesian estimates of the parameters in the BKwW regression model, respectively. We can note that the MSE values decrease when the sample size increases in agreement with first-order asymptotics.

Table 1: Mean estimates, Bias and MSEs of the MLEs of the parameters in the bivariate Kumaraswamy Weibull regression model

Parameter (true value)	N = 50			N = 100			N = 150		
	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
$\alpha$ (0.75)	0.8291	0.0791	0.0153	0.8265	0.0765	0.0105	0.8237	0.0737	0.0085
$\sigma_1$ (1.00)	1.4484	0.4484	0.2878	1.3876	0.3876	0.1908	1.3754	0.3754	0.1655
$\beta_{01}$ (4.00)	4.4953	0.4953	0.3299	4.5588	0.5588	0.3568	4.5528	0.5528	0.3350
$\beta_{11}$ (0.75)	0.7097	-0.0403	0.1078	0.6589	-0.0911	0.0639	0.6473	-0.1027	0.0484
$\sigma_2$ (1.25)	1.5893	0.3393	0.2878	1.5172	0.2672	0.1908	1.5064	0.2564	0.1655
$\beta_{02}$ (3.25)	2.8722	-0.3778	0.1935	2.8918	-0.3582	0.1549	2.9043	-0.3457	0.1355
$\beta_{12}$ (1.75)	1.7191	-0.0309	0.0742	1.7039	-0.0461	0.0386	1.6916	-0.0584	0.0274
$a$ (1.00)	0.5705	-0.4295	0.2116	0.5710	-0.4290	0.1977	0.5669	-0.4331	0.1963
$b$ (0.75)	0.9049	0.1549	0.0707	0.9256	0.1756	0.0517	0.9127	0.1627	0.0413

MSE = biases and the mean square error; MLE = maximum likelihood estimator.

Table 2: Mean estimates, biases and MSEs of the Bayesian estimates of the parameters in the bivariate Kumaraswamy Weibull regression model

Parameter (true value)	N = 50			N = 100			N = 150		
	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
$\alpha$ (0.75)	0.9027	0.1527	0.0433	0.8615	0.1115	0.0263	0.8466	0.0966	0.0187
$\sigma_1$ (1.00)	1.4323	0.4323	0.3582	1.3811	0.3811	0.2231	1.3689	0.3689	0.1859
$\beta_{01}$ (4.00)	4.5309	0.5309	0.4626	4.5699	0.5699	0.4057	4.5582	0.5582	0.3619
$\beta_{11}$ (0.75)	0.7189	-0.0311	0.2439	0.6658	-0.0842	0.1183	0.6542	-0.0958	0.0815
$\sigma_2$ (1.25)	1.5642	0.3142	0.3582	1.5061	0.2561	0.2231	1.4923	0.2423	0.1859
$\beta_{02}$ (3.25)	2.8727	-0.3773	0.2455	2.8966	-0.3534	0.1716	2.9022	-0.3478	0.1515
$\beta_{12}$ (1.75)	1.7407	-0.0093	0.1619	1.7200	-0.0300	0.0731	1.6987	-0.0513	0.0493
$a$ (1.00)	0.5730	-0.4270	0.2515	0.5732	-0.4268	0.2140	0.5645	-0.4355	0.2084
$b$ (0.75)	0.9153	0.1653	0.2987	0.9190	0.1690	0.1058	0.8973	0.1473	0.0559

MSE = biases and the mean square error; MLE = maximum likelihood estimator.

### 5.1. Influence of outlying observations

A goal of this study is to show the need for robust models to deal with the presence of outliers in the data. We generate a sample of length 200 with fixed parameters  $\alpha = 0.5$ ,  $\sigma_1 = 3.0$ ,  $\beta_{01} = -2.5$ ,  $\beta_{11} = -3.5$ ,  $\sigma_2 = 1.5$ ,  $\beta_{02} = -2.0$ ,  $\beta_{12} = -3.0$ ,  $a = 1.25$ , and  $b = 0.5$ .

We select cases 35, 95, and 132 for perturbation. The perturbation scheme is structured as following. To create influential observations artificially in the dataset, we choose one, two or three of these selected cases. For each case, we perturb one or both lifetimes as follows  $\tilde{t}_i = t_i + 5S_t$ ,  $i = 1, 2$ , where  $S_t$  is the standard deviation of the  $t_i$ 's. For case 35, we perturb only the lifetime  $t_1$  and, for case 132, the lifetime  $t_2$ , and for case 95, both lifetimes are perturbed.

In this study, we consider eight setups. Dataset a: original dataset without outliers; Dataset b: data with outlier 35; Dataset c: data with outlier 95; Dataset d: data with outlier 132; Dataset e: data with outliers 35 and 95; Dataset f: data with outliers 35 and 132; Dataset g: data with outliers 95 and 132; and dataset h: data with outliers 35, 95, and 132. The MCMC computations are made similar to those in the last subsection using the Geweke's convergence diagnostic (Geweke, 1992) to monitor the convergence of the Gibbs samples as well as trace plots.

Table 3 reveals that the posterior inferences about the parameters are sensitive to the perturbation of the selected case(s), except the  $\beta$  parameters. Table 4 gives the Monte Carlo estimates of the measures DIC, EAIC, EBIC, and LPML for each perturbed version of the original data. As expected, the original simulated data (dataset a) yields the best fit.

**Table 3:** Posterior means and standard deviations (SDs) for the bivariate Kumaraswamy Weibull regression model according to different perturbation schemes

Dataset	Perturbed case	$\alpha$	$\sigma_1$	$\beta_{01}$	$\beta_{11}$	$\sigma_2$	$\beta_{02}$	$\beta_{12}$	a	b
		Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
a	None	0.6508 (0.0578)	3.5380 (0.3078)	-2.2298 (0.0458)	-3.6836 (0.0525)	1.3794 (0.1332)	-2.6428 (0.1139)	-3.3939 (0.1386)	0.8366 (0.1470)	0.6921 (0.0917)
b	35	0.6238 (0.0636)	3.0476 (0.2550)	-2.2270 (0.0557)	-3.7769 (0.0542)	1.2228 (0.1147)	-2.6994 (0.1314)	-3.5492 (0.1447)	0.9812 (0.1800)	0.6296 (0.0881)
c	95	0.6244 (0.0610)	3.0751 (0.2511)	-2.2256 (0.0532)	-3.7723 (0.0544)	1.2334 (0.1124)	-2.6925 (0.1282)	-3.5439 (0.1432)	0.9677 (0.1709)	0.6288 (0.0879)
d	132	0.4392 (0.0487)	2.5259 (0.2528)	-2.2664 (0.0592)	-3.6404 (0.0606)	0.8930 (0.0827)	-2.6394 (0.1666)	-3.3198 (0.1710)	1.5083 (0.2647)	1.1779 (0.1709)
e	{35, 95}	0.6121 (0.0640)	2.8289 (0.2238)	-2.2190 (0.0606)	-3.8397 (0.0565)	1.1567 (0.1060)	-2.7151 (0.1367)	-3.6568 (0.1509)	1.0416 (0.1908)	0.5838 (0.0829)
f	{35, 132}	0.3922 (0.0460)	2.1309 (0.2052)	-2.2725 (0.0712)	-3.7082 (0.0615)	0.7735 (0.0700)	-2.7021 (0.1956)	-3.4645 (0.1741)	1.8419 (0.3199)	1.2165 (0.1852)
g	{95, 132}	0.3958 (0.0458)	2.1574 (0.2071)	-2.2769 (0.0718)	-3.7204 (0.0586)	0.7797 (0.0694)	-2.7047 (0.1908)	-3.5108 (0.1730)	1.8169 (0.3232)	1.1387 (0.1637)
h	{35, 95, 132}	0.3803 (0.0456)	2.0017 (0.1878)	-2.2713 (0.0799)	-3.7651 (0.0641)	0.7451 (0.0672)	-2.7274 (0.2074)	-3.5925 (0.1747)	1.9430 (0.3573)	1.1302 (0.1907)

**Table 4:** Bayesian criteria for each perturbed version by fitting the bivariate Kumaraswamy Weibull regression model according to different perturbation schemes

Data names	Bayesian criteria			
	EAIC	EBIC	DIC	LPML
a	-2740.684	-2710.999	-2750.117	1374.758
b	-2717.665	-2687.981	-2726.491	1361.560
c	-2718.284	-2688.599	-2727.412	1362.171
d	-2693.735	-2664.050	-2703.276	1344.087
e	-2700.677	-2670.993	-2709.585	1353.458
f	-2666.795	-2637.110	-2675.774	1330.980
g	-2665.915	-2636.230	-2675.220	1332.170
h	-2650.958	-2621.273	-2660.200	1324.159

EAIC = expected Akaike information criterion; EBIC = expected Bayesian information criterion; DIC = deviance information criterion; LPML = log pseudo marginal likelihood.

We consider the sample from the posterior distributions of the parameters of the BKwW regression model to calculate the  $\psi$ -divergence measures in (4.3). The results in Table 5 reveal, before perturbation (Dataset a), that the selected cases are not influential according to all  $\psi$ -divergence measures. However, after perturbation (Dataset b–h), the measures increase, which indicates that the perturbed cases are influential.

Figures 2–4 display the four  $\psi$ -divergence measures for cases (a), (f), and (h), respectively. We can note that all measures identify influential case(s) and provide larger  $\psi$ -divergence measures in comparison to the other cases.

### 6. Application: renal insufficiency data

In this section, we perform a statistical analysis on the renal insufficiency study originally presented by McGilchrist and Aisbett (1991) using the BKwW regression model. Recently, these data are analyzed by Barriga *et al.* (2010) by considering a location scale model for bivariate survival times

Table 5: Case influence diagnostics for the simulated data

Data names	Case number	$\psi$ -divergence measure		
		$d_{K-L}$	$d_J$	$d_{L_1}$
a	35	0.0048	0.0096	0.0387
	95	0.0013	0.0025	0.0198
b	132	0.3832	0.7914	0.3426
	35	2.0466	4.3206	0.7223
c	95	1.7511	3.6743	0.6782
d	132	8.2482	15.1319	0.9833
e	35	0.7872	1.6328	0.4880
f	95	1.0060	2.0804	0.5427
	35	1.1026	2.4538	0.5679
g	132	6.7799	13.475	0.9618
	95	1.1436	2.4212	0.5735
h	132	5.3909	9.9707	0.9179
	35	0.5931	1.3437	0.4240
	95	1.1146	3.1001	0.5759
	132	5.3933	10.4104	0.9240

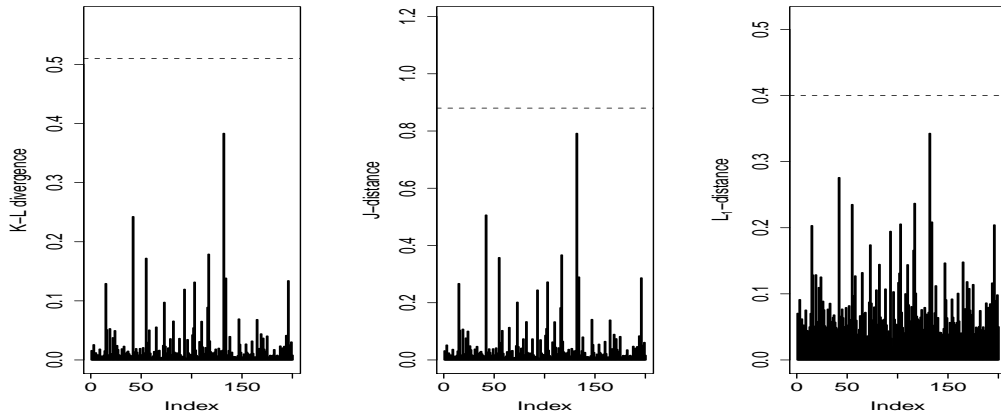


Figure 2:  $\psi$ -divergence measures from Dataset (a).

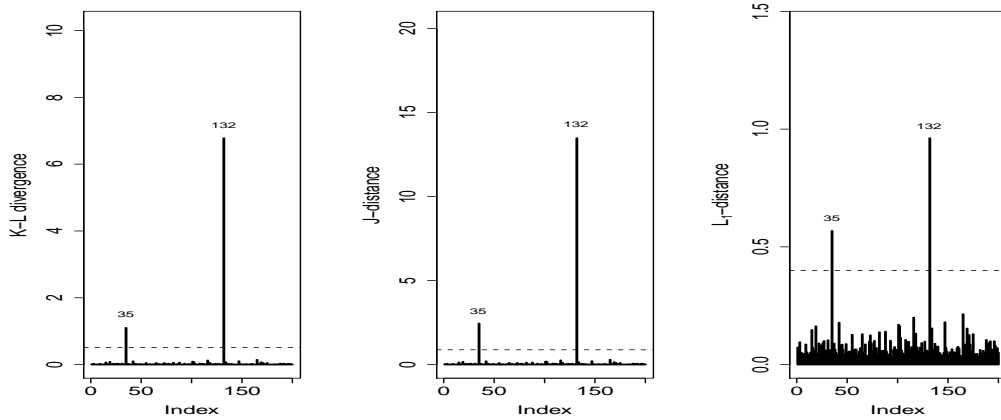


Figure 3:  $\psi$ -divergence measures from Dataset (f).

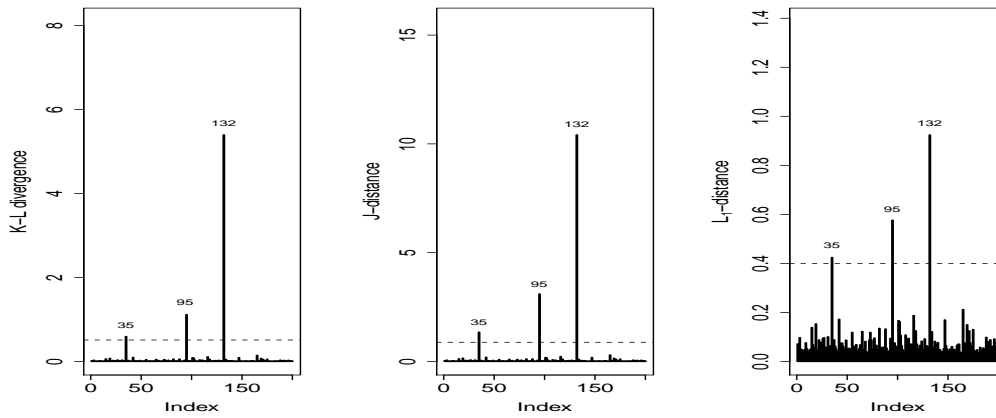
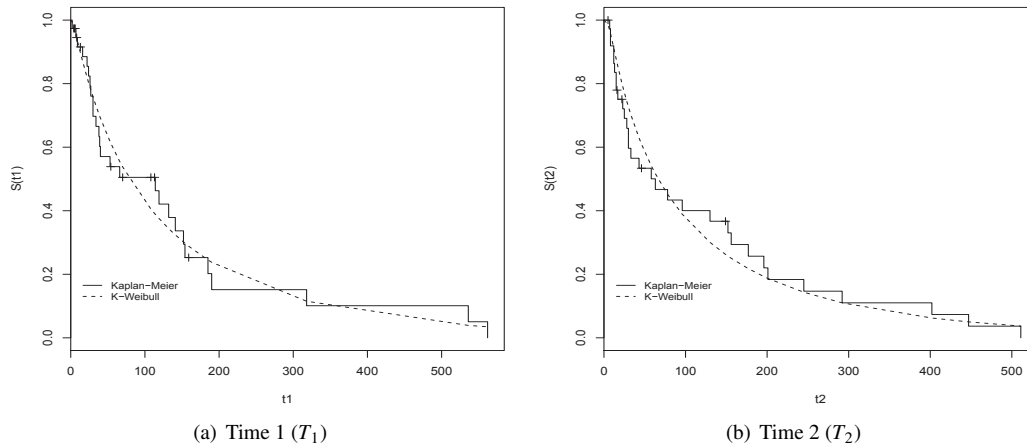
Figure 4:  $\psi$ -divergence measures from Dataset (h).

Figure 5: Marginal survival estimates by Kaplan-Meier and Kumaraswamy Weibull distribution based on the renal insufficiency data.

based on the proposal of a copula to model the dependence of bivariate survival data. The dataset refers to the occurrence times of two distinct infection events in patients suffering from renal insufficiency. The patients use portable dialysis machines and the occurrence of infection at the catheter insertion point is considered. When this occurs, the catheter must be removed until the infection is cured, when the catheter is reinserted. The times between insertion of the catheter and occurrence of an infection are recorded. Therefore, various infections can occur in the patients, but only two are recorded in this dataset. There are also situations when the catheter must be removed for reasons other than infection, in which case there is right censoring of the data. The response variables of interest are the times between catheter insertion and infection. For each patient  $i$ ,  $i = 1, 2, \dots, 38$ , the associated variables are:  $t_{i1}$ : time to occurrence of the first infection, in weeks;  $t_{i2}$ : time to occurrence of the second infection, in weeks;  $\delta_{i1}$ : censoring indicator of event 1;  $\delta_{i2}$ : censoring indicator of event 2;  $x_{i1}$ : patients sex (0: male, 1: female).

We perform an exploratory analysis, mainly considering the variables referring to the occurrence times of the two events of interest. The estimated Kendall correlation coefficient is  $\hat{\rho}_K = 0.03$ , thus

Table 6: MLEs for the bivariate Kumaraswamy Weibull regression model based on the renal insufficiency data

Parameter	Estimate	Standard error	<i>p</i> -value
$\beta_{01}$	2.5204	1.6159	0.1188
$\beta_{11}$	0.7075	0.4981	0.1555
$\beta_{02}$	1.7274	1.2462	0.1657
$\beta_{12}$	1.9164	0.3659	<0.0001
$\sigma_1$	1.6826	0.8100	-
$\sigma_2$	1.4891	0.6823	-
<i>a</i>	2.9146	1.1558	-
<i>b</i>	0.5782	0.1739	-
$\alpha$	0.9999	0.3743	-

MLEs = maximum likelihood estimators.

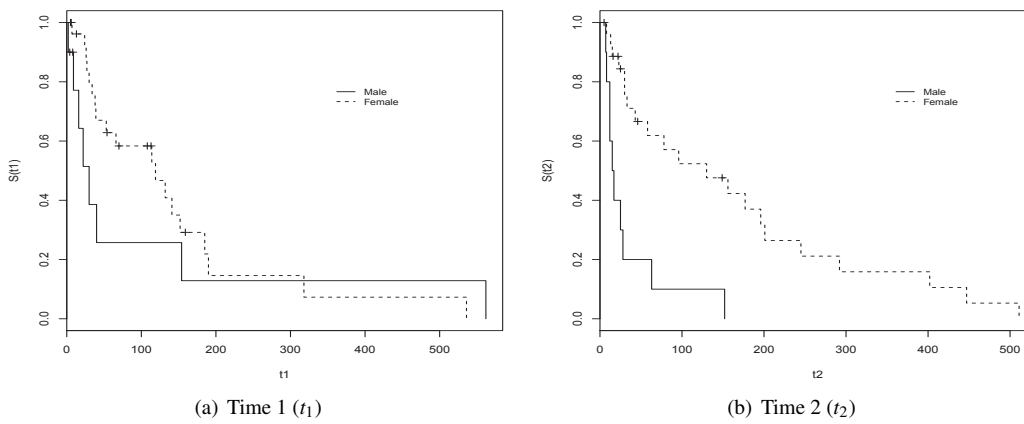


Figure 6: Survival curves marginally estimated by the Kaplan-Meier method for the renal insufficiency data for each event by means of the sex variable.

indicating an absence of correlation between the times of the events of interest. Despite this very low correlation, we adopt these data to illustrate the application of the BKwW regression model. The parameter referring to the association between the responses in the regression model must detect this independence, without impairing the model’s adjustment to these data. The Kaplan-Meier survival estimates and the marginal KwW distributions fitted to both events 1 and 2 are displayed in Figure 5. So, it is adequate to assume the BKwW model for the survival times. Since the KwW distribution is adequate for the times to events 1 and 2, it is coherent, for the case of bivariate data, to consider the bivariate Weibull cumulative distribution (Hougaard, 1986).

6.1. Results: classical inference

Based on the results of Sections 2 and 3, Table 6 lists the MLEs of the parameters for the BKwW regression model, their standard errors and *p*-values. The results indicate that male patients have accelerated times for the second infection than female patients. This fact can be confirmed visually by means of Figure 6. The estimated association parameter indicates that the lifetimes are independent and confirm the result obtained at the beginning of this section, when calculating  $\rho_K$ .

Diagnostic analysis

After modeling, it is important to verify whether there are observations that influence the fit of the BKwW regression model. To investigate this, we calculate global influence measures, likelihood

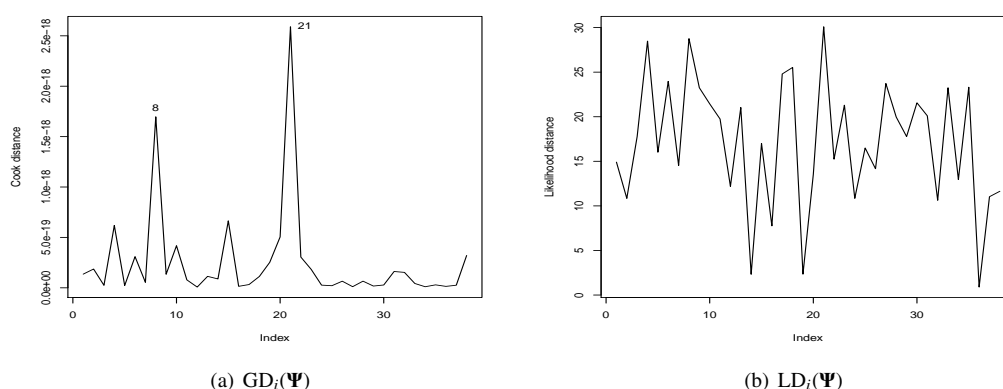


Figure 7: (a) Index plot of  $GD_i(\Psi)$  (generalized Cook distance). (b) Index plot of  $LD_i(\Psi)$  (likelihood distance) to the renal insufficiency data.

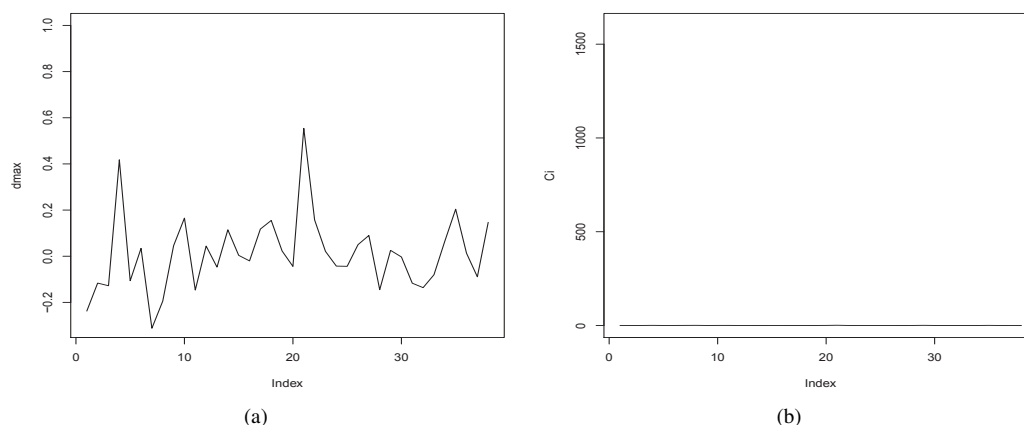


Figure 8: (a) Index plot of  $d_{\max}$  for  $\Psi$  (case-weight perturbation) and (b) total local influence for  $\Psi$  (case-weight perturbation) based on the fit of the BKwW model to the renal insufficiency data.

distance ( $LD_i(\Psi)$ ) and generalized Cook distance ( $GD_i(\Psi)$ ), as defined in Section 3, using the R software. The results of these influence measure index plots are displayed in Figure 7. Based on these plots, we note that case #21 is possibly an influential observation.

We apply the local influence theory developed in Section 3.2, where case-weight perturbation is used, and obtain the value of the maximum curvature  $C_{d_{\max}}(\Psi) = 2.83$ . Figure 8 shows the plot of the eigenvector corresponding to  $d_{\max}$  and  $C_i$  versus the observation index.

The influence of perturbations on the observed survival times is now analyzed (response variable perturbation). The value of the maximum curvature is  $C_{d_{\max}}(\Psi) = 1677.50$ . In Figure 9, we plot  $d_{\max}$  and  $C_i$  versus the observation index. Figure 9 indicates that the case #29 is distinguished from the others.

### Impact of the detected influential observations

We conclude that the diagnostic analysis (global influence and local influence) detect as potentially influential observations, the following two cases: #21 and #29. The observation #21 corresponds to



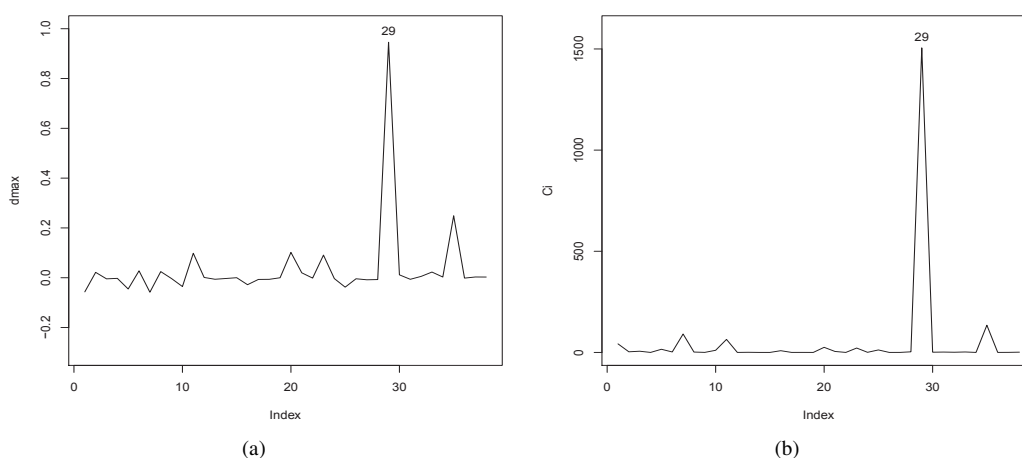


Figure 9: (a) Index plot of  $d_{\max}$  for  $\Psi$  (simultaneous response perturbation) and (b) total local influence for  $\Psi$  (simultaneous response perturbation) based on the model fitted to the renal insufficiency data.

Table 7: Relative changes, estimates, and the corresponding  $p$ -values in parentheses

Set	$I_{\text{complete}}$	$I - \{21\}$	$I - \{29\}$	$I - \{21 \text{ and } 29\}$
$\beta_{01}$	[-]	[1.692]	[6.9591]	[0.0000]
	2.5204	2.4777	2.3450	2.5204
	(0.1188)	(0.0846)	(0.1649)	(0.0213)
$\beta_{11}$	[-]	[-67.7746]	[25.0070]	[0.0000]
	0.7075	1.1871	0.5306	0.7075
	(0.1555)	(0.0291)	(0.2141)	(0.1854)
$\beta_{02}$	[-]	[-10.9615]	[29.7890]	[0.0000]
	1.7274	1.9168	1.2128	1.7274
	(0.1657)	(0.1098)	(0.4268)	(0.0421)
$\beta_{12}$	[-]	[-1.9063]	[-0.8373]	[0.0000]
	1.9164	1.9529	1.9324	1.9164
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\sigma_1$	[-]	[9.1365]	[-4.1007]	[0.0000]
	1.6826	1.5289	1.7516	1.6826
$\sigma_2$	[-]	[0.3287]	[-14.7761]	[0.0000]
	1.4891	1.4842	1.7091	1.4891
$a$	[-]	[6.1817]	[-41.5784]	[0.0000]
	2.9146	2.7344	4.1264	2.9146
$b$	[-]	[-31.6078]	[5.2288]	[0.0000]
	0.5782	0.7609	0.5479	0.5782
$\alpha$	[-]	[0.0000]	[0.0000]	[0.0000]
	0.9999	1.0000	1.0000	1.0000

a male with a longer time to first occurrence of the infection and a median time to occurrence of infection 2. The observation #29 is related to a male with less time until the occurrence of infection 2. In order to reveal the impact of the two observations on the parameter estimates, we refit the model under some situations. First, we individually eliminate each one of these two observations. Next, we remove from the original dataset the totality of potentially influential observations.

Table 7 gives the relative changes (in percentages) of the estimates defined by  $RC_{\Psi_v} = [(\hat{\Psi}_v -$

Table 8: Case influence diagnostics for renal insufficiency data

Parameter	Complete data set				Dropped observations 15 and 21			
	Mean	SD	CI (95%)		Mean	SD	CI (95%)	
$\alpha$	0.9989	0.0142	0.9999	1.0000	0.9992	0.0106	0.9998	1.0000
$\sigma_1$	0.9572	0.1867	0.6394	1.3806	1.1916	0.2685	0.7447	1.7832
$\beta_{01}$	4.1414	0.4844	3.1348	5.0698	3.8239	0.4435	2.9181	4.6951
$\beta_{11}$	0.7127	0.4501	-0.1110	1.6324	0.9632	0.4367	0.1139	1.8319
$\sigma_2$	1.0956	0.1828	0.7657	1.4866	1.1331	0.2079	0.7687	1.5826
$\beta_{02}$	3.1853	0.3614	2.4386	3.8682	3.0973	0.3728	2.3768	3.8408
$\beta_{12}$	1.8701	0.3513	1.1885	2.5791	1.9809	0.3921	1.1779	2.7060
$a$	0.8438	0.2722	0.4398	1.4704	0.8618	0.2862	0.4328	1.5458
$b$	0.6876	0.1340	0.4161	0.9312	0.8132	0.1433	0.5117	1.0495

SD = standard deviation; CI = credible intervals.

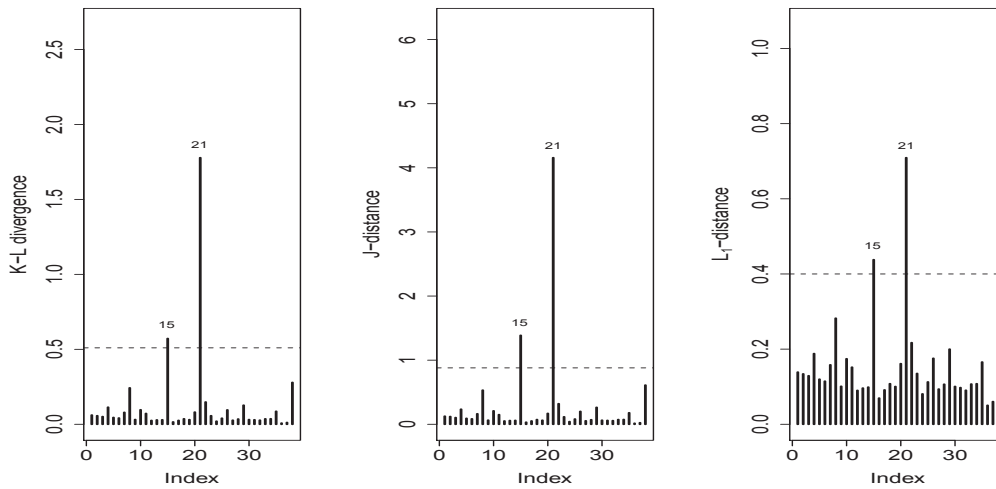


Figure 10: Index plots of  $\psi$ -divergence measures for the renal insufficiency data.

$\hat{\Psi}_v(I)/\hat{\Psi}_v \times 100$ , and the corresponding  $p$ -values, where  $\hat{\Psi}_v(I)$  is the MLE of  $\Psi_v$  after the “set  $I$ ” of observations is removed.

The figures in Table 7 reveal that there are substantial changes in the estimated parameter values as well as changes in the set of parameters that show the model’s significance. Despite this sensitivity of the model, inclusion and exclusion of these points do not imply changes in the interpretation of the results, since there is no change in the sign of the coefficient referring to the sex variable, which indicates that men tend to suffer infections sooner than women.

### 6.2. Results: Bayesian analysis

Table 8 provides the means, standard deviations (SDs) and 95% Bayesian credible intervals (CI (95%)) for the estimates of the parameters in the BKwW regression model fitted to the original dataset and when two observations (#15 and #21) are dropped.

The sample considered for the posterior distributions of the parameters of the BKwW regression model and the  $\psi$ -divergence measures in (4.3) are calculated. Figure 10 displays the index plot of the three  $\psi$ -divergence measures. The cases #15 and #21 are possible influential observations in the posterior distribution. The three influence diagnostic measures for these two observations are given in

Table 9: Case influence diagnostics for the renal insufficiency data

Case number	$\psi$ -divergence measure		
	$d_{K-L}$	$d_J$	$d_{L_1}$
15	0.5722	1.3869	0.4380
21	1.7783	4.1576	0.7093

Table 10: Bayesian criteria

Dropped observation	Criterion			
	EAIC	EBIC	DIC	LPML
None	683.4028	698.1411	673.4582	-337.3232
15	664.0772	678.5755	654.2425	-327.3600
21	647.4933	661.9915	637.0273	-318.9128
15 and 21	628.4176	642.6693	618.2174	-309.4217

EAIC = expected Akaike information criterion; EBIC = expected Bayesian information criterion; DIC = deviance information criterion; LPML = log pseudo marginal likelihood.

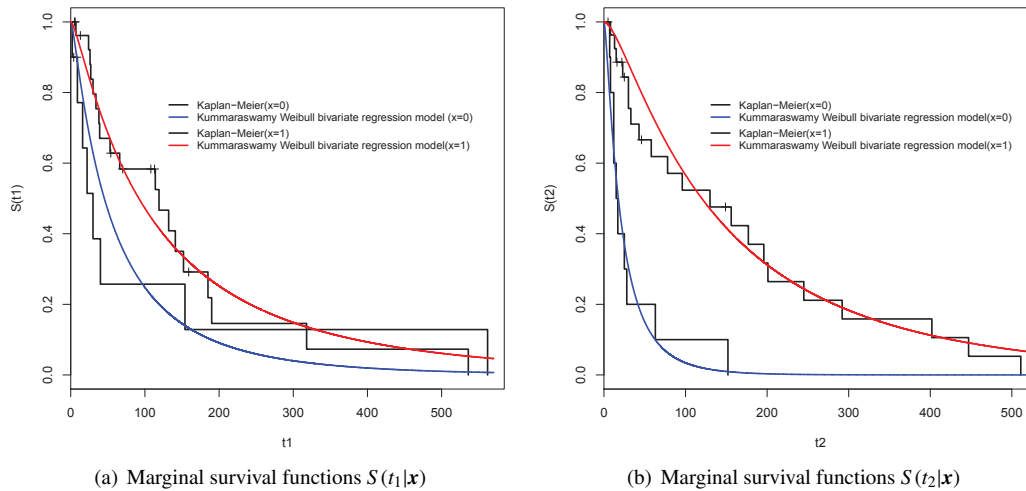


Figure 11: Kaplan-Meier curves stratified by gender (0 = masculine, 1 = feminine) and estimated survival functions for the renal insufficiency data.

Table 9.

Also, we obtain the DIC, EAIC, EBIC, and LPML values to compare the impact of the influential points. Table 10 presents the values for the complete data, without observation #15, #21 and both of two cases. By dropping the observation #21, we obtain a better fit compared to the case without the observation #15.

### 6.3. Goodness-of-fit

We check the quality of the fitted regression model by means of plots of the Kaplan-Meier survival function and the estimated marginal survival functions for the BKwW regression model. Figure 11 shows that the proposed model is well adjusted, because its fitted survival function follows the Kaplan-Meier curve closely.

## 7. Conclusions and remarks

We propose the BKwW regression model as an extension of the bivariate Weibull distribution. We estimate the model parameters using the maximum likelihood methodology subject to linear restrictions in the parameters. We also perform a sensitivity analysis to assess the robustness of the results. We carry out the estimation method using the R software. Several functions are implemented as well as new functions with the partial derivatives of the log-likelihood function.

The optimization process is sensitive to the choice of the initial values used in the algorithms. To choose the initial regression parameters, we obtain the estimates considering marginal models for each response variable, while we choose the values  $a = 1$  and  $b = 1$  for the other parameters that therefore represent the special bivariate Weibull distribution. By using the global influence, local influence and total local influence techniques, we identify some possible influential points. We examine these points to detect possible errors in managing the dataset, but discard this possibility. Further, we discuss the use of MCMC methods as an alternative way to perform Bayesian inference for lifetime data that are supposed to follow the BKwW regression model. We also adopt Bayesian case influence diagnostics based on the K-L divergence to study the sensitivity of the Bayesian estimates under perturbations in the model/data. We verify the robustness of the results after re-estimating the model parameters after individual and joint exclusion of the identified points. Despite small changes in the estimates, the results and their interpretations are not influenced by the possible influential points indicated by the diagnostic techniques. The functions and techniques presented allow the utility of the BKwW regression model to analyze survival data with two response variables and covariates, and to perform sensitivity analysis to confirm the adequacy of the model assumptions and validate the obtained results.

## Appendix:

Expressions for  $A(\cdot)$ ,  $B(\cdot)$ ,  $C(\cdot)$ , and  $G(\cdot)$  described in Section 2 are given below:

$$\begin{aligned}
 A(t_1, t_2|\mathbf{x}) &= -\frac{a(b-1)G^a(t_1, t_2|\mathbf{x})D(t_1, t_2|\mathbf{x})E(t_1, t_2|\mathbf{x})}{1-G^a(t_1, t_2|\mathbf{x})}, \\
 B(t_1, t_2|\mathbf{x}) &= (a-1)D(t_1, t_2|\mathbf{x})E(t_1, t_2|\mathbf{x}), \\
 C(t_1, t_2|\mathbf{x}) &= G(t_1, t_2|\mathbf{x})g(t_1, t_2|\mathbf{x}), \\
 G(t_1, t_2|\mathbf{x}) &= \exp\left\{-\left[\left(\exp\{-\mathbf{x}^T\boldsymbol{\beta}_1\}t_1\right)^{\frac{1}{\sigma_1^a}} + \left(\exp\{-\mathbf{x}^T\boldsymbol{\beta}_2\}t_2\right)^{\frac{1}{\sigma_2^a}}\right]^\alpha\right\} - \exp\left[-\left(\exp\{-\mathbf{x}^T\boldsymbol{\beta}_1\}t_1\right)^{\frac{1}{\sigma_1}}\right] \\
 &\quad - \exp\left[-\left(\exp\{-\mathbf{x}^T\boldsymbol{\beta}_2\}t_2\right)^{\frac{1}{\sigma_2}}\right] + 1, \\
 D(t_1, t_2|\mathbf{x}) &= -\left[\left(\exp\{-\mathbf{x}^T\boldsymbol{\beta}_1\}t_1\right)^{\frac{1}{\sigma_1^a}} + \left(\exp\{-\mathbf{x}^T\boldsymbol{\beta}_2\}t_2\right)^{\frac{1}{\sigma_2^a}}\right]^{\alpha-1} t_1^{\frac{1}{\sigma_1^a}-1} \exp\{-\mathbf{x}^T\boldsymbol{\beta}_1\}^{\frac{1}{\sigma_1^a}} \\
 &\quad \frac{1}{\sigma_1} \exp\left\{-\left[\left(\exp\{-\mathbf{x}^T\boldsymbol{\beta}_1\}t_1\right)^{\frac{1}{\sigma_1^a}} + \left(\exp\{-\mathbf{x}^T\boldsymbol{\beta}_2\}t_2\right)^{\frac{1}{\sigma_2^a}}\right]^\alpha\right\} + \exp\{-\mathbf{x}^T\boldsymbol{\beta}_1\}^{\frac{1}{\sigma_1}} t_1^{\frac{1}{\sigma_1}-1} \\
 &\quad \frac{1}{\sigma_1} \exp\left[-\left(\exp\{-\mathbf{x}^T\boldsymbol{\beta}_1\}t_1\right)^{\frac{1}{\sigma_1}}\right], \\
 E(t_1, t_2|\mathbf{x}) &= -\left[\left(\exp\{-\mathbf{x}^T\boldsymbol{\beta}_1\}t_1\right)^{\frac{1}{\sigma_1^a}} + \left(\exp\{-\mathbf{x}^T\boldsymbol{\beta}_2\}t_2\right)^{\frac{1}{\sigma_2^a}}\right]^{\alpha-1} t_2^{\frac{1}{\sigma_2^a}-1} \exp\{-\mathbf{x}^T\boldsymbol{\beta}_2\}^{\frac{1}{\sigma_2^a}}
 \end{aligned}$$

$$\frac{1}{\sigma_2} \exp \left\{ - \left[ \left( \exp \{ -\mathbf{x}^T \boldsymbol{\beta}_1 \} t_1 \right)^{\frac{1}{\sigma_1^\alpha}} + \left( \exp \{ -\mathbf{x}^T \boldsymbol{\beta}_2 \} t_2 \right)^{\frac{1}{\sigma_2^\alpha}} \right]^\alpha \right\} + \exp \{ -\mathbf{x}^T \boldsymbol{\beta}_2 \}^{\frac{1}{\sigma_2}} t_2^{\frac{1}{\sigma_2}-1} \frac{1}{\sigma_2} \exp \left[ - \left( \exp \{ -\mathbf{x}^T \boldsymbol{\beta}_2 \} t_2 \right)^{\frac{1}{\sigma_2}} \right],$$

and

$$g(t_1, t_2 | \mathbf{x}) = \frac{1}{\sigma_1} \exp \{ -\mathbf{x}^T \boldsymbol{\beta}_1 \} \left( \exp \{ -\mathbf{x}^T \boldsymbol{\beta}_1 \} t_1 \right)^{\frac{1}{\sigma_1^\alpha}-1} \frac{1}{\sigma_2} \exp \{ -\mathbf{x}^T \boldsymbol{\beta}_2 \} \left( \exp \{ -\mathbf{x}^T \boldsymbol{\beta}_2 \} t_2 \right)^{\frac{1}{\sigma_2^\alpha}-1} \left[ \left( \exp \{ -\mathbf{x}^T \boldsymbol{\beta}_1 \} t_1 \right)^{\frac{1}{\sigma_1^\alpha}} + \left( \exp \{ -\mathbf{x}^T \boldsymbol{\beta}_2 \} t_2 \right)^{\frac{1}{\sigma_2^\alpha}} \right]^{\alpha-2} \left\{ \left[ \left( \exp \{ -\mathbf{x}^T \boldsymbol{\beta}_1 \} t_1 \right)^{\frac{1}{\sigma_1^\alpha}} + \left( \exp \{ -\mathbf{x}^T \boldsymbol{\beta}_2 \} t_2 \right)^{\frac{1}{\sigma_2^\alpha}} \right]^\alpha + \frac{1}{\alpha} - 1 \right\} \exp \left\{ - \left[ \left( \exp \{ -\mathbf{x}^T \boldsymbol{\beta}_1 \} t_1 \right)^{\frac{1}{\sigma_1^\alpha}} + \left( \exp \{ -\mathbf{x}^T \boldsymbol{\beta}_2 \} t_2 \right)^{\frac{1}{\sigma_2^\alpha}} \right]^\alpha \right\}.$$

### References

- Barriga GDC, Louzada-Neto F, Ortega EMM, and Cancho VG (2010). A bivariate regression model for matched paired survival data: local influence and residual analysis, *Statistics Methods and Applications*, **19**, 477–496.
- Brooks SP (2002). Discussion on the paper by Spiegelhalter, Best, Carlin, and van der Linde, *Journal of the Royal Statistical Society Series B*, **64**, 616–618.
- Carlin BP and Louis TA (2001). *Bayes and Empirical Bayes Methods for Data Analysis* (2nd ed), Chapman & Hall/CRC, Boca Raton.
- Cook RD (1977). Detection of influential observations in linear regression, *Technometrics*, **19**, 15–18.
- Cook RD (1986). Assessment of local influence (with discussion), *Journal of the Royal Statistical Society B*, **48**, 133–169.
- Cordeiro GM and de Castro M (2011). A new family of generalized distributions, *Journal of Statistical Computation and Simulation*, **81**, 883–898.
- Cordeiro GM, Ortega EMM, and Nadarajah S (2010). The Kumaraswamy Weibull distribution with application to failure data, *Journal of the Franklin Institute*, **347**, 1399–1429.
- da Cruz JN, Ortega EMM, and Cordeiro GM (2016). The log-odd log-logistic Weibull regression model: modelling, estimation, influence diagnostics and residual analysis, *Journal of Statistical Computation and Simulation*, **86**, 1516–1538.
- Dey D and Birmiwal L (1994). Robust Bayesian analysis using divergence measures, *Statistics and Probability Letters*, **20**, 287–294.
- Geweke J (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: Bernardo JM, Berger JO, Dawid AP, Smith AFM (eds), *Bayesian Statistics* (4th ed), Oxford University Press, 169–188.
- Hashimoto EM, Ortega EMM, Cancho VG, and Cordeiro GM (2013). On estimation and diagnostics analysis in log-generalized gamma regression model for interval-censored data, *Statistics*, **47**, 379–398.
- He W and Lawless JF (2005). Bivariate location-scale models for regression analysis, with applications to lifetime data, *Journal of the Royal Statistical Society*, **67**, 63–78.
- Hougaard, P. (1986). A class of multivariate failure time distributions, *Biometrika*, **73**, 671–678.
- Ibrahim JG, Chen MH, and Sinha D (2001). *Bayesian Survival Analysis*, Springer, New York.
- Lange K (1999). *Numerical Analysis for Statisticians*, Springer, New York.
- Lawless JF (2003). *Statistical Models and Methods for Lifetime Data*, Wiley, New York.

- McGilchrist CA and Aisbett CW (1991). Regression with frailty in survival analysis, *Biometrics*, **47**, 461–466.
- Mudholkar GS, Srivastava DK, and Friemer M (1995). The exponentiated Weibull family: a reanalysis of the bus-motor-failure data, *Technometrics*, **37**, 436–445.
- Ortega EMM, Cordeiro GM, and Kattan MW (2013). The log-beta Weibull regression model with application to predict recurrence of prostate cancer, *Statistical Papers*, **54**, 113–132.
- Ortega EMM, Cordeiro GM, Campelo AK, Kattan MW, and Cancho VG (2015). A power series beta Weibull regression model for predicting breast carcinoma, *Statistics in Medicine*, **34**, 1366–1388.
- Peng F and Dey D (1995). Bayesian analysis of outlier problems using divergence measures, *Canadian Journal of Statistics*, **23**, 199–213.
- R Core Team (2016). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org>
- Silva GO, Ortega EMM, and Cordeiro GM (2010). The beta modified Weibull distribution, *Lifetime Data Analysis*, **16**, 409–430.
- Spiegelhalter DJ, Best NG, Carlin BP, and van der Linde A (2002). Bayesian measures of model complexity and fit, *Journal of the Royal Statistical Society Series B*, **64**, 583–639.

*Received March 20, 2018; Revised June 23, 2018; Accepted August 14, 2018*