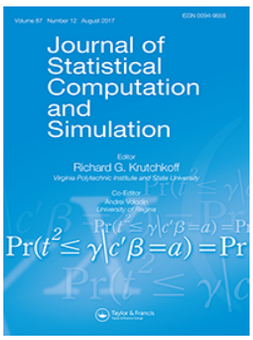

Taylor & Francis
Taylor & Francis Group

Journal of Statistical Computation and Simulation

ISSN: 0094-9655 (Print) 1563-5163 (Online) Journal homepage: http://www.tandfonline.com/loi/gscs20

Bayesian inference for Birnbaum–Saunders distribution and its generalization

Naijun Sha & Tun Lee Ng

To cite this article: Naijun Sha & Tun Lee Ng (2017) Bayesian inference for Birnbaum–Saunders distribution and its generalization, Journal of Statistical Computation and Simulation, 87:12, 2411-2429, DOI: 10.1080/00949655.2017.1334145

To link to this article: http://dx.doi.org/10.1080/00949655.2017.1334145

Published online: 31 May 2017.

Submit your article to this journal

Article views: 15

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=gscs20

Download by: [Naijun Sha] **Date:** 07 June 2017, At: 16:02

JOURNAL OF STATISTICAL COMPUTATION AND SIMULATION, 2017
VOL. 87, NO. 12, 2411–2429
https://doi.org/10.1080/00949655.2017.1334145

Taylor & Francis
Taylor & Francis Group

Check for updates

Bayesian inference for Birnbaum–Saunders distribution and its generalization

Naijun Sha[a] and Tun Lee Ng[b]

[a]Department of Mathematical Sciences, University of Texas at El Paso, El Paso, TX, USA; [b]Department of Statistics, University of Wisconsin, Madison, WI, USA

ABSTRACT

We present a Bayesian approach for parameter inference of the Birnbaum–Saunders distribution [Birnbaum ZW, Saunders SC. A new family of life distributions. J Appl Probab. 1969;6:319–327], as well as the generalized Birnbaum–Saunders distribution developed by Owen [A new three-parameter extension to the Birnbaum–Saunders distribution. IEEE Trans Reliab. 2006;55:475–479], in the presence of random right-censored data. To handle the instance of commonly occurred censored observations, we utilize the data augmentation technique [Tanner MA, Wong WH. The calculation of posterior distributions by data augmentation. J Amer Statist Assoc. 1987;82(398):528–540] to circumvent the arduous expressions involving the censored data in posterior inferences. Simulation studies are carried out to assess performance of these methods under different parameter values, with small and large sample sizes, as well as various degrees of censoring. Two real data are analysed for illustrative purpose.

ARTICLE HISTORY
Received 31 August 2016
Accepted 20 May 2017

KEYWORDS
Birnbaum–Saunders distribution; generalization; data augmentation; Bayesian method; MCMC sampling; estimation

AMS SUBJECT CLASSIFICATION
62F15

1. Introduction

Birnbaum and Saunders [1] developed a family of two-parameter distributions (henceforth abbreviated as BS) to model failure time of a specimen under a cyclic loading of 'stress'. The ultimate failure is due to the growth of a dominant crack in the material, where, at each increment of load, the dominant crack extends by a random amount. The BS random failure time T with parameters α, β, denoted as $T \sim \text{BS}(\alpha, \beta)$, has the distribution function given by

$$F_T(t) = \Phi\left[\frac{1}{\alpha}\left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right)\right], \tag{1}$$

where $\alpha > 0, \beta > 0$ are the shape and scale parameters, and $\Phi(\cdot)$ denotes the distribution function of the standard normal distribution. Clearly, β is also the median of the BS distribution since $F_T(\beta) = \Phi(0) = 0.5$. Additionally, the BS distribution exhibits the well-known reciprocal property, that is, $T^{-1} \sim \text{BS}(\alpha, \beta^{-1})$, and so it is in the same distribution family [2]. It is easily seen that the transformation $Z = (\sqrt{T/\beta} - \sqrt{\beta/T})/\alpha$ leads to a standard normal variate. This relation is useful for random number generation, integer moments derivation as well as the development of the sampling procedure presented in this article.

CONTACT Naijun Sha nsha@utep.edu Department of Mathematical Sciences, University of Texas at El Paso, El Paso, TX 79968, USA

© 2017 Informa UK Limited, trading as Taylor & Francis Group

The BS distribution, with its generalizations, has seen vast areas of practical applications including engineering, business, environmental, medical and many others [3]. Many authors have contributed extensively to the development and parameter inference of the BS distribution. For instance, Desmond [4] provided a more general and robust proof for the BS distribution in a biological random environment context. Subsequently, Desmond [5] proved that the BS distribution is in fact an equal mixture of an inverse normal distribution and its reciprocal. Later, Rieck [6] derived the moment generating function of the sinh-normal distribution in order to obtain both integer and non-integer moments of the BS distribution. In the context of complete data, the maximum likelihood estimators (MLEs) were originally obtained by Birnbaum and Saunders [7] with their asymptotic distributions derived by Engelhardt et al. [8]. Dupuis and Mills [9] utilized the optimal bias-robust estimator to estimate the parameters and quantiles of the BS distribution. Later, Ng et al. [10] proposed modified moment estimators (MMEs) with two methods of bias correction. Subsequently, Wu and Wong [11] improved interval estimation of the parameters using the signed log-likelihood ratio method introduced by Barndorff-Nielsen [12]. Recently, Balakrishnan and Zhu [13] obtained another form of estimator with smaller bias compared to the MLE and MME proposed by Ng et al. [10]. Lemonte et al. [14] derived modified MLEs which are bias-free to second order, and obtained the analytical expression of the Fisher information matrix only involving the standard normal distribution function, which is an improvement over the expression given by Engelhardt et al. [8]. Other important efforts dealing with censored samples can be found in [15–18]. Bayesian parameter estimation based on Jeffrey's and reference priors [19] were developed in [20–22], where the Laplace or Lindley's [23] methods were applied in the integral approximation. Recently, Wang et al. [24] proposed a sampling algorithm using the generalized ratio of uniforms method for Bayesian inference. In the context of accelerated life testing, Rieck and Nedelman [25] constructed a log-linear model, and Owen and Padgett [26] developed the BS inverse power accelerated life model for BS distribution.

Different generalizations and applications of the BS distribution have been introduced by various authors over the years. Some generalizations are constructed by introducing different kernels or additional parameters, or by using extreme value and non-centrality arguments, see, for example, [3,27–30], and many others. One generalization was introduced by Owen [31] through relaxing the independence assumption of the crack extensions in the material. From the practical point of view, the current crack extension depends on and is built up from previous crack extensions in the specimen. It is probable that the specimen wears out even faster once there are existing internal cracks that constitute the 'weak point' in the body of the specimen. This leads to the phenomenon in which the crack extension tends to be smaller at the beginning of the experiment and gradually becomes larger as the specimen is getting close to failure or breakdown. It validates the idea of modelling the sequence of crack extensions as a long memory process [32]. Upon this observation, Owen [31] derived to obtain a generalized BS distribution, referred as $T \sim \mathrm{GBS}(\kappa, \alpha, \beta)$, whose distribution function is given by

$$F_T(t) = \Phi\left[\frac{1}{\alpha}\left(\frac{t^{1-\kappa}}{\sqrt{\beta}} - \frac{\sqrt{\beta}}{t^{\kappa}}\right)\right], \tag{2}$$

where the additional parameter $0 < \kappa < 1$ is the rate of variation decay of total crack extensions. When $\kappa = 0.5$, it corresponds to the independence assumption of crack extensions, resulting in the original BS distribution. On the other hand, $\kappa > 0.5$ corresponds to a long memory process with slow rate of decay, whereas $\kappa < 0.5$ to a short memory process with the fast rate of decay. Figure 1 adequately captures the fact that, as experiment time increases, a short memory process ($\kappa < 0.5$) always has a higher hazard rate, followed by an independent process ($\kappa = 0.5$) and finally a long memory process ($\kappa > 0.5$), under different values of α and β.

From the distribution function given by Equation (2), it is clear that β is still the median but no longer a scale parameter of the generalized Birnbaum–Saunders (GBS) distribution. The reciprocal property becomes $T^{-1} \sim \mathrm{GBS}(1-\kappa, \alpha, \beta^{-1})$, that it, if T is a long memory process, then T^{-1} is the short memory analog. Additionally, the normal transformation is $Z = (T^{1-\kappa}/\sqrt{\beta} - \sqrt{\beta}/T^{\kappa})/\alpha$.

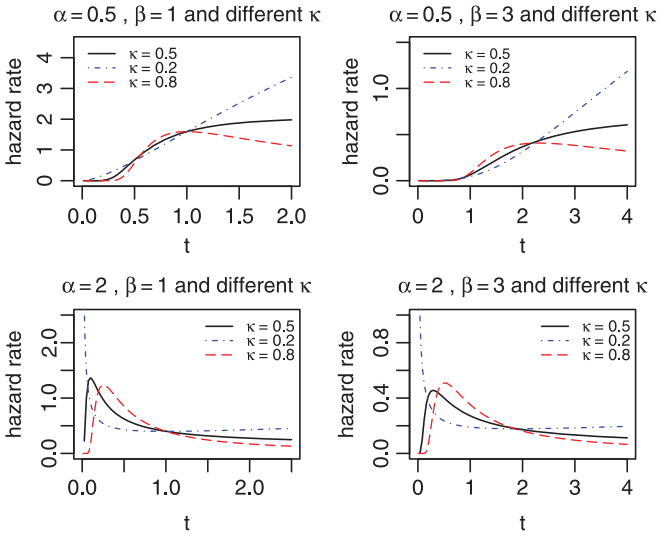

Figure 1. Hazard function of GBS distribution.

Due to the complex distribution function, there is no tractable expression to be obtained for the moments, and a numerical method was used for MLE estimation by Owen [31] under the case of complete data, and thus far, no other research was done to further investigate parameter inference for the GBS distribution. In this paper, our primary research focuses on Bayesian methods for parameter inference by data augmentation method [33]. We also wish to contribute to the relatively neoteric body of research on Bayesian method for the BS and GBS distributions.

The rest of the article is arranged as follows. A Bayesian methodology is presented for estimation procedure in Section 2. Subsequently, we carry out simulation studies to investigate the performance of proposed methods in Section 3. For illustrative purpose, two real data sets are analysed in Section 4, followed by some discussions and concluding remarks in Section 5.

2. Bayesian analysis

Suppose we have n observed samples $t_1, t_2, \ldots, t_n$ and m right-censored values $c_1, \ldots, c_m$ from an experiment following BS or GBS distribution. We denote the m unobserved failure times as t_{n+j} satisfying $t_{n+j} > c_j$, $j = 1, 2, \ldots, m$. Note that the subscripts here do not refer to order statistics, but only serve the purpose of simplifying notation. Throughout our discussion, we assume the censoring mechanism is independent of the event process [34]. It is easily seen that our inferences can be readily applied to the Type-I and Type-II right-censored data under the case $c_1 = c_2 = \cdots = c_m$. The observed failure data $(t_1, \ldots, t_n)$ together with unobserved latent values $(t_{n+1}, \ldots, t_{n+m})$ constitute the augmented data, denoted by $\boldsymbol{t}$. In our Bayesian approach, the latent variables will be sampled together with the parameters in the Markov chain Monte Carlo (MCMC) procedures.

2.1. The BS likelihood

The density function of BS distribution in Equation (1) is

$$f_T(t) = \frac{t+\beta}{2\sqrt{2\pi}t^{3/2}\alpha\sqrt{\beta}} \exp\left\{-\frac{1}{2\alpha^2}\left(\frac{t}{\beta}+\frac{\beta}{t}-2\right)\right\}, \tag{3}$$

and then the augmented likelihood is given by

$$L(\alpha, \beta|\boldsymbol{t}) \propto (\alpha^2\beta)^{-(n+m)/2} \exp\left\{-\frac{\varphi_3(\boldsymbol{t})}{2\alpha^2}\right\} \prod_{i=1}^{n}(t_i+\beta) \prod_{j=1}^{m}(t_{n+j}+\beta), \tag{4}$$

where $\varphi_1(\boldsymbol{t}) = \sum_{i=1}^{n} t_i + \sum_{j=1}^{m} t_{n+j}, \varphi_2(\boldsymbol{t}) = \sum_{i=1}^{n} t_i^{-1} + \sum_{j=1}^{m} t_{n+j}^{-1}, \varphi_3(\boldsymbol{t}) = \varphi_1(\boldsymbol{t})/\beta + \beta\varphi_2(\boldsymbol{t}) - 2(n+m)$. We refrain from lumping together the notations into $t_1, \ldots, t_{n+m}$ in order to distinguish between the observed samples $t_1, \ldots, t_n$ and the latent variables $t_{n+1}, \ldots, t_{n+m}$. This allows us to properly specify the variance–covariance of our proposal distributions in the MCMC procedure. By sampling these latent variables, we draw our inference procedures as if we manage to 'observe' all failures. In other words, we overcome the limitations of the censored likelihood by working with the augmented likelihood. This method was first popularized by Tanner and Wong [33].

2.1.1. Prior specification

From Equation (4), it is easily discernible that an inverse gamma is a conjugate prior for α based on the conditional likelihood $L(\alpha \mid \beta, \boldsymbol{t})$. Also from the original assumption made in the derivation of the BS distribution [1], we notice the proportional relation $\alpha^2 \propto \beta$, and so we construct a joint prior $\pi(\alpha, \beta) = \pi(\alpha \mid \beta) \times \pi(\beta)$, such that the conditional prior mean $E(\alpha^2 \mid \beta) \propto \beta$. Therefore, we specify that

$$\alpha^2 \mid \beta \sim \mathcal{IG}\left(\frac{a_0}{2}, \frac{a_0\beta}{2a_1}\right), \tag{5}$$

with $a_1 > 0$ and $a_0 > 4$ to ensure existence of the variance. It is also clear that there is no conjugate prior for β. However, we may consider a prior distribution $\pi(\beta)$ which has a similar functional form as its conditional likelihood function. In this case, we pick the prior of β to be an inverse gamma distribution

$$\beta \sim \mathcal{IG}\left(\frac{b_0}{2}, \frac{b_0}{2b_1}\right), \tag{6}$$

with $b_1 > 0$ and $b_0 > 4$ to ensure existence of the variance. The hyperparameter values can be specified based on the following factors: (i) since β is the median of the BS distribution, we can refer to the sample median in our attempt to specify b_0 and b_1; (ii) the MLEs $\hat{\alpha}_{\text{MLE}}, \hat{\beta}_{\text{MLE}}$ can be used to determine a_0 and a_1.

2.1.2. Posterior inference

From Equations (4) to (6), the joint posterior distribution of the parameters is given by

$$\pi(\alpha, \beta \mid \boldsymbol{t}) \propto L(\alpha, \beta)\pi(\alpha \mid \beta)\pi(\beta) \propto (\alpha^2)^{-(\nu_0+1)} \beta^{-(\tau_0+1)} \, \mathrm{e}^{-\varphi_4(\boldsymbol{t})} \prod_{i=1}^{n}(t_i+\beta) \prod_{j=1}^{m}(t_{n+j}+\beta), \tag{7}$$

where $\nu_0 = (a_0 + n + m)/2, \tau_0 = (b_0 - a_0 + n + m)/2$, and $\varphi_4(\boldsymbol{t}) = [\varphi_3(\boldsymbol{t}) + a_0\beta/a_1]/(2\alpha^2) + b_0/(2b_1\beta)$. It follows that the full conditional posteriors are

$$\alpha^2 \mid \beta, \sim \mathcal{IG}\left(\nu_0, \frac{1}{2}\left[\varphi_3(\boldsymbol{t}) + \frac{a_0\beta}{a_1}\right]\right), \tag{8}$$

$$\pi(\beta \mid \alpha^2, \boldsymbol{t}) \propto \beta^{-(\tau_0+1)} \, \mathrm{e}^{\varphi_4(\boldsymbol{t})} \prod_{i=1}^{n}(t_i+\beta) \prod_{j=1}^{m}(t_{n+j}+\beta). \tag{9}$$

We draw posterior samples by adopting an MCMC algorithm including Gibbs sampling (see, e.g. [35]) and Metropolis–Hastings (MH) procedure (see, e.g. [36]). We will implement two different sampling procedures for comparisons: the first is to sample the parameters individually (henceforth called conditional sampling), and the second to sample the parameters jointly (called joint sampling).

Conditional sampling

Set initial values (such as MLEs) for the parameters α and β, and repeat the following steps M times among which given the values α_k, β_k and $\boldsymbol{t}$ at the kth iteration, the $(k+1)$th iteration is as follows:

(i) Sample the latent variables $t_{n+1}, \ldots, t_{n+m}$ to obtain the updated augmented data $\boldsymbol{t}$. The increasing transformation to standard normal variate $z = (\sqrt{t/\beta} - \sqrt{\beta/t})/\alpha$ provides us a convenient way to sample the truncated BS random variates $t_{n+j} > c_j$. We first draw truncated standard normal variates z_{n+j} with $z_{n+j} > d_j$, where $d_j = (\sqrt{c_j/\beta_k} - \sqrt{\beta_k/c_j})/\alpha_k$, and then convert the sampled z_{n+j} to t_{n+j} using the following relationship:

$$t_{n+j} = \beta_k \left(1 + \frac{\alpha_k^2 z_{n+j}^2}{2} + \alpha_k z_{n+j} \sqrt{1 + \frac{\alpha_k^2 z_{n+j}^2}{4}} \right), \quad j = 1, 2, \ldots, m. \tag{10}$$

(ii) Draw $\alpha_{k+1}^2 \mid \beta_k, \boldsymbol{t} \sim \mathcal{IG}(\nu_0, [\varphi_3(\boldsymbol{t})_{k+1} + a_0\beta_k/a_1]/2)$, where $\varphi_3(\boldsymbol{t})_{k+1}$ denotes the function value of $\varphi_3(\cdot)$ evaluated at the $(k+1)$th augmented data $\boldsymbol{t}$ from step (i) and the kth iteration value β_k.

(iii) Draw β_{k+1} from $\pi(\beta \mid \alpha_{k+1}^2, \boldsymbol{t})$ via an MH procedure, where $\boldsymbol{t}$ is the augmented data updated in step (i). We propose β_p from a lognormal distribution with centred at previous value, that is, $\log \beta_p \sim N(\log \beta_k, c^2\sigma_k^2)$, where $c > 0$ is a tuning parameter. For the purpose of sampling efficiency, we intend to specify a proposal distribution which closely resembles the conditional posterior of β. This consideration prompts us to specify the variance term σ_k^2 as the reciprocal of Fisher information of the conditional posterior of $\log \beta$, whose log-likelihood is $\log \pi(\beta \mid \alpha_{k+1}^2, \boldsymbol{t}) + \log |J|$. The Jacobian term $J = \beta$ is needed here as we make a log transformation on β. Working with the augmented data which 'removes' the troublesome censored variables, the existing closed form expression of the Fisher information [14] can be easily modified and implemented to obtain the variance in Equation (11), where the first term is obtained from the likelihood for complete data, and the second and third terms are resultant derivatives from the prior component in the conditional posterior distribution (see the derivation in the appendix)

$$\begin{aligned} \sigma_k^2 &= \left(-E \left\{ \frac{\partial^2 (\log \pi(\beta \mid \alpha_{k+1}^2, \boldsymbol{t}) + \log |J|)}{\partial (\log \beta)^2} \right\} \Bigg|_{\alpha=\alpha_{k+1}, \beta=\beta_k} \right)^{-1} \\ &= \left(\frac{(n+m)[\alpha_{k+1}(2\pi)^{-1/2} h(\alpha_{k+1}) + 1]}{\alpha_{k+1}^2} + \frac{b_0}{2b_1\beta_k} + \frac{a_0\beta_k}{2a_1\alpha_{k+1}^2} \right)^{-1}, \end{aligned} \tag{11}$$

where $h(\alpha_{k+1}) = \alpha_{k+1}\sqrt{\pi/2} - \pi\, \mathrm{e}^{2/\alpha_{k+1}^2}[1 - \Phi(2/\alpha_{k+1})]$. For each data set, the tuning parameter c is calibrated such that the acceptance rate of the MH step lies within the preferable range of around 40% [37]. After obtaining $\log \beta_p$, exponentiate it to get β_p, and take

$$\beta_{k+1} = \begin{cases} \beta_p & \text{with probability } \lambda_\beta, \\ \beta_k & \text{with probability } 1 - \lambda_\beta, \end{cases}$$

where

$$\lambda_\beta = \min\left\{1, \frac{\pi(\beta_p \mid \alpha_{k+1}^2, \boldsymbol{t}) \times q_\beta(\beta_k \mid \beta_p)}{\pi(\beta_k \mid \alpha_{k+1}^2, \boldsymbol{t}) \times q_\beta(\beta_p \mid \beta_k)}\right\}, \tag{12}$$

with $q_\beta(\cdot)$ being the lognormal density, and so

$$\frac{q_\beta(\beta_k \mid \beta_p)}{q_\beta(\beta_p \mid \beta_k)} = \frac{(\sigma_p \beta_k \sqrt{2\pi})^{-1} \exp\{-\frac{(\log \beta_k - \log \beta_p)^2}{2\sigma_p^2}\}}{(\sigma_k \beta_p \sqrt{2\pi})^{-1} \exp\{-\frac{(\log \beta_p - \log \beta_k)^2}{2\sigma_k^2}\}}, \tag{13}$$

where σ_p^2 is shown in Equation (11) with β_k replaced by β_p.

Alternatively, we may combine the above steps (ii) and (iii) together to sample the parameters α and β jointly (called joint sampling) as explained below. To make the notation simple, let $\boldsymbol{\theta} = (\alpha, \beta)^{\mathrm{T}}, \log \boldsymbol{\theta} = (\log \alpha, \log \beta)^{\mathrm{T}}$, and so that the log-likelihood of joint posterior density of $\log \boldsymbol{\theta}$ is $\log \pi(\boldsymbol{\theta} \mid \boldsymbol{t}) + \log |J|$, where $J = \alpha\beta$. This joint sampling also requires an MH procedure. First, we propose $\boldsymbol{\theta}_p$ from a bivariate lognormal distribution, that is, $\log \boldsymbol{\theta}_p \sim N_2(\log \boldsymbol{\theta}_k, c^2 \Sigma_k)$, where $\log \boldsymbol{\theta}_k = (\log \alpha_k, \log \beta_k)^{\mathrm{T}}$, and some tuning parameter $c > 0$. Using the closed form in [14], the covariance matrix Σ_k is specified by the Fisher information matrix of the joint posterior likelihood (see the appendix for detailed derivation)

$$\begin{aligned}\Sigma_k &= \left[-E\left\{\frac{\partial^2(\log \pi(\boldsymbol{\theta} \mid \boldsymbol{t}) + \log |J|)}{\partial(\log \boldsymbol{\theta})\partial(\log \boldsymbol{\theta})^{\mathrm{T}}}\right\}_{\boldsymbol{\theta}=\boldsymbol{\theta}_k}\right]^{-1} \\ &= \begin{bmatrix} 2(n+m) + \dfrac{2a_0\beta_k}{a_1\alpha_k^2} & -\dfrac{a_0\beta_k}{a_1\alpha_k^2} \\ -\dfrac{a_0\beta_k}{a_1\alpha_k^2} & g(\alpha_k) + \dfrac{b_0}{2b_1\beta_k} + \dfrac{a_0\beta_k}{2a_1\alpha_k^2} \end{bmatrix}^{-1},\end{aligned} \tag{14}$$

where $g(\alpha_k) = (n+m)[\alpha_k(2\pi)^{-1/2}h(\alpha_k) + 1]/\alpha_k^2$. Note that, even though α and β are asymptotically independent in the absence of censored data, the off-diagonal terms in the covariance are non-zero because they are the resultant derivative terms from the joint prior $\pi(\alpha, \beta)$ that we specified earlier. The tuning parameter $c = 2.4/\sqrt{2}$, as suggested by Gelman et al. [37], is calibrated such that the acceptance rate of the MH step lies within the preferable range of around 30%. Finally, take $\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_p$ with probability

$$\lambda_{\boldsymbol{\theta}} = \min\left\{1, \frac{\pi(\boldsymbol{\theta}_p \mid \boldsymbol{t}) \times q_{\boldsymbol{\theta}}(\boldsymbol{\theta}_k \mid \boldsymbol{\theta}_p)}{\pi(\boldsymbol{\theta}_k \mid \boldsymbol{t}) \times q_{\boldsymbol{\theta}}(\boldsymbol{\theta}_p \mid \boldsymbol{\theta}_k)}\right\}, \tag{15}$$

where $q_{\boldsymbol{\theta}}(\cdot)$ is the bivariate lognormal density function, and so

$$\frac{q_{\boldsymbol{\theta}}(\boldsymbol{\theta}_k \mid \boldsymbol{\theta}_p)}{q_{\boldsymbol{\theta}}(\boldsymbol{\theta}_p \mid \boldsymbol{\theta}_k)} = \frac{(2\pi\alpha_k\beta_k)^{-1}|\Sigma_p|^{-1/2} \exp\{-\frac{1}{2}(\log \boldsymbol{\theta}_k - \log \boldsymbol{\theta}_p)^{\mathrm{T}} \Sigma_p^{-1} (\log \boldsymbol{\theta}_k - \log \boldsymbol{\theta}_p)\}}{(2\pi\alpha_p\beta_p)^{-1}|\Sigma_k|^{-1/2} \exp\{-\frac{1}{2}(\log \boldsymbol{\theta}_p - \log \boldsymbol{\theta}_k)^{\mathrm{T}} \Sigma_k^{-1} (\log \boldsymbol{\theta}_p - \log \boldsymbol{\theta}_k)\}}, \tag{16}$$

where Σ_p is the form in Equation (11) with α_k and β_k replaced by α_p and β_p, respectively.

2.2. The GBS distribution

Since the density function of GBS distribution in Equation (2) is given by

$$f_T(t) = \frac{(1-\kappa)t + \kappa\beta}{\sqrt{2\pi}\alpha\sqrt{\beta}t^{1+\kappa}} \exp\left\{-\frac{(t-\beta)^2}{2\alpha^2\beta t^{2\kappa}}\right\}, \tag{17}$$

the augmented likelihood function, by letting $\boldsymbol{\theta} = (\kappa, \alpha, \beta)^{\mathrm{T}}$, becomes

$$L(\boldsymbol{\theta} \mid \boldsymbol{t}) \propto (\alpha^2\beta)^{-(n+m)/2} \exp\left\{-\frac{\varphi_4(\boldsymbol{t})}{2\alpha^2}\right\} \prod_{i=1}^{n} \varphi_5(t_i) \prod_{j=1}^{m} \varphi_5(t_{n+j}), \tag{18}$$

where $\varphi_1(\boldsymbol{t}) = \sum_{i=1}^{n} t_i^{2-2\kappa} + \sum_{j=1}^{m} t_{n+j}^{2-2\kappa}, \varphi_2(\boldsymbol{t}) = \sum_{i=1}^{n} t_i^{-2\kappa} + \sum_{j=1}^{m} t_{n+j}^{-2\kappa}, \varphi_3(\boldsymbol{t}) = \sum_{i=1}^{n} t_i^{1-2\kappa} + \sum_{j=1}^{m} t_{n+j}^{1-2\kappa}, \varphi_4(\boldsymbol{t}) = \varphi_1(\boldsymbol{t})/\beta + \beta\varphi_2(\boldsymbol{t}) - 2\varphi_3(\boldsymbol{t}), \varphi_5(t) = t^{1+\kappa}[(1-\kappa)t + \kappa\beta]$. From the model development of the GBS distribution by Owen [31], we notice that the rate κ is independent of other two parameters α and β. Hence we propose a joint prior $\pi(\kappa, \alpha, \beta) = \pi(\kappa)\pi(\alpha \mid \beta)\pi(\beta)$, where the same priors of α and β are specified as in Equations (5) and (6), and a Beta prior is adopted for κ

$$\kappa \sim \mathrm{Beta}(d_0, d_1). \tag{19}$$

The uniform distribution with $d_0 = d_1 = 1$ can be chosen if no prior knowledge is available. The joint posterior distribution of the parameters then is given by

$$\begin{aligned}\pi(\boldsymbol{\theta} \mid \boldsymbol{t}) \propto (\alpha^2)^{-(\nu_0+1)}\beta^{-(\tau_0+1)} \exp\left\{-\frac{1}{2\alpha^2}\left[\varphi_4(\boldsymbol{t}) + \frac{a_0\beta}{a_1}\right] - \frac{b_0}{2b_1\beta}\right\} \\ \times \kappa^{d_0-1}(1-\kappa)^{d_1-1} \prod_{i=1}^{n} \varphi_5(t_i) \prod_{j=1}^{m} \varphi_5(t_{n+j}),\end{aligned} \tag{20}$$

where $\nu_0 = (a_0 + n + m)/2, \tau_0 = (b_0 - a_0 + n + m)/2$. It follows that the full conditional posteriors are

$$\pi(\kappa \mid \alpha^2, \beta, \boldsymbol{t}) \propto \kappa^{d_0-1}(1-\kappa)^{d_1-1} \exp\left\{-\frac{\varphi_4(\boldsymbol{t})}{2\alpha^2}\right\} \prod_{i=1}^{n} \varphi_5(t_i) \prod_{j=1}^{m} \varphi_5(t_{n+j}), \tag{21}$$

$$\alpha^2 \mid \beta, \kappa, \boldsymbol{t} \sim \mathcal{IG}\left(\nu_0, \frac{1}{2}\left[\varphi_4(\boldsymbol{t}) + \frac{a_0\beta}{a_1}\right]\right), \tag{22}$$

$$\pi(\beta \mid \alpha^2, \kappa, \boldsymbol{t}) \propto \beta^{-(\tau_0+1)} \exp\left\{-\frac{1}{2\alpha^2}\left[\varphi_4(\boldsymbol{t}) + \frac{a_0\beta}{a_1}\right] - \frac{b_0}{2b_1\beta}\right\} \prod_{i=1}^{n} \varphi_5(t_i) \prod_{j=1}^{m} \varphi_5(t_{n+j}). \tag{23}$$

We implement a Gibbs sampling algorithm to sample the parameters from their full conditional posteriors. First, set initial values for the parameters κ, α and β, and then repeat the following steps M times among which given the values at the kth iteration, the $(k+1)$th iteration is as follows:

(i) Sample the censored latent variables $t_{n+j} > c_j$, $j = 1, \ldots, m$. By the monotone transformation to standard normal random variable $z = (t^{1-\kappa}/\sqrt{\beta} - \sqrt{\beta}/t^{\kappa})/\alpha$, we first draw truncated standard normal random variates $z_{n+j} > d_j$, where $d_j = (c_j^{1-\kappa_k}/\sqrt{\beta_k} - \sqrt{\beta_k}/c_j^{\kappa_k})/\alpha_k$, and then convert the sampled z_{n+j} to t_{n+j} by solving the following non-linear equations:

$$\alpha_k\sqrt{\beta_k}z_{n+j}t_{n+j}^{\kappa_k} - t_{n+j} + \beta_k = 0, \quad j = 1, 2, \ldots, m. \tag{24}$$

(ii) Draw κ from $\pi(\kappa \mid \alpha_k^2, \beta_k, \boldsymbol{t})$ using an MH procedure. We first propose $\kappa_p \sim \text{Beta}(\gamma\kappa_k, \gamma(1-\kappa_k))$, where γ is a tuning parameter, and then take $\kappa_{k+1} = \kappa_p$ with probability

$$\lambda_\kappa = \min\left\{1, \frac{\pi(\kappa_p \mid \alpha_k^2, \beta_k, \boldsymbol{t}) \times q_\kappa(\kappa_k \mid \kappa_p)}{\pi(\kappa_k \mid \alpha_k^2, \beta_k, \boldsymbol{t}) \times q_\kappa(\kappa_p \mid \kappa_k)}\right\}, \tag{25}$$

with the Beta proposal density function $q_\kappa(\cdot)$, and so

$$\frac{q_\kappa(\kappa_k \mid \kappa_p)}{q_\kappa(\kappa_p \mid \kappa_k)} = \frac{\Gamma(\gamma\kappa_k)\Gamma[\gamma(1-\kappa_k)]}{\Gamma(\gamma\kappa_p)\Gamma[\gamma(1-\kappa_p)]} \times \frac{\kappa_k^{\gamma\kappa_p - 1}(1-\kappa_k)^{\gamma(1-\kappa_p)-1}}{\kappa_p^{\gamma\kappa_k - 1}(1-\kappa_p)^{\gamma(1-\kappa_k)-1}}. \tag{26}$$

(iii) Draw $\alpha^2 \sim \mathcal{IG}(\nu_0, \frac{1}{2}[\varphi_4(\boldsymbol{t})_{k+1} + a_0\beta_k/a_1])$.

(iv) Draw β from $\pi(\beta \mid \alpha_{k+1}^2, \kappa_{k+1}, \boldsymbol{t})$ using an MH procedure. First draw β_p from a lognormal proposal density, $\log\beta_p \sim N(\log\beta_k, c^2\hat{\sigma}^2)$, where c is the tuning parameter. Since there is no analytic form under GBS distribution, we rely on the 'observed' Fisher information to specify the variance $\hat{\sigma}^2$. This 'approximate' data-dependent variance may not be positive for the entire parameter space $\{\kappa \in (0,1), \alpha \in (0,\infty), \beta \in (0,\infty)\}$. To overcome this problem, we calculate the observed Fisher information through estimating the parameters by the MLE $\hat{\boldsymbol{\theta}}$ under the updated augmented data $\mathbf{t}$ to obtain the variance term

$$\begin{aligned}\hat{\sigma}^2 &= \left[-\widehat{E}\left\{\frac{\partial^2(\log\pi(\beta|\kappa_{k+1}, \alpha_{k+1}^2, \boldsymbol{t}) + \log|J|)}{\partial(\log\beta)^2}\right\}\right]^{-1}\Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \\ &= \left(-\beta^2\frac{\partial^2\ell(\boldsymbol{\theta} \mid \mathbf{t})}{\partial\beta^2} + \frac{b_0}{2b_1\beta} + \frac{a_0\beta}{2a_1\alpha^2}\right)^{-1}\Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}},\end{aligned} \tag{27}$$

where $\ell(\boldsymbol{\theta} \mid \mathbf{t})$ is the log-likelihood function and the Jacobian term $J = \beta$. Finally, take $\beta_{k+1} = \beta_p$ with probability

$$\lambda_\beta = \min\left\{1, \frac{\pi(\beta_p \mid \kappa_{k+1}, \alpha_{k+1}^2, \boldsymbol{t}) \times q_\beta(\beta_k \mid \beta_p)}{\pi(\beta_k \mid \kappa_{k+1}, \alpha_{k+1}^2, \boldsymbol{t}) \times q_\beta(\beta_p \mid \beta_k)}\right\}, \tag{28}$$

with the proposal density $q_\beta(\cdot)$, and so $q_\beta(\beta_k \mid \beta_p)/q_\beta(\beta_p \mid \beta_k) = \beta_p/\beta_k$.

It is worthwhile to point out that the presented algorithm can be easily extended to other common censoring schemes such as interval and progressive censorings, where the latent variables can be drawn through left or/and right, or interval truncated normal variates. For example, in the case of interval censoring for the BS distribution, assuming that an unobserved failure time t_{n+j} with $a_j < t_{n+j} < b_j$, $1 \le j \le m$, we first draw interval truncated standard normal variate z_{n+j} with $d_j < z_{n+j} < e_j$, where $d_j = (\sqrt{a_j/\beta_k} - \sqrt{\beta_k/a_j})/\alpha_k$, $e_j = (\sqrt{b_j/\beta_k} - \sqrt{\beta_k/b_j})/\alpha_k$, and then convert the sampled z_{n+j} to t_{n+j} using the relationship in Equation (10). Similar procedure can be applied in the sampling latent failure times for the GBS distribution.

Table 1. BS estimation results for parameter setting 1: $\alpha = 0.5, \beta = 1.0$.

	Parameter		α				β			
n	Method	CEP (%)	Estimate	SRMSE	AL	CP (%)	Estimate	AL	SRMSE	CP (%)
20	MLE	10	0.5111	0.0838	0.3127	93.56	1.0466	0.4374	0.1243	93.20
		20	0.5281	0.0893	0.3345	91.77	1.0955	0.4739	0.1556	90.25
		30	0.5452	0.0974	0.3650	85.20	1.1544	0.5215	0.2028	84.18
		40	0.5679	0.1102	0.4069	75.88	1.2293	0.5858	0.2701	73.44
	Cond	10	0.5061	0.0710	0.3181	94.43	1.0353	0.4679	0.1166	94.18
		20	0.5130	0.0779	0.3446	92.09	1.0849	0.5223	0.1461	92.62
		30	0.5283	0.0875	0.3778	86.66	1.1450	0.5952	0.1932	86.58
		40	0.5458	0.1015	0.4205	77.68	1.2220	0.7002	0.2619	75.11
	Joint	10	0.5065	0.0693	0.3064	94.97	1.0605	0.4746	0.1297	94.30
		20	0.5134	0.0747	0.3329	92.07	1.1134	0.5343	0.1677	92.09
		30	0.5284	0.0836	0.3677	86.90	1.1784	0.6178	0.2229	86.06
		40	0.5469	0.0978	0.4148	76.09	1.2634	0.7452	0.3020	74.84
50	MLE	10	0.5048	0.0540	0.2030	94.15	1.0462	0.2844	0.0859	94.10
		20	0.5193	0.0613	0.2181	93.73	1.0996	0.3106	0.1259	93.22
		30	0.5361	0.0732	0.2395	88.01	1.1656	0.3456	0.1853	87.95
		40	0.5572	0.0918	0.2703	83.12	1.2497	0.3937	0.2656	78.00
	Cond	10	0.5011	0.0520	0.2081	95.29	1.0417	0.2959	0.0826	94.14
		20	0.5125	0.0610	0.2274	93.66	1.0953	0.3317	0.1220	93.70
		30	0.5261	0.0746	0.2517	90.26	1.1618	0.3803	0.1814	88.62
		40	0.5382	0.0944	0.2842	85.25	1.2466	0.4520	0.2625	80.55
	Joint	10	0.5015	0.0506	0.2041	95.32	1.0519	0.2967	0.0889	94.72
		20	0.5136	0.0586	0.2233	94.16	1.1070	0.3336	0.1319	92.72
		30	0.5278	0.0715	0.2476	91.13	1.1754	0.3845	0.1944	88.48
		40	0.5385	0.0912	0.2805	86.25	1.2633	0.4605	0.2790	80.14

Table 2. BS estimation results for parameter setting 3: $\alpha = 2.0, \beta = 1.0$.

	Parameter		α				β			
n	Method	CEP (%)	Estimate	SRMSE	AL	CP (%)	Estimate	AL	SRMSE	CP (%)
20	MLE	10	1.9464	0.3410	1.2440	93.17	1.1570	1.2808	0.4041	93.44
		20	2.0619	0.3646	1.3266	92.08	1.2979	1.4457	0.5208	92.43
		30	2.1722	0.4071	1.4480	91.18	1.4859	1.6895	0.6976	88.51
		40	2.3094	0.4771	1.6212	90.18	1.7664	2.0803	0.9846	83.65
	Cond	10	1.9688	0.3487	1.1678	93.66	1.2176	1.3525	0.4313	93.80
		20	2.0418	0.3573	1.2654	92.20	1.3553	1.6059	0.5520	92.82
		30	2.1005	0.3753	1.3919	89.85	1.5331	1.9608	0.7226	89.57
		40	2.1453	0.4058	1.5580	85.29	1.7867	2.5154	0.9827	83.96
	Joint	10	1.9593	0.3559	1.1915	93.68	1.4126	1.6506	0.5978	93.22
		20	2.0420	0.3648	1.3387	92.14	1.6058	2.0774	0.7865	92.45
		30	2.1010	0.3930	1.5689	90.40	1.8808	2.8299	1.0713	88.95
		40	2.1447	0.4695	1.9856	85.33	2.3726	4.9964	2.6583	82.87
50	MLE	10	2.0256	0.2147	0.8095	94.78	1.1331	0.8028	0.2547	94.35
		20	2.0785	0.2457	0.8685	92.12	1.2830	0.9147	0.3786	92.73
		30	2.1569	0.3029	0.9537	91.85	1.4885	1.0854	0.5721	89.78
		40	2.2651	0.3982	1.0803	90.35	1.7875	1.3556	0.8705	83.92
	Cond	10	1.9834	0.2100	0.8073	94.91	1.1613	0.8597	0.2712	94.84
		20	2.0300	0.2277	0.8879	93.58	1.3110	1.0394	0.3999	93.46
		30	2.0884	0.2685	0.9954	92.22	1.5131	1.3074	0.5922	89.74
		40	2.1300	0.3405	1.1535	90.98	1.8023	1.7446	0.8814	84.77
	Joint	10	1.9832	0.2116	0.8044	94.77	1.2419	0.9262	0.3360	94.38
		20	2.0289	0.2287	0.8961	93.48	1.4136	1.1423	0.4945	93.48
		30	2.0949	0.2754	1.0250	92.07	1.6527	1.4854	0.7284	89.09
		40	2.1326	0.3667	1.2351	90.37	2.0102	2.1022	1.0901	84.30

3. Simulation study

3.1. The BS distribution

We conduct a simulation study to assess performance of the presented Bayesian method. Rieck and Nedelman [25] pointed out that in practice, the shape parameter α would usually not exceed one for the BS-distributed fatigue lifetime of a metal specimen subjected to cyclical stress loading. Therefore, in our simulated data, we take the shape parameter $\alpha = 0.5, 1.0$ and 2.0. Without loss of generality, the scale parameter β is fixed at 1. In addition, we apply four or six random right censoring percentages (CEP) at 10%(10%)60% for each simulated data. For each of these parameter settings with two sample sizes $n = 20$ and 50, we generate 10,000 data sets for the Bayesian analysis, where we choose the hyperparameter values $a_0 = a_1 = b_0 = b_1 = 5$, such that the prior distributions are rather 'flat' or 'less informative' to reflect little prior knowledge about the parameters. We find that the rule of thumb $c = 2.4$ as the tuning parameter is adequate in ensuring the acceptance rates to hover around 41–47 %. For both conditional and joint samplings with each simulated data, we run five MCMC chains with fairly different initial values and each with a burn-in period of 2000 followed by 8000 iterations. The scale reduction factor estimate $\sqrt{\hat{R}} = \sqrt{\mathrm{Var}(\psi)/W}$ is used to monitor convergence of MCMC simulations [37], where ψ is the estimand of parameter interest, $\mathrm{Var}(\psi) = (N-1)W/N + B/N$ with the iteration number N for each chain, the between- and within-sequence variances B and W. The scale factors for the sequences of α and β are within 1.00–1.02 for all five MCMC chains, indicating their convergence. The remaining 8000 samples are used to compute the average estimates, square root of mean squared error (SRMSE) of the estimates, average lengths (AL) of the 95% credible

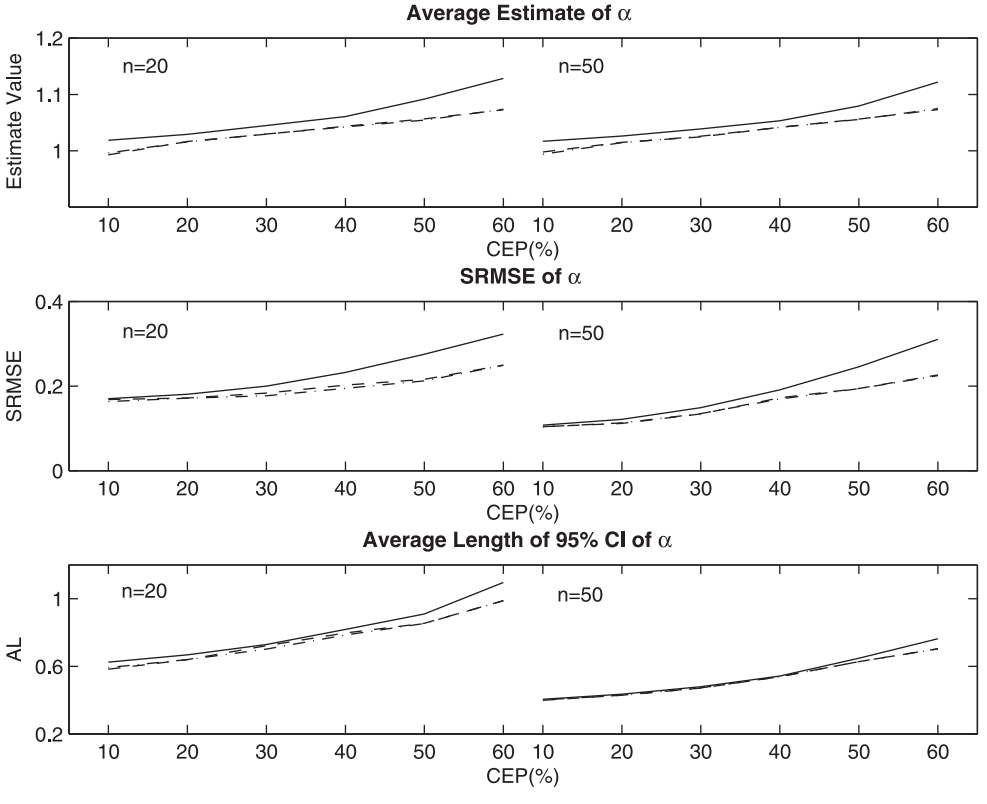

Figure 2. BS parameter setting 2 ($\alpha = 1.0, \beta = 1.0$): average estimate, SRMSE and AL for α with MLE (solid line –), conditional (dashed line - -), joint (dashdot line -.).

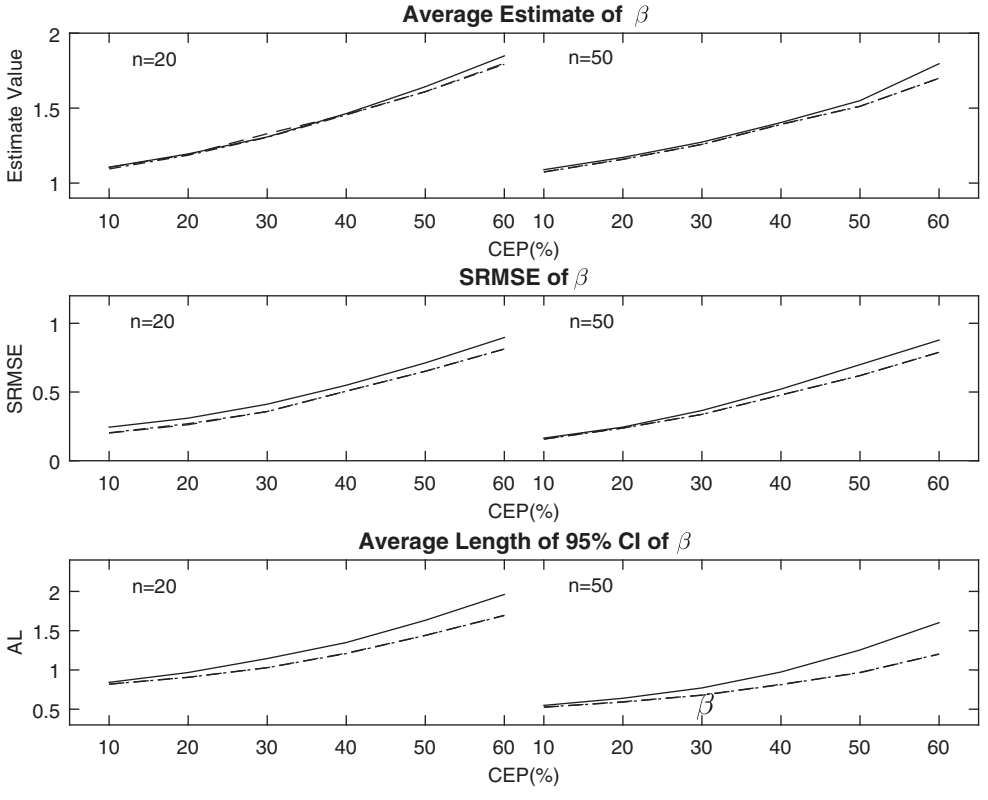

Figure 3. BS parameter setting 2 ($\alpha = 1.0, \beta = 1.0$): average estimate, SRMSE and AL for β with MLE (solid line –), conditional (dashed line - -), joint (dashdot line -.).

intervals, and coverage probability (CP) for the parameters. For comparison purpose, we also calculate the MLE estimates, AL of the 95% confidence intervals. The results are shown in Tables 1 and 2 for the parameter setting 1 ($\alpha = 0.5$) and setting 3 ($\alpha = 2.0$), respectively. To aid visualization of the results, we show the plot of estimates under the parameter setting 2 ($\alpha = 1.0$) in Figures 2 and 3, where we did more simulations up to 50% and 60% of CEP. The figures have similar patterns of results under other two-parameter settings. From these tables and plots, some features are summarized as follows: (i) bias, SRMSE and AL of 95% CI (confidence or credible intervals) decrease with sample size n but increase with CEP; (ii) with larger sample size, the performances of the three algorithms, in terms of bias, SRMSE and AL, are similar to one another; (iii) the effect of priors in posterior inference is more prevalent with smaller sample size. For the choice of relatively 'flat' priors, the Bayesian approach produces smaller bias and SRMSE as compared to MLE, especially at larger CEP, but slightly wider credible intervals than the confidence intervals and (iv) the conditional and joint samplings obtain similar estimations for the parameters, but it seems that the former algorithm run faster than the later one. For this reason, we would not consider joint sampling for the case of GBS distribution.

3.2. The GBS distribution

Recall that for the GBS distribution, β is no longer a scale parameter. For two sample sizes $n = 20$ and 50, we generate 10,000 data sets for each of the following three-parameter settings: $(\kappa, \alpha, \beta) = (0.5, 0.5, 1.0), (0.8, 1.0, 5.0), (0.2, 2.0, 5.0)$ under four or six CEP at 10%(10%)60%. For the Bayesian analysis, we adopt the same hyperparameters $a_0 = a_1 = b_0 = b_1 = 5, d_0 = d_1 = 1$. Five chains are

Table 3. GBS estimation results for parameter setting 1: $\kappa = 0.5, \alpha = 0.5, \beta = 1.0$, PM = parameter.

			MLE				Bayesian			
n	PM	CEP(%)	10	20	30	40	10	20	30	40
20	κ	Estimate	0.4911	0.4861	0.4816	0.4780	0.4916	0.4893	0.4820	0.4803
		SRMSE	0.2155	0.2251	0.2422	0.2570	0.1331	0.1299	0.1260	0.1204
		AL	0.9563	1.0308	1.1256	1.2469	0.7163	0.7375	0.7515	0.7733
		CP(%)	93.59	91.34	90.16	87.13	94.64	93.44	92.30	90.04
	α	Estimate	0.4861	0.5188	0.5419	0.5644	0.5056	0.5182	0.5333	0.5521
		SRMSE	0.0882	0.0919	0.0993	0.1124	0.0699	0.0756	0.0849	0.0988
		AL	0.3205	0.3498	0.3898	0.4474	0.3227	0.3502	0.3851	0.4323
		CP(%)	93.69	91.86	90.49	88.64	95.92	94.67	94.37	93.85
	β	Estimate	1.0483	1.0971	1.1552	1.2284	1.0362	1.0855	1.1465	1.2232
		SRMSE	0.1361	0.1659	0.2123	0.2769	0.1212	0.1491	0.1967	0.2649
		AL	0.4856	0.5322	0.5914	0.6739	0.4850	0.5360	0.6053	0.7064
		CP(%)	92.94	91.40	87.92	82.54	95.24	93.40	90.79	85.04
50	κ	Estimate	0.4939	0.4909	0.4901	0.4871	0.5002	0.4980	0.4927	0.4902
		SRMSE	0.1330	0.1412	0.1510	0.1650	0.1206	0.1245	0.1277	0.1316
		AL	0.5595	0.5995	0.6487	0.7096	0.5286	0.5564	0.5868	0.6189
		CP(%)	95.69	95.19	94.29	93.10	96.58	95.27	94.48	94.76
	α	Estimate	0.4978	0.5115	0.5275	0.5482	0.5093	0.5114	0.5221	0.5448
		SRMSE	0.0536	0.0594	0.0701	0.0877	0.0507	0.0591	0.0724	0.0924
		AL	0.2087	0.2293	0.2563	0.2952	0.2097	0.2302	0.2572	0.2949
		CP(%)	95.68	95.17	94.73	93.90	95.77	95.65	94.81	94.11
	β	Estimate	1.0427	1.0925	1.1522	1.2215	1.0328	1.0774	1.1424	1.2218
		SRMSE	0.0947	0.1339	0.1922	0.2711	0.0887	0.1267	0.1848	0.2648
		AL	0.3217	0.3562	0.4009	0.4631	0.3196	0.3548	0.4018	0.4696
		CP(%)	94.82	93.76	91.53	86.73	95.31	94.87	93.06	90.19

Table 4. GBS estimation results for parameter setting 3: $\kappa = 0.2, \alpha = 2.0, \beta = 5.0$, PM = parameter.

		Method	MLE				Bayesian			
n	PM	CEP(%)	10	20	30	40	10	20	30	40
20	κ	Estimate	0.2195	0.2215	0.2244	0.2315	0.2181	0.2198	0.2226	0.2306
		SRMSE	0.0957	0.0991	0.1060	0.1156	0.1204	0.1250	0.1333	0.1423
		AL	0.3519	0.3698	0.3933	0.4256	0.3452	0.3633	0.3855	0.4137
		CP(%)	92.65	92.47	91.73	90.42	93.45	93.32	91.39	90.65
	α	Estimate	1.8653	1.8521	1.8453	1.8333	1.8701	1.8642	1.8484	1.8384
		SRMSE	0.4061	0.4259	0.4509	0.4836	0.4070	0.4199	0.4343	0.4495
		AL	1.3058	1.3766	1.4674	1.5957	1.2160	1.2748	1.3495	1.4516
		CP(%)	90.48	88.80	87.12	85.84	92.09	90.13	88.51	86.77
	β	Estimate	5.4049	5.9894	6.6844	7.5644	4.9604	5.4408	6.0163	6.7359
		SRMSE	1.5724	1.8921	2.4005	3.1405	1.3320	1.4870	1.8287	2.3796
		AL	5.8884	6.4589	7.1767	8.1791	5.4311	6.0066	6.7627	7.8288
		CP(%)	92.62	91.39	87.29	80.84	94.71	92.80	88.01	82.25
50	κ	Estimate	0.2058	0.1907	0.1812	0.1783	0.2046	0.2087	0.2174	0.2210
		SRMSE	0.0507	0.0513	0.0547	0.0588	0.0569	0.0569	0.0596	0.0629
		AL	0.1956	0.2019	0.2095	0.2191	0.1951	0.2020	0.2100	0.2202
		CP(%)	95.15	94.36	93.01	90.87	95.89	95.31	93.69	93.24
	α	Estimate	2.0218	1.9801	1.9544	1.9324	2.0102	1.9817	1.9605	1.9425
		SRMSE	0.2297	0.2382	0.2543	0.2757	0.2307	0.2377	0.2514	0.2695
		AL	0.8382	0.8847	0.9442	1.0242	0.8193	0.8633	0.9198	0.9953
		CP(%)	94.84	93.78	92.20	91.03	95.46	95.09	93.85	92.16
	β	Estimate	5.2927	5.4477	6.5505	7.3653	5.2707	5.4322	6.5438	7.3331
		SRMSE	1.0921	1.5411	2.2372	3.1877	0.9846	1.3559	1.9869	2.8669
		AL	3.8895	4.2911	4.7939	5.4800	3.7713	4.1841	4.7070	5.4342
		CP(%)	94.11	92.08	90.83	85.48	95.11	94.83	93.86	89.97

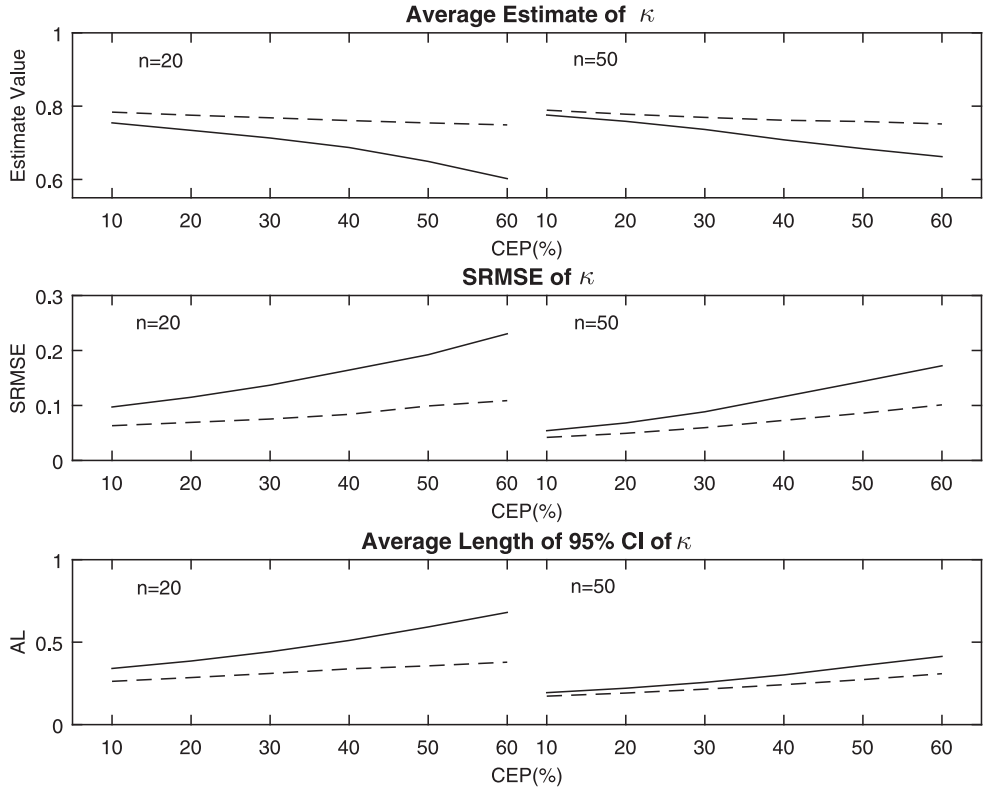

Figure 4. GBS parameter setting 2 ($\kappa = 0.8, \alpha = 1.0, \beta = 5.0$): average estimate, SRMSE and AL for κ with MLE (solid line –), Bayesian (dashed line - -).

run for the Gibbs sampling each with 10,000 iterations and 2000 burn-in. The scale factors for the sequences of all parameters are within 1.00–1.01, showing the convergence of all chains. The estimation results are summarized in Tables 3 and 4 for the parameter settings 1 and 3, and the visualized plots under parameter setting 2 are shown in Figures 4–6, where more simulations are conducted up to 50% and 60% of CEP. Similar features as in the BS distribution can be seen from the results. Mainly, the Bayesian approach produces smaller bias, SRMSE and AL as compared to MLE at higher CEPs. Other features are summarized as follows: (i) under parameter setting 1 with $\kappa = 0.5$, a few lengths of the confidence intervals for κ is larger than 1 due to the interval exceeding the range (0, 1), whereas the credible intervals are inside parameter space due to the choice of priors and proposals; (ii) under the parameter setting 2 with larger value of κ, MLE tends to overestimate both α and β and underestimate κ. However, the Bayesian approach obtains less bias for all three parameters. Overall, even with relatively 'less' informative priors, the Bayesian method outperforms MLE for all parameter settings, especially when sample size is small.

4. Real data analysis

In this section, we present the application for two real data sets. The first data set was given by Birnbaum and Saunders [7] about the fatigue life of 6061-T6 aluminium coupons cut parallel to the direction of rolling and oscillated at 18 cycles /s. This is a complete data set, as presented in Table 5, consisting of 101 failure observations with maximum stress per cycle of 21,000 psi. Fitting the GBS distribution for the data, we take the prior $\kappa \sim U(0, 1)$, and choose the hyperparameters b_0 and b_1 such that the prior mean of β is close to the sample median 1420 of the data, a_0 and a_1 such that

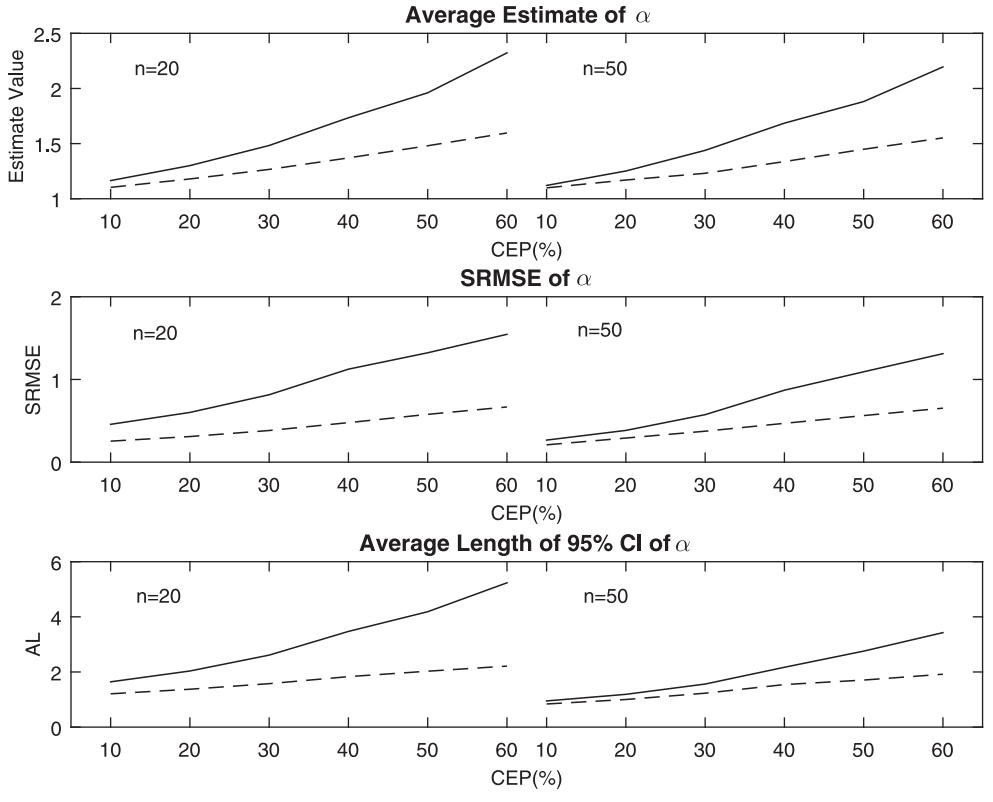

Figure 5. GBS parameter setting 2 ($\kappa = 0.8, \alpha = 1.0, \beta = 5.0$): average estimate, SRMSE and AL for α with MLE (solid line –), Bayesian (dashed line - -).

the conditional prior mean $E(\alpha^2 \mid \beta = 1420)$ is close to the MLE $\hat{\alpha}^2_{\text{MLE}} = 32.49$, and the CV (coefficient of variation) of the inverse gamma priors is close to the CV of standard uniform distribution $(1/\sqrt{3})$. Hence, we have the hyperparameter values as followings: $a_0 = 10, a_1 = 55, b_0 = 10, b_1 = 0.00088, d_0 = d_1 = 1$. We run a chain of 20,000 iterations with a burn-in period of 5000. To reduce the correlation among the samples, every 5th sample of the remaining 15,000 samples are used for posterior inference. The results are tabulated in Table 6, where one can see that the point estimates obtained by both MLE and Bayesian methods are close to each other due to the large sample size. However, the 95% confidence intervals contain negative values for κ and α, whereas the credible intervals do not, and are much narrower. Compared to the results obtained by Owen [31], whose MLE estimates $\hat{\kappa}_{\text{MLE}} = 0.064, \hat{\alpha}_{\text{MLE}} = 6.605$ and $\hat{\beta}_{\text{MLE}} = 1393.42$, we believe that our MLEs are more accurate, since the values of the normal equations are much closer to zero by plugging in our MLEs than his MLEs. Owen [31] did not proceed with interval estimation, nor did he consider the case of right-censored data. Finally, we show the fitted reliability curves based on the MLEs and posterior estimates, and the Kaplan–Meier 's plot [38] in Figure 7. It is worth noting that the fitted curves are very similar to the Kaplan–Meier 's plot, indicating efficiency of our estimation approach.

The second data set was presented in [39] regarding the lifetimes (in months) of 20 cancer patients receiving a new treatment: 3, 5, 6, 7, 8, 9, 10, 10+, 12, 15, 15+, 18, 19, 20, 22, 25, 28, 30, 40, 45+, where symbol '+' denotes a right-censored observation. Assuming an underlying GBS distribution for the data, we adopt the same procedure as discussed to specify the hyperparameter values by using the sample median 15 of the data and $\hat{\alpha}^2_{\text{MLE}} = 1$. In summary, we choose the following hyperparameter values: $a_0 = 10, a_1 = 19, b_0 = 10, b_1 = 0.083, d_0 = d_1 = 1$. An MCMC chain of 20,000 iterations

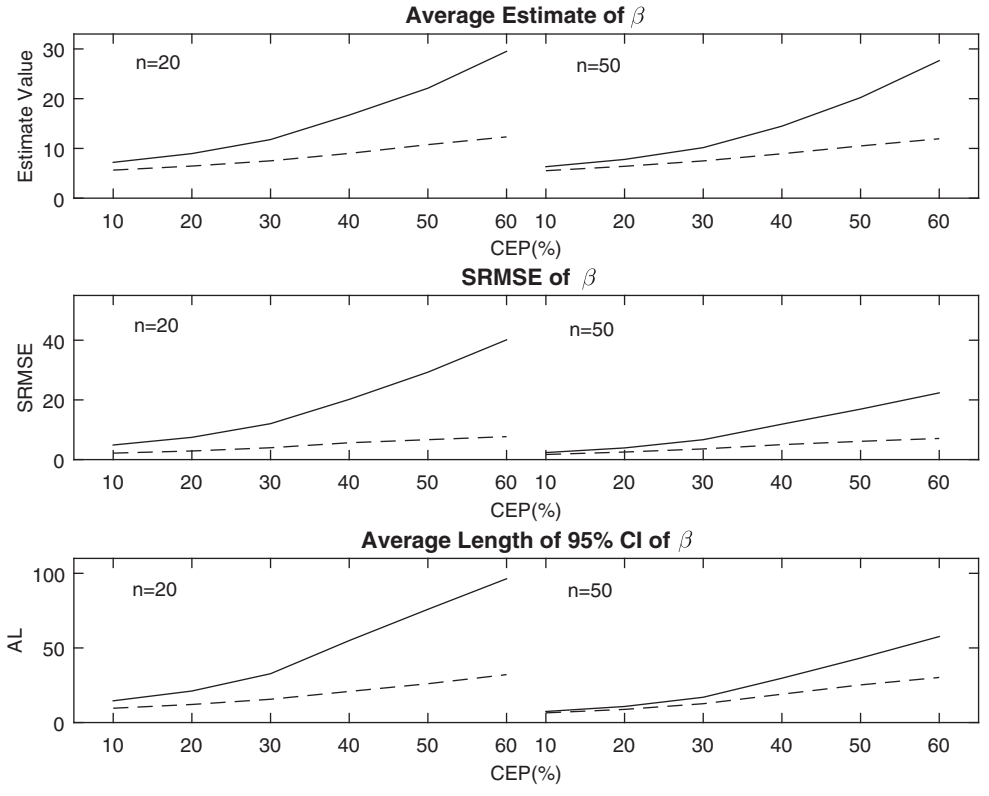

Figure 6. GBS parameter setting 2 ($\kappa = 0.8, \alpha = 1.0, \beta = 5.0$): average estimate, SRMSE and AL for β with MLE (solid line –), Bayesian (dashed line - -).

Table 5. Fatigue life of 6061-T6 aluminium coupons exerted with maximum stress per cycle of 21,000 psi.

370	706	716	746	785	797	844	855	858	886	886	930	960
988	990	1000	1010	1016	1018	1020	1055	1085	1102	1102	1108	1115
1120	1134	1140	1199	1200	1200	1203	1222	1235	1238	1252	1258	1262
1269	1270	1290	1293	1300	1310	1313	1315	1330	1355	1390	1416	1419
1420	1420	1450	1452	1475	1478	1481	1485	1502	1505	1513	1522	1522
1530	1540	1560	1567	1578	1594	1602	1604	1608	1630	1642	1674	1730
1750	1750	1763	1768	1781	1782	1792	1820	1868	1881	1890	1893	1895
1910	1923	1940	1945	2023	2100	2130	2215	2268	2440			

Table 6. Estimation results for fatigue life data.

Parameter	Method	Estimate	95% CI	CI length
κ	MLE	0.0844	(−0.1569, 0.3257)	0.4826
	Bayesian	0.1005	(0.0347, 0.1563)	0.1215
α	MLE	5.7112	(−4.0904, 15.5127)	19.6031
	Bayesian	5.2754	(3.5531, 8.0930)	4.5399
β	MLE	1391.1037	(1309.5219, 1472.6856)	163.1637
	Bayesian	1387.7813	(1309.7578, 1467.5915)	157.8337

with a burn-in period of 5000 produces the estimation results of our method, together with the results by Achcar and Moala [39], tabulated in Table 7. For this data with relatively small sample size ($n = 17, m = 3$), the point estimates obtained by both MLE and Bayesian methods are very close to each other, but the credible intervals are narrower than the corresponding confidence intervals, especially for κ and α. Our results are also close to the posterior estimates obtained by Achcar and

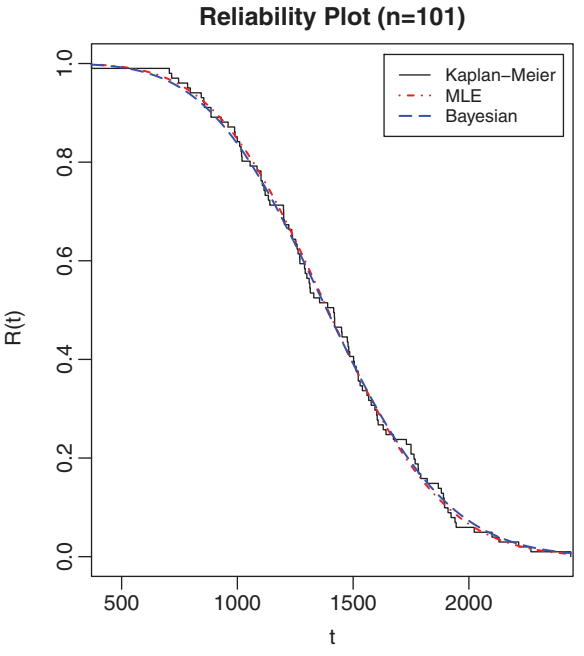

Figure 7. Reliability plots for fatigue life data.

Table 7. Estimation results for cancer data.

Parameter	Method	Estimate	95% CI	CI length
κ	MLE	0.4195	(0.0833, 0.7558)	0.6725
	Bayesian	0.4558	(0.2472, 0.6736)	0.4264
α	MLE	0.9740	(0.1273, 1.8207)	1.6934
	Bayesian	0.9619	(0.6035, 1.5103)	0.9068
	A-M (2010)	0.8854	(0.6101, 1.295)	0.6849
β	MLE	15.6289	(9.6137, 21.6441)	12.0304
	Bayesian	15.4105	(10.4887, 21.6960)	11.2073
	A-M (2010)	16.03	(10.93, 24.36)	13.43

Moala [39], who fitted a BS distribution for this data. Recall that $\text{BS}(\alpha, \beta) = \text{GBS}(0.5, \alpha, \beta)$, we notice that our 95% credible interval for κ includes 0.5, and their posterior estimates also lie within our 95% credible intervals for α and β, respectively. Additionally, our credible interval is slightly wider for α, as we adopt a 'flatter' prior for α, and narrower for β, since they specified a uniform prior for β in [39]. Finally, the fitted reliability curves and the Kaplan–Meier 's reliability estimate are shown in Figure 8. The similarities of these curves demonstrate the efficiency of our estimation procedure.

5. Discussions and concluding remarks

We presented a Bayesian parameter inference of the BS distribution, as well as the GBS distribution [31], in the presence of random right-censored data. We utilized the data augmentation technique to circumvent the arduous expressions involving the censored data and adopted innovative sampling schemes in the posterior inference. The simulation study demonstrated the efficient and impressive performance of our Bayesian method, where the accuracy of parameter estimation improves with larger sample size but deteriorates in the presence of more censored observations. Through the comparison, the Bayesian method outperformed the traditional likelihood-based approach in terms of

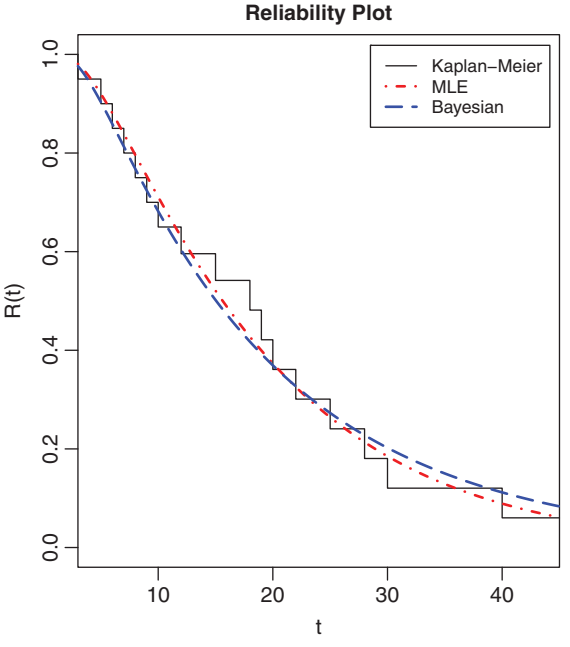

Figure 8. Reliability plots for cancer data.

bias, precision and interval width in parameter inference, especially when sample size is small. We have also illustrated, with two real data sets, that our Bayesian method can be readily applied to Type-II censored data, as a special case, for efficient, reliable and precise inference. Finally, we may perform a Bayesian hypothesis testing of parameters similar to a frequentist approach in [40], who conducted the likelihood ratio and gradient tests for the BS and GBS distributions. Although their GBS formed from an elliptical model is different from the GBS in this article, it is very straightforward to apply such tests in our Bayesian framework. For example, based on the approach of gradient test, we may replace the likelihood function and MLE by the posterior function and 'posterior' MLE, respectively. In other words, in the expression of gradient test statistic, the score function becomes the 'posterior' score which is the summation of scores of likelihood and priors.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Sha's work was partially supported by NSF CMM-0654417 and NIH NIMHD-2G12MD007592.

References

[1] Birnbaum ZW, Saunders SC. A new family of life distributions. J Appl Probab. 1969;6:319–327.
[2] Saunders SC. A family of random variables closed under reciprocation. J Amer Statist Assoc. 1974;69(346):533–539.
[3] Fierro R, Leiva V, Ruggeri F, et al. On a Birnbaum–Saunders distribution arising from a non-homogeneous Poisson process. Statist Probab Lett. 2012;83(4):1233–1239.
[4] Desmond AF. Stochastic models of failure in random environments. Canad J Statist. 1985;13(3):171–183.
[5] Desmond AF. On the relationship between two fatigue-life models. IEEE Trans Reliab. 1986;35(2):167–169.
[6] Rieck JR. A moment generating function with applications to the Birnbaum–Saunders distributions. Comm Statist Theory Methods. 1999;28(9):2213–2222.

[7] Birnbaum ZW, Saunders SC. Estimation for a family of life distributions with applications to fatigue. J Appl Probab. 1969;6(2):328–347.
[8] Engelhardt M, Bain LJ, Wright FT. Inferences on the parameters of the Birnbaum–Saunders fatigue life distribution based on maximum likelihood estimation. Technometrics. 1981;23(3):251.
[9] Dupuis DJ, Mills JE. Robust estimation of the Birnbaum–Saunders distribution. IEEE Trans Reliab. 1998;47(1):88–95.
[10] Ng HKT, Kundu D, Balakrishnan N. Modified moment estimation for the two-parameter Birnbaum–Saunders distribution. Comput Statist Data Anal. 2003;43:283–298.
[11] Wu J, Wong ACM. Improved interval estimation for the two-parameter Birnbaum–Saunders distribution. Comput Statist Data Anal. 2004;47:809–821.
[12] Barndorff-Nielsen OE. Modified signed log likelihood ratio. Biometrika. 1991;78:557–563.
[13] Balakrishnan N, Zhu X. An improved method of estimation for the parameters of the Birnbaum–Saunders distribution. J Stat Comput Simul. 2014;84(10):2285–2294.
[14] Lemonte AJ, Cribari-Neto F, Vasconcellos KLP. Improved statistical inference for the two-parameter Birnbaum–Saunders distribution. Comput Statist Data Anal. 2007;51:4656–4681.
[15] Hirose H. Bias correction for the maximum likelihood estimates in the two-parameter Weibull distribution. IEEE Trans Dielectr Electr Insul. 1999;6(1):66–68.
[16] Ng HKT, Kundu D, Balakrishnan N. Point and interval estimation for the two-parameter Birnbaum–Saunders distribution based on Type-II censored samples. Comput Statist Data Anal. 2006;50:3222–3242.
[17] Steven G, Li L. Estimation of the parameters of the Birnbaum–Saunders distribution. Comm Statist Theory Methods. 2006;35(12):2157–2169.
[18] Wang Z, Desmond AF, Lu X. Modified censored moment estimation for the two-parameter Birnbaum–Saunders distribution. Comput Statist Data Anal. 2006;50:1033–1051.
[19] Sun D, Berger JO. Reference priors with partial information. Biometrika. 1998;85(1):55–71.
[20] Achcar JA. Inferences for the Birnbaum–Saunders fatigue life model using Bayesian methods. Comput Statist Data Anal. 1993;15(4):367–380.
[21] Xu A, Tang Y. Reference analysis for Birnbaum–Saunders distribution. Comput Statist Data Anal. 2010;54:185–192.
[22] Xu A, Tang Y. Bayesian analysis of Birnbaum–Saunders distribution with partial information. Comput Statist Data Anal. 2011;55:2324–2333.
[23] Lindley DV. Approximate Bayesian methods. Trabajos de Estadistica Y de Investigacion Operativa. 1980;31(1):223–245.
[24] Wang M, Sun X, Park C. Bayesian analysis of Birnbaum–Saunders distribution via the generalized ratio-of-uniforms method. Comput Statist. 2016;31(1):207–225.
[25] Rieck JR, Nedelman JR. A log-linear model for the Birnbaum–Saunders distribution. Technometrics. 1991;33(1):51–60.
[26] Owen WJ, Padgett WJ. A Birnbaum–Saunders accelerated life model. IEEE Trans Reliab. 2000;49(2):224–229.
[27] Diaz-Garcia JA, Leiva-Sanchez V. A new family of life distributions based on the elliptically contoured distributions. J Statist Plan Inference. 2005;128(2):445–457.
[28] Gomez HW, Olivares-Pacheco JF, Bolfarine H. An extension of the generalized Birnbaum–Saunders distribution. Statist Probab Lett. 2009;79(3):331–338.
[29] Leiva V, Riquelme M, Balakrishnan N, et al. Lifetime analysis based on the generalized Birnbaum–Saunders distribution. Comput Statist Data Anal. 2008;52(4):2079–2097.
[30] Sanhueza A, Leiva V, Balakrishnan N. The generalized Birnbaum–Saunders distribution and its theory, methodology, and application. Comm Statist Theory Methods. 2008;37(5):645–670.
[31] Owen WJ. A new three-parameter extension to the Birnbaum–Saunders distribution. IEEE Trans Reliab. 2006;55:475–479.
[32] Beran J. Statistics for long memory processes. London: Chapman & Hall; 1994.
[33] Tanner MA, Wong WH. The calculation of posterior distributions by data augmentation. J Amer Statist Assoc. 1987;82(398):528–540.
[34] Lawless JF. Statistical models and methods for lifetime data. 2nd ed. New York, NJ: John Wiley & Sons; 2003.
[35] Casella G, George EI. Explaining the Gibbs sampler. Am Stat. 1992;46(3):167–174.
[36] Chib S, Greenberg E. Understanding the Metropolis–Hastings algorithm. Am Stat. 1995;49(4):327–335.
[37] Gelman A, Carlin JB, Stern HS, et al. Bayesian data analysis. 2nd ed. London: Chapman & Hall; 2004.
[38] Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Amer Statist Assoc. 1958;53(282):457–481.
[39] Achcar JA, Moala FA. Use of MCMC methods to obtain Bayesian inferences for the Birnbaum–Saunders distribution in the presence of censored data and covariates. Adv Appl Stat. 2010;17:1–27.
[40] Lemonte AJ, Ferrari SLP. Testing hypotheses in the Birnbaum–Saunders distribution under type-II censored samples. Comput Statist Data Anal. 2011;55:2388–2399.

Appendix

We present the detailed derivation for Equation (14). Since the Jacobian $J = \alpha\beta$, the Fisher information matrix for the joint posterior of $\boldsymbol{\theta} = (\alpha, \beta)^{\mathrm{T}}$ is

$$-E\left\{\frac{\partial^2(\log\pi(\boldsymbol{\theta}\mid\boldsymbol{t})+\log|J|)}{\partial(\log\boldsymbol{\theta})\partial(\log\boldsymbol{\theta})^{\mathrm{T}}}\right\} = -E\left\{\frac{\partial^2\log\pi(\boldsymbol{\theta}\mid\boldsymbol{t})}{\partial(\log\boldsymbol{\theta})\partial(\log\boldsymbol{\theta})^{\mathrm{T}}}\right\}. \tag{A1}$$

We denote the log-likelihood function $\ell(\alpha,\beta) = \log L(\alpha,\beta\mid\boldsymbol{t})$, and so the log joint posterior, apart from a constant, becomes $\log\pi(\alpha,\beta\mid\boldsymbol{t}) = \ell(\alpha,\beta) + \log\pi(\alpha,\beta)$, and then

$$\frac{\partial^2\log\pi(\alpha,\beta\mid\boldsymbol{t})}{\partial(\log\alpha)^2} = \frac{\partial^2\ell(\alpha,\beta)}{\partial\alpha^2}\times\alpha^2 + \frac{\partial\ell(\alpha,\beta)}{\partial\alpha}\times\alpha + \frac{\partial^2\log\pi(\alpha,\beta)}{\partial(\log\alpha)^2}, \tag{A2}$$

$$\frac{\partial^2\log\pi(\alpha,\beta\mid\boldsymbol{t})}{\partial(\log\alpha)\partial(\log\beta)} = \frac{\partial^2\ell(\alpha,\beta)}{\partial\alpha\partial\beta}\times\alpha\beta + \frac{\partial^2\log\pi(\alpha,\beta)}{\partial(\log\alpha)\partial(\log\beta)}, \tag{A3}$$

$$\frac{\partial^2\log\pi(\alpha,\beta\mid\boldsymbol{t})}{\partial(\log\beta)^2} = \frac{\partial^2\ell(\alpha,\beta)}{\partial\beta^2}\times\beta^2 + \frac{\partial\ell(\alpha,\beta)}{\partial\beta}\times\beta + \frac{\partial^2\log\pi(\alpha,\beta)}{\partial(\log\beta)^2}. \tag{A4}$$

From the priors in Equations (5) and (6), we have

$$\frac{\partial^2\log\pi(\alpha,\beta)}{\partial(\log\alpha)^2} = -\frac{2a_0\beta}{a_1\alpha^2}, \frac{\partial^2\log\pi(\alpha,\beta)}{\partial(\log\alpha)\partial(\log\beta)} = \frac{a_0\beta}{a_1\alpha^2}, \tag{A5}$$

$$\frac{\partial^2\log\pi(\alpha,\beta)}{\partial(\log\beta)^2} = -\left(\frac{a_0\beta}{2a_1\alpha^2} + \frac{b_0}{2b_1\beta}\right). \tag{A6}$$

By the fact that $E\{\partial\ell(\alpha,\beta)/\partial\alpha\} = E\{\partial\ell(\alpha,\beta)/\partial\beta\} = 0$, and $-E\{\partial^2\ell(\alpha,\beta)/\partial\alpha^2\} = 2(n+m)/\alpha^2, -E\{\partial^2\ell(\alpha,\beta)/\partial\alpha\partial\beta\} = 0, -E\{\partial^2\ell(\alpha,\beta)/\partial\beta^2\} = (n+m)[\alpha(2\pi)^{-1/2}h(\alpha)+1]/(\alpha^2\beta^2)$ from [14], we obtain the elements of the Fisher information matrix in Equation (29)

$$-E\left\{\frac{\partial^2\log\pi(\alpha,\beta\mid\boldsymbol{t})}{\partial(\log\alpha)^2}\right\} = 2(n+m) + \frac{2a_0\beta}{a_1\alpha^2}, \tag{A7}$$

$$-E\left\{\frac{\partial^2\log\pi(\alpha,\beta\mid\boldsymbol{t})}{\partial(\log\alpha)\partial(\log\beta)}\right\} = -\frac{a_0\beta}{a_1\alpha^2}, \tag{A8}$$

$$-E\left\{\frac{\partial^2\log\pi(\alpha,\beta\mid\boldsymbol{t})}{\partial(\log\beta)^2}\right\} = g(\alpha) + \frac{a_0\beta}{2a_1\alpha^2} + \frac{b_0}{2b_1\beta}, \tag{A9}$$

with $g(\alpha) = (n+m)[\alpha(2\pi)^{-1/2}h(\alpha)+1]/(\alpha^2\beta^2), h(\alpha) = \alpha\sqrt{\pi/2} - \pi\,\mathrm{e}^{2/\alpha^2}[1-\Phi(2/\alpha)]$.

Additionally, the log conditional posterior of $\log\beta$ is $\log\pi(\beta\mid\alpha,\boldsymbol{t}) + \log|J|$ with $J = \beta$, and $\pi(\beta\mid\alpha,\boldsymbol{t}) \propto \pi(\alpha,\beta\mid\boldsymbol{t})$, and so the Fisher information in Equation (11) is

$$-E\left\{\frac{\partial^2(\log\pi(\beta\mid\alpha,\boldsymbol{t})+\log|J|)}{\partial(\log\beta)^2}\right\} = -E\left\{\frac{\partial^2\log\pi(\alpha,\beta\mid\boldsymbol{t})}{\partial(\log\beta)^2}\right\} = g(\alpha) + \frac{a_0\beta}{2a_1\alpha^2} + \frac{b_0}{2b_1\beta}.$$