

Bayesian accelerated life testing under competing log-location-scale family of causes of failure

Chiranjit Mukhopadhyay¹ · Soumya Roy²

Received: 9 April 2014 / Accepted: 24 June 2015 / Published online: 21 July 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract This article provides Bayesian analyses of data arising from multi-stress accelerated life testing of series systems. The component log-lifetimes are assumed to independently belong to some log-concave location-scale family of distributions. The location parameters are assumed to depend on the stress variables through a linear stress translation function. Bayesian analyses and associated predictive inference of reliability characteristics at usage stresses are performed using Gibbs sampling from the joint posterior. The developed methodology is numerically illustrated by analyzing a real data set through Bayesian model averaging of the two popular cases of Weibull and log-normal, with the later getting a special focus in this article as a slightly easier example of the log-location-scale family. A detailed simulation study is also carried out to compare the performance of various Bayesian point estimators for the log-normal case.

Keywords Data augmentation · Gibbs sampling · Log-concave densities · Model averaging · Predictive inference · Reliability · Stress translation function · Series system

✉ Soumya Roy
soumya@iimk.ac.in

Chiranjit Mukhopadhyay
cm@mgmt.iisc.ernet.in

¹ Department of Management Studies, Indian Institute of Science, Bangalore 560012, India

² Indian Institute of Management Kozhikode, Kozhikode 673570, India

1 Introduction

Testing modern, reliable hardware systems at normal usage condition for failures can be highly time consuming. To induce early failures, it is therefore common practice in the manufacturing and defense industry to subject the developed systems to more extreme stress levels compared to stress levels at the normal usage condition. For instance, consider the example of Class-B electrical insulation system in electric motors discussed in Nelson (1990, p. 418). The insulation system is designed to operate at a usage temperature of 403.15 K (Kelvin). Since it is extremely unlikely for the insulation systems to fail within any reasonable time at 403.15 K, they are tested at temperatures 423.15, 443.15, 463.15 and 493.15 K to produce quick breakdown data. Such tests are commonly known as Accelerated Life Tests (ALT henceforth). The observations collected on the failure times of the systems at accelerated conditions (such as the four different temperature levels mentioned above) are then used to infer about its reliability at normal usage condition (like the 403.15 K above), which is also called usage stress. Nelson (1990) provides a comprehensive account of the relevant statistical models and methods for analyzing such accelerated data.

In practice, a system often has more than one failure modes, which are the causes of system failures. For example, the insulation system mentioned above has three failure modes, called turn, phase and ground. Each of these modes of failures occurs at different parts of the insulation system, but the insulation system as a whole fails as soon as one of these causes takes place. Such systems which fail as soon as one of their causes or modes of failures occurs, are called series systems. A series system ALT data thus also consist of the causes or failure modes, along with the (observed or censored) times to system failures at the elevated stress levels. The failure modes responsible for system failures are also called components of the system and through-out this article we use these two terms interchangeably.

A number of studies in the literature has considered the maximum likelihood inference of series system ALT data with a single stress variable (see Klein and Basu 1981, 1982a, b; Nelson 1990; Kim and Bai 2002; Jiang 2011, for example). These studies exploit large-sample properties of the MLEs to infer on the component and system lifetime distributions at usage stress. However in an ALT, the available sample size is expected to be small. In such cases, Bayesian procedures are known to provide more reasonable results (see Berger 1985, p. 125). There are also situations where important prior information are available in an ALT (see Zhang and Meeker 2006). Bayesian methods allow easy and formal incorporation of such prior information into decision-making. Furthermore, as explained in the first paragraph, the goal of performing an ALT is to predict the component and system lifetime distributions at usage stress. Since the Bayesian paradigm handles the problem of prediction in a more coherent manner (see Berger 1985, pp. 67, 157), development of such Bayesian methodologies for analyzing ALT data seems imperative and important.

A considerable amount of literature exists on the Bayesian analysis of ALT data for systems with a single component (see Pathak et al. 1991; Tojeiro et al. 2004; Dorp and Mazzuchi 2004, 2005, for example). However, the literature on Bayesian analysis of accelerated data for series systems is comparatively less expansive. Bunea and Mazzuchi (2006), Tan et al. (2009) and Xu and Tang (2011) have provided Bayesian

analyses of single-stress series system ALT data with independent exponential or Weibull component lives.

In these Bayesian analyses, first the Bayes' estimates of the exponential or Weibull model parameters are obtained at each elevated stress level. The model parameters at usage stress are then estimated using the method of least squares through an assumed Stress Translation Function (STF henceforth). Thus this approach is not fully Bayesian. Roy and Mukhopadhyay (2013, 2014) have developed full Bayesian solutions to this problem assuming that the component lifetimes are independent exponentials and Weibulls with identical shape parameters (see Fan and Hsu (2014) for the non-identical Weibull case). Generalizing these analyses to a more flexible log-location-scale family, this article now presents full Bayesian analyses of fixed multi-stress ALT data of series systems.

Briefly, it is assumed that the component log-lifetimes of a series system are independent and belong to some log-concave location-scale family of distributions. The families of these distributions may be different, like some are extreme-value while the others are normal, but the location parameters of these distributions are assumed to be connected with the multiple stress variables through a linear STF. The combined choices of this log-location-scale family for the component lifetime distributions and the form of the proposed linear STF, provide a rich and comprehensive model covering all the practical and standard cases of fixed multi-stress ALT data of series systems. A detailed discussion of this model is carried out in Sect. 2.

Section 3 next discusses the issue of Bayesian analyses of such models. It is first shown that the one-dimensional conditional posterior distributions of all the model parameters given the rest are log-concave, if so are the corresponding priors. This fact is then used to Gibbs sample from the joint posterior which can be subsequently utilized for standard Bayesian predictive inference for component and system lifetime distributions at usage stresses under a fixed assumed model. Thus this approach is fully Bayesian in the sense that it takes care of the uncertainty in the model parameters through their joint posterior density, resulting in a coherent set of predictions.

Although log-normally distributed component lives is an important special case of the model introduced in Sect. 2, here the general solution can be eased (in the sense of avoiding log-concave sampling) and made more elegant through data-augmentation and choice of priors. Unlike Weibull, the other important component life distribution of the log-location-scale family, log-normal allows objective Bayesian analysis in this problem. Thus Sect. 4 is devoted to the special case of log-normal distribution demonstrating these issues and a detailed simulation study comparing different optimal (with respect to various loss functions) Bayesian point estimates.

Section 5 illustrates Bayesian analyses of the data of the example of the electrical insulation system introduced in the first paragraph. First an attempt is made to choose an "appropriate" model (among exponential, Weibull and log-normal) for the component lifetimes by computing exact Bayes' Factor (BF henceforth) through Markov Chain Monte Carlo (MCMC henceforth). But the BF based model selections being inconclusive, the standard fixed model predictive Bayesian inference developed in Sect. 3, is further enhanced by Bayesian model averaging for obtaining appropriate Bayesian point and interval estimates of reliability characteristics at usage stress. Section 6 concludes the article by summarizing its contributions and providing a few

pointers towards future research needs. The proofs are deferred to Appendix 1, and the abbreviations and notations used in the article are listed in Appendix 2.

2 Log-location-scale model with linear STF

Consider a series system with J independent causes of failure, with the j th cause having a potential lifetime X_j and log-lifetime $Y_j = \log X_j$. Suppose such a system is put on an ALT with K life acceleration factors or stress variables. Let $\mathbf{Z} = (Z_1, \dots, Z_K)'$ denote the vector of such K stress variables acting on the system. Given $\mathbf{Z} = \mathbf{z}$, it is assumed that the Y_j s are mutually independent with Y_j having p.d.f.

$$f_{Y_j}(y|\boldsymbol{\theta}_j, \tau_j, \mathbf{z}) = \tau_j f_j(\tau_j(y - \mu_j(\boldsymbol{\theta}_j, \mathbf{z}))), \tag{1}$$

where $f_j(\cdot)$ is a univariate log-concave p.d.f., $\boldsymbol{\theta}_j = (\theta_{1j}, \dots, \theta_{Kj})' \in \mathfrak{R}^K$ is a parameter vector of stress coefficients determining the location $\mu_j(\boldsymbol{\theta}_j, \mathbf{z})$ of the distribution of Y_j , and $\tau_j > 0$ is a scale parameter. Thus as is standard in ALT (see Nelson 1990, p. 80, for example), Eq. (1) assumes that any change in the stresses results in just a translation or shift in location of the distribution of the component log-lifetimes (without distorting its exact shape or scaling). This is why $\mu_j(\boldsymbol{\theta}_j, \mathbf{z})$ s are more commonly known as the STFs in the literature (see Klein and Basu 1981, for example). With $\bar{F}_j(t) = \int_t^\infty f_j(u)du$ and $h_j(t) = f_j(t)/\bar{F}_j(t)$, let $\bar{F}_{Y_j}(y|\boldsymbol{\theta}_j, \tau_j, \mathbf{z}) = \bar{F}_j(\tau_j(y - \mu_j(\boldsymbol{\theta}_j, \mathbf{z})))$ and $h_{Y_j}(y|\boldsymbol{\theta}_j, \tau_j, \mathbf{z}) = \tau_j h_j(\tau_j(y - \mu_j(\boldsymbol{\theta}_j, \mathbf{z})))$ respectively denote the s.f. and h.r. of Y_j at stress $\mathbf{Z} = \mathbf{z}$.

Equation (1) as a basic model for analyzing multi-stress ALT data is very flexible in two respects. First, it gives the user a choice of component life-distributions in terms of the specification of $f_j(\cdot)$, and second, it can accommodate any multi-stress STF by an appropriate choice of $\mu_j(\boldsymbol{\theta}_j, \mathbf{z})$. Though well-known (see Meeker and Escobar 1998, p. 79, for example), for the sake of easy reference and the additional requirement of log-concavity (which is easily checked via $\frac{d^2}{dt^2} \log f_j(t) \leq 0$ in most cases), Table 1 lists the standard life-distributions used in the literature that can be modeled by an appropriate choice of log-concave $f_j(\cdot)$, with the corresponding expressions for the location μ_j and scale τ_j in terms of the original parameters. It should be noted that for ease of recognition, Table 1 gives the names and p.d.f.s of the life-distributions in terms of that of lifetime X_j .

Equation (1) now requires specification of the location μ_j through the STF $\mu_j(\boldsymbol{\theta}_j, \mathbf{z})$. As introduced in Roy and Mukhopadhyay (2014), here also we assume a linear STF given by:

$$\mu_j(\boldsymbol{\theta}_j, \mathbf{z}) = \sum_{k=1}^K \theta_{kj} g_{kj}(z_k) = \sum_{k=1}^K \theta_{kj} s_{kj} = \mathbf{s}_j' \boldsymbol{\theta}_j, \tag{2}$$

where $\mathbf{s}_j = (s_{1j}, \dots, s_{Kj})'$ with $s_{kj} = g_{kj}(z_k)$, for some real valued functions $g_{kj}(\cdot)$ s.

Though already explained somewhat briefly in Roy and Mukhopadhyay (2014, 2015), the reasoning behind choosing Eq. (2) as the STF is as follows. Typically in

Table 1 Log-concave log-location-scale family of life distributions

Name	p.d.f.	$f_j(t)$	μ_j	τ_j
Exponential	$\frac{1}{\eta} e^{-(x/\eta)}$	$\exp(t - e^t)$	$\log \eta$	1
Weibull	$\frac{\beta}{\eta} (\frac{x}{\eta})^{\beta-1} e^{-(x/\eta)^\beta}$	$\exp(t - e^t)$	$\log \eta$	β
Generalized Gamma	$\frac{\beta}{\eta \Gamma(\alpha)} x^{\alpha\beta-1} e^{-(x/\eta)^\beta}$	$\frac{1}{\Gamma(\alpha)} \exp(\alpha t - e^t)$	$\log \eta$	β
Log-normal	$\frac{1}{\sqrt{2\pi\sigma^2}x} e^{-\frac{1}{2}(\frac{\log x - \mu}{\sigma})^2}$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$	μ	$1/\sigma$
Log-Laplace	$\frac{1}{2\sigma x} e^{-\frac{ \log x - \mu }{\sigma}}$	$\frac{1}{2} e^{- t }$	$\log \eta$	$1/\sigma$
Log-logistic	$\frac{\beta}{\alpha} \frac{(x/\alpha)^{\beta-1}}{(1+(x/\alpha)^\beta)^2}$	$\frac{e^t}{(1+e^t)^2}$	$\log \alpha$	β
Log-sech	$\frac{\beta}{\pi\alpha} \frac{(x/\alpha)^{\frac{\beta}{2}-1}}{1+(x/\alpha)^\beta}$	$\frac{1}{\pi} \operatorname{sech}(t)$	$\log \alpha$	$\beta/2$

an ALT there are several choices for stress variables such as temperature, voltage, pressure, humidity, dust, vibration, etc. and are often used in combinations. Since the final aim of an ALT requires extrapolation to usage stresses, it is imperative that the underlying models that do so, have some engineering basis (see [Meeker and Escobar 1998](#), p. 495). In this regard there are several multi-stress models that exist in the engineering literature,¹ which express X_j , the time to failure (of the j th component in this case) as a non-linear function of the applied stresses. A natural way to accommodate such (non-linear) models is to let the logarithm of these functions equal the STF $\mu_j(\boldsymbol{\theta}_j, \mathbf{z})$, which in Eq. (1) is the location parameter of $Y_j = \log X_j$. Now it turns out that quite a few such standard engineering choices of STF, especially in the context of multiple stress variables, can be expressed as in Eq. (2), and are exemplified with further details in Table 2.

It needs to be remarked that as Eq. (1) allows one to assume different life distributions, Eq. (2) lets one choose different STFs for different components. This feature of the model is particularly important and attractive for the problem of multi-stress ALT of multi-component systems. This is because sometimes certain particular life distributions get associated with a specific STF for a common underlying physical or engineering reasoning of failures. For example, suppose a mechanical component of a system fails due to the “weakest link” theory leading to a Weibull, and an electrical component of the same system fails due to the “multiplicative degradation” theory leading to a log-normal, as their respective life distributions. Then with temperature as the common stress variable, while the Eyring interaction model with humidity as the second stress variable is a reasonable physical choice of STF for the Weibull component, the electromigration STF with current density as the third stress variable becomes the most natural physical choice for the log-normal component (since the same ionic movement phenomenon leads to electromigration as well as “multiplica-

¹ <http://www.itl.nist.gov/div898/handbook/apr/section1/apr153.htm>.

Table 2 Different linear stress translation functions

Model	Stress	$g_{kj}(\cdot)$	Stress translation function
Arrhenius model	z_1	$g_{1j}(\cdot) = 1, g_{2j}(z_1) = 1/z_1$	$\mu_j(\theta_j, z_1) = \theta_{1j} + \theta_{2j}/z_1$
	z_2	$g_{1j}(\cdot) = 1, g_{2j}(z_2) = \log z_2$	$\mu_j(\theta_j, z_2) = \theta_{1j} + \theta_{2j} \log z_2$
Log-linear model	z_3	$g_{1j}(\cdot) = 1, g_{2j}(z_3) = z_3$	$\mu_j(\theta_j, z_3) = \theta_{1j} + \theta_{2j} z_3$
	z_1, z_4	$g_{1j}(\cdot) = 1, g_{2j}(z_1) = 1/z_1,$ $g_{3j}(z_4) = \log z_4$	$\mu_j(\theta_j, z_1, z_4) = \theta_{1j} + \theta_{2j}/z_1 + \theta_{2j} \log z_4$
Three-Stress model	z_1, z_2, z_5	$g_{1j}(\cdot) = 1, g_{2j}(z_1) = 1/z_1,$ $g_{3j}(z_2) = \log z_2,$ $g_{4j}(z_5) = \log z_5$	$\mu_j(\theta_j, z_1, z_2, z_5) = \theta_{1j} + \theta_{2j}/z_1 + \theta_{3j} \log z_2 + \theta_{4j} \log z_5$
	z_6, z_7, z_8	$g_{1j}(\cdot) = 1, g_{2j}(z_6) = \log z_6,$ $g_{3j}(z_7) = \log z_7,$ $g_{4j}(z_8) = 1/z_8$	$\mu_j(\theta_j, z_6, z_7, z_8) = \theta_{1j} + \theta_{2j} \log z_6 + \theta_{3j} \log z_7 + \theta_{4j}/z_8$
	z_1, z_5	$g_{1j}(\cdot) = 1, g_{2j}(z_1) = (1/z_1),$ $g_{3j}(z_5) = \log z_5, g_{4j}(z_9) = z_9,$ where $z_9 = g_{2j}(z_1) * g_{3j}(z_5)$	$\mu_j(\theta_j, z_1, z_5) = \theta_{1j} + \theta_{2j}/z_1 + \theta_{3j} \log z_5 + \theta_{4j} \frac{\log z_5}{z_1}$

Z_1 : temperature (in Kelvin), Z_2 : voltage, Z_3 : mechanical stress, Z_4 : current density, Z_5 : humidity, Z_6 : cycling frequency, Z_7 : temperature range during a cycle (in Kelvin), Z_8 : maximum temperature in each cycle (in Kelvin)

tive degradation” theory of failures in ICs). Such situations are easily accommodated by the proposed model.

Now consider modeling the lifetime of such J -component series systems. Since a series system fails as soon as one of its constituent J components fails, the system log-lifetime T at stress $\mathbf{Z} = \mathbf{z}$ is given by $\{T|\mathbf{Z} = \mathbf{z}\} = \text{Min}\{Y_1|\mathbf{Z} = \mathbf{z}, \dots, Y_J|\mathbf{Z} = \mathbf{z}\}$. Now since the J components are also assumed to be independent, the system s.f. $\overline{F}_T(t|\boldsymbol{\theta}, \boldsymbol{\tau}, \mathbf{z})$ is given by

$$\overline{F}_T(t|\boldsymbol{\theta}, \boldsymbol{\tau}, \mathbf{z}) = \prod_{j=1}^J \overline{F}_j(\tau_j(t - \mathbf{s}_j'\boldsymbol{\theta}_j)), \tag{3}$$

where $\boldsymbol{\theta} = (\boldsymbol{\theta}'_1, \dots, \boldsymbol{\theta}'_J)'$ and $\boldsymbol{\tau} = (\tau_1, \dots, \tau_J)'$. Let the discrete random variable I , taking values in $\{1, \dots, J\}$, denote the component causing system failure. Then, clearly $I = \arg \min_j \{Y_j|\mathbf{Z} = \mathbf{z}\}$. For a series system that has failed after being put on an ALT, one typically observes (T, I) . The joint p.d.f. of (T, I) is given by

$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} P(t \leq T \leq t + \Delta t, I = j|\boldsymbol{\theta}, \boldsymbol{\tau}, \mathbf{z}) = p_j(t|\boldsymbol{\theta}, \boldsymbol{\tau}, \mathbf{z}), \text{ say,}$$

where

$$p_j(t|\boldsymbol{\theta}, \boldsymbol{\tau}, \mathbf{z}) = \tau_j h_j(\tau_j(t - \mathbf{s}_j'\boldsymbol{\theta}_j)) \prod_{j=1}^J \overline{F}_j(\tau_j(t - \mathbf{s}_j'\boldsymbol{\theta}_j)). \tag{4}$$

The complete set of model parameters is then given by $\boldsymbol{\psi} = (\boldsymbol{\theta}', \boldsymbol{\tau}')$. Also, let $\boldsymbol{\psi}_j = (\boldsymbol{\theta}'_j, \tau_j)'$.

3 Bayesian analysis

3.1 Likelihood function

Suppose N J -component series systems as in Sect. 2 are put on an ALT at different levels of the K stress variables. For $i = 1, \dots, N$, let $\mathbf{z}^i = (z^i_1, \dots, z^i_K)'$ denote the value of the stress vector \mathbf{Z}^i applied to the i th system. Let the number of systems that fail in this ALT be denoted by n . For $i = 1, \dots, n$, let T^i and I^i respectively denote the log-failure time and cause of failure of i th such system. For the remaining $N - n$ systems that do not fail in the ALT, denote the future log-failure time of the r th system by T^{n+r} , for $r = 1, \dots, N - n$. $(T^1, I^1), \dots, (T^n, I^n), T^{n+1}, \dots, T^N$ are assumed to be mutually independent of each other. For $i = 1, \dots, n$, let t^i denote the observed log-failure time of the i th failed system, and for $r = 1, \dots, N - n$, let t^{n+r} denote the observed log-censoring time of the r th running system. Thus the observed data is given by

$$\mathcal{D} = \left\{ (T^1 = t^1, I^1, \mathbf{Z}^1 = \mathbf{z}^1), \dots, (T^n = t^n, I^n, \mathbf{Z}^n = \mathbf{z}^n), \right. \\ \left. (T^{n+1} > t^{n+1}, \mathbf{Z}^{n+1} = \mathbf{z}^{n+1}), \dots, (T^N > t^N, \mathbf{Z}^N = \mathbf{z}^N) \right\}. \tag{5}$$

Now by using Eqs. (3) and (4), the likelihood function of ψ given the observed data \mathcal{D} , is given by

$$L(\psi|\mathcal{D}) = \left[\prod_{i=1}^n \left\{ \frac{\tau_{I^i} f_{I^i}(\tau_{I^i}(t^i - \mathbf{s}_{I^i}'\theta_{I^i}))}{\bar{F}_{I^i}(\tau_{I^i}(t^i - \mathbf{s}_{I^i}'\theta_{I^i}))} \right\} \right] \times \left[\prod_{i=1}^N \prod_{j=1}^J \bar{F}_j(\tau_j(t^i - \mathbf{s}_j'\theta_j)) \right], \tag{6}$$

where $\mathbf{s}_j^i = (s_{1j}^i, \dots, s_{Kj}^i)'$ with $s_{kj}^i = g_{kj}(z_k^i)$. Now it is easy to see that

$$L(\psi|\mathcal{D}) = \prod_{j=1}^J L_j(\psi_j|\mathcal{D}) \tag{7}$$

where

$$L_j(\psi_j|\mathcal{D}) = \tau_j^{n_j} \times \prod_{\substack{1 \leq i \leq n \\ \& I^i = j}} f_j(\tau_j(t^i - \mathbf{s}_j'\theta_j)) \times \prod_{\substack{1 \leq i \leq n \\ \& I^i \neq j}} \bar{F}_j(\tau_j(t^i - \mathbf{s}_j'\theta_j)) \times \prod_{i=n+1}^N \bar{F}_j(\tau_j(t^i - \mathbf{s}_j'\theta_j)), \tag{8}$$

and $n_j = \#\{1 \leq i \leq n : I^i = j\}$ denotes the number of failures due to cause j so that $\sum_{j=1}^J n_j = n$.

Bayesian inference is based on $\pi(\psi|\mathcal{D})$, the posterior p.d.f. of ψ given \mathcal{D} , which is proportional to the product of $L(\psi|\mathcal{D})$, the likelihood function of ψ given in Eq. (6), and $\pi(\psi)$, the joint prior of ψ .

3.2 Prior distribution

Equation (7) reveals a natural independence structure among the ψ_j s which is respected in the prior distribution as well, that is, ψ_1, \dots, ψ_J are assumed to be mutually independent a priori, resulting in the joint prior p.d.f. of ψ as

$$\pi(\psi) = \prod_{j=1}^J \pi(\psi_j), \tag{9}$$

where $\pi(\psi_j)$ is the joint prior p.d.f. of $\psi_j = (\theta_j, \tau_j)'$. Furthermore, it is also assumed a priori that $\theta_{1j}, \dots, \theta_{Kj}$ and τ_j are all mutually independent. Thus

$$\pi(\boldsymbol{\psi}_j) = \prod_{j=1}^J \left\{ \prod_{k=1}^K \pi(\theta_{kj}) \right\} \pi(\tau_j) = \pi(\boldsymbol{\theta}_j) \pi(\tau_j), \tag{10}$$

where $\pi(\boldsymbol{\theta}_j) = \prod_{k=1}^K \pi(\theta_{kj})$ and $\pi(\theta_{kj})$, $\pi(\tau_j)$ and $\pi(\boldsymbol{\theta}_j)$ are the prior p.d.fs of θ_{kj} , τ_j and $\boldsymbol{\theta}_j$ respectively. In principle Bayesian computation only needs the ability of generating observations from $\pi(\boldsymbol{\psi}_j|\mathcal{D})$. Towards that end as required in Theorem 1 below, it is further assumed that $\pi(\theta_{kj})$ and $\pi(\tau_j)$ are log-concave functions of their respective arguments.

Under such circumstances, we propose the following simple method of choosing subjective proper priors for θ_{kjs} and τ_js . First note that as defined in Sect. 2, θ_{kjs} are unrestricted real numbers. Thus any log-concave density such as the $f_j(\cdot)$ s listed in Table 1 with the entire real line as its support, may be chosen as a proper prior p.d.f. for θ_{kj} . In case of $f_j(\cdot)$ s of Table 1 (or even otherwise) as priors, their location and scale hyper-parameters may be chosen by specifying a guess-interval and one’s subjective probability of that guess-interval containing the unknown parameter value. The choice of the particular $f_j(\cdot)$ among the ones listed in Table 1 may now be used to better model one’s subjective belief, such as an additional subjective probability consistency check, both inside but more importantly outside the guess-interval. Since $\tau_j > 0$, one may choose any log-concave density with $(0, \infty)$ as its support (such as gamma or Weibull with shape parameter ≥ 1) as a proper prior for τ_j .

Along with the discussion of subjective proper priors, in Bayesian analysis it is also very important to address the issue of non-informative priors (see Berger 2006). For the location-scale regression model of Eqs. (1) and (2), a common non-informative prior is given by (see Fernández and Steel 1999, for example)

$$\pi(\boldsymbol{\theta}_j, \tau_j) \propto \frac{1}{\tau_j}. \tag{11}$$

Note that the above non-informative prior is improper and thus as in any other situation, care must be exercised ensuring the propriety of the resulting joint posterior before employing such improper priors. For the problem at hand, if for instance, the X_js are Weibull or equivalently Y_js are extreme-value, the posterior resulting from the prior in (11) may be improper (see Kim and Ibrahim 2000; Roy and Mukhopadhyay 2014).

3.3 Posterior analysis

As mentioned briefly at the end of Sect. 3.1, the joint posterior $\pi(\boldsymbol{\psi}|\mathcal{D})$ is proportional to the product of $L(\boldsymbol{\psi}|\mathcal{D})$ given in Eq. (6), and $\pi(\boldsymbol{\psi})$ given in Eq. (9). But, by Eqs. (7) and (10), one gets

$$\pi(\boldsymbol{\psi}|\mathcal{D}) \propto \prod_{j=1}^J L_j(\boldsymbol{\psi}_j|\mathcal{D})\pi(\boldsymbol{\theta}_j)\pi(\tau_j) \propto \prod_{j=1}^J \pi(\boldsymbol{\psi}_j|\mathcal{D}), \tag{12}$$

where $\pi(\boldsymbol{\psi}_j|\mathcal{D}) \propto L_j(\boldsymbol{\psi}_j|\mathcal{D})\pi(\boldsymbol{\theta}_j)\pi(\tau_j)$ is the joint posterior p.d.f. of $\boldsymbol{\psi}_j$ given the observed data \mathcal{D} . Thus posterior analysis of $\boldsymbol{\psi}$ can be performed by independently drawing samples from $\pi(\boldsymbol{\psi}_j|\mathcal{D})$ for $j = 1, \dots, J$.

Now as usual, one can Gibbs sample from the joint posterior $\pi(\boldsymbol{\psi}_j|\mathcal{D})$, if it is easy to generate observations from the univariate conditional posterior of each component of $\boldsymbol{\psi}_j$, given the rest. The following theorem states why that is the case in this problem.

Theorem 1 *If $f_j(\cdot)$ [in Eq. (1)], $\pi(\theta_{1j}), \dots, \pi(\theta_{Kj})$ and $\pi(\tau_j)$ [in Eq. (10)] are log-concave, then so are the univariate conditional posteriors of θ_{kj} s and τ_j , given the rest, in Eq. (12).*

The proof of the theorem is provided in Appendix 1. Using this result one can now employ any efficient algorithm (such as the adaptive rejection sampling technique of Gilks and Wild (1992)) to iteratively draw samples from these univariate log-concave conditional posteriors, to obtain a sample $\{\boldsymbol{\psi}_j^{(l)}\}_{l=1}^L$ from $\pi(\boldsymbol{\psi}_j|\mathcal{D})$ in the limit.

The main objective of performing an ALT is to infer on component and system reliability metrics at usage stress $\mathbf{Z} = \mathbf{z}_u = (z_{u1}, \dots, z_{uK})'$, such as MTTF, component and system s.f.s, etc. Given $\mathbf{Z} = \mathbf{z}_u$, all these are functions of $\boldsymbol{\psi}$ and some are also of t . Let us generically denote any one of these functions by $\Delta(\boldsymbol{\psi}, t)$, for brevity.

Given a sample $\{\boldsymbol{\psi}^{(l)} = \{\boldsymbol{\psi}_1^{(l)}, \dots, \boldsymbol{\psi}_J^{(l)}\}\}_{l=1}^L$ from the joint posterior p.d.f. $\pi(\boldsymbol{\psi}|\mathcal{D})$, the problem is to estimate $\Delta(\boldsymbol{\psi}, t)$ over a sufficiently large time interval $[0, t_{\max}]$. Towards this goal, the interval $[0, t_{\max}]$ is first partitioned into a grid of $(P + 1)$ equispaced points given by $\{0 = t_0, t_1, \dots, t_P = t_{\max}\}$. Next, $\forall l = 1, \dots, L$ and $p = 0, \dots, P$, one computes $\Delta(\boldsymbol{\psi}^{(l)}, t_p)$ denoted by $\Delta^{(l)}(t_p)$, yielding $\{(\Delta^{(1)}(t_0), \dots, \Delta^{(1)}(t_P)), \dots, (\Delta^{(L)}(t_0), \dots, \Delta^{(L)}(t_P))\}$ as a sample of size L drawn from the joint posterior of $(\Delta(\boldsymbol{\psi}, t_0), \dots, \Delta(\boldsymbol{\psi}, t_P))$. This sample may be now used for computing the usual Bayesian point and interval estimates of $\Delta(\boldsymbol{\psi}, t)$, under J fixed models $f_1(\cdot), \dots, f_J(\cdot)$ for the J component lifetimes.

4 Log-normal model

The special case of an X_j having a log-normal distribution, or equivalently the corresponding Y_j having a normal distribution deserves particular attention for several reasons. First, there are strong theoretical and empirical evidences for occurrence of log-normal as a life distribution in many physical failure mechanisms, such as, for example, the electromigration phenomenon mentioned in Sect. 2 (see Hauschildt et al. 2007 for a more recent account). Even otherwise also from a purely statistical perspective, it is usually required to be tested against the other popular alternative of Weibull (see Kundu and Manglick 2004; Kim and Yum 2008, for example), which require Bayesian analyses of both.

As mentioned in the Introduction, the analyses of the Weibull model through log-concave sampling as in Sect. 3.3, are available in Roy and Mukhopadhyay (2014) and Fan and Hsu (2014) (though note that the former's model is slightly different from the one being considered here owing to the equality of the τ_j s). Obviously as a special

case of Eq. (1), the same can also be carried out for the log-normal distribution as in Sect. 3.3. However in this case, we propose a simple data-augmentation technique and a conjugate prior [which unlike Eq. (10) allows a priori correlation among the θ_{kj} s and τ_j], to circumvent the computationally expensive log-concave sampling.

4.1 Data augmentation

In the rest of this section, j is assumed to be fixed and consideration is given to the special case of Y_j , a given j th component log-lifetime, having a normal distribution with mean $\mu_j(\boldsymbol{\theta}_j, \mathbf{z})$ and variance σ_j^2 , given $\mathbf{Z} = \mathbf{z}$. Based on the observed data \mathcal{D} given in Eq. (5) with $\tau_j = 1/\sigma_j$ (see Table 1) in Eq. (8), it then follows that

$$\begin{aligned}
 L_j(\boldsymbol{\psi}_j|\mathcal{D}) \propto & \sigma_j^{-nj} \times \prod_{\substack{1 \leq i \leq n \\ \& I^i=j}} \phi\left(\frac{t^i - \mathbf{s}_j^{i'}\boldsymbol{\theta}_j}{\sigma_j}\right) \times \prod_{\substack{1 \leq i \leq n \\ \& I^i \neq j}} \bar{\Phi}\left(\frac{t^i - \mathbf{s}_j^{i'}\boldsymbol{\theta}_j}{\sigma_j}\right) \\
 & \times \prod_{i=n+1}^N \bar{\Phi}\left(\frac{t^i - \mathbf{s}_j^{i'}\boldsymbol{\theta}_j}{\sigma_j}\right), \tag{13}
 \end{aligned}$$

where $\boldsymbol{\psi}_j = (\boldsymbol{\theta}_j, \sigma_j^2)'$ and $\phi(\cdot)$ and $\bar{\Phi}(\cdot)$ are standard normal p.d.f. and s.f. respectively. In this case, the need for log-concave sampling arises due to the presence of the $\bar{\Phi}(\cdot)$ terms in (13), which might be thought ensuing due to incomplete information. Such situations are typically remedied by invoking a standard Bayesian tool for handling missing data called data augmentation (see Tanner and Wong 1987). Note that if needed, computations may be further accelerated using refinements such as parameter expanded data augmentation (PX-DA) algorithm (see Liu and Wu 1999, for example).

For $i = 1, \dots, N$, let Y_j^i denote the log-lifetime of the j th component of the i th system put on the ALT. Then for $i = 1, \dots, n$, if $I^i = j$, one has $Y_j^i = t^i$ as part of \mathcal{D} , while if $I^i \neq j$, its associated t^i is its survival time from cause j , and thus its Y_j^i , which is latent, is imputed subject to the constraint $Y_j^i > t^i$. Similarly, for $r = 1, \dots, N - n$, all the censored Y_j^{n+r} values are imputed subject to the constraint $Y_j^{n+r} > t^{n+r}$. For $i = 1, \dots, N$, let y_j^i denote this observed or imputed value of Y_j^i and $\mathbf{y}_j = (y_j^1, \dots, y_j^N)'$. That is, the observed data set \mathcal{D} is proposed to be augmented with an auxiliary part-latent data set \mathcal{Y}_j to form the complete data set $\mathcal{Y} = \{\mathcal{D}, \mathcal{Y}_j\}$, where $\mathcal{Y}_j = \{Y_j^1 = y_j^1, \dots, Y_j^N = y_j^N\}$. It should be mentioned that this data augmentation scheme is similar to that in Mukhopadhyay and Basu (2007).

This data augmentation is done to change the problem of sample generation from $\pi(\boldsymbol{\psi}_j|\mathcal{D})$ to that from $\pi(\mathbf{y}_j, \boldsymbol{\psi}_j|\mathcal{D})$, since the $\boldsymbol{\psi}_j$ values generated from $\pi(\mathbf{y}_j, \boldsymbol{\psi}_j|\mathcal{D})$ can also be regarded as a sample from $\pi(\boldsymbol{\psi}_j|\mathcal{D})$. The later task is much easier as a Gibbs sample from $\pi(\mathbf{y}_j, \boldsymbol{\psi}_j|\mathcal{D})$ can be generated by iteratively sampling from $\pi(\mathbf{y}_j|\boldsymbol{\psi}_j, \mathcal{D})$ and $\pi(\boldsymbol{\psi}_j|\mathbf{y}_j, \mathcal{D}) = \pi(\boldsymbol{\psi}_j|\mathcal{D})$ as follows.

4.2 Gibbs sampling

For $i = 1, \dots, N$, since Y_j^i s are mutually independent,

$$\pi(\mathbf{y}_j | \boldsymbol{\psi}_j, \mathcal{D}) = \prod_{i=1}^N \pi(y_j^i | \boldsymbol{\psi}_j, \mathcal{D}),$$

where for $i = 1, \dots, n$, if $I^i = j$, $\pi(y_j^i | \boldsymbol{\psi}_j, \mathcal{D})$ is degenerate at t^i and if $I^i \neq j$ and $1 \leq i \leq n$, or if $n < i \leq N$,

$$\pi(y_j^i | \boldsymbol{\psi}_j, \mathcal{D}) = \begin{cases} \frac{\frac{1}{\sigma_j} \phi\left(\frac{y_j^i - s_j^i \theta_j}{\sigma_j}\right)}{\Phi\left(\frac{t^i - s_j^i \theta_j}{\sigma_j}\right)}, & \text{if } y_j^i > t^i \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

Thus for $i = 1, \dots, n$, if $I^i = j$, fix $y_j^i = t^i$, otherwise for all other $1 \leq i \leq N$, one needs to generate an y_j^i according to the p.d.f. given in Eq. (14), which can be done in at least two different ways. The first is a straight-forward rejection method in which one keeps generating observations from $N(s_j^i \theta_j, \sigma_j^2)$ until it is greater than t^i . The second is a univariate inversion method in which one numerically inverts the integral of $\pi(y_j^i | \boldsymbol{\psi}_j, \mathcal{D})$ given in Eq. (14). The later method is numerically more efficient for large t^i .

This generated \mathbf{y}_j is augmented to \mathcal{D} to form the complete data \mathcal{Y} , so that samples may be drawn from $\pi(\boldsymbol{\psi}_j | \mathcal{Y})$, which is given by

$$\pi(\boldsymbol{\psi}_j | \mathcal{Y}) \propto L_j(\boldsymbol{\psi}_j | \mathcal{Y}) \pi(\boldsymbol{\psi}_j),$$

where $L_j(\boldsymbol{\psi}_j | \mathcal{Y})$ is the likelihood function of $\boldsymbol{\psi}_j$ based on the complete data \mathcal{Y} and $\pi(\boldsymbol{\psi}_j)$ is the prior p.d.f. of $\boldsymbol{\psi}_j$. By Eq. (1),

$$\begin{aligned} L_j(\boldsymbol{\psi}_j | \mathcal{Y}) &= \prod_{i=1}^N \frac{1}{\sigma_j} \phi\left(\frac{y_j^i - s_j^i \theta_j}{\sigma_j}\right) \\ &\propto \sigma_j^{-N} \exp\left\{-\frac{1}{2\sigma_j^2}(\mathbf{y}_j - \mathbf{S}_j \boldsymbol{\theta}_j)'(\mathbf{y}_j - \mathbf{S}_j \boldsymbol{\theta}_j)\right\}, \end{aligned} \tag{15}$$

where $\mathbf{S}_j = [s_j^1, \dots, s_j^N]'$ is the $N \times K$ stress matrix corresponding to component j .

Now the prior $\pi(\boldsymbol{\psi}_j)$ is specified as follows. It is first assumed that $\sigma_j^2 \sim$ Inverse Gamma (α_j, β_j) a priori, i.e.,

$$\pi(\sigma_j^2) \propto (\sigma_j^2)^{-\alpha_j-1} \exp\left(-\frac{\beta_j}{\sigma_j^2}\right), \quad \sigma_j^2 > 0.$$

Then it is assumed that $\boldsymbol{\theta}_j|\sigma_j^2 \sim N_K(\boldsymbol{\zeta}_j, \sigma_j^2 \mathbf{P}_j^{-1})$ a priori, i.e.,

$$\pi(\boldsymbol{\theta}_j|\sigma_j^2) \propto (\sigma_j^2)^{-\frac{K}{2}} \exp\left\{-\frac{1}{2\sigma_j^2} [(\boldsymbol{\theta}_j - \boldsymbol{\zeta}_j)' \mathbf{P}_j (\boldsymbol{\theta}_j - \boldsymbol{\zeta}_j)]\right\}.$$

Thus it is assumed that the joint prior p.d.f. of $\boldsymbol{\psi}_j$ is given by

$$\pi(\boldsymbol{\psi}_j) \propto (\sigma_j^2)^{-\alpha_j - \frac{K}{2} - 1} \exp\left\{-\frac{1}{2\sigma_j^2} [2\beta_j + (\boldsymbol{\theta}_j - \boldsymbol{\zeta}_j)' \mathbf{P}_j (\boldsymbol{\theta}_j - \boldsymbol{\zeta}_j)]\right\}. \tag{16}$$

Apart from conjugacy with (15), (16) also allows a user to reasonably specify his or her prior opinion about the unknown model parameters (see Gelman et al. 2013, pp. 376–378). The prior hyper-parameters may be chosen as outlined in Sect. 3.2 with the restriction of normal distribution, but here one also has an opportunity to model a priori dependence among the model parameters. With $\tau_j = (\sigma_j^2)^{-1/2}$, the non-informative prior (11) in this case [also as is standard with the likelihood in Eq. (15), see Gelman et al. (2013, p. 355)] becomes

$$\pi(\boldsymbol{\psi}_j) \propto \frac{1}{\sigma_j^2}, \tag{17}$$

which can be easily modeled by choosing $\alpha_j = -K/2$, $\beta_j = 0$ and $\mathbf{P}_j = 0$ in (16). Note that unlike the extreme-value case, here the posterior of $\boldsymbol{\psi}_j$ given \mathcal{D} with the non-informative prior in Eq. (17) is proper (i.e., $\int \pi(\boldsymbol{\psi}_j|\mathcal{D})d\boldsymbol{\psi}_j < \infty$) provided (i) $n_j > K$ and (ii) the rank of \mathbf{S}_j is K [this is because the $\overline{\Phi}(\cdot)$ terms in Eq. (13) are bounded by 1 and the rest follows as in Gelman et al. (2013, p. 356)].

Now by Eqs. (15) and (16),

$$\pi(\boldsymbol{\psi}_j|\mathcal{D}) \propto (\sigma_j^2)^{-\frac{2\alpha_j+N+K}{2}-1} \exp\left\{-\frac{1}{2\sigma_j^2} [(\boldsymbol{\theta}_j - \boldsymbol{\eta}_j)' \mathbf{V}_j (\boldsymbol{\theta}_j - \boldsymbol{\eta}_j) + C_j(\mathbf{y}_j)]\right\},$$

where $\mathbf{V}_j = (\mathbf{S}_j' \mathbf{S}_j + \mathbf{P}_j)$, $\boldsymbol{\eta}_j = \mathbf{V}_j^{-1} (\mathbf{S}_j' \mathbf{y}_j + \mathbf{P}_j \boldsymbol{\zeta}_j)$ and $C_j(\mathbf{y}_j) = 2\beta_j + \mathbf{y}_j' \mathbf{y}_j + \boldsymbol{\zeta}_j' \mathbf{P}_j \boldsymbol{\zeta}_j - \boldsymbol{\eta}_j' \mathbf{V}_j \boldsymbol{\eta}_j$. Thus we have $\boldsymbol{\theta}_j|\sigma_j^2, \mathcal{D} \sim N_K(\boldsymbol{\eta}_j, \sigma_j^2 \mathbf{V}_j^{-1})$ and $\sigma_j^2|\mathcal{D} \sim$ Inverse Gamma $(\frac{2\alpha_j+N}{2}, \frac{1}{2} C_j(\mathbf{y}_j))$. In order to generate a $\boldsymbol{\psi}_j$ from $\pi(\boldsymbol{\psi}_j|\mathcal{D})$, first generate a σ_j^2 from the inverse gamma distribution with parameters $(2\alpha_j + N)/2$ and $C_j(\mathbf{y}_j)/2$;

then with that σ_j^2 , generate a θ_j from the K -variate normal distribution with mean η_j and variance–covariance matrix $\sigma_j^2 V_j^{-1}$.

Implementation of Gibbs sampling involves iterative sampling from $\pi(\mathbf{y}_j | \boldsymbol{\psi}_j, \mathcal{D})$ and $\pi(\boldsymbol{\psi}_j | \mathcal{D})$. Starting with an initial value $\boldsymbol{\psi}_j^{(0)}$, for $l \geq 0$, one iteratively draws (i) $\mathbf{y}_j^{(l+1)}$ from $\pi(\mathbf{y}_j | \boldsymbol{\psi}_j^{(l)}, \mathcal{D})$ and (ii) $\boldsymbol{\psi}_j^{(l+1)}$ from $\pi(\boldsymbol{\psi}_j | \mathcal{D}^{(l+1)})$, where $\mathcal{D}^{(l+1)} = \{\mathbf{y}_j^{(l+1)}, \mathcal{D}\}$. This yields a Markov chain $\{(\mathbf{y}_j^{(l)}, \boldsymbol{\psi}_j^{(l)}), l \geq 1\}$, which produces an observation from $\pi(\mathbf{y}_j, \boldsymbol{\psi}_j | \mathcal{D})$ in the limit. Thus generated $\{\boldsymbol{\psi}_j^{(l)}\}_{l=1}^L$ can now be used for further Bayesian analyses as detailed in Sect. 3.3.

Since from this point on the discussion is in numerical lines, at the outset it is worth commenting on one detail regarding the collection of (Gibbs) sample from $\pi(\boldsymbol{\psi} | \mathcal{D})$ or $\pi(M | \mathcal{D})$ (where M is a model variable). In the numerical implementations we employ a long chain strategy (of typically 1,001,000 elements) with a burn-in (of typical size 1000) and a thinning interval (of typical size 200) resulting in a Gibbs sample (of size 5000). The exact values of these parameters are determined following the MCMC convergence diagnostic tests reviewed by Cowles and Carlin (1996), which are readily available in the R-library CODA.

4.3 Simulation study

This sub-section presents a simulation study examining the performance of different Bayesian point estimates of lifetime parameters of the independent log-normal model. The basic set up of the simulation study, such as the stress variable, its levels, the number of systems allocated at each level of the stress variable in the ALT, the STF, etc. is borrowed from the electrical insulation system example of Sect. 1.

For convenience, it is assumed that the series system has two failure modes. Now suppose ($N=$) 40 such two-component series systems are put on an ALT at four elevated temperature levels of 423.15, 443.15, 463.15 and 493.15 K with equal number of systems allocated at each level of temperature. It is assumed that none of the ALT observations is censored, implying that the exact failure times are available for all the systems. As in the insulation system example (see Nelson 1990, p. 394), the following Arrhenius model is used as the STF for both the failure modes:

$$\mu_j(\theta_{1j}, \theta_{2j}, T_K) = \theta_{1j} + \theta_{2j} \left(\frac{1000}{T_K} \right), \quad (18)$$

where T_K is temperature measured in Kelvin (K). The simulation study is carried out with two different sets of parameter values namely

- Set 1: $\theta_{11} = -24.00$, $\theta_{21} = 11.00$, $\theta_{12} = -22.00$, $\theta_{22} = 10.00$, $\sigma_1 = 1.00$ and $\sigma_2 = 1.00$.
- Set 2: $\theta_{11} = -20.00$, $\theta_{21} = 9.00$, $\theta_{12} = -18.00$, $\theta_{22} = 8.00$, $\sigma_1 = 1.00$ and $\sigma_2 = 1.00$.

Bayesian analyses are performed using six different Inverse-Gamma–Normal priors of Eq. (16) with their means, variances and correlations as listed in Table 3. The

Table 3 Prior parameters for simulation study ($j = 1, 2$)

Set	Prior	$E(\theta_{11})$	$E(\theta_{21})$	$E(\theta_{12})$	$E(\theta_{22})$	$E(\sigma_j^2)$	Corr(θ_{1j}, θ_{2j})
		($V(\theta_{11})$)	($V(\theta_{21})$)	($V(\theta_{12})$)	($V(\theta_{22})$)	($V(\sigma_j^2)$)	
1	1	Non-informative prior given in Eq. (17)					
	2	-24.0 (4.0)	11.0 (2.0)	-22.0 (4.0)	10.0 (2.0)	1.0 (2.0)	0.0
	3	-24.0 (4.0)	11.0 (2.0)	-22.0 (4.0)	10.0 (2.0)	1.0 (2.0)	0.7
	4	-24.0 (9.0)	11.0 (4.0)	-22.0 (9.0)	10.0 (4.0)	1.0 (3.0)	0.7
	5	-28.8 (9.0)	13.2 (4.0)	-26.4 (9.0)	12.0 (4.0)	1.4 (3.0)	0.7
	6	-19.2 (9.0)	8.8 (4.0)	-17.6 (9.0)	8.0 (4.0)	0.6 (3.0)	0.7
2	1	Non-informative prior given in Eq. (17)					
	2	-20.0 (4.0)	9.0 (2.0)	-18.0 (4.0)	8.0 (2.0)	1.0 (2.0)	0.0
	3	-20.0 (4.0)	9.0 (2.0)	-18.0 (4.0)	8.0 (2.0)	1.0 (2.0)	0.7
	4	-20.0 (9.0)	9.0 (4.0)	-18.0 (9.0)	8.0 (4.0)	1.0 (3.0)	0.7
	5	-24.0 (9.0)	10.8 (4.0)	-21.6 (9.0)	9.6 (4.0)	1.4 (3.0)	0.7
	6	-16.0 (9.0)	7.2 (4.0)	-14.4 (9.0)	6.4 (4.0)	0.6 (3.0)	0.7

Table 4 Average Bayesian estimates (MSEs) of θ_{21}

Set	Prior	Avg mean (MSE)	Avg median (MSE)	Avg mode (MSE)	LINEX	
					$a = -1$	$a = 1$
TV = 11	1	10.99 (2.62)	10.98 (2.52)	10.92 (2.34)	13.11 (10.58)	9.01 (10.19)
	2	11.01 (0.09)	11.00 (0.09)	10.99 (0.10)	11.29 (0.20)	10.74 (0.18)
	3	11.01 (0.02)	11.00 (0.02)	11.00 (0.02)	11.11 (0.03)	10.91 (0.02)
	4	11.01 (0.05)	11.00 (0.05)	10.99 (0.05)	11.21 (0.10)	10.81 (0.09)
	5	12.99 (3.98)	12.98 (3.95)	12.98 (3.95)	13.15 (4.67)	12.83 (3.38)
	6	9.26 (3.12)	9.26 (3.14)	9.25 (3.17)	9.56 (2.21)	8.97 (4.21)
TV = 9	1	8.99 (2.74)	8.97 (2.65)	8.94 (2.56)	10.99 (10.71)	7.31 (6.67)
	2	9.00 (0.10)	8.99 (0.09)	8.99 (0.10)	9.26 (0.18)	8.75 (0.17)
	3	9.00 (0.01)	9.00 (0.01)	9.00 (0.02)	9.10 (0.03)	8.91 (0.02)
	4	9.00 (0.05)	8.99 (0.05)	8.99 (0.05)	9.19 (0.10)	8.82 (0.08)
	5	10.64 (2.71)	10.63 (2.69)	10.63 (2.67)	10.79 (3.23)	10.49 (2.26)
	6	7.56 (2.18)	7.56 (2.19)	7.55 (2.21)	7.84 (1.49)	7.30 (3.01)

prior hyper-parameters are chosen to have the prior means fixed at the respective true values (TVs) of the parameters for priors 2, 3 and 4 (with varying degree of prior variance), and misspecified by 20% for priors 5 and 6. For the sake of brevity, here the simulation results are discussed only for the slope parameters θ_{2j} s and the scale parameters σ_j s. The findings for the intercept parameters θ_{1j} s are similar to that of the slope parameters.

The simulation is carried out for 200 times. Tables 4, 5, 6 and 7 respectively report the average posterior mean (Avg mean), median (Avg median) and mode (Avg mode)

Table 5 Average Bayesian estimates (MSEs) of θ_{22}

Set	Prior	Avg mean (MSE)	Avg median (MSE)	Avg mode (MSE)	LINEX	
					$a = -1$	$a = 1$
1 TV = 10	1	9.89 (1.77)	9.89 (1.74)	9.90 (1.71)	11.22 (3.65)	8.50 (4.69)
	2	9.98 (0.08)	9.98 (0.08)	9.99 (0.08)	10.23 (0.14)	9.74 (0.15)
	3	9.99 (0.01)	9.99 (0.01)	9.99 (0.01)	10.08 (0.02)	9.90 (0.02)
	4	9.99 (0.04)	9.99 (0.04)	9.99 (0.04)	10.16 (0.08)	9.81 (0.08)
	5	11.76 (3.12)	11.76 (3.11)	11.75 (3.10)	11.90 (3.63)	11.62 (2.65)
	6	8.44 (2.54)	8.43 (2.54)	8.43 (2.57)	8.70 (1.80)	8.17 (3.43)
2 TV = 8	1	7.88 (2.58)	7.89 (2.51)	7.91 (2.39)	9.28 (4.86)	6.31 (6.77)
	2	7.99 (0.11)	7.99 (0.11)	8.00 (0.11)	8.25 (0.19)	7.74 (0.18)
	3	8.00 (0.02)	8.00 (0.02)	8.00 (0.02)	8.10 (0.03)	7.91 (0.02)
	4	8.00 (0.06)	8.00 (0.06)	8.00 (0.06)	8.18 (0.10)	7.81 (0.09)
	5	9.43 (2.08)	9.43 (2.08)	9.43 (2.08)	9.57 (2.51)	9.29 (1.71)
	6	6.74 (1.72)	6.74 (1.71)	6.74 (1.71)	7.01 (1.14)	6.47 (2.46)

Table 6 Average Bayesian estimates (MSEs) of σ_1

Set	Prior	Avg mean (MSE)	Avg median (MSE)	Avg mode (MSE)	LINEX	
					$a = -1$	$a = 1$
1 TV = 1	1	1.08 (0.06)	1.05 (0.05)	1.17 (0.06)	1.11 (0.09)	1.06 (0.05)
	2	1.02 (0.04)	0.99 (0.04)	1.13 (0.04)	1.04 (0.05)	1.00 (0.04)
	3	1.02 (0.04)	1.00 (0.04)	1.13 (0.04)	1.04 (0.05)	1.00 (0.04)
	4	1.02 (0.04)	0.99 (0.04)	1.11 (0.04)	1.04 (0.05)	1.00 (0.04)
	5	1.09 (0.06)	1.06 (0.05)	1.19 (0.06)	1.11 (0.07)	1.06 (0.05)
	6	1.03 (0.03)	1.01 (0.02)	1.15 (0.05)	1.06 (0.04)	1.02 (0.03)
2 TV = 1	1	1.06 (0.04)	1.03 (0.04)	1.15 (0.05)	1.08 (0.06)	1.04 (0.04)
	2	1.00 (0.03)	0.98 (0.03)	1.12 (0.04)	1.02 (0.03)	0.99 (0.03)
	3	1.00 (0.03)	0.98 (0.03)	1.12 (0.04)	1.02 (0.03)	0.99 (0.03)
	4	1.00 (0.03)	0.98 (0.03)	1.12 (0.04)	1.02 (0.03)	0.99 (0.03)
	5	1.05 (0.04)	1.03 (0.03)	1.16 (0.05)	1.07 (0.04)	1.03 (0.03)
	6	1.01 (0.03)	0.99 (0.03)	1.13 (0.04)	1.03 (0.03)	0.99 (0.03)

of θ_{21} , θ_{22} , σ_1 and σ_2 along with their mean squared errors (MSEs) over the 200 simulations. Note that all these posterior point estimators are optimal under different symmetric loss functions. However in practice, sometimes overestimation of a lifetime parameter may be more serious than underestimation and vice versa. In such situations, one may use an asymmetric loss function like the LINEX loss function introduced by Zellner (1986). Thus Tables 4, 5, 6 and 7 also provide the average Bayes' estimates under the LINEX loss function and the corresponding MSEs for $a = -1$ and 1, where a is as in Zellner (1986).

Table 7 Average Bayesian estimates (MSEs) of σ_2

Set	Prior	Avg mean (MSE)	Avg median (MSE)	Avg mode (MSE)	LINEX	
					$a = -1$	$a = 1$
1 TV = 1	1	1.04 (0.03)	1.02 (0.03)	1.15 (0.04)	1.06 (0.03)	1.02 (0.03)
	2	1.00 (0.02)	0.98 (0.02)	1.12 (0.03)	1.01 (0.02)	0.99 (0.02)
	3	1.00 (0.02)	0.98 (0.02)	1.12 (0.03)	1.01 (0.02)	0.99 (0.02)
	4	1.00 (0.02)	0.98 (0.02)	1.12 (0.03)	1.01 (0.02)	0.99 (0.02)
	5	1.04 (0.03)	1.02 (0.02)	1.16 (0.04)	1.05 (0.03)	1.02 (0.02)
	6	1.02 (0.02)	1.00 (0.02)	1.14 (0.04)	1.03 (0.02)	1.01 (0.02)
2 TV = 1	1	1.06 (0.03)	1.04 (0.03)	1.18 (0.05)	1.08 (0.04)	1.05 (0.03)
	2	1.02 (0.02)	1.00 (0.02)	1.14 (0.04)	1.03 (0.02)	1.01 (0.02)
	3	1.02 (0.02)	1.00 (0.02)	1.15 (0.04)	1.04 (0.02)	1.01 (0.02)
	4	1.02 (0.02)	1.00 (0.02)	1.15 (0.04)	1.03 (0.03)	1.01 (0.02)
	5	1.05 (0.03)	1.02 (0.02)	1.16 (0.04)	1.06 (0.03)	1.03 (0.02)
	6	1.03 (0.02)	1.01 (0.02)	1.16 (0.04)	1.04 (0.03)	1.02 (0.02)

Observe that the performances of the posterior mean, median and mode of the slope parameters are almost identical under all the six priors for both sets of parameter values. When the prior means are correctly specified (i.e., for priors 2, 3 and 4), the Bayes’ estimates of the slope parameters under the LINEX loss function are clearly outperformed by the posterior mean, median and mode in all the cases. This is true even for the non-informative prior (i.e., prior 1). However, the Bayes’ estimates of the slope parameters under the LINEX loss function are clearly better than the corresponding posterior mean, median and mode, when the prior means are misspecified. When the prior means of θ_{21} and θ_{22} are over-specified as in prior 5, the Bayes’ estimates under the LINEX loss function for $a = 1$ is better compared to others. This is expected since the LINEX loss function with $a = 1$ penalizes overestimation more heavily than underestimation. Analogous observations may be made in case of prior 6 and $a = -1$. The biases and MSEs of all the Bayesian estimators of θ_{21} and θ_{22} are smaller under prior 3 compared to prior 2. Thus multivariate normal priors with non-zero correlation seem to perform better than independent normal priors for estimating the stress coefficients. Also as expected, the MSEs of the Bayesian estimators of θ_{21} and θ_{22} tend to go up with the increase in prior variances.

From Tables 6 and 7, it is evident that the posterior modes perform poorly compared to other point estimates of the population SDs. Unlike the slope (and intercept) parameters, curiously enough, here neither the priors nor the loss functions seem to exhibit the same expected (or even otherwise) effects on the posterior estimates.

The main purpose of performing an ALT is to predict reliability at usage stress, which is 403.15 K for the underlying example of this simulation study. Thus we also study the behavior of the posterior estimates of the mean component log-lifetimes μ_j as in (18) with $T = 403.15$. The results are reported in Tables 8 and 9 from which conclusions similar to the ones for the slope parameters may be drawn. The simulation study is also repeated for $N = 80$, resulting in qualitatively similar findings.

Table 8 Average Bayesian estimates (MSEs) of μ_1

Set	Prior	Avg mean (MSE)	Avg median (MSE)	Avg mode (MSE)	LINEX	
					$a = -1$	$a = 1$
1 TV = 3.29	1	3.39 (0.30)	3.36 (0.27)	3.32 (0.25)	3.60 (0.57)	3.23 (0.25)
	2	3.34 (0.08)	3.32 (0.07)	3.33 (0.07)	3.40 (0.10)	3.29 (0.07)
	3	3.34 (0.07)	3.32 (0.06)	3.35 (0.06)	3.38 (0.09)	3.30 (0.06)
	4	3.34 (0.07)	3.32 (0.07)	3.34 (0.07)	3.39 (0.10)	3.30 (0.06)
	5	3.94 (0.50)	3.92 (0.47)	3.93 (0.48)	3.99 (0.60)	3.89 (0.44)
	6	2.87 (0.25)	2.85 (0.26)	2.85 (0.25)	2.93 (0.22)	2.82 (0.29)
2 TV = 2.32	1	2.40 (0.28)	2.37 (0.26)	2.34 (0.24)	2.60 (0.51)	2.25 (0.25)
	2	2.35 (0.06)	2.33 (0.05)	2.35 (0.05)	2.41 (0.08)	2.31 (0.05)
	3	2.35 (0.05)	2.34 (0.05)	2.37 (0.05)	2.39 (0.06)	2.32 (0.05)
	4	2.36 (0.05)	2.34 (0.05)	2.36 (0.05)	2.40 (0.07)	2.31 (0.05)
	5	2.84 (0.33)	2.82 (0.30)	2.85 (0.32)	2.89 (0.39)	2.80 (0.28)
	6	1.96 (0.19)	1.94 (0.20)	1.95 (0.19)	2.01 (0.16)	1.91 (0.22)

Table 9 Average Bayesian estimates (MSEs) of μ_2

Set	Prior	Avg mean (MSE)	Avg median (MSE)	Avg mode (MSE)	LINEX	
					$a = -1$	$a = 1$
1 TV = 2.81	1	2.80 (0.16)	2.79 (0.16)	2.78 (0.16)	2.92 (0.19)	2.69 (0.17)
	2	2.80 (0.05)	2.79 (0.05)	2.80 (0.05)	2.84 (0.05)	2.76 (0.05)
	3	2.80 (0.04)	2.79 (0.04)	2.82 (0.04)	2.83 (0.05)	2.77 (0.04)
	4	2.80 (0.05)	2.79 (0.05)	2.80 (0.05)	2.83 (0.05)	2.77 (0.05)
	5	3.30 (0.29)	3.29 (0.28)	3.31 (0.30)	3.33 (0.33)	3.26 (0.26)
	6	2.39 (0.22)	2.38 (0.23)	2.39 (0.22)	2.43 (0.19)	2.35 (0.25)
2 TV = 1.84	1	1.86 (0.21)	1.85 (0.20)	1.84 (0.20)	1.98 (0.24)	1.74 (0.21)
	2	1.86 (0.04)	1.85 (0.04)	1.86 (0.04)	1.90 (0.05)	1.82 (0.04)
	3	1.86 (0.04)	1.85 (0.03)	1.88 (0.04)	1.90 (0.04)	1.84 (0.03)
	4	1.86 (0.04)	1.85 (0.04)	1.87 (0.04)	1.90 (0.04)	1.83 (0.04)
	5	2.26 (0.21)	2.25 (0.20)	2.27 (0.22)	2.30 (0.25)	2.23 (0.18)
	6	1.53 (0.14)	1.52 (0.15)	1.53 (0.14)	1.57 (0.12)	1.49 (0.17)

5 Data analysis

In this section, the methodologies developed in the previous two sections are illustrated by analyzing the real data set pertaining to the electrical insulation system introduced in Sect. 1. Though this system has three failure modes, due to lack of observations, like Pascual (2008, 2010) we drop the failure mode Phase and pay attention only to the remaining two failure modes Turn ($j = 1$) and Ground ($j = 2$). For computational ease, the system lifetimes are re-calibrated to 10^4 hours. As in Sect. 4.3, here also the

Arrhenius model given in (18) is used as the STF for both the failure modes. Now instead of assuming a particular lifetime model $f_j(\cdot)$ for the j th failure mode at the outset, since the developed methodology allows one to choose different probability distributions for different failure modes, we first attempt to resolve the issue of model selection for both Turn and Ground, based on the data \mathcal{D} .

5.1 Bayesian model selection

Though any one of the distributions of X_j listed in Table 1 can serve as a possible candidate for an appropriate lifetime model for Turn or Ground, here for illustrative purpose, we consider only the three popular alternatives in the reliability literature, viz., exponential, Weibull and log-normal. Among the standard model selection methods that are currently available in the literature, two of the most popular ones are the Akaike information criterion (AIC) and the Bayes information criterion (BIC) (see Kadane and Lazar 2004, for example). Table 10 reports the maximized value of the log-likelihood, AIC and BIC for the three models for the two failure modes. Based on these preliminary numbers, though exponential can be clearly ruled out in favor of the other two, evidences for Weibull against log-normal are not so convincing. Thus the exponential model is dropped from further consideration.

The two criteria reported in Table 10 are “preliminary” in the sense that they have their own asymptotic interpretations for their validity as “goodness-of-fit” measures for model selection, despite possessing the attractive property of being independent of priors for the individual model parameters. The standard Bayesian criterion for assessing this is off course the BF (see Kass and Raftery 1995, for example), and thus concentration is now focused on obtaining the values of these BFs. Computation of BF has two related issues, namely the choice of priors and the integrals of the likelihoods with respect to them. While the second issue, though important, is essentially computational, the first issue of choice of prior is more critical (see Berger and Pericchi 1996, for example).

First of all the priors need to be proper. Since they also subsume the degree of knowledge about individual parameters of an underlying model, while comparing two models using a BF, it is crucial that the same level of information (be they correct or not, precise or not) is conveyed by the priors of the parameters of both the models. Towards this goal, the level of prior knowledge about a parameter value is simlized to its first two prior moments. Now in order to change the level of this knowledge,

Table 10 Preliminary comparison of goodness-of-fit for different models

Model	Turn			Ground		
	Log-likelihood	AIC	BIC	Log-likelihood	AIC	BIC
Exponential	-0.57	5.14	6.42	-8.73	21.45	23.12
Weibull	18.32	-30.65	-28.73	-2.38	10.77	13.27
Log-normal	18.22	-30.45	-28.54	-3.13	12.28	14.78

the prior means and SDs of all the involved parameters (of both the models—Weibull and log-normal for both Turn and Ground) are respectively varied through the three values of their MLEs and MLEs \pm SEs, and SEs and SEs*/2, resulting in nine different choices for the prior hyper-parameters.

Once one has the prior means and SDs of the model parameters, the prior specification now proceeds as explained in Sect. 3.2. Specifically, independent normal priors for the stress coefficients of both the Weibull and log-normal models, and gamma and inverse-gamma priors respectively for the Weibull shape (β_j) and log-normal variance (σ_j^2), are used with their respective hyper-parameters matching their first two prior moments. Thus one now needs to compute nine BFs corresponding to the nine different priors (at any given instance, the priors of all the parameters for both the models should be fixed at the same level for fair comparison as explained above).

Laplace approximations of the integrals of products of likelihoods and priors, expressed in terms of the posterior modes and the Hessian of the log-posteriors, as in equation (4) of Kass and Raftery (1995), may be readily used to compute a BF for a given set of priors. Like the AIC and BIC, this approximation is theoretically valid only for large N , but 40, its value in the present example, may not be adequate for that (with 6 unknown parameters). Currently the most popular method of computing a BF to an arbitrary degree of accuracy (limited only to time and computing resources) for any N , is the reversible jump MCMC introduced by Green (1995). In this method, if one has already developed MCMC algorithms for the individual models (as in this article), the best way to exploit them would be to use them as part of the generation of proposal samples in the Metropolis–Hastings iteration of the reversible jump MCMC (see Godsill 2001, for further details). But an easier approach, in the sense of avoiding “rejection” in the Metropolis–Hastings, is offered by Carlin and Chib (1995) (CC henceforth) through a Gibbs sampling, which also incorporates the developed MCMC algorithms for the individual models within it. However the CC approach requires one to specify, what they call “pseudo-priors”. Following their suggestion as well as that of Godsill (2001), here we use multivariate normal approximations (with posterior moments) of the joint posteriors of the parameters of the “other” model as pseudo-priors in their equation (2), together with equal prior probabilities for both the Weibull and log-normal models, denoted by $\pi(W) = \pi(L) = 1/2$. Equation (1) of CC is same as sampling from the posteriors of the Weibull and log-normal model parameters, which can be done following the methods developed in Sects. 3, 4 and elsewhere.

We compute the nine BFs using this CC method, which we call CCBF henceforth, for all the nine priors mentioned above. Along with the CCBF for each of the nine priors, BF is also calculated (with $\pi(W) = \pi(L) = 1/2$ for both Turn and Ground) using the Laplace approximation mentioned above. Both these BFs for log-normal versus Weibull for Turn and Ground are reported in Table 11.

Note that there is a wide discrepancy between the CCBF and the Laplace approximation of the corresponding BF. Though very different, even in terms of the direction of the evidence in most cases, none of these BFs actually even “positively” supports one model over the other in case of Turn. Here we use words like “positive” or “strong” in accordance to the second table in Section 3.2 of Kass and Raftery (1995). This interpretation of the BF values in case of Ground however leads one to two different

Table 11 CCBF (Laplace approximated BF) for log-normal versus Weibull

Prior mean	Prior SD					
	Turn			Ground		
	0.5 × SE	SE	2 × SE	0.5 × SE	SE	2 × SE
MLE−SE	0.72 (0.87)	0.58 (0.86)	0.45 (0.78)	46.32 (0.41)	17.21 (0.39)	12.54 (0.36)
MLE	0.65 (1.05)	0.58 (1.05)	0.50 (1.04)	19.27 (0.47)	10.80 (0.47)	11.02 (0.46)
MLE+SE	0.69 (1.04)	0.58 (1.10)	0.53 (1.15)	29.64 (0.48)	12.75 (0.50)	12.19 (0.51)

Table 12 Posterior summary measures for turn

Parameter	Prior	Mean	Median	Mode	LINEX		SD	95 % HPDCS
					a = −1	a = 1		
θ_{11}	1	−24.65	−24.65	−24.44	−24.06	−25.29	1.11	(−26.79, −22.47)
	2	−24.66	−24.64	−24.52	−24.10	−25.27	1.08	(−26.69, −22.51)
θ_{21}	1	10.82	10.81	10.81	10.96	10.69	0.52	(9.79, 11.81)
	2	10.82	10.81	10.84	10.95	10.70	0.50	(9.83, 11.79)
β_1	1	5.42	5.38	5.24	5.68	5.18	0.71	(4.06, 6.82)
	2	5.43	5.42	5.42	5.57	5.30	0.52	(4.43, 6.47)

conclusions. While according to the Laplace approximations, there isn't any evidence "worth more than a bare mention", in favor of Weibull over log-normal; the CCBFs say that there is "positive" to "strong" evidence in favor of log-normal over Weibull. Thus according to the Laplace approximations, there is no clear choice between the two models for either of the failure modes, while the CCBFs suggest the log-normal model for Ground. Hence for the purpose of illustration, we present Bayesian analyses for Turn and Ground under Weibull and log-normal model respectively.

5.2 Bayesian analyses

Through out this sub-section and the next, we work with normal priors for the stress coefficients of both the Weibull and log-normal models. The means and SDs of these normal priors are assumed to be respectively same as the MLE and SE of the corresponding parameters (i.e. the same prior as case (2,2) of Table 11). The corresponding priors for the Weibull- β_j and log-normal- σ_j^2 are also assumed to be same as case (2,2) of Table 11. In the sequel, this prior is referred as prior-2 for both the Weibull and log-normal models. For the log-normal model, prior-1 is same as the non-informative prior given in (17), while for Weibull, prior-1 refers to $\pi(\theta_{11}) \propto 1$, $\pi(\theta_{21}) \propto 1$ and $\beta_1 \sim \text{Gamma}(29.80, 5.47)$.

Table 12 presents posterior mean, median, mode and Bayes' estimates under the LINEX loss function for θ_{11} , θ_{21} and β_1 , the Weibull lifetime parameters for Turn. Table 12 also provides the posterior SDs and 95 % Highest Posterior Density Credible

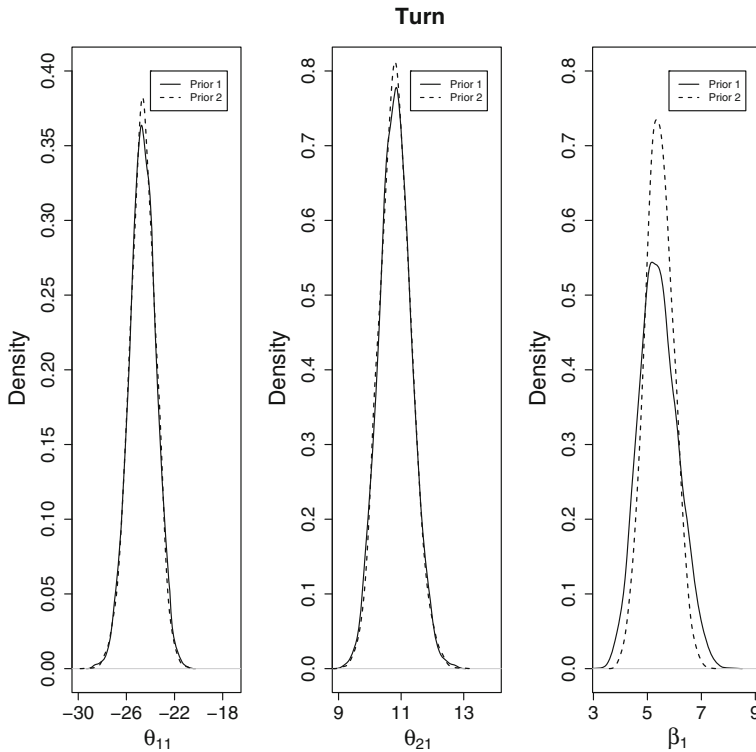


Fig. 1 Density estimates of marginal posteriors of the lifetime parameters for turn

Sets (HPDCS). The posterior p.d.f.s of the model parameters are plotted in Fig. 1. Note that for both the priors, the Bayesian point estimates of the slope parameter θ_{21} is positive and their 95 % HPDCS exclude 0. This is suggestive of the fact that the mean log-lifetime of Turn reduces as the temperature increases.

Similarly the posterior p.d.f.s of the log-normal parameters θ_{12} , θ_{22} and σ_2 for Ground are plotted in Fig. 2, and their posterior mean, median, mode, Bayes' estimates under the LINEX loss function, SD and 95 % HPDCS are reported in Table 13. Note that here also, as expected, temperature seems to reduce the mean log-lifetime.

Inferences on particular distributional parameters are meaningful if and only if the distribution is adequately supported by the data. In this case as seen in Sect. 5.1, neither the Weibull nor the log-normal model is “positively” supported by the data against each other for Turn. Though there is “positive” to “strong” evidence in favor of log-normal against Weibull for Ground, it will not be correct to behave as if one is “hundred percent certain” that it is so. That is, in either case there are uncertainties regarding the model—in case of Turn it is acute, much less for Ground. In such situations where there are model uncertainties, proper Bayesian inferences can still be drawn about different model-independent aspects of life distributions, which are indeed the quantities of actual interest in an ALT rather than the model parameters themselves, through Bayesian Model Averaging or BMA (see Hoeting et al. 1999, for example).

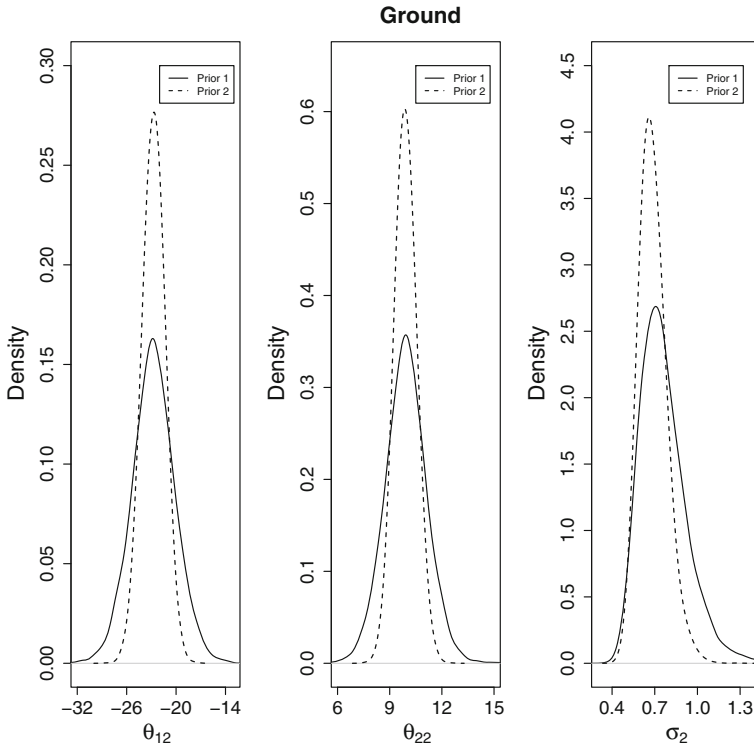


Fig. 2 Density estimates of marginal posteriors of the lifetime parameters for ground

Table 13 Posterior summary measures for ground

Parameter	Prior	Mean	Median	Mode	LINEX		SD	95 % HPDCS
					$a = -1$	$a = 1$		
θ_{12}	1	-22.75	-22.77	-23.17	-17.86	-27.73	2.63	(-27.95, -17.63)
	2	-22.75	-22.73	-22.62	-21.86	-23.67	1.34	(-25.29, -20.13)
θ_{22}	1	9.93	9.92	9.82	10.75	9.17	1.19	(7.55, 12.24)
	2	9.91	9.89	9.84	10.10	9.72	0.61	(8.73, 11.09)
σ_2	1	0.76	0.73	0.94	0.77	0.75	0.16	(0.47, 1.07)
	2	0.69	0.68	0.83	0.69	0.68	0.09	(0.52, 0.87)

5.3 Bayesian model averaging

In the context of ALT of series systems, where one is finally concerned with drawing inferences about reliability characteristics at usage stress, examples of such model-independent quantities of interest include component MTTF and component and system failure s.f. at usage stress. The component MTTFs for the j th component at usage stress are given by $e^{\mu_j(\theta_{1j}, \theta_{2j}, 403.15)} \Gamma(1 + 1/\beta_j)$ and $e^{\mu_j(\theta_{1j}, \theta_{2j}, 403.15) + \frac{1}{2}\sigma_j^2}$ for

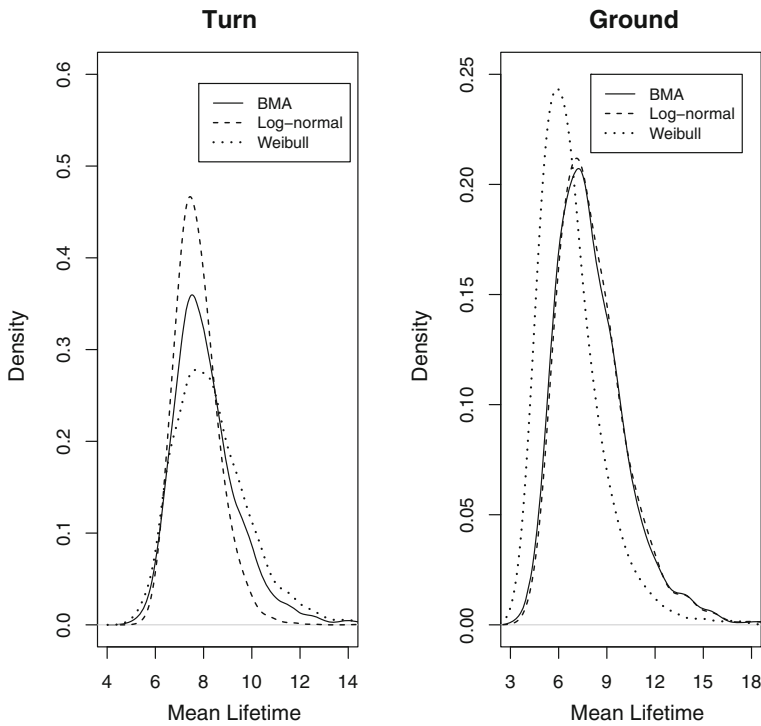


Fig. 3 Posterior densities of MTTF at usage stress for Turn and Ground

Table 14 Posterior summary measures of MTTF at usage stress for Turn and Ground

	Turn			Ground		
	Log-normal	Weibull	BMA	Log-normal	Weibull	BMA
Mean	7.72	8.33	8.11	8.17	6.75	8.04
SD	0.91	1.57	1.39	2.26	1.99	2.26
HPDCS	(6.01, 9.46)	(5.69, 11.58)	(5.81, 10.73)	(4.56, 12.49)	(3.63, 10.64)	(4.27, 12.22)

the Weibull and log-normal models respectively (where $\mu_j(\theta_{1j}, \theta_{2j}, 403.15)$ is as in (18)). The expressions for component s.f. at usage stress in log-scale (for convenience of plotting) are similarly obtained from Table 1. Denoting these quantities by $\Delta(\psi_j, t)$ in Sect. 3.3, it was shown how to obtain a sample from the posterior distributions of such $\Delta(\psi_j, t)$ s for a given fixed model, like Weibull (W) and log-normal (L). Now it will be convenient to denote them by $\Delta_W(t)$ and $\Delta_L(t)$ respectively. The uncertainties between W and L for drawing inference about $\Delta(t)$ are accounted through BMA as follows.

In the Bayesian paradigm, if one starts with prior model uncertainties given by $\pi(W)$ and $\pi(L)$ (here taken to be 1/2 each), in light of the observed data (and off course also the assumed priors on the individual model parameters), these uncertainties get

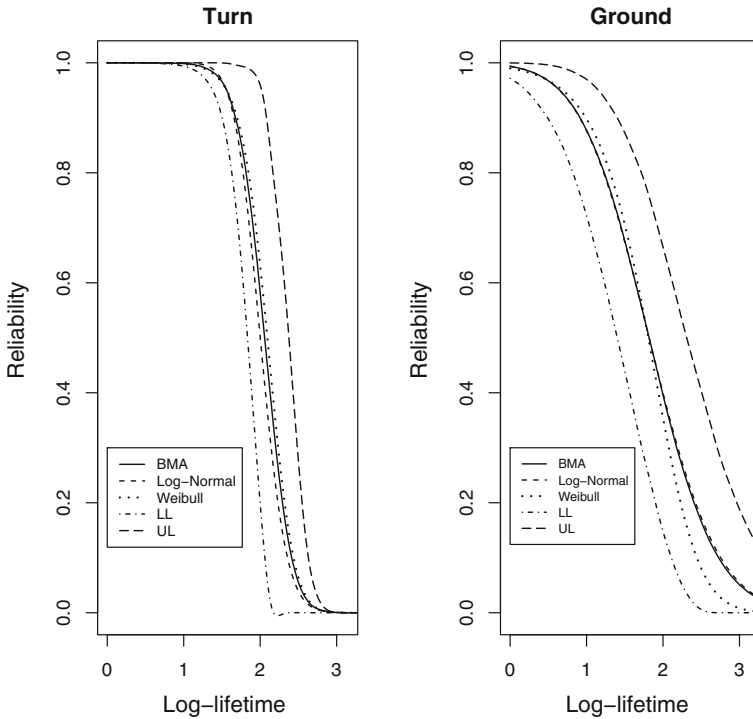


Fig. 4 Predictive survival function at usage stress for Turn and Ground

naturally quantified by their respective posterior probabilities given by $\pi(W|\mathcal{D})$ and $\pi(L|\mathcal{D})$. The BF's reported in Table 11 are nothing but $\frac{\pi(L|\mathcal{D})/\pi(W|\mathcal{D})}{\pi(L)/\pi(W)}$ and thus here $\pi(L|\mathcal{D}) = BF/(1 + BF)$ and $\pi(W|\mathcal{D}) = 1 - \pi(L|\mathcal{D})$. With model uncertainties thus quantified, it is now easy to see that this model-uncertainty-adjusted posterior of $\Delta(t)$ is just a mixture of posteriors of $\Delta_W(t)$ and $\Delta_L(t)$ with respective mixing proportions $\pi(W|\mathcal{D})$ and $\pi(L|\mathcal{D})$ (see, for example, equation (1) of Hoeting et al. (1999)). Thus a BMA sample of a $\Delta(t)$ (such as a component MTTF or s.f.) is easily obtained by first drawing a 0–1 valued random variate with $\Pr(1) = \pi(W|\mathcal{D})$, and then drawing a sample from the posteriors of either $\Delta_W(t)$ or $\Delta_L(t)$, if it is 1 or 0 respectively. Such samples are needed for computation of a BMA HPDCS for $\Delta(t)$ for instance, but are not required if one is just interested in the first two BMA moments of $\Delta(t)$ (by using equations (18) and (19) of Kass and Raftery (1995), for example). For the later case it is enough to just have the value of $\pi(W|\mathcal{D})$ and samples from the posteriors of $\Delta_W(t)$ and $\Delta_L(t)$. The BMA mean, the Bayes' estimate with respect to the squared error loss, is an important Bayesian estimate of $\Delta(t)$. For example, if $\Delta(t)$ denotes a component s.f. (or p.d.f. or c.d.f.), its BMA mean is same as its BMA predictive s.f. (or p.d.f. or c.d.f.).

For the data set at hand, for illustrative purpose, we use the aforementioned prior-2 for both Weibull and log-normal. Recall that Gibbs samples from the respective posteriors of ψ_j s under this prior for both $j = 1, 2$ have already been obtained for

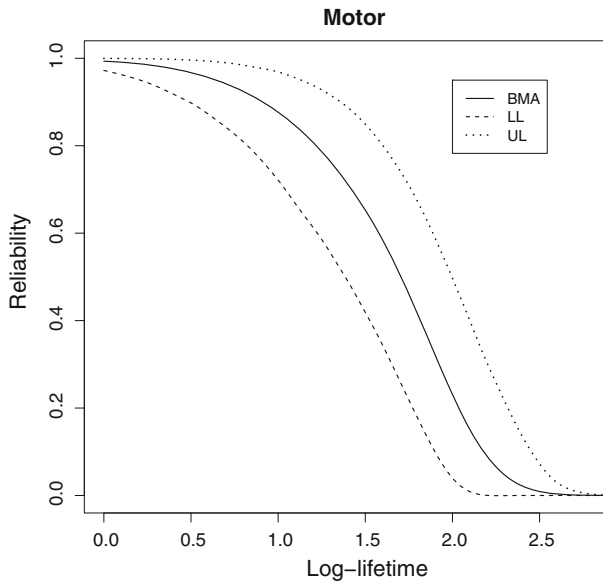


Fig. 5 Predictive survival function of motor at usage stress

computing the (2,2) entries of Table 11. At each of these sampled values we now compute $\Delta_W(t)$ and $\Delta_L(t)$ as in Sect. 3.3 to obtain samples from the posteriors of $\Delta_W(t)$ and $\Delta_L(t)$. Now the BMA sample generations and moment calculations, as explained above, are done using the (2,2) CCBFs of Table 11. The posterior densities of MTTF at usage stress for both Turn and Ground for both the models and their BMA counterparts are plotted in Fig. 3. The corresponding posterior summary measures are presented in Table 14. Figure 4 depicts the Weibull, log-normal and BMA predictive s.f. together with the BMA 95 % HPDCS band for the s.f. at usage stress for Turn and Ground.

We finish this section by estimating the system s.f. at usage stress. In this case, because of the assumed independence between the two failure modes, the BMA predictive system s.f. at usage stress is simply their product for each failure mode (which are already depicted in Fig. 4). A BMA sample of the system s.f. $\bar{F}_T(t)$ is generated by repeatedly, independently and randomly drawing one element each from the BMA samples of $\bar{F}_1(t)$ and $\bar{F}_2(t)$ and forming their product. This BMA sample on $\bar{F}_T(t)$ is then used to compute a BMA 95 % HPDCS band for $\bar{F}_T(t)$, which is plotted in Fig. 5, together with the BMA predictive system s.f. computed as explained above.

6 Concluding remarks

This article presents full Bayesian analyses of fixed multi-stress ALT data of J -component series systems. It is assumed that the component log-lifetimes belong to independent arbitrary log-concave log-location-scale family of distributions. The location parameters of the component log-lifetimes are assumed to depend on the

stress variables through a linear STF, which covers almost all the standard ones used in the engineering literature and practice, as special cases.

Next, a Bayesian methodology is developed for such generic log-concave log-location-scale family of component lifetime models. It is proved that the one-dimensional univariate conditional posteriors of each lifetime parameter given the rest is log-concave, provided the corresponding priors of these lifetime parameters are also log-concave. This fact is then utilized to carry out Bayesian analyses by drawing samples from the joint posterior of the model parameters via the ubiquitous Gibbs sampling. It is shown how the sample generated from the joint posterior of the lifetime parameters can be used for predictive inferences on the component and system reliability metrics at the usage stress, given fixed assumed models for component lifetimes.

Later this article focuses on the special case of log-normally distributed component lifetime, for its importance as a theoretical lifetime model, especially in the context of ALT. It is observed that for the log-normal model one can avoid computationally intensive log-concave sampling by adopting a data-augmentation scheme. A detailed simulation of the log-normal model reveal that there is hardly any difference in the performances of the posterior mean, median and mode of the stress coefficients and MTTF, provided the prior means are correctly specified. However if the prior means are misspecified, the Bayes' estimates under LINEX loss may be a better choice. Furthermore it is also found that, in case of the SDs of the component log-lifetimes, the posterior modes perform poorly compared to other Bayesian point estimators, and the LINEX loss does not seem to exhibit the same expected (or otherwise) influence like in case of the location parameters.

Next the developed methodology is illustrated by analyzing a real data set pertaining to the electrical insulation system of a motor with two failure modes. First as is tacit in the rest of the article, an effort is made to choose a "correct" model for the two failure mode lifetimes based on the BFs. After not being able to conclusively arrive at such a "correct" model for at least one of the failure-modes (while for the other failure mode, curiously enough arriving at the log-normal model, having a special section in this article), subsequent Bayesian analyses of quantities of interest from a series system ALT are carried out using Bayesian model averaging.

Thus this article now provides methods of performing Bayesian analyses of fixed multi-stress series system ALT data, with standard models that are expected to be encountered in practice. However there are still issues such as objective criteria, utilization of prior information etc. for optimally planning or designing such fixed multi-stress ALTs. There are both design and analyses issues with other types of ALTs which are not fixed stress such as the step-stress ALT (Nelson 1980). These are a few future directions, research towards which will be fruitful and appreciated by reliability practitioners, apart from off course theoretical developments with dependent component lives or more complex systems.

Acknowledgments The authors would like to gratefully thank three anonymous referees for their constructive comments, which greatly helped enhance both the content and presentation of the article.

7 Appendix 1: Proof of Theorem 1

It is enough to show that $\log L_j(\boldsymbol{\psi}_j|\mathcal{D}) (= \ell_j(\boldsymbol{\psi}_j|\mathcal{D}))$, say) is a concave function of θ_{kj} and τ_j . From Eq. (8), one gets

$$\begin{aligned} \ell_j(\boldsymbol{\psi}_j|\mathcal{D}) &= n_j \log \tau_j + \sum_{\substack{1 \leq i \leq n \\ \&I^i=j}} \log f_j(\tau_j(t^i - \mathbf{s}_j^{i'}\boldsymbol{\theta}_j)) \\ &\quad + \sum_{\substack{1 \leq i \leq n \\ \&I^i \neq j}} \log \bar{F}_j(\tau_j(t^i - \mathbf{s}_j^{i'}\boldsymbol{\theta}_j)) \\ &\quad + \sum_{i=n+1}^N \log \bar{F}_j(\tau_j(t^i - \mathbf{s}_j^{i'}\boldsymbol{\theta}_j)), \end{aligned}$$

which implies

$$\begin{aligned} \frac{\partial^2 \ell_j(\boldsymbol{\psi}_j|\mathcal{D})}{\partial \theta_{kj}^2} &= \sum_{\substack{1 \leq i \leq n \\ \&I^i=j}} (-\tau_j s_{kj}^i)^2 \frac{\partial^2 \log f_j(u_j^i)}{\partial u_j^{i2}} + \sum_{\substack{1 \leq i \leq n \\ \&I^i \neq j}} (-\tau_j s_{kj}^i)^2 \frac{\partial^2 \log \bar{F}_j(u_j^i)}{\partial u_j^{i2}} \\ &\quad + \sum_{i=n+1}^N (-\tau_j s_{kj}^i)^2 \frac{\partial^2 \log \bar{F}_j(u_j^i)}{\partial u_j^{i2}}, \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2 \ell_j(\boldsymbol{\psi}_j|\mathcal{D})}{\partial \tau_j^2} &= -\frac{n_j}{\tau_j^2} + \sum_{\substack{1 \leq i \leq n \\ \&I^i=j}} \left(\frac{u_j^i}{\tau_j}\right)^2 \frac{\partial^2 \log f_j(u_j^i)}{\partial u_j^{i2}} + \sum_{\substack{1 \leq i \leq n \\ \&I^i \neq j}} \left(\frac{u_j^i}{\tau_j}\right)^2 \frac{\partial^2 \log \bar{F}_j(u_j^i)}{\partial u_j^{i2}} \\ &\quad + \sum_{i=n+1}^N \left(\frac{u_j^i}{\tau_j}\right)^2 \frac{\partial^2 \log \bar{F}_j(u_j^i)}{\partial u_j^{i2}}, \end{aligned}$$

where $u_j^i = \tau_j(t^i - \mathbf{s}_j^{i'}\boldsymbol{\theta}_j)$. Note that $\frac{\partial^2 \log f_j(u_j^i)}{\partial u_j^{i2}} < 0$, since $f_j(\cdot)$ is log-concave. Furthermore, the log-concavity of $f_j(\cdot)$ implies the same for $\bar{F}_j(\cdot)$ (see Basu et al. 2003). Thus, $\frac{\partial^2 \log \bar{F}_j(u_j^i)}{\partial u_j^{i2}} < 0$. Therefore, $\frac{\partial^2 \ell_j(\boldsymbol{\psi}_j|\mathcal{D})}{\partial \theta_{kj}^2} < 0$ and $\frac{\partial^2 \ell_j(\boldsymbol{\psi}_j|\mathcal{D})}{\partial \tau_j^2} < 0$, which completes the proof.

8 Appendix 2: Acronyms and notations

BF	Bayes' factor
BMA	Bayesian Model Averaging
h.r.	Hazard rate
MCMC	Markov Chain Monte Carlo
MLE	Maximum likelihood estimate
MTTF	Mean time to failure
p.d.f.	Probability density function
SD	Standard deviation
SE	Standard error
s.f.	Survival function
J, j	Number of components in series system and dummy in $\{1, \dots, J\}$
K, k	Number of stress variables and and dummy in $\{1, \dots, K\}$
N, i	Number of systems put on an ALT and dummy in $\{1, \dots, N(n)\}$
n	Number of failed systems in an ALT
X_j, Y_j	Lifetime and log-lifetime of the j th component
T, I	System log-lifetime and cause of failure
$\mathbf{Z} = (Z_1, \dots, Z_K)'$	$K \times 1$ vector of stress variables
U, u	A variable U and its observed value with or without an i in the super-script (like U^i or u^i) for the i th system
$f(\cdot \cdot), \bar{F}(\cdot \cdot), h(\cdot \cdot)$	p.d.f., s.f., and h.r. of random variables
$g_{kj}(\cdot)$	Known transformation of the k th stress variable acting on the j th component of the system
$\mathbf{s}_j = (s_{1j}, \dots, s_{Kj})'$ with $s_{kj} = g_{kj}(z_k)$	$K \times 1$ vector of transformed stress variables for the j th component
$\mathbf{S}_j = [\mathbf{s}_j^1, \dots, \mathbf{s}_j^N]'$	$N \times K$ stress matrix corresponding to component j
$\phi(\cdot), \bar{\Phi}(\cdot)$	p.d.f. and s.f. of the standard normal distribution
$\boldsymbol{\theta}_j = (\theta_{1j}, \dots, \theta_{Kj})', \boldsymbol{\theta} = (\theta'_1, \dots, \theta'_j)'$	$K \times 1$ and $KJ \times 1$ vectors of STF parameters for the j th component and system
$\mu_j(\boldsymbol{\theta}_j, \mathbf{z}), \tau_j, \sigma_j^2$	STF (location) and scale parameters of Y_j and its variance when normal
$\boldsymbol{\psi}_j = (\theta'_j, \tau_j)', \boldsymbol{\Psi} = (\boldsymbol{\psi}'_1, \dots, \boldsymbol{\psi}'_j)'$	$(K + 1) \times 1$ and $(KJ + J) \times 1$ vectors of lifetime parameters for the j th component and system

Δ	A model-independent quantity
$\mathcal{D}, \mathcal{Y}_j, \mathcal{Y}$	Observed, augmented and complete data
$L_j(\psi_j \mathcal{D}), L_j(\psi_j \mathcal{Y}), L(\psi \mathcal{D})$	Likelihood function of ψ_j and ψ given \mathcal{D} or \mathcal{Y}
$\pi(\cdot), \pi(\cdot \mathcal{D}), \pi(\cdot \mathcal{Y})$	Prior and posterior of relevant quantities usually clear from the context

References

- Basu S, Sen A, Banerjee M (2003) Bayesian analysis of competing risks with partially masked cause of failure. *J R Stat Soc Ser C (Appl Stat)* 52:77–93
- Berger J (2006) The case for objective Bayesian analysis. *Bayesian Anal* 1:385–402
- Berger JO (1985) *Statistical decision theory and Bayesian analysis*, 2nd edn. Springer, New York
- Berger JO, Pericchi LR (1996) The intrinsic Bayes factor for model selection and prediction. *J Am Stat Assoc* 91:109–122
- Bunea C, Mazzuchi TA (2006) Competing failure modes in accelerated life testing. *J Stat Plan Inference* 136:1608–1620
- Carlin BP, Chib S (1995) Bayesian model choice via Markov chain Monte Carlo methods. *J R Stat Soc Ser B (Methodol)* 57:473–484
- Cowles MK, Carlin BP (1996) Markov chain Monte Carlo convergence diagnostics: a comparative review. *J Am Stat Assoc* 91:883–904
- Fan Th, Hsu Tm (2014) Constant stress accelerated life test on a multiple-component series system under Weibull lifetime distributions. *Commun Stat Theory Methods* 43:2370–2383
- Fernández C, Steel MF (1999) Reference priors for the general location-scale modelm. *Stat Probab Lett* 43:377–384
- Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB (2013) *Bayesian data analysis*, 3rd edn. Chapman and Hall/CRC, Boca Raton
- Gilks WR, Wild P (1992) Adaptive rejection sampling for Gibbs sampling. *Appl Stat* 41:337–348
- Godsill SJ (2001) On the relationship between Markov chain Monte Carlo methods for model uncertainty. *J Comput Graph Stat* 10:230–248
- Green PJ (1995) Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* 82:711–732
- Hauschildt M, Gall M, Thrasher S, Justison P, Hernandez R, Kawasaki H, Ho PS (2007) Statistical analysis of electromigration lifetimes and void evolution. *J Appl Phys* 101(043):523
- Hoeting JA, Madigan D, Raftery AE, Volinsky CT (1999) Bayesian model averaging: a tutorial. *Stat Sci* 14:382–401
- Jiang R (2011) Analysis of accelerated life test data involving two failure modes. *Adv Mater Res* 211–212:1002–1006
- Kadane JB, Lazar NA (2004) Methods and criteria for model selection. *J Am Stat Assoc* 99:279–290
- Kass RE, Raftery AE (1995) Bayes factors. *J Am Stat Assoc* 90:773–795
- Kim CM, Bai DS (2002) Analyses of accelerated life test data under two failure modes. *Int J Reliab Qual Saf Eng* 9:111–125
- Kim JS, Yum B (2008) Selection between Weibull and lognormal distributions: a comparative simulation study. *Comput Stat Data Anal* 53:477–485
- Kim SW, Ibrahim JG (2000) On Bayesian inference for proportional hazards models using noninformative priors. *Lifetime Data Anal* 6:331–341
- Klein JP, Basu AP (1981) Weibull accelerated life tests when there are competing causes of failure. *Commun Stat Theory Methods* 10:2073–2100
- Klein JP, Basu AP (1982a) Accelerated life testing under competing exponential failure distributions. *IAPQR Trans* 7:1–20
- Klein JP, Basu AP (1982b) Accelerated life tests under competing Weibull causes of failure. *Commun Stat Theory Methods* 11:2271–2286

- Kundu D, Manglick A (2004) Discriminating between the Weibull and log-normal distributions. *Nav Res Logist* 51(6):893–905
- Liu JS, Wu YN (1999) Parameter expansion for data augmentation. *J Am Stat Assoc* 94:1264–1274
- Meeker WQ, Escobar LA (1998) *Statistical methods for reliability data*. Wiley, New York
- Mukhopadhyay C, Basu S (2007) Bayesian analysis of masked series system lifetime data. *Commun Stat Theory Methods* 36:329–348
- Nelson W (1980) Accelerated life testing-step-stress models and data analyses. *Reliab IEEE Trans* 29:103–108
- Nelson WB (1990) *Accelerated testing: statistical models, test plans and data analysis*. Wiley, New York
- Pascual F (2008) Accelerated life test planning with independent Weibull competing risks. *Reliab IEEE Trans* 57:435–444
- Pascual F (2010) Accelerated life test planning with independent lognormal competing risks. *J Stat Plan Inference* 140:1089–1100
- Pathak PK, Singh AK, Zimmer WJ (1991) Bayes estimation of hazard & acceleration in accelerated testing. *Reliab IEEE Trans* 40:615–621
- Roy S, Mukhopadhyay C (2013) Bayesian accelerated life testing under competing exponential causes of failure. In: John B, Acharya UH, Chakraborty AK (eds) *Quality and reliability engineering—recent trends and future directions*. Allied Publishers Pvt. Ltd., New Delhi, pp 229–245
- Roy S, Mukhopadhyay C (2014) Bayesian accelerated life testing under competing Weibull causes of failure. *Commun Stat Theory Methods* 43:2429–2451
- Roy S, Mukhopadhyay C (2015) Maximum likelihood analysis of multi-stress ALT data of series systems with competing log-normal causes of failure. *J Risk Reliab* 229:119–130
- Tan Y, Zhang C, Chen X (2009) Bayesian analysis of incomplete data from accelerated life testing with competing failure modes. In: *8th international conference on reliability, maintainability and safety*, pp 1268–1272
- Tanner MA, Wong WH (1987) The calculation of posterior distributions by data augmentation. *J Am Stat Assoc* 82:528–540
- Tojeiro CAV, Louzada-Neto F, Bolfarine H (2004) A Bayesian analysis for accelerated lifetime tests under an exponential power law model with threshold stress. *J Appl Stat* 31:685–691
- Van Dorp JR, Mazzuchi TA (2004) A general Bayes exponential inference model for accelerated life testing. *J Stat Plan Inference* 119:55–74
- Van Dorp JR, Mazzuchi TA (2005) A general Bayes weibull inference model for accelerated life testing. *Reliab Eng Syst Saf* 90:140–147
- Xu A, Tang Y (2011) Objective Bayesian analysis of accelerated competing failure models under Type-I censoring. *Comput Stat Data Anal* 55:2830–2839
- Zellner A (1986) Bayesian estimation and prediction using asymmetric loss functions. *J Am Stat Assoc* 81:446–451
- Zhang Y, Meeker WQ (2006) Bayesian methods for planning accelerated life tests. *Technometrics* 48:49–60