Sains Malaysiana 53(9)(2024): 3215-3227
http://doi.org/10.17576/jsm-2024-5309-24

Bayesian Approach for Estimating the Parameters of the Time-to-Failure Distribution and Its Percentiles under Power Degradation Model

(Pendekatan Bayesian untuk Menganggar Parameter Taburan Masa Kegagalan dan Peratusannya di bawah Model Penurunan Kuasa)

Laila Naji Ba Dakhn[1,2], Mohd Aftar Abu Bakar[1], Razik Ridzuan Mohd Tajuddin[1,*] & Kamarulzaman Ibrahim[1]

[1]*Department of Mathematical Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia*

[2]*Department of Mathematics, Faculty of Science, Hadhramout University, Mukalla, Hadhramout, Yemen*

Received: 8 January 2024/Accepted: 9 July 2024

ABSTRACT

The degradation models are often applied on the degradation data for studying time-to-failure distribution. In this study, the Bayesian approach is applied on the power degradation model for estimating the parameters of the time-to-failure distribution and its percentiles. Two different distributions are assumed for the degradation parameter of the model. The degradation parameter is firstly assumed to follow the skew-normal distribution with three jointly independently distributed parameters such that the gamma prior is assumed for the shape parameter, while the scale and the location parameters are assumed uniform. The second distribution assumed for the degradation parameter is the log-logistic distribution with two jointly independent random parameters where the shape parameter is assumed gamma, while the scale parameter is assumed uniform. Based on the Gibbs sampling method carried out under the JAGS platform, the models considered are applied on the simulated data and the NASA turbofan Jet engine dataset and the results found are compared. In modeling the time-to-failure distribution, it is shown that based on the simulated data and real data, the Bayesian approach for the power degradation model with the skew-normal degradation parameter outperformed the Bayesian approach for the power degradation model with the log-logistic degradation parameter.

Keywords: Bayesian approach; log-logistic distribution; power degradation model; skew-normal distribution; time-to-failure distribution

ABSTRAK

Model degradasi sering digunakan pada data degradasi untuk mempelajari taburan masa kegagalan. Dalam kajian ini, pendekatan Bayes digunakan pada model degradasi kuasa untuk menganggarkan parameter taburan masa kegagalan dan persentilnya. Dua taburan yang berbeza diandaikan untuk parameter degradasi model. Parameter degradasi pertama diandaikan mengikut taburan normal terpencong dengan tiga parameter yang tertabur secara tak bersandar dengan diandaikan prior gamma untuk parameter bentuk, sementara parameter skala dan lokasi diandaikan tertabur secara seragam. Taburan kedua yang diandaikan untuk parameter degradasi adalah taburan log-logistik dengan dua parameter rawak bercantum yang tak bersandar dengan andaian parameter bentuk tertabur secara gamma, sementara parameter skala tertabur secara seragam. Berdasarkan kaedah pensampelan Gibbs yang dilaksanakan di platform JAGS, model yang dipertimbangkan digunakan pada data simulasi dan set data enjin jet kipas turbo NASA dan hasil yang ditemui dibandingkan. Dalam pemodelan taburan masa kegagalan, didapati bahawa berdasarkan data yang disimulasikan dan data sebenar, prestasi pendekatan Bayes untuk model degradasi kuasa dengan parameter degradasi normal terpencong mengatasi prestasi pendekatan Bayes untuk model degradasi kuasa dengan parameter degradasi log-logistik.

Kata kunci: Masa kegagalan; model degradasi kuasa; pendekatan Bayes; taburan log-logistik; taburan normal terpencong

INTRODUCTION

In the reliability study, degradation data are more commonly applied instead of the lifetime data which involve records of failure time of a particular product. Reliability has been clearly defined in many studies (Meeker & Escobar 1998). Degradation data plays an important role in helping to monitor equipment reliability, contributing to the benefit of less frequent maintenance and lower costs. In many recent studies, degradation data are considered for estimating the time-to-failure distribution (Eidous, Ebrahem & Dakhn 2017; Oliveira, Loschi & Freitas 2018; Siju & Kumar 2018).

The time-to-failure distribution is determined using two different methods, model-based and data-driven (Wu et al. 2019). Based on the degradation data, model-based methods involve the application of mathematical models to represent the failures in the form of degradation path models. Most of these models include certain mathematical functions such as linear, exponential, power and logarithm. In the degradation path model, the degradation parameter follows either a finite combination of two or more probability distributions or a particular probability distribution (Sen, Maiti & Chandra 2016).

In modeling lifetime data involving the degradation path model, the exponential, log-logistic, Weibull, gamma, normal and xgamma probability distributions are often applied (Dakhn, Ebrahem & Eidous 2017; Gonzalez et al. 2010; Yadav et al. 2021) for describing the distribution of the degradation parameter. It has been reported that a family of flexible distributions known as skew-symmetric distributions are found to be adequate in modeling unusual characteristics in the degradation data such as exhibiting a high level of skewness and the existence of multimodality (Alhamidie et al. 2019; Ghaderinezhad, Christophe & Nicola 2020). Indeed, the skew-normal distribution which can be reduced to another symmetric distribution, depending on the value of the skewness parameter, such as the normal distribution if the skewness parameter equals 0 and a half normal distribution if the skewness parameter approaches infinity, as reported by Pan, Liu and Yang (2018), have been found by Dakhn, Mohd Aftar and Kamarulzaman (2023) to be quite flexible in the degradation modeling. Dakhn, Mohd Aftar and Kamarulzaman (2023) have determined the posterior distribution based on the linear degradation model under the assumption that the degradation parameter follows the skew-normal distribution.

In this paper, the skew-normal distribution that has been introduced by Azzalini (1985) is again applied for modelling the degradation parameter, but now for the power degradation model instead of the linear degradation model for estimating the parameters of the time-to-failure distribution and its percentiles as in the study by Dakhn, Mohd Aftar and Kamarulzaman (2023). In addition, the degradation parameter is assumed to follow the log-logistic distribution with random shape and scale parameters. Under these two assumptions on the distribution of the degradation parameter, a prior sensitivity analysis is carried out involving informative and non-informative priors. Several works on the Bayesian approach is as such (Guure & Noor Akma 2014; Puggard, Niwitpong & Niwitpong 2022; Shafiq, Alamgir & Muhammad 2016). Apart from the simulation study, a real data application involving the NASA Turbofan Jet Engine dataset (NTJE) is illustrated using the two models of interest.

The outline of this paper is summarized as follows. Next section explains the power degradation model and the derivation of the time-to-failure distribution when the degradation parameter follows the skew-normal and log-logistic distributions. After that, the Bayesian modeling which involves the degradation parameter following either skew-normal or log-logistic distribution will be discussed. Assessment of the estimated parameters for the simulated data found using the Bayesian approach for the power degradation model involving the different choices of the priors and different choices of the degradation parameter distribution is carried out subsequently by using the JAGS or Just Another Gibbs Sampler. In the section that follows, a real application that consists of NTJE is considered for the two models and the different choices of the priors. Discussion on the convergence diagnostic found is provided particularly for the case of the model of the skew-normal degradation parameter. Finally, the summary of the study is given in last section.

MATERIALS AND METHODS

POWER DEGRADATION MODEL

The general path degradation model is expressed as

$$y_{ij} = D\left(t_{ij}\,; \boldsymbol{\varphi}, X_i\right) + \varepsilon_{ij}\ \ , i = 1, \dots, n, j = 1, \dots, m_i \quad (1)$$

where y_{ij} denotes the observed degradation measurement of i th unit at time t_{ij}, $D(t_{ij};\ \varphi,\ X_i)$ is the actual path of i th unit at time t_{ij}, the term φ is a vector of fixed-effect parameters, X_i is the degradation parameter for i th unit, $\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$ is a random error term where σ_ε^2 is a constant, n is the number of units that are tested and m_i is the total number of observations on i th unit. Here, $\{\varepsilon_{ij}\}$ and $\{X_i\}$ are assumed independent and $X_i's$ are independent.

The actual path in Equation (1) can be written in several forms of mathematical expressions. A nonlinear degradation model which consists of the power function that is commonly applied in many studies can be given as the following:

$$y_{ij} = \boldsymbol{\varphi}\, t_{ij}^{\frac{1}{X_i}} + \varepsilon_{ij}\,, \qquad \boldsymbol{\varphi}, X_i \text{ and } t > 0 \tag{2}$$

In this section, the time-to-failure distribution and its percentiles are derived based on the power degradation model in Equation (2).

Suppose that D_f is a critical degradation value at the time T. Then, the actual path of the power degradation model in Equation (2) can be written as

$$D_f = \boldsymbol{\varphi}\, T^{\frac{1}{X}}$$

then,

$$T = \left(\frac{D_f}{\boldsymbol{\varphi}}\right)^X$$

It can be shown that the cumulative distribution function and the probability density function of the time-to-failure distribution can respectively be given as

$$F_T(t) = G_X\left(\frac{\ln t}{\ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}\right), \qquad \frac{D_f}{\boldsymbol{\varphi}} > 1 \text{ and } t > 1 \tag{3}$$

and

$$f_T(t) = \frac{1}{t\,\ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}\, g_X\left(\frac{\ln t}{\ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}\right) \tag{4}$$

As shown in Equations (3) and (4), $t > 1$ will ensure the positivity of ln t, which is important so that G_X can be defined. In addition, the cumulative distribution function and the probability density function of the time-to-failure distribution are dependent on the distribution of the assumed degradation parameter X. In the following two subsections, the derivation of the time-to-failure distribution is given when the degradation parameter is supposed to follow either the skew-normal or log-logistic distributions.

POWER DEGRADATION MODEL WITH SKEW-NORMAL DEGRADATION PARAMETER

Azzalini (1985) defined that a random variable X follows the standard skew-normal distribution if the probability density function of X is given by:

$$g_{\boldsymbol{X}}(x) = 2\,\phi(x)\,\Phi(\beta x)\ , x \in \mathbb{R}$$

where φ (.) is the probability density function and φ (.) is the cumulative distribution function of a standard normal distribution and $\boldsymbol{\beta} \in \mathbb{R}$ is the shape parameter.

Assumed that X follows the skew-normal distribution with location, scale and shape parameters denoted as μ, σ and β, respectively. Here, we denote:

$$X \sim SN(\mu, \sigma^2, \beta)$$

The cumulative distribution function of the skew-normal distribution is given by

$$G_X(x;\, \mu, \sigma^2, \beta) = \Phi\left(\frac{x-\mu}{\sigma}\right) - 2\boldsymbol{T}\left(\frac{x-\mu}{\sigma}, \beta\right) \tag{5}$$

and the probability density function of skew-normal distribution is given as

$$g_X(x;\, \mu, \sigma^2, \beta) = \frac{2}{\sigma}\,\phi\left(\frac{x-\mu}{\sigma}\right)\Phi\left(\beta\left(\frac{x-\mu}{\sigma}\right)\right) \tag{6}$$

where x, β and $\boldsymbol{\mu} \in \mathbb{R}$, $\boldsymbol{\sigma}^2 \in \mathbb{R}^+$ and $\boldsymbol{T}(h, a)$ is Owen's $\boldsymbol{T}$ function defined as

$$\boldsymbol{T}(h, a) = \frac{1}{2\pi}\int_0^a \frac{e^{-\frac{h^2(1+u^2)}{2}}}{1+u^2}\, du \quad , \ -\infty < a, h < +\infty$$

By modifying Equation (3) based on the definition of Equation (5), the cumulative distribution function and the probability density function of the time-to-failure distribution can be shown as

$$F_T(t) = \Phi\left(\frac{\ln(t) - \mu\ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}{\sigma\,\ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}\right) - 2T\left(\frac{\ln(t) - \mu\,\ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}{\sigma\ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}, \beta\right) \tag{7}$$

and

$$f_T(t) = \frac{2}{\sigma\, t\,\ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}\,\phi\left(\frac{\ln(t) - \mu\ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}{\sigma\,\ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}\right)\Phi\left(\beta\left(\frac{\ln(t) - \mu\ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}{\sigma\,\ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}\right)\right) \tag{8}$$

To determine the 100 p th percentile of the time-to-failure distribution, denoted as t_p, we solve the Equation (3) for t_p as follows:

$$t_p = e^{ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right) G_X^{-1}(p)} \quad (9)$$

Equation (9) is solved numerically to find the value of t_p.

Suppose a random sample of size n from the time-to-failure distribution with parameters μ, σ and β, is denoted as $t_1, t_2, ..., t_n$. Then, the likelihood function for the time-to-failure distribution, denoted as $L_{SN}(\mu, \sigma, \beta; t)$ based on Equation (8) can be shown as the following:

$$L_{SN}(\mu, \sigma, \beta;\ \boldsymbol{t})$$

$$= \left(\frac{2}{\sigma\ ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}\right)^n \prod_{i=1}^{n} \frac{1}{t_i} \phi\left(\frac{\ln t_i - \mu\ ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}{\sigma\ ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}\right) \Phi\left(\beta\left(\frac{\ln t_i - \mu\ ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}{\sigma\ ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}\right)\right) \quad (10)$$

Equation (10) is considered in determining the posterior distribution of the parameters for the time-to-failure distribution.

POWER DEGRADATION MODEL WITH LOG-LOGISTIC DEGRADATION PARAMETER

In this subsection, the time-to-failure distribution is derived based on the power degradation model in Equation (2) by using the degradation parameter X which is assumed to follow the log-logistic distribution with shape parameter ω and scale parameter α.

The cumulative distribution function and the probability density function of the log-logistic distribution are respectively given as the following:

$$G_X(X) = \frac{1}{1 + \left(\frac{x}{\alpha}\right)^{-\omega}} \quad (11)$$

$$g_X(x) = \frac{\left(\frac{\omega}{\alpha}\right)\left(\frac{x}{\alpha}\right)^{\omega-1}}{\left(1 + \left(\frac{x}{\alpha}\right)^{\omega}\right)^2} \quad (12)$$

Following the same procedure as in the previous subsection, the cumulative distribution function and the probability density function of the time-to-failure distribution, denoted as $F_{TL}(t_L)$ and $f_{TL}(t_L)$, are given respectively as the following:

$$F_{T_L}(t_L) = \frac{1}{1 + \left(\frac{\ln t_L}{\alpha \ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}\right)^{-\omega}} \quad (13)$$

$$f_{T_L}(t_L) = \frac{\left(\frac{\omega}{\alpha\, t_L \ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}\right)\left(\frac{\ln t_L}{\alpha \ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}\right)^{\omega-1}}{\left(1 + \left(\frac{\ln t_L}{\alpha \ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}\right)^{\omega}\right)^2} \quad (14)$$

The 100 p^{th} percentile of the time-to-failure distribution, denoted as t_{L_p}, can be determined by solving Equation (13) for t_{L_p}. Thus, we have:

$$t_{L_p} = e^{\alpha\, ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)\left(\frac{1-p}{p}\right)^{\frac{-1}{\omega}}} \quad (15)$$

Suppose $t_{L_1}, t_{L_2}, ..., t_{L_n}$ is a random sample of size n from the time-to-failure distribution with parameters α and ω. Based on Equation (14), the likelihood function of the time-to-failure distribution is given as follows:

$$L_{LL}(\alpha, \omega\ ; \boldsymbol{t_L})$$

$$= \frac{\omega^n}{\left(\alpha \ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)\right)^{2n+\omega-1}} \prod_{i=1}^{n} \frac{\left(\ln t_{L_i}\right)^{\omega-1}}{t_{L_i}\left(1 + \left(\frac{\ln t_{L_i}}{\alpha \ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}\right)^{\omega}\right)^2} \quad (16)$$

In the Bayesian approach, the likelihood function in Equation (16) is considered to obtain the posterior distribution of the parameters α, ω and φ.

BAYESIAN MODELLING OF THE TIME-TO-FAILURE DISTRIBUTION

The Bayesian approach is a popular method that can be applied to estimate the parameters of the time-to-failure distribution and its percentiles. The sensitivity of the choice of prior distribution for estimating the parameters and the percentiles of the time-to-failure distribution is studied when the degradation parameter is assumed to follow the skew-normal or log-logistic distribution by considering both informative and non-informative prior distributions. The parameters of the time-to-failure distribution, say $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_i)$, are assumed independent. Thus, the posterior distribution $\pi(\boldsymbol{\theta}|t)$ can be given by

$$\pi(\boldsymbol{\theta}|t) = \frac{L(\boldsymbol{\theta};t)\, q_1(\theta_1) \dots q_i(\theta_i)}{\int \dots \int L(\boldsymbol{\theta};t)\, q_1(\theta_1) \dots q_i(\theta_i) d\theta_1 \dots d\theta_i} \tag{17}$$

where $L(\boldsymbol{\theta};t)$ is the likelihood function of the time-to-failure distribution for the parameters $\boldsymbol{\theta}$ and $q_i(\theta_i)$ is the prior distribution of θ_i.

POSTERIOR DENSITY WITH SKEW-NORMAL DEGRADATION PARAMETER

Based on the likelihood function given in Equation (10), it is evident that the time-to-failure distribution based on the skew-normal degradation parameter, using a power degradation model, has four parameters where $\boldsymbol{\theta} = (\theta_1 = \mu,\ \theta_2 = \sigma,\ \theta_3 = \beta,\ \theta_4 = \varphi)$. To apply the Bayesian approach, the following prior distributions for the parameters μ, σ, β and φ are considered: $\mu \sim Uniform(0, b)$, $\sigma \sim Uniform(0, b)$, $\beta \sim Uniform(r, s)$ and $\varphi \sim Uniform(0, b)$.

Here, r and s are the shape and the scale hyperparameters for the gamma prior distribution and their values and b are assumed constant.

Based on the Bayesian mechanism, the joint posterior density function $\pi(\mu, \sigma, \beta, \varphi\, |t)$ is found to be proportional to

$$\frac{(\beta)^{r-1}\, e^{\frac{-\beta}{s}}}{b^3\, s^r\, \Gamma(r)} \left(\frac{2}{\sigma\, ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)} \right)^n$$

$$\prod_{i=1}^{n} \frac{1}{t_i} \phi\left(\frac{\ln t_i - \mu\, ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}{\sigma\, ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)} \right) \Phi\left(\beta \left(\frac{\ln t_i - \mu\, ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}{\sigma\, ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)} \right) \right) \tag{18}$$

The joint posterior distribution of the parameters μ, σ, β and φ is analytically intractable. So, the parameters are estimated by implementing the Markov chain Monte Carlo method available in JAGS software.

POSTERIOR DENSITY WITH LOG-LOGISTIC DEGRADATION PARAMETER

The prior distributions for the parameters of the time-to-failure distribution for the power degradation model involving the log-logistic degradation parameter are assumed the same as in the previous subsection. Accordingly, we have $\alpha \sim Uniform(0, b)$, $\omega \sim Gamma(r, s)$ and $\varphi \sim Uniform(0, b)$.

Then, the posterior density $\pi(\alpha, \omega, \varphi\, |t_L)$ is proportional to:

$$\frac{\omega^{n+r-1}\, e^{\frac{-\omega}{s}}}{b^2\, s^r\, \Gamma(r) \left(\alpha \ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)\right)^{2n+\omega-1}}$$

$$\prod_{i=1}^{n} \frac{\left(\ln t_{L_i}\right)^{\omega-1}}{t_{L_i} \left(1 + \left(\frac{\ln t_{L_i}}{\alpha \ln\left(\frac{D_f}{\boldsymbol{\varphi}}\right)}\right)^{\omega}\right)^2} \tag{19}$$

The posterior distribution of the time-to-failure distribution given in Equation (19) is also not analytically tractable. Thus, the estimation of the parameters α, ω and φ is carried out based on the Markov chain Monte Carlo method in the JAGS platform.

RESULTS AND DISCUSSION

JAGS PROCEDURE

In this study, the Markov chain Monte Carlo (MCMC) method in the JAGS platform is carried out to estimate the parameters of the time-to-failure distribution and its percentiles for both models. To implement the JAGS algorithm successfully, two parts of the JAGS model have to be correctly identified: the likelihood function and the prior distributions. For this study, the likelihood functions have been given in Equations (10) and (16) while the prior distributions are provided in the previous section. For more illustration of the simulation process and determination of the posterior densities, the following flowchart is provided for the case of the power degradation model with the skew-normal degradation parameter (Figure 1).

SKEW NORMAL MODEL BASED ON DIFFERENT CHOICES OF PRIORS

In this subsection, the sensitivity of the estimated parameters and percentiles of the time-to-failure distribution based on the skew-normal model is presented under different choices of the prior distributions which are informative, weakly informative and non-informative priors. Under the Bayesian approach, the comparison between different choices of the priors is provided in terms of point estimated (PE), and standard deviation (SD). The data are simulated randomly by using Equation (8) based on different sample sizes $n = 30$, 60 and 200. The true values of the parameters μ, σ, β and φ are, respectively, assumed 1, 2, 3 and 6. Also, the critical degradation level is assumed to be 20.

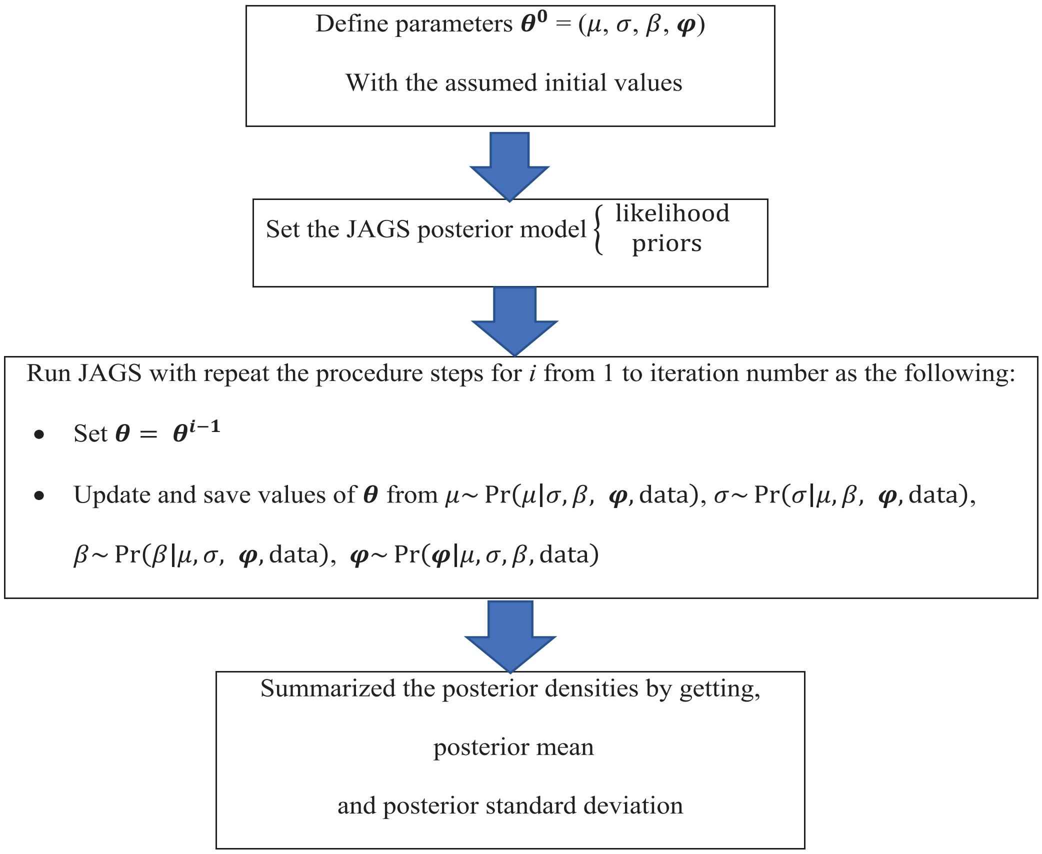

FIGURE 1. Flowchart for the posterior distribution of the skew-normal degradation parameter in estimation of the parameters using JAGS

Under the choices of the prior distribution, the informative prior for the parameters are assumed as $\mu \sim unif\ (0,\ 1)$, $\sigma \sim unif\ (0,\ 2)$, $\beta \sim gamma\ (3,\ 3)$ and $\varphi \sim unif\ (0,\ 6)$ while under weakly informative gamma prior, the hyper prior parameters r and s are assumed equal to 0.01 and 100, respectively, and the others as same as the informative prior. For non-informative priors, the distributions for all parameters are assumed uniform as in the case of informative priors instead for shape parameter is assumed to be $\beta \sim unif\ (0,\ 3)$. The results are provided in Table 1.

From the results found in Table 1, the following points are noted: a) 17 of 27 cases of the SD values of the estimated parameters and the percentiles for different choices of the prior of the skew-normal model decrease as n increases in different sample sizes, b) In different sample sizes, the PE of μ and β parameters of the skew-normal model under non-informative gamma prior are closer to the true values rather than the other choices of priors while the PE of σ and φ parameters of the informative gamma prior have closer values, c) 7 cases out of 15 of the PE of the percentiles of the skew-normal model under weakly informative gamma prior is closer to the true values rather than the other choices of priors in different sample sizes, d) 9 cases out of 12 for different sample sizes, the SD of the estimated parameters of the time-to-failure distribution by using the Bayesian approach of the non-informative prior are found less than those under the other choices of priors, and e) 9 cases out of 15 for different sample sizes, the SD of the estimated percentiles of the time to failure distribution by using Bayesian approach of the weakly informative prior are found less than those under the other choices of priors. According to these points, the Bayesian approach based on the weakly informative gamma prior is more precise in estimating the percentiles of the time-to-failure distribution under skew-normal distribution.

COMPARISON BETWEEN SKEW NORMAL AND LOG-LOGISTIC MODELS

Comparison between the posterior densities given in Equations (18) and (19) is made in terms of the point estimate, standard error (SE) and the deviance information criterion (DIC). For this comparison, the weakly informative gamma prior will be used for the skew-normal distribution based on the previous results. The data of the time-to-failure distributions are simulated randomly based on Equations (8) and (14) for the sample size n = 30, 60 and 200. In the case of degradation parameter following the skew-normal distribution, it is assumed that $\mu = 0$, $\sigma = 0.2$, $\beta = 3$ and $\varphi = 6$ when generating the data. The critical degradation level D_f is assumed equal to 20. Under the stated assumptions, JAGS software is iterated 100000 times where the first half of the iterations are treated as burn-in. To assess the overall goodness of fit of the models, values of DIC are determined. Spiegelhalter et al. (2002) define the DIC as the following:

$$DIC = \bar{D}(\boldsymbol{\theta}) + P_D$$

where $\boldsymbol{\theta}$ is the parameter interest; D is the deviance which is defined by – 2log (l); l is the likelihood function; $\bar{D}$ is the posterior deviance mean which is given based on the values of the estimated parameters of the posterior distribution; and P_D is the effective number of parameters of the model which is defined as $\bar{D}(\theta) - D(\bar{\theta})$ where $D(\bar{\theta})$ is the deviance which is found by finding the mean of each value of the estimated parameters in the posterior distribution. The smaller value of DIC indicates a better model. The results are provided in Table 2.

TABLE 1. PE and SD for the parameter and the percentiles of the skew-normal model based on informative, weakly informative and non-informative gamma prior for n = 30, 60 and 200

Parameter	Gamma prior	Sample size					
		n = 30		n = 60		n = 200	
		PE	SD	PE	SD	PE	SD
$\mu = 1$	Informative	0.658	0.207	0.691	0.189	0.732	0.178
	Weakly informative	0.809	0.152	0.738	0.175	0.588	0.172
	Non-informative	0.755	0.174	0.773	0.164	0.772	0.171
$\sigma = 2$	Informative	1.430	0.359	1.369	0.368	1.326	0.349
	Weakly informative	0.545	0.373	1.257	0.388	1.505	0.342
	Non-informative	1.203	0.375	1.065	0.348	1.205	0.306
$\beta = 3$	Informative	4.731	2.163	4.342	2.050	3.420	0.753
	Weakly informative	0.842	2.060	3.715	1.676	4.665	1.607
	Non-informative	2.232	0.558	2.099	0.616	2.578	0.314
$\varphi = 6$	Informative	2.838	1.250	3.444	1.393	3.580	1.455
	Weakly informative	1.248	1.261	3.188	1.431	3.765	1.402
	Non-informative	2.520	1.253	2.740	1.325	3.352	1.337
$t_{0.05} = 1.506$	Informative	2.129	0.694	1.934	0.427	1.848	0.242
	Weakly informative	1.579	0.652	2.047	0.529	1.631	0.213
	Non-informative	1.554	0.668	1.585	0.443	1.677	0.231
$t_{0.2} = 5.395$	Informative	7.148	1.973	5.793	1.054	5.702	0.528
	Weakly informative	6.349	1.748	6.288	1.072	4.748	0.492
	Non-informative	7.291	2.197	6.260	1.196	5.826	0.555
$t_{0.05} = 16.812$	Informative	25.646	8.538	17.402	3.560	16.309	1.662
	Weakly informative	18.14	4.51	18.531	3.548	14.444	1.647
	Non-informative	26.854	8.470	19.224	3.721	16.833	1.671
$t_{0.75} = 53.19$	Informative	103.618	53.981	55.696	15.268	48.423	6.501
	Weakly informative	44.68	13.39	56.804	14.914	47.944	7.057
	Non-informative	94.634	41.663	55.143	13.430	47.568	6.061
$t_{0.9} = 174.98$	Informative	464.18	415.601	189.019	78.177	150.057	29.335
	Weakly informative	106.66	53.22	182.953	72.552	167.649	36.310
	Non-informative	349.559	237.349	160.914	55.443	138.463	24.356

From the results found in Table 2, the following points are noted: a) 5 cases out of 9 of different sample sizes, under the power degradation model PE of the parameters of the time-to-failure distribution found based on the assumption of skew-normal degradation parameter is closer to the true parameter values than those found based on the assumption of the log-logistic degradation parameter. On the other hand, 9 cases out of 15 of the PE of the estimated percentiles is closer to the true value based on the assumption of log-logistic degradation parameter, b) Based on the different sample sizes, the SE of the estimated parameters and percentiles of the time-to-failure distribution found for both models decrease as n increases, c) For most of the sample sizes, the SE of the percentiles of the time-to-failure distribution for both models increase as p increases, d) 9 cases out of 24 of the different sample sizes, the SE of the estimated parameters and percentiles of the skew-normal model is less than the SE of those found for the log-logistic model while 7 cases out of 24 of the SE in both models are found equals, e) For all sample sizes, the DIC and the BIC for both models decrease as n increases, and f) For all sample sizes, the DIC and the BIC of the skew-normal model are found to be smaller than those found for the log-logistic model. According to these notes, it is clear that the Bayesian approach for the power degradation model with the skew-normal degradation parameter outperformed the power degradation model with the log-logistic degradation parameter.

NASA TURBOFAN JET ENGINE DATASET (NTJE) - DESCRIBING THE NTJE DATA

The data considered in this study is run-to-failure simulated data of turbofan jet engines which are available from NASA. The experiment of this NASA study is implemented for four different types of engines, denoted as FD001, FD002, FD003, and FD004, where each engine set contains a fleet of engines under different manufacturing operations, degrees of initial wear and run operational settings. In the assessment of the faults, 21 sensor measurements are made for each engine. The sensor measurements consist of two sub-data sets which are training and test data. The training data are obtained by letting all engines follow the process of the operation until failure while the test data are found until all engines arrive at a certain point of time before the failure. Accordingly, the NTJE data consists of 26 columns of numbers which are organized as engine number, cycles, operation settings 1,2 and 3 and 21 columns for sensor measurements. In addition, the vector of Remaining Useful Life (RUL) for the 100 engines were used in this study. More details on the data can be found in Chen et al. (2020) and Saxena et al. (2008).

CONVERGENCE ANALYSIS

The total number of MCMC iterations involving the three chains is set to be 100000 where the first 50000 iterations are treated as burn-in. Following Kundu (2008), the values of the hyperparameters b, r and s are all assumed equal to 2. To check for the convergence of the series produced based on the MCMC simulation, graphs of the trace plots, posterior density functions and autocorrelation of the parameters μ, σ, β and φ are determined and presented in Figure 2.

As shown in Figure 2, the results found based on the proposed model converge satisfactorily well. Also, it appears that the autocorrelation graph with thinning of 30 reaches zero rather quickly, indicating that the chain is mixing adequately. In addition, the values of the potential scale reduction factor (psrf) of Gelman-Rubin and Geweke's stationarity test for convergence assessment which are provided in Table 3 further support the case of convergence.

Generally, if the values of Geweke's stationarity test are all between -1.96 and 1.96 and the values of psrf for all the parameters are found less than 1.1, the chain is said to converge to a stationary distribution (Brooks & Gelman 1998; Gelman & Rubin 1992; Geweke 1991). It is apparent that the results from Table 3 satisfy the desired properties of convergence as outlined by those authors. Thus, the chains produced by the MCMC algorithm based on the model studied have successfully converged to a stationary distribution. Finally, a summary of the posterior distribution for the parameters with convergence statistics $\hat{R}$ based on the NTJE data in terms of mean and standard deviations is provided in Table 4. The convergence statistics $\hat{R}$ for all the parameters are found to be less than 1.1, which indicates that the resultant posterior distributions attained convergence (Gelman & Hill 2007).

DATA ANALYSIS

Here, the NTJE data are analyzed based on the sensitivity of the different choices of the priors and the choices of the distributions of degradation parameters as in simulated data. For the first case, the comparison is conducted to determine the type of prior which will give more precise results in terms of PE and SD. All the hyperparameter values were set equal to 2 for informative and non-informative priors, while for the weakly informative, the value of the hyperparameters of the gamma prior is similar as in the simulated data. The results found are provided in Table 5.

TABLE 2. True values, PE and SE for the parameters and percentiles of the time-to-failure distribution based on skew-normal and log-logistic parameters under the power degradation model based on simulated data for n = 30, 60, 200

Parameters	Skew-normal							Log-logistic						
	True values	n = 30		n = 60		n = 200		True values	n = 30		n = 60		n = 200	
		PE	SE	PE	SE	PE	SE		PE	SE	PE	SE	PE	SE
Scale	0.200	0.212	0.013	0.198	0.008	0.186	0.004	0.200	0.187	0.012	0.193	0.008	0.176	0.004
Shape	3.000	3.268	0.202	3.180	0.103	2.917	0.031	3.000	2.745	0.073	3.047	0.042	2.769	0.012
φ	6.000	4.964	0.385	4.961	0.273	4.878	0.152	6.000	4.843	0.395	5.006	0.268	5.014	0.148
$t_{0.05}$	0.924	0.908	0.007	0.916	0.004	0.914	0.001	1.068	1.064	0.003	1.075	0.002	1.060	0.001
$t_{0.2}$	1.049	1.057	0.004	1.055	0.002	1.050	0.001	1.164	1.168	0.005	1.179	0.002	1.155	0.001
$t_{0.5}$	1.176	1.214	0.006	1.198	0.003	1.188	0.001	1.272	1.296	0.008	1.298	0.003	1.269	0.001
$t_{0.75}$	1.319	1.396	0.012	1.364	0.005	1.343	0.001	1.415	1.480	0.015	1.457	0.006	1.425	0.002
$t_{0.9}$	1.486	1.613	0.020	1.559	0.008	1.525	0.002	1.650	1.818	0.039	1.724	0.014	1.697	0.004
DIC	-	-3.943		-18.357		-77.359		-	-3.683		-16.577		-66.687	
BIC	-	-8.001		-22.257		-136.013		-	-5.564		-15.710		-112.804	

TABLE 3. Geweke's stationarity test and the potential scale reduction factor to assess convergence of the MCMC chains for the parameters μ, σ, β and φ of the posterior distribution based on NTJE data under skew-normal model

Parameters	Geweke's stationarity test			psrf
	Chain 1	Chain 2	Chain 3	
μ	-0.26	1.24	-1.22	1.01
σ	-0.68	0.94	-1.29	1.00
β	-0.85	-0.98	0.15	1.01
φ	-0.16	0.63	-1.09	1.01

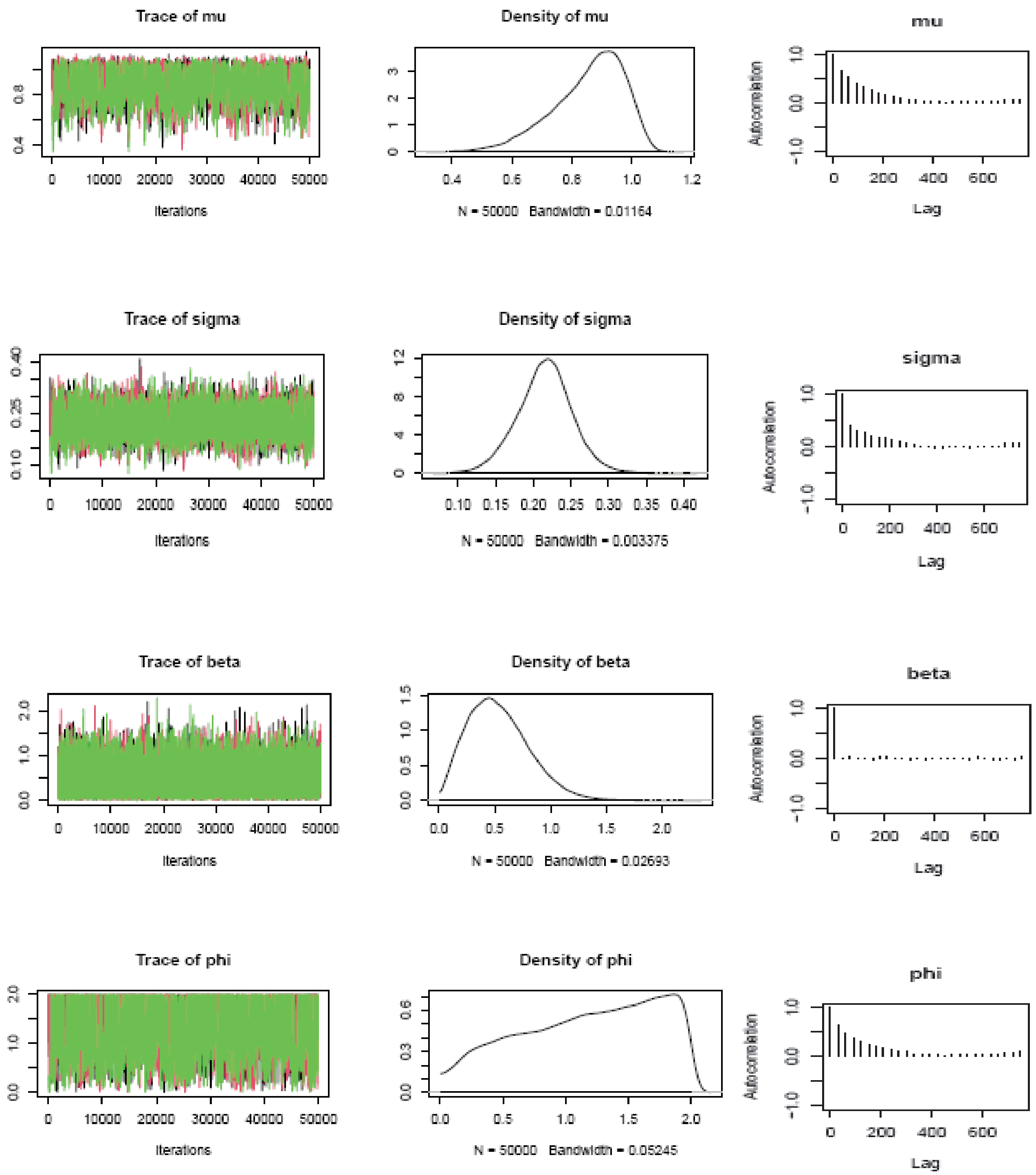

FIGURE 2. Trace plot, posterior density function and autocorrelation of the parameters μ, σ, β and φ based on the Bayesian analysis of NTJE data under skew-normal model

Table 5 shows that the results found are close for informative, weakly informative and non-informative priors. Even for larger data sizes, there are no big different in the values of the estimated parameters and the percentiles of the time to failure distribution found for the NTJE data, but the SD of the estimated parameters and the percentiles based on informative prior is slightly smaller. Comparison between the adequacy of the models for describing NTJE data is presented in terms of PE, SE, and DIC where all hyperparameters *r; s* and *b* of the prior distributions are all assumed equal 2. The results are provided in Table 6.

As noted in Table 6, the estimated parameters of the skew-normal model are more precise based on the analysis of NTJE data because the results indicate smaller SE for the parameter estimates found for the skew-normal model as opposed to those found for the log-logistic model. In addition, the skew-normal model has a smaller value of DIC. So, it is a better model in terms of goodness of fit for modeling the NTJE data.

TABLE 4. Summary of the posterior distribution based on the skew-normal degradation parameter with NTJE data

Parameters	Mean	SD	Quantiles					$\hat{R}$
			2.5%	25%	50%	75%	97.5%	
μ	0.855	0.119	0.580	0.783	0.876	0.944	1.033	1.006
σ	0.215	0.036	0.141	0.192	0.216	0.238	0.287	1.003
β	0.537	0.276	0.100	0.331	0.507	0.708	1.149	1.002
φ	1.181	0.537	0.151	0.754	1.240	1.647	1.966	1.007

TABLE 5. PE and SE of the parameters and the percentiles of the skew-normal model based on the NTJE data

Parameters	Informative prior		Weakly informative prior		Non-informative prior	
	PE	SD	PE	SD	PE	SD
μ	0.853	0.118	0.853	0.119	0.879	0.119
σ	0.216	0.037	0.218	0.037	0.208	0.033
β	0.550	0.279	0.578	0.289	0.381	0.270
φ	1.179	0.527	1.196	0.534	1.187	0.525
$t_{0.05}$	14.454	1.865	14.471	1.875	14.514	1.887
$t_{0.2}$	28.430	2.815	28.445	2.838	28.580	2.829
$t_{0.5}$	58.328	5.053	58.366	5.085	58.507	5.015
$t_{0.75}$	104.680	10.312	104.842	10.318	104.475	10.124
$t_{0.9}$	178.615	22.326	179.173	22.343	176.920	21.526

TABLE 6. PE and SE of the parameters and the percentiles of the time-to-failure distribution based on skew-normal model and log-logistic model and DIC found for the models using the NTJE data

Parameters	Skew-normal		Log-logistic	
	PE	SE	PE	SE
Scale	0.216	0.004	0.916	0.014
Shape	0.550	0.028	7.633	0.066
φ	1.179	0.053	1.055	0.057
$t_{0.05}$	14.450	0.187	16.620	0.201
$t_{0.2}$	28.430	0.282	31.410	0.320
$t_{0.5}$	58.330	0.505	62.610	0.593
$t_{0.75}$	104.700	1.031	119.600	1.421
$t_{0.9}$	178.600	2.233	257.600	4.626
DIC	1066.900		1085.300	

CONCLUSIONS

Time-to-failure distribution for the power degradation model is derived based on two assumptions of the distribution of the degradation parameter which are the skew-normal and the log-logistic distributions. By using the Bayesian technique, the parameters and percentiles of the time-to-failure distribution are estimated. Based on some convergence tests, the chains are found mixing well and attaining convergence for the posterior distribution of the parameters of the skew-normal distribution and the fixed-effect parameter. This is also the case for the parameters of the log-logistic distribution and the fixed-effect parameter. The Bayesian estimator of the parameters and the percentiles of time-to-failure distribution based on the skew-normal and the log-logistic models are compared using the simulated data and NTJE data. Based on point estimate, standard deviation, standard error, and deviance information criteria, the Bayesian technique for the power degradation model with the skew-normal distribution outperformed the Bayesian for the other model because the parameters of the time-to-failure distribution and its percentiles are estimated with higher precision.

ACKNOWLEDGEMENTS

The authors are thankful for the financial support funded for this research by Universiti Kebangsaan Malaysia in the form of grants Geran Galakan Penyelidik Muda (no. GGPM-2023-052) and Tabung Agihan Penyelidikan (no. TAP-K017073).

REFERENCES

Alhamidie, A.A., Kamarulzaman Ibrahim, Alodat, M.T. & Wan Zawiah Wan Zin. 2019. Bayesian inference for linear regression under alpha-skew-normal prior. *Sains Malaysiana* 48(1): 227-235.

Azzalini, A. 1985. A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics* 12(2): 171-178.

Dakhn, L.N.B., Mohd Aftar Abu Bakar & Kamarulzaman Ibrahim. 2023. Bayesian estimation of time to failure distributions based on skew normal degradation model: An application to GaAs laser degradation data. *Sains Malaysiana* 52(2): 641-653. https://doi.org/http://doi.org/10.17576/jsm-2023-5202-24

Dakhn, L.N.B., Ebrahem, M.A-Haj. & Eidous, O. 2017. Semi-parametric method to estimate the time-to-failure distribution and its percentiles for simple linear degradation model. *Journal of Modern Applied Statistical Methods* 16(2): 322-346. https://doi.org/10.22237/jmasm/1509495420

Brooks, S.P. & Gelman, A. 1998. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 7(4): 434-455. https://doi.org/10.1080/10618600.1998.10474787

Chen, S., Wang, M., Huang, D., Wen, P., Wang, S. & Zhao, S. 2020. Remaining useful life prediction for complex systems with multiple indicators based on particle filter and parameter correlation. *IEEE Access* 8: 215145-215156. https://doi.org/10.1109/ACCESS.2020.3041682

Eidous, O., Ebrahem, M.A-Haj. & Dakhn, L.N.B. 2017. Estimating the time-to-failure distribution and its percentiles for simple linear degradation model using double kernel method. *Journal of Probability and Statistical Science* 15(1): 121-134.

Gelman, A. & Hill, J. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press.

Gelman, A. & Rubin, D.B. 1992. Inference from iterative simulation using multiple sequences. *Statistical Science* 7(4): 457-511.

Geweke, J. 1991. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. *The Fourth Valencia International Meeting on Bayesian Statistics,* Peniscola, Spain, April 15-20.

Ghaderinezhad, F., Christophe, L. & Nicola, L. 2020. Bayesian inference for skew-symmetric distributions. *Symmetry* 12(4): 491. https://doi.org/10.3390/sym12040491

Gonzalez, J.R., Vazquez, M., Algora, C. & Nunez, N. 2010. Real-time reliability test for a CPV module based on a power degradation model. *Progress in Photovoltaics: Research and Applications* 20(1): 6-11. https://doi.org/10.1002/pip

Guure, C.B. & Noor Akma Ibrahim. 2014. Approximate Bayesian estimates of Weibull parameters with Lindley's method. *Sains Malaysiana* 43(9): 1433-1437.

Kundu, D. 2008. Bayesian inference and life testing plan for the Weibull distribution in presence of progressive censoring. *Technometrics* 50(2): 144-154. https://doi.org/10.1198/004017008000000217

Meeker, W.Q. & Escobar, L.A. 1998. *Statistical Method for Reliability Data*. New York: John Wiley & Sons, Inc.

Oliveira, R.P.B., Loschi, R.H. & Freitas, M.A. 2018. Skew-heavy-tailed degradation models: An application to train wheel degradation. *IEEE Transactions on Reliability* 67(1): 129-141. https://doi.org/10.1109/TR.2017.2765485

Pan, D., Liu, J-B. & Yang, W. 2018. A new result on lifetime estimation based on Skew-Wiener degradation model. *Statistics and Probability Letters* 138: 157-164. https://doi.org/10.1016/j.spl.2018.03.009

Puggard, W., Niwitpong, S-A. & Niwitpong, S. 2022. Comparison analysis on the coefficients of variation of two independent Birnbaum-Saunders distributions by constructing confidence intervals for the ratio of coefficients of variation. *Sains Malaysiana* 51(7): 2265-2281.

Saxena, A., Goebel, K., Simon, D. & Eklund, N. 2008. Damage propagation modeling for aircraft engine run-to-failure simulation. *International Conference on Prognostics and Health Management, PHM 2008*. https://doi.org/10.1109/PHM.2008.4711414

Sen, S., Maiti, S.S. & Chandra, N. 2016. The Xgamma distribution: Statistical properties and application. *Journal of Modern Applied Statistical Methods* 15(1): 774-788. https://doi.org/10.22237/jmasm/1462077420

Shafiq, M., Alamgir & Muhammad Atif. 2016. On the estimation of three parameters lognormal distribution based on fuzzy life time data. *Sains Malaysiana* 45(11): 1773-1777.

Siju, K.C. & Kumar, M. 2018. Bayesian estimation of reliability using time-to-failure distribution of parametric degradation models. *Journal of Statistical Computation and Simulation* 88(9): 1717-1748. https://doi.org/10.1080/00949655.2018.1445743

Spiegelhalter, D.J., Best, N.G., Carlin, B.P. & Van Der Linde, A. 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society* 64(4): 583-639.

Wu, J., Wu, C., Cao, S., Or, S.W., Deng, C. & Shao, X. 2019. Degradation data-driven time-to-failure prognostics approach for rolling element bearings in electrical machines. *IEEE Transactions on Industrial Electronics* 66(1): 529-539. https://doi.org/10.1109/TIE.2018.2811366

Yadav, A.S., Saha, M., Tripathi, H. & Kumar, S. 2021. The exponentiated Xgamma distribution: A new monotone failure rate model and its applications to lifetime data. *Statistica* 81(3): 303-334. https://doi.org/10.6092/issn.1973-2201/10578

*Corresponding author; email: rrmt@ukm.edu.my