Journal of Data Science and Modeling, Vol. 3, No. 1, 255-280, December 2024

Research Manuscript

Bayesian Analysis of the Weighted Marshall-Olkin Bivariate Exponential Model

Ali Sakhaei, Iman Makhdoom*

Department of Statistics, Payame Noor University, Tehran, Iran.

Recieved: 20/05/2025 Accepted: 31/03/2026

Abstract: The Weighted Marshall-Olkin Bivariate Exponential (WMOBE) distribution was first proposed by Jamalizadeh and Kundu (2013), who examined its different characteristics and properties. Bayesian estimation of the model parameters is carried out using both the squared error loss (SEL) function, which is symmetric, and the linear-exponential (LINEX) loss function, which is asymmetric. These estimators are derived under both informative and non-informative gamma priors. Given the complexity of the four-parameters model, explicit analytical solutions for the Bayesian estimators are not attainable, making it necessary to employ the Gibbs sampling procedure. Markov Chain Monte Carlo (MCMC) methods are widely utilized to compute and implement these estimates. Furthermore, the convergence behavior of the Markov chain toward a stationary distribution is carefully analyzed. Credible intervals, particularly the highest posterior density (HPD) intervals for the unknown parameters, are also constructed. To assess and compare the effectiveness of these estimation approaches, Monte Carlo simulations are performed. Finally, the methodology is applied to a real-world dataset for illustrative purposes.

Keywords: Bayesian estimation, Markov Chain Monte Carlo, Gibss sampling, Weighted Marshall-Olkin Bivariate Exponential distribution.

Mathematics Subject Classification (2010): 62F15, 62N05,; 62C10, 65C60

*Corresponding Author: makhdoom@pnu.ac.ir

1. Introduction

In a recent study, Gupta and Kundu (2009) introduced a new shape parameter specifically for the exponential distribution, naming it the "weighted exponential (WE) distribution." Building upon the framework established by Azzalini (1985), their work examined various characteristics of this distribution. They discovered several interesting features of the WE distribution, such as its unimodal probability density function (PDF) and the monotonic increase in its corresponding hazard function (HF). The two-parameter WE distribution shares similarities with other well-known two-parameter distributions, like the Weibull, gamma, and generalized exponential distributions. It has several desirable properties and can also be viewed as a hidden truncation model. In some cases, it offers a better fit than the previously mentioned two-parameter distributions. Basikhasteh and Makhdoom (2022) have presented the Bayesian inference of bivariate Weibull geometric model based on LINEX and quadratic loss functions.

Supporting the findings of Gupta and Kundu (2009), Shahbaz et al. (2010) proposed a three-parameter weighted Weibull (WW) model, which generalizes the WE distribution. Al-Mutairi et al. (2011) reported the existence of a continuous bivariate distribution with WE marginals. More recently, Jamalizadeh and Kundu (2013) introduced the weighted Marshall-Olkin bivariate exponential (WMOBE) distribution, utilizing a method inspired by Azzalini (1985), which significantly differs from the approach of Al-Mutairi et al. (2011). The WMOBE distribution, which has four parameters, is singular and can also be considered a hidden truncation model, similar to the one proposed by Arnold and Beaver (2000). The interpretation of multivariate hidden truncation models was explored by Arnold et al. (2002), which helped substantiate the WMOBE model proposed by Jamalizadeh and Kundu (2013). This contributed significantly to the development of their model. Recently, Makhdoom and Sakhaei (2023) modeled the dependence structure between the margins of the WMOBE distribution using copula functions, providing a flexible framework for capturing complex dependencies. However, their analysis was limited to classical inference methods, and no attempt has been made to study this model within a Bayesian framework.

Additionally, Kotz et al. (2000) highlighted that among singular bivariate distributions, the three-parameter Marshall-Olkin bivariate exponential (MOBE) distribution and the four-parameter Marshall-Olkin bivariate Weibull (MOBW) distribution are the most commonly utilized. Considering the finding that the WE distribution can offer a superior fit compared to the Weibull or exponential distributions in specific scenarios (refer to Gupta and Kundu (2009)), it is reasonable to expect that the WMOBE model may also outperform the MOBE or MOBW model in certain cases. This expectation has indeed been validated by Jamalizadeh and Kundu (2013). They acquired estimates of the models using the expectation-maximization (EM) algorithm. It is worth noting that in certain cases, the EM algorithm may yield highly biased estimates, resulting in negative variances for the maximum likelihood estimates (MLEs) according to the observed Fisher information. Consequently, the MLEs can be somewhat misleading. To address this issue and overcome these challenges, adopting Bayesian inference as a viable alternative becomes an attractive option. Therefore, the problem addressed in this paper primarily stems from the aforementioned concerns. Interestingly, none of the aforementioned researchers have attempted the Bayesian approach, which motivates us to explore Bayesian inference in this paper. The primary objective of this study is to provide a comprehensive investigation within the Bayesian framework. To achieve this, we consider the WMOBE distribution as the focal model, as proposed by Jamalizadeh and Kundu (2013), who thoroughly examined its properties. The structure of this paper is as follows: **Section 2** introduces the WMOB model and explains the likelihood formulation. **Section 3** provides an overview of the Bayesian framework and discusses prior distribution selection. In **Section 4**, we outline the Bayesian estimation process using the MCMC technique to estimate the unknown parameters. **Section 5** presents an extensive simulation study to evaluate the accuracy of the estimates. The application of the proposed model to a real-world dataset is discussed in **Section 6**, followed by a summary of key findings and conclusions in **Section 7**.

2. Model definition: problem formulation and assumptions

This section provides a precise definition and characterization of the WMOBE model. The random vector (Y_1, Y_2) is said to follow the WMOBE distribution when the joint probability density function (PDF) of (Y_1, Y_2) is expressed as presented by Jamalizadeh and Kundu (2013):

$$g(y_1, y_2) = \begin{cases} g_1(y_1, y_2) & \text{if} \quad y_2 > y_1 > 0, \\ g_2(y_1, y_2) & \text{if} \quad y_1 > y_2 > 0\ , \\ g_0(y_1, y_2) & \text{if} \quad y_1 = y_2 = y. \end{cases} \tag{2.1}$$

In which

$$g_1(y_1, y_2) = \frac{\alpha + \lambda}{\alpha} \lambda_1 \exp(-\lambda_1 y_1)(\lambda_2 + \lambda_3) \exp(-(\lambda_2 + \lambda_3) y_2)(1 - \exp(-y_1 \alpha)),$$

$$g_2(y_1, y_2) = \frac{\alpha + \lambda}{\alpha} (\lambda_1 + \lambda_3) \exp(-(\lambda_1 + \lambda_3) y_1) \lambda_2 \exp(-(\lambda_2 y_2)(1 - \exp(-y_2 \alpha)),$$

$$g_0(y) = \frac{\alpha+\lambda}{\alpha}\lambda_3 \exp(-\lambda y)(1-\exp(-y\alpha)),$$

where $\lambda = \lambda_1 + \lambda_2 + \lambda_3$, and $\theta = (\alpha, \lambda_1, \lambda_2, \lambda_3)$ is also the vector parameter of the model. Moving forward, we will use the notation WMOBE $(\alpha, \lambda_1, \lambda_2, \lambda_3)$ to represent a WMOBE distribution with the probability density function (PDF) (2.1). The joint survival function and moment generating function (MGF) of the WMOBE$(\alpha, \lambda_1, \lambda_2, \lambda_3)$ have also been demonstrated in Jamalizadeh and Kundu (2013). We show the perspective plot of the absolute continuous part of the joint PDF of (2.1) in Figure 1 for the various parametes.

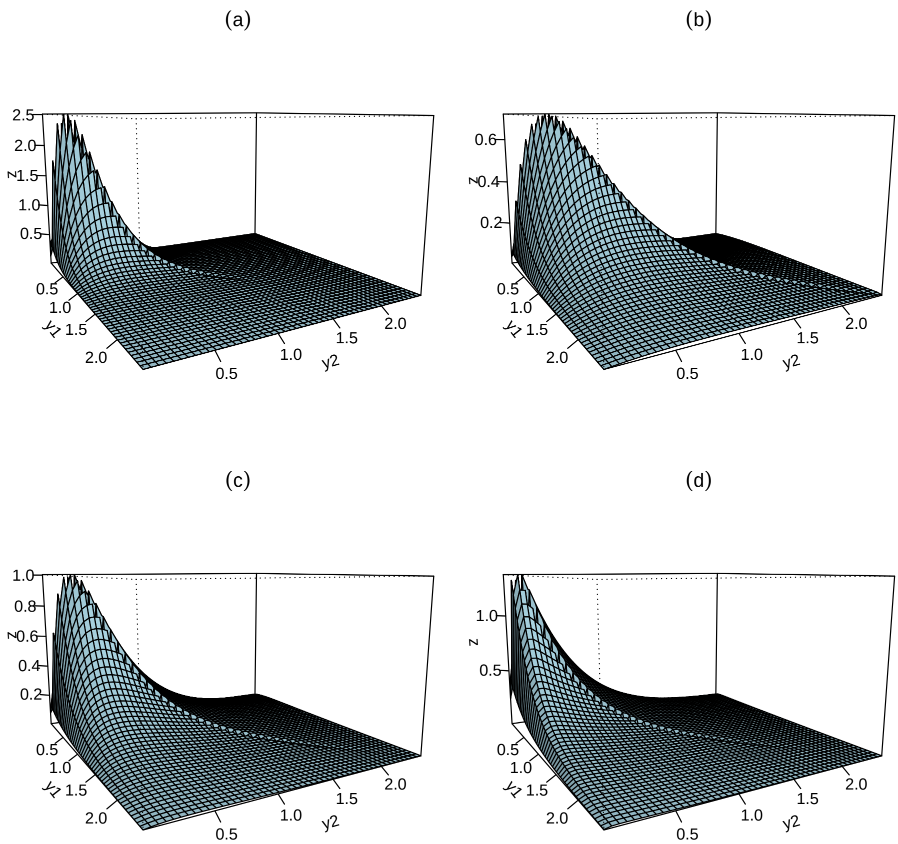

Figure 1: The perspective plots of the joint PDF of (1) (a) $\alpha = 0.5, \lambda_1 = \lambda_2 = \lambda_3 = 2$ (b) $\alpha = 0.5, \lambda_1 = \lambda_2 = \lambda_3 = 1$ (c) $\alpha = 5, \lambda_1 = \lambda_2 = \lambda_3 = 1$ (d) $\alpha = 25, \lambda_1 = \lambda_2 = \lambda_3 = 1$

2.1 Generation method of the WMOBE model

The generation a random sample from the WMOBE $(\alpha, \lambda_1, \lambda_2, \lambda_3)$ can be easily accomplished by following these steps:

- Step 1: Start by generating three independent exponential random variables, denoted as U_1,U_2 and U_0, with parameters λ_1, λ_2 and λ_3 respectively. Then, define
$$(X_1, X_2) \stackrel{d}{=} (\min\{U_0, U_1\}, \min\{U_0, U_1\})$$
where $\stackrel{d}{=}$ means equal in distribution.

- Step 2: Let $W \sim \exp(\alpha)$, then
$$Y_1 \stackrel{d}{=} X_1 | W < \min\{X_1, X_2\} \qquad and \qquad Y_2 \stackrel{d}{=} X_2 | W < \min\{X_1, X_2\}$$

Subsequently, the random vector (Y_1, Y_2) is characterized as having a WMOBE distribution parameterized by $\boldsymbol{\eta} = (\alpha, \lambda_1, \lambda_2, \lambda_3)$. It is noteworthy that as α approaches infinity, the distribution of (Y_1, Y_2) converges in distribution to (X_1, X_2),ie $(Y_1, Y_2) \stackrel{d}{\to} (X_1, X_2)$ where the symbol $\stackrel{d}{\to}$ denotes convergence. Therefore, the MOBE distribution can be obtained as a limiting distribution within the WMOBE family.

Theorem 2.1 (Jamalizadeh and Kundu (2013))**.** *Let* $(Y_1, Y_2) \sim WMOBE(\boldsymbol{\eta})$*, then*

- *1 :* $Y_1 \sim WE((\alpha + \lambda_2)/(\lambda_1 + \lambda_3), \lambda_1 + \lambda_3)$*,*
- *2 :* $Y_2 \sim WE((\alpha + \lambda_1)/(\lambda_2 + \lambda_3), \lambda_2 + \lambda_3)$*, and*
- *3 :* $\min\{Y_1, Y_2\} \sim WE(\alpha/\lambda, \lambda)$*.*

2.2 The likelihood structure

Within this subsection, we present the likelihood function of the model. Let $\{(y_{1i}, y_{2i}), i = 1, \ldots, n\}$ represent a random sample of size n drawn from the WMOBE$(\alpha, \lambda_1, \lambda_2, \lambda_3)$. The subsequent notations will be utilized:

$$I = \{1, \ldots, n\}, I_0 = \{i \in I; y_{1i} = y_{2i} = y_i\},$$
$$I_1 = \{i \in I; y_{1i} < y_{2i}\}, I_2 = \{i \in I; y_{1i} > y_{2i}\},$$

also, n_0, n_1, and n_2 define the number of elements in I_0, I_1, and I_2, respectively. We obtain the likelihood function based on the observed sample as follows:

$$L(\alpha, \lambda_1, \lambda_2, \lambda_3 | data) = \prod_{i=1}^{n} g(y_{1i}, y_{2i}) = \prod_{i \in I_1} g_1(y_{1i}, y_{2i}) \prod_{i \in I_2} g_2(y_{1i}, y_{2i}) \prod_{i \in I_0} g_0(y_i, y_i), \tag{2.2}$$

in which g_0,g_1 and g_2 are introduced in Eq.(2.1).

2.3 Maximum-likelihood estimation

Based on the likelihood function presented in Eq.(2.2), the contribution to the log-likelihood can be expressed as follows:

$$\begin{aligned}\ell(\alpha,\lambda_1,\lambda_2,\lambda_3) &= \sum_{I_0}\ln g_0(y_i,y_i)+\sum_{I_1}\ln g_1(y_{1i},y_{2i})+\sum_{I_2}\ln g_2(y_{1i},y_{2i})\\ &= \sum_{i\in l_0}\ln(1-\exp(-\alpha y_i))+\sum_{i\in I_1}\ln(1-\exp(-\alpha y_{1i}))\\ &+\sum_{i\in I_2}\ln(1-\exp(-\alpha y_{2i}))+n_0\ln\lambda_3+n_1\ln\lambda_1+n_2\ln\lambda_2\\ &+n_1\ln(\lambda_2+\lambda_3)+n_2\ln(\lambda_2+\lambda_3)-\lambda_1\sum_{i\in I_0}y_i\\ &-\lambda_2\left(\sum_{i\in I_0}y_i+\sum_{i\in I_1\cup I_2}\right)-\lambda_2\left(\sum_{i\in I_0}y_i+\sum_{i\in I_1\cup I_2}y_{2i}\right)\\ &-\lambda_3\left(\sum_{i\in I_0}y_i+\sum_{i\in I_2}y_{1i}+\sum_{i\in I_1}y_{2i}\right)+n\ln(\alpha+\lambda)-n\ln\alpha \qquad (2.3)\end{aligned}$$

In this method, the estimates are consistent and asymptotically efficient. However, it can be computationally demanding in high-dimensional vector parameters. To address this issue, Makhdoom and Sakhaei (2023) proposed an alternative approach to parameter estimation. Jamalizadeh and Kundu (2013) obtained the MLEs of the model through an example, there is a lack of simulation studies to collectively evaluate the estimates. Hence, in this paper, we were motivated to enhance their work by a Bayesian approach.

3. Bayesian notation and prior choice

In recent decades, the Bayesian approach has emerged as a robust and credible alternative to traditional statistical methods. This perspective treats model parameters as random variables that follow a specific distribution, referred to as the prior distribution. Bayesian inference has gained significant attention in statistical analysis.

Consider a random sample $\mathbf{y}=(\mathbf{y_1},\mathbf{y_2})$ of size n drawn from the WMOBE($\boldsymbol{\eta}$). The initial step in Bayesian analysis involves formulating the joint posterior function of the parameters. Given that $L(\mathbf{y}|\boldsymbol{\eta})$ represents the likelihood function and $\pi(\boldsymbol{\eta})$ denotes the prior density of the parameter $\boldsymbol{\eta}$, the corresponding posterior density function, $\pi(\boldsymbol{\eta}|\mathbf{y})$, is expressed as

$$\pi(\boldsymbol{\eta}|\mathbf{y}) \propto \pi(\boldsymbol{\eta})L(\mathbf{y}|\boldsymbol{\eta}).$$

For the model under consideration, the joint posterior distribution of the parameters given the observed data is written as:

$$\pi(\alpha, \lambda_1, \lambda_2, \lambda_3 \mid \mathbf{y}) \propto \pi(\alpha, \lambda_1, \lambda_2, \lambda_3)\, L(\mathbf{y} \mid \alpha, \lambda_1, \lambda_2, \lambda_3), \tag{3.4}$$

where $\mathbf{y} = (\mathbf{y_1}, \mathbf{y_2})$ denotes the observed data.

Assuming that the parameters are mutually independent and follow Gamma prior distributions, i.e., $(\alpha, \lambda_1, \lambda_2, \lambda_3) \sim$ independent Gamma(a_i, b_i), $(i = 1, 2, 3, 4)$, the joint prior distribution of $\boldsymbol{\eta} = (\alpha, \lambda_1, \lambda_2, \lambda_3)$ can be written as

$$\pi(\boldsymbol{\eta}) = \pi_1(\alpha)\pi_2(\lambda_1)\pi_3(\lambda_2)\pi_4(\lambda_3), \tag{3.5}$$

where

$$\begin{aligned}
\pi_1(\alpha) &\propto \alpha^{a_1-1}\exp(-b_1\alpha), \quad \alpha > 0,\ a_1, b_1 > 0,\\
\pi_2(\lambda_1) &\propto \lambda_1^{a_2-1}\exp(-b_2\lambda_1), \quad \lambda_1 > 0,\ a_2, b_2 > 0,\\
\pi_3(\lambda_2) &\propto \lambda_2^{a_3-1}\exp(-b_3\lambda_2), \quad \lambda_2 > 0,\ a_3, b_3 > 0,\\
\pi_4(\lambda_3) &\propto \lambda_3^{a_4-1}\exp(-b_4\lambda_3), \quad \lambda_3 > 0,\ a_4, b_4 > 0.
\end{aligned}$$

By rewriting Eq.(3.4), the joint posterior distribution is given by

$$\begin{aligned}
\pi(\alpha, \lambda_1, \lambda_2, \lambda_3)|data) \propto\ &\{\prod_{i\in I_1} g_1(y_{1i}, y_{2i}) \prod_{i\in I_2} g_2(y_{1i}, y_{2i}) \prod_{i\in I_0} g_0(y_{1i}, y_{2i})\}\\
&\times \{\alpha^{a_1-1}\exp(-b_1\alpha)\}\{\lambda_1^{a_2-1}\exp(-b_2\lambda_1)\}\\
&\times \{\lambda_2^{a_3-1}\exp(-b_3\lambda_2)\}\{\lambda_3^{a_4-1}\exp(-b_4\lambda_3)\}
\end{aligned} \tag{3.6}$$

3.1 Loss functions

In Bayesian inference, unknown parameters are considered random variables and are characterized by prior and posterior distributions. During estimation, we approximate the parameter θ with $\hat{\theta}$, which may not perfectly match the true value of θ. To account for any discrepancies in this approximation, we introduce a loss function $l(\hat{\theta}, \theta)$ to quantify the loss incurred when θ is estimated by $\hat{\theta}$. The objective during model fitting is to minimize this loss function. If $\hat{\theta}$ equals θ, no loss occurs; if $\hat{\theta}$ falls below θ, it is considered underestimation, while a value above θ is deemed overestimation. Instead of calculating loss from a single sample, we consider the expected loss to obtain a more reliable measure of the true loss. Less sensitive loss functions can produce more robust parameter estimates (see Terven et al. (2023)). As noted in Dey (2010), the behavior of the Bayes estimate can vary with the choice of loss function in specific situations. Consequently, our Bayesian analysis examines both symmetric and asymmetric loss functions, as detailed in Sections 4.1, 4.2, and 4.3. The parameter estimate $\hat{\theta}$ that minimizes the expected posterior loss is known as the Bayes estimator.

3.1.1 The SEL function

The square error loss (SEL) function is a symmetric approach for evaluating loss. It is often central in discussions about the Bayes estimator because it aims to minimize the expected squared error. In this framework, the Bayes estimator is taken as the mean of the posterior distribution. This loss function treats overestimation and underestimation with equal significance and places more weight on larger errors. The function is defined as:

$$l(\theta, \hat{\theta}) = (\hat{\theta} - \theta)^2, \tag{12}$$

where $\hat{\theta}$ is the estimate of the parameter θ. According to this loss function, shown in equation (12), the posterior mean of $\theta|x$ minimizes the expected squared error. Thus, the Bayes estimator is given in equation (13) as:

$$\hat{\theta}_{BS} = E_{\theta}(\theta|x) \tag{13}$$

where E_{θ} denotes the posterior expectation of θ, which is assumed to be finite and well-defined.

Under the SEL function hence, $l(\theta, \hat{\theta}) = (\theta - \hat{\theta})^2$, the Bayes estimate of a function $\Phi(\boldsymbol{\eta})$, namely $\hat{\Phi}_{SEL}$ is the posterior mean as follows:

$$\hat{\Phi}_{SEL} = E[\Phi(\boldsymbol{\eta})|data] = \frac{\int \Phi(\boldsymbol{\eta})\pi(\theta, \boldsymbol{\eta})|data)d\boldsymbol{\eta}}{\int \pi(\boldsymbol{\eta})|data)d\boldsymbol{\eta}}, \tag{3.7}$$

where $\boldsymbol{\eta} = (\alpha, \lambda_1, \lambda_2, \lambda_3)$.

3.1.2 The linear exponential loss function

The Linear Exponential (LINEX) loss function is an asymmetric measure of estimation error that has been widely used in statistical applications. It is defined as:

$$l(\theta, \hat{\theta}) \propto e^{c(\hat{\theta} - \theta)} - c(\hat{\theta} - \theta) - 1, \quad c \neq 0. \tag{3.8}$$

Studies by Zellner (1986), Calabria and Pulcini (1996), and Chang and Hung (2007) highlight that the LINEX function exhibits different growth patterns on either side of zero, with one side increasing nearly linearly and the other side growing exponentially. As discussed in Nassar et al. (2022), this function is useful for addressing biases in both underestimation and overestimation. The parameter c sets the direction and magnitude of this asymmetry: a positive c results in a stronger penalty for overestimation, while a negative c does the same for underestimation. As c approaches zero, the LINEX function becomes more balanced and gradually approximates the symmetric Square Error Loss Function (SELF).

The Bayes estimate of a function $\Phi(\alpha, \lambda_1, \lambda_2, \lambda_3)$ under the LINEX loss function, namely $L(\theta, \hat{\theta}) = b(\exp(a(\hat{\theta} - \theta)) - a(\hat{\theta} - \theta) - 1), a \neq 0, b > 0$ is presented by the below form:

$$\begin{aligned}\hat{\Phi}_{LINEX} &= -\frac{1}{C} \times \log(E(e^{-C\Phi(\boldsymbol{\eta})}|data)) \\ &= \frac{-1}{C} \times \log \frac{\int \exp(-C\phi(\theta, \boldsymbol{\eta}))\pi(\boldsymbol{\eta})|data)d\boldsymbol{\eta}}{\int \pi(\boldsymbol{\eta})|data)d\boldsymbol{\eta}},\end{aligned} \tag{3.9}$$

where $\boldsymbol{\eta} = (\alpha, \lambda_1, \lambda_2, \lambda_3)$.
Since solving the integral ratios in Eq. (3.7) and Eq. (3.9) analytically is not feasible, numerical integration techniques must be employed. Some common approaches include Lindley's approximation Lindley (1980), the $T-K$ approximation Tierney and Kadane (1986), and Markov Chain Monte Carlo (MCMC) methods. In this study, we focus on obtaining the Bayes estimates using the MCMC approach.

4. Bayesian inferences: MCMC

Markov Chain Monte Carlo (MCMC) methods comprise a set of algorithms designed for sampling from probability distributions by creating Markov chains that converge to the target distribution. In these methods, a Markov chain is produced step-by-step, with each state relying solely on the preceding state, directed by a proposal distribution. By repeating this process, the Markov chain ultimately achieves a stationary distribution that closely approximates the desired target distribution. MCMC has transformed statistical inference, allowing practitioners to conduct Bayesian analysis, estimate intricate models, and effectively explore high-dimensional parameter spaces. Some of the widely used MCMC techniques include the Metropolis-Hastings algorithm, Gibbs sampling, and Hamiltonian Monte Carlo (HMC) (Robert et al. (1999)). In the Bayesian paradigm, MCMC techniques rely on computer simulations to construct a Markov sequence that exhibits ergodic properties, guaranteeing convergence to a stationary distribution. The implementation of MCMC methods assumes that, conditioned on the explanatory variables and the full set of parameters, the observations are independent, and the prior distributions of all parameters are likewise mutually independent. The posterior density function is derived by refining the prior distribution through the incorporation of the likelihood function.

If the squared error loss function is considered, the Bayesian estimates for the parameters $\boldsymbol{\eta} = (\alpha, \lambda_1, \lambda_2, \lambda_3)$ correspond to their respective posterior means. However, extracting samples directly from the posterior function is computationally challenging. To address this, the widely recognized **Gibbs sampling** technique is employed. This method forms a significant category of MCMC approaches commonly used in Bayesian inference. Additionally, Gibbs sampling represents a specific instance of the broader Metropolis-Hastings (MH) algorithm Hanagal and Ahmadi (2009), making it a multivariate extension of the MH method.

In the subsequent section, the Gibbs sampling approach is applied to numerically compute Bayesian estimates. Since the posterior probability density function (PDF) of $\boldsymbol{\eta} = (\alpha, \lambda_1, \lambda_2, \lambda_3)$ in Eq. (3.6) is unknown and its full conditional densities cannot be derived explicitly, the MH method is required to generate samples from the joint posterior distributions. The Gibbs sampling algorithm operates as follows:

Step 1 : Start with an initial guess value $\boldsymbol{\eta}^{(0)} = (\alpha^{(0)}, \lambda_1^{(0)}, \lambda_2^{(0)}, \lambda_3^{(0)})$.

Step 2 : Generate $\alpha^{(i)}$ from its proposal density and set

$$\boldsymbol{\eta}^{(i)} = (\alpha^{(i)}, \lambda_1^{(i-1)}, \lambda_2^{(i-1)}, \lambda_3^{(i-1)})$$

then accept $\alpha^{(i)}$ with probability:

$$\rho_1 = \min\left\{\frac{\pi(\boldsymbol{\eta}^{(i)}|\mathbf{y_1}, \mathbf{y_2})}{\pi(\boldsymbol{\eta}^{(i-1)}|\mathbf{y_1}, \mathbf{y_2})} \frac{q(\alpha^{(i-1)})}{q(\alpha^{(i)})}, 1\right\}.$$

Step 3 : Generate $\lambda_0^{(i)}$ from its proposal density and set

$$\boldsymbol{\eta}^{(i)} = (\theta^{(i)}, \alpha^{(i)}, \lambda_0^{(i)}, \lambda_1^{(i-1)}, \lambda_2^{(i-1)}),$$

then accept $\lambda_1^{(i)}$ with probability:

$$\rho_2 = \min\left\{\frac{\pi(\boldsymbol{\eta}^{(i)}|\mathbf{y_1}, \mathbf{y_2})}{\pi(\boldsymbol{\eta}^{(i-1)}|\mathbf{y_1}, \mathbf{y_2})} \frac{q({\lambda_1}^{(i-1)})}{q({\lambda_1}^{(i)})}, 1\right\}.$$

Step 4 : Generate $\lambda_1^{(i)}$ from its proposal density and set

$$\boldsymbol{\eta}^{(i)} = (\alpha^{(i)}, \lambda_1^{(i)}, \lambda_2^{(i-1)}, \lambda_3^{(i-1)}),$$

then accept $\lambda_2^{(i)}$ with probability:

$$\rho_3 = \min\left\{\frac{\pi(\boldsymbol{\eta}^{(i)}|\mathbf{y_1}, \mathbf{y_2})}{\pi(\boldsymbol{\eta}^{(i-1)}|\mathbf{y_1}, \mathbf{y_2})} \frac{q({\lambda_2}^{(i-1)})}{q({\lambda_2}^{(i)})}, 1\right\}.$$

Step 5 : Generate $\lambda_2^{(i)}$ from its proposal density and set

$$\boldsymbol{\eta}^{(i)} = (\theta^{(i)}, \alpha^{(i)}, \lambda_1^{(i)}, \lambda_2^{(i)}, \lambda_3^{(i-1)}),$$

then accept $\lambda_3^{(i)}$ with probability:

$$\rho_4 = \min\left\{ \frac{\pi(\boldsymbol{\eta}^{(i)}|\mathbf{y_1}, \mathbf{y_2})}{\pi(\boldsymbol{\eta}^{(i-1)}|\mathbf{y_1}, \mathbf{y_2})} \frac{q({\lambda_3}^{(i-1)})}{q({\lambda_3}^{(i)})}, 1 \right\}.$$

Step 6 : Set $i = i + 1$.

Step 7 : Repeat steps 2-7, N times.

By considering a truncated normal proposal distribution for θ, defined as

$$q_\theta \propto N(\theta^{(i-1)}, \tau^2) I(0 \leq \theta \leq 1),$$

and a gamma distribution $q \propto \text{Gamma}(v_1, v_2)$ for the remaining parameters, the values of τ, v_1, and v_2 must be appropriately selected to ensure that the optimized values of $\rho_1, \rho_2, \rho_3, \rho_4$, and ρ_5 fall within the interval $(0.1, 0.7)$.

With these conditions in place, the Bayesian estimations for the parameters $\alpha, \lambda_1, \lambda_2, \lambda_3$, as well as the reliability parameter R, can be derived using both the squared error loss (SEL) and linear-exponential (LINEX) loss functions as follows:

$$\hat{\alpha} = \hat{E}(\alpha|data) = \frac{1}{N-M} \sum_{i=M+1}^{N} \alpha^{(i)}, \tag{4.10}$$

$$\tilde{\alpha} = -\frac{1}{C} \ln[\hat{E}(e^{-C\alpha}|data)] = -\frac{1}{C} \ln\left[\frac{1}{N-M} \sum_{i=M+1}^{N} e^{-C\alpha^{(i)}}\right], \tag{4.11}$$

$$\hat{\lambda_1} = \hat{E}(\lambda_0|data) = \frac{1}{N-M} \sum_{i=M+1}^{N} {\lambda_0}^{(i)}, \tag{4.12}$$

$$\tilde{\lambda_1} = -\frac{1}{C} \ln[\hat{E}(e^{-C\lambda_0}|data)] = -\frac{1}{C} \ln\left[\frac{1}{N-M} \sum_{i=M+1}^{N} e^{-C{\lambda_0}^{(i)}}\right], \tag{4.13}$$

$$\hat{\lambda_2} = \hat{E}(\alpha|data) = \frac{1}{N-M} \sum_{i=M+1}^{N} {\lambda_1}^{(i)}, \tag{4.14}$$

$$\tilde{\lambda_2} = -\frac{1}{C} \ln[\hat{E}(e^{-C\lambda_1}|data)] = -\frac{1}{C} \ln\left[\frac{1}{N-M} \sum_{i=M+1}^{N} e^{-C{\lambda_1}^{(i)}}\right], \tag{4.15}$$

$$\hat{\lambda_3} = \hat{E}(\alpha|data) = \frac{1}{N-M} \sum_{i=M+1}^{N} {\lambda_2}^{(i)}, \tag{4.16}$$

$$\tilde{\lambda_3} = -\frac{1}{C}\ln[\hat{E}(e^{-C\lambda_2}|data)] = -\frac{1}{C}\ln\left[\frac{1}{N-M}\sum_{i=M+1}^{N} e^{-C\lambda_2{}^{(i)}}\right], \tag{4.17}$$

Here, M represents the burn-in period during the generation of $\alpha^{(i)}, \lambda_1^{(i)}, \lambda_2^{(i)}, \lambda_3^{(i)}$. As suggested in Basikhasteh and Makhdoom (2022), the highest posterior density (HPD) confidence interval for the parameter vector $\boldsymbol{\eta} = (\alpha, \lambda_1, \lambda_2, \lambda_3)$ can be constructed as follows:

- First order $\alpha_1^{(i)}, \ldots, \alpha_{N-M}^{(i)}$ as $\alpha_{(1)}^{(i)} < \ldots < \alpha_{(N-M)}^{(i)}$, then construct all the $(100(1-\eta)\%)$ confidence intervals of α is given by

$$\left(\alpha^{(i)}_{[(\frac{\eta}{2})(N-M)]}, \alpha^{(i)}_{[(1-\frac{\eta}{2})(N-M)]}\right), \tag{4.18}$$

  where $[x]$ is symbolized as the largest integer less than or equal to x and $\alpha^{(i)}_{[(\frac{\eta}{2})(N-M)]}$ and $\alpha^{(i)}_{[(1-\frac{\eta}{2})(N-M)]}$ are the $[(\frac{\eta}{2})(N-M)]$th smallest integer and the $[(1-\frac{\eta}{2})(N-M)]$th smallest integer of $\{\alpha_j, j = M+1, \ldots, N$.

- The average length of the highest posterior density (HPD) confidence interval for α is determined by computing the mean of the interval lengths given in Eq. (4.18).

Similarly, we can make a $100(1-\eta)\%$ HPD confidence interval of λ_1, λ_2 and λ_3 as follows:

$$\left(\lambda^{(i)}_{1[(\frac{\eta}{2})(N-M)]}, \lambda^{(i)}_{1[(1-\frac{\eta}{2})(N-M)]}\right), \tag{4.19}$$

$$\left(\lambda^{(i)}_{2[(\frac{\eta}{2})(N-M)]}, \lambda^{(i)}_{2[(1-\frac{\eta}{2})(N-M)]}\right), \tag{4.20}$$

$$\left(\lambda^{(i)}_{3[(\frac{\eta}{2})(N-M)]}, \lambda^{(i)}_{3[(1-\frac{\eta}{2})(N-M)]}\right). \tag{4.21}$$

4.1 Convergence measurements

Two criteria for measuring the convergence of the Markov chain are described below:

The Geweke test, proposed by Geweke (1992), is a statistical method used to assess the convergence of MCMC simulations, particularly in Bayesian inference. It partitions the MCMC chain into two segments, typically the early and late parts, computes a chosen statistic (often the mean) for each segment, and tests for significant differences between them. If the means of the two segments are significantly different, it suggests potential issues with convergence or other problems in the MCMC sampling process, while similar means indicate adequate convergence. While the Geweke test offers a straightforward diagnostic approach, it's typically used alongside other convergence diagnostics to ensure the reliability of MCMC results.

The Gelman-Rubin diagnostic, commonly known as the Gelman test, is a statistical method used to assess the convergence of MCMC simulations. Introduced by Gelman and Rubin (1992), this diagnostic helps practitioners determine if multiple chains of an MCMC algorithm have converged to the same target distribution. It compares the variance between chains to the variance within chains, providing a ratio known as the potential scale reduction factor (PSRF) or Gelman-Rubin statistic. A PSRF close to 1 indicates convergence, implying that the chains have explored the target distribution adequately. A value greater than 1 suggests that further exploration may be needed. This diagnostic has become a standard tool in Bayesian inference and MCMC-based analyses. Gelman and Rubin recommend running multiple chains from different starting points to ensure convergence and using the Gelman test to verify it.

5. Simulation study

In this section, we conduct an extensive Monte Carlo simulation study to assess the effectiveness of Bayesian estimators. Two different simulation scenarios are considered. The first scenario utilizes simulated data to analyze the convergence behavior of the generated Markov chains. The second scenario focuses on estimating Bayesian parameters using MCMC methods and evaluating their performance through Monte Carlo simulations. All numerical computations are carried out using `R` version 4.4.0. The results are shown in Tables (2), (3) and (4).

5.1 Numerical Study 1: Assessing Convergence

The MCMC methods require thorough convergence diagnostics to ensure reliable posterior estimates. Trace plots provide a visual tool for inspecting convergence by illustrating parameter evolution over iterations. These plots help assess whether the MCMC algorithm has reached a stable distribution(Makhdoom et al. (2016) and Basikhasteh and Makhdoom (2022)).

5.2 Statistical Diagnostics for Convergence

To assess convergence reliably, we employed multiple statistical diagnostics:

5.2.1 Trace Plots

Trace plots track the values of sampled parameters over iterations, allowing identification of stationarity and mixing properties of the chain.

5.2.2 Coupling from the Past and Running Mean Plots

These plots provide insights into the stabilization speed of the chain and detect erratic behavior.

5.2.3 Gelman–Rubin (G-R) Convergence Statistic

The G-R diagnostic estimates the corrected scale reduction factor, which helps determine whether multiple chains have adequately mixed.

5.2.4 Geweke Test

Geweke (1992) introduced a test that evaluates the convergence of a Markov chain by comparing the means of its initial and final segments. If the chain has attained stationarity, these means should be approximately equal, and Geweke's statistic is expected to follow an asymptotically standard normal distribution.

5.3 Implementation of Convergence Diagnostics

We computed the Gelman–Rubin statistic to rigorously evaluate convergence. The analysis involved:

- Running two parallel MCMC chains with different starting values.
- Using the Metropolis-Hastings (MH) algorithm and the Gibbs sampler.
- Employing truncated normal distributions as proposal functions.
- Iterating each chain 20,000 times to ensure a sufficiently long run.

5.3.1 Prior Distributions

An informative Bayesian framework was used for prior distributions:

- Baseline parameters (α, λ_1, λ_2, λ_3) followed Gamma distributions with hyperparameters (10,10), incorporating prior knowledge.
- The additional parameter θ was assigned a Uniform(0,1) distribution, reflecting a non-informative prior belief.

5.4 Results and Interpretation

Since both MCMC chains produced similar results, we focus on reporting the analysis for Chain I. The Table (1) presents the Gelman–Rubin (G-R) statistic values for various sample sizes alongside results from the Heidelberg and Welch diagnostic test.

5.4.1 Heidelberg and Welch Test

This test uses the Cramer-von Mises test statistic to determine whether a Markov chain has reached a stationary distribution. It evaluates the null hypothesis that the chain originates from a stationary distribution and makes decisions based on a significance level.

5.4.2 Hypothesis Testing

Five statistical hypothesis tests were conducted for different model parameters. Based on computed P-values, the null hypothesis was accepted for all parameters, confirming that the Markov chain successfully reached its stationary distribution. The results confirm the reliability of the MCMC process, ensuring robust posterior estimates. These diagnostic checks provide strong evidence that the MCMC chains have converged properly.

Table 1: The Gelman-Rubin statistic values for various sample sizes.

Parameter	**Point Estimate**	**Upper Confidence Interval**
α	0.98	1.12
λ_1	1.01	1.02
λ_2	1.05	2.95
λ_3	1.03	1.14

The point estimates and upper confidence intervals for individual parameters (α, λ_1, λ_2, λ_3) give you an idea of the convergence behavior of each parameter. The multivariate PSRF (Potential Scale Reduction Factor) value combines all of the parameters and provides an overall assessment of convergence. A PSRF value closer to 1.0 indicates that the MCMC chains have likely converged, while values above 1.0 (like the multivariate value of 0.97) suggest without issues with convergence. For these MCMC chains, the multivariate PSRF is 0.97, which is very close to 1. This indicates that the between-chain and within-chain variances are nearly equal, suggesting that there is no evidence of lack of convergence.

Table 2: Scenario 1 for Posterior Statistics for Different Sample Sizes (n).

Parameter	**n=10**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
2gray!10!whitewhite α	1.1377348	1.0801405	1.2500750	0.2012991
λ_1	1.4494644	1.4273035	1.5479265	0.1801030
λ_2	1.6407139	1.6132897	1.7425181	0.1888556
λ_3	0.8062105	0.7664033	0.8682813	0.1095262
Parameter	**n=20**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	1.1031928	1.0512090	1.2131302	0.1952268
λ_1	1.8315364	1.8170073	1.9023510	0.1356295
λ_2	1.5098600	1.4840359	1.5718622	0.1172541
λ_3	0.6164228	0.5864659	0.6492519	0.0606993
Parameter	**n=40**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	1.0812236	1.0276825	1.1814010	0.1790262
λ_1	1.5404535	1.5291819	1.5713652	0.0601862
λ_2	1.9139510	1.8982649	1.9557032	0.0803201
λ_3	0.3433431	0.3230673	0.3536594	0.0198796
Parameter	**n=50**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	1.0041934	0.9620250	1.0892349	0.1563211
λ_1	1.7568142	1.7422153	1.7935964	0.0713095
λ_2	2.1729010	2.1611528	2.2190596	0.0902964
λ_3	0.3875464	0.3667422	0.4056455	0.0260195
Parameter	**n=80**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	1.3572014	1.3039759	1.4982955	0.2515800
λ_1	2.2941323	2.2660275	2.3274887	0.0651675
λ_2	2.0596635	2.0721525	2.0850739	0.0503641
λ_3	0.2296535	0.2131387	0.2363236	0.0107941
Parameter	**n=120**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	0.9911122	0.9456906	1.0715924	0.1439836
λ_1	2.3975854	2.3857655	2.4219051	0.0479813
λ_2	1.9419576	1.9353831	1.9590895	0.0343086
λ_3	0.1818217	0.1705877	0.1943129	0.0079955
Parameter	**n=150**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	1.3208188	1.2680483	1.4430728	0.2164945
λ_1	1.9269779	1.9309968	1.9415048	0.0280175
λ_2	2.4466167	2.4349297	2.4660474	0.0388233
λ_3	0.1291541	0.1240931	0.1313137	0.0034718
Parameter	**n=200**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	0.8813835	0.8364845	0.9481716	0.1235075
λ_1	2.3115141	2.3028568	2.3226891	0.0221587
λ_2	2.1775382	2.1847369	2.1894051	0.0236323
λ_3	0.1175985	0.1105472	0.1193460	0.0030655

Table 3: Scenario 2 for Posterior Statistics for Different Sample Sizes (n).

Parameter	**n=10**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
2gray!10!whitewhite α	1.0849622	1.0297650	1.1873501	0.18423940
λ_1	1.5320105	1.5033349	1.6364989	0.19313894
λ_2	1.3319028	1.3127526	1.4041901	0.13528974
λ_3	0.7487524	0.7144214	0.8001293	0.09475585
Parameter	**n=20**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	1.0815593	1.0312747	1.1849725	0.18562359
λ_1	1.6991254	1.6728333	1.7610432	0.11811410
λ_2	1.3926047	1.3705380	1.4471117	0.10480075
λ_3	0.5627398	0.5343269	0.5901542	0.05159591
Parameter	**n=40**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	1.3114720	1.2529393	1.4566886	0.25562916
λ_1	2.4314942	2.4354508	2.4973873	0.12809131
λ_2	1.8588121	1.8514591	1.8985548	0.07863569
λ_3	0.3710458	0.3526856	0.3836128	0.02397197
Parameter	**n=50**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	1.3608115	1.3065885	1.5092700	0.26183577
λ_1	2.6654666	2.6338338	2.7277930	0.12103012
λ_2	2.3636639	2.3594826	2.4250250	0.11857033
λ_3	0.3926508	0.3706328	0.4069120	0.02687767
Parameter	**n=80**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	1.0972401	1.0452201	1.1964250	0.17859100
λ_1	2.4255231	2.4216805	2.4583580	0.06478210
λ_2	1.8690524	1.8645883	1.8954077	0.05106787
λ_3	0.2601793	0.2419849	0.2673105	0.01263048
Parameter	**n=120**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	1.2845271	1.2297873	1.4100692	0.22690551
λ_1	2.7772018	2.7719871	2.8068586	0.05857808
λ_2	2.3247866	2.3240269	2.3516989	0.05313933
λ_3	0.1913377	0.1785497	0.1948694	0.00667259
Parameter	**n=150**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	1.2762573	1.2284086	1.3973509	0.21693899
λ_1	2.6680889	2.6722628	2.6921925	0.04701114
λ_2	2.2667459	2.2762536	2.2856165	0.03763501
λ_3	0.1509933	0.1412605	0.1554865	0.00522826
Parameter	**n=200**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	1.2092379	1.1661182	1.3083665	0.18435302
λ_1	2.3795963	2.3651374	2.3925951	0.02576585
λ_2	2.4710638	2.4694416	2.4876320	0.03050239
λ_3	0.1165567	0.1095203	0.1181287	0.00302951

Table 4: Scenario 3 for Posterior Statistics for Different Sample Sizes (n).

Parameter	n=10			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
2gray!10!whitewhite α	1.0678556	1.0102195	1.1671521	0.17841242
λ_1	1.3512795	1.3413274	1.4190811	0.12945023
λ_2	1.1942457	1.1654145	1.2619486	0.12522379
λ_3	0.6496863	0.6204313	0.6856036	0.06757489
Parameter	**n=20**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	1.1752013	1.1201252	1.2927067	0.20971199
λ_1	0.9336936	0.9194073	0.9593332	0.04838076
λ_2	1.8194797	1.8034674	1.8856342	0.12592938
λ_3	0.3930706	0.3730762	0.4100179	0.02701957
Parameter	**n=40**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	1.179139	1.115613	1.3051392	0.21981936
λ_1	1.336545	1.327768	1.3668171	0.05601283
λ_2	2.526596	2.513069	2.5926295	0.12533778
λ_3	0.392289	0.374053	0.4069696	0.02548173
Parameter	**n=50**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	1.3289379	1.2735229	1.4688721	0.24774147
λ_1	1.6454337	1.6420141	1.6780687	0.06220585
λ_2	3.0272898	2.9916299	3.1124342	0.16687495
λ_3	0.3548697	0.3398216	0.3664927	0.02185163
Parameter	**n=80**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	1.0273899	0.9799850	1.1190751	0.16538449
λ_1	2.0380665	2.0434484	2.0619525	0.04773740
λ_2	1.8212890	1.8150313	1.8443754	0.04599423
λ_3	0.2455979	0.2274215	0.2529896	0.01128327
Parameter	**n=120**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	1.5466847	1.4969741	1.7145402	0.29952721
λ_1	1.4537510	1.4456143	1.4643219	0.02085335
λ_2	2.8522425	2.8610710	2.8820645	0.05954783
λ_3	0.1496542	0.1392597	0.1536812	0.00516160
Parameter	**n=150**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	1.3385168	1.2840987	1.473284	0.24027583
λ_1	1.6963770	1.6820915	1.707627	0.02220597
λ_2	2.5278816	2.5104725	2.549701	0.04307295
λ_3	0.1326791	0.1259961	0.134811	0.00382175
Parameter	**n=200**			
	Posterior Mean	**Posterior Median**	**Linex Loss**	**MSE**
α	1.46326042	1.41566944	1.60534311	0.254256033
λ_1	1.55308091	1.55461703	1.55958825	0.013090916
λ_2	2.59849747	2.60156955	2.61507151	0.033656732
λ_3	0.09406016	0.08565018	0.09852859	0.003240932

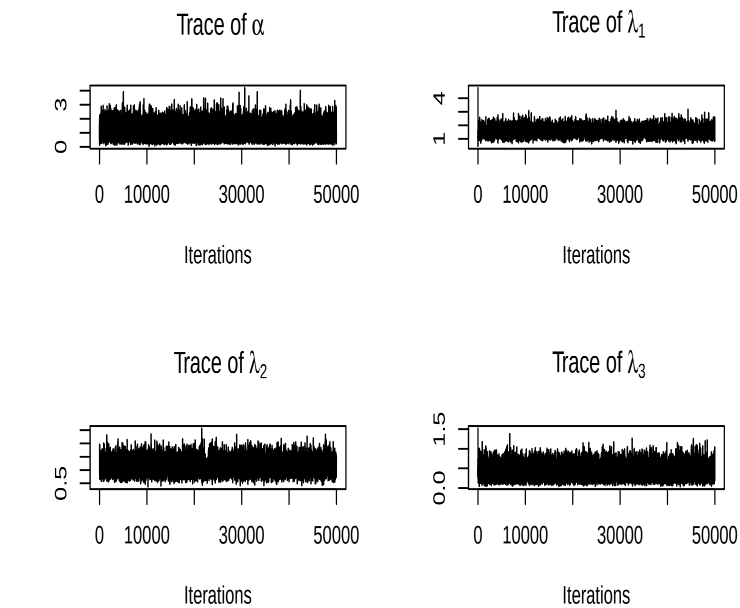

Figure 2: Trace plots for the parameters $\alpha, \lambda_1, \lambda_2, \lambda_3$ in case of $n = 20$ for chain 1

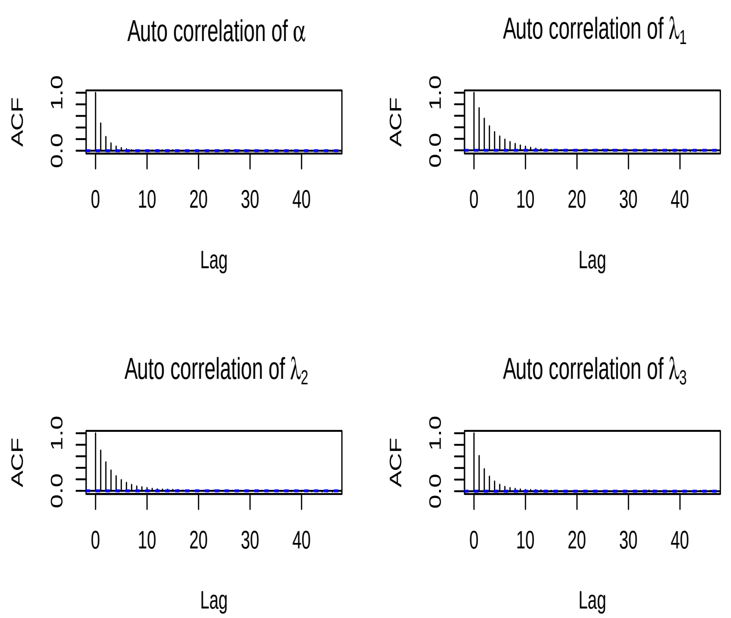

Figure 3: Sample autocorrelation plots for the parameters $\alpha, \lambda_1, \lambda_2, \lambda_3$ in case of $n = 15$ for chain 1

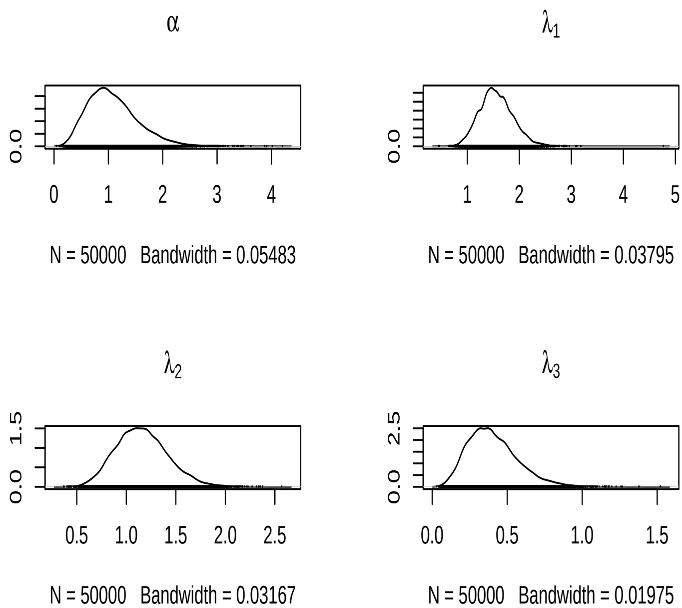

Figure 4: Posterior density estimates for the parameters $\alpha, \lambda_1, \lambda_2, \lambda_3$ in case of $n = 15$ for chain 1

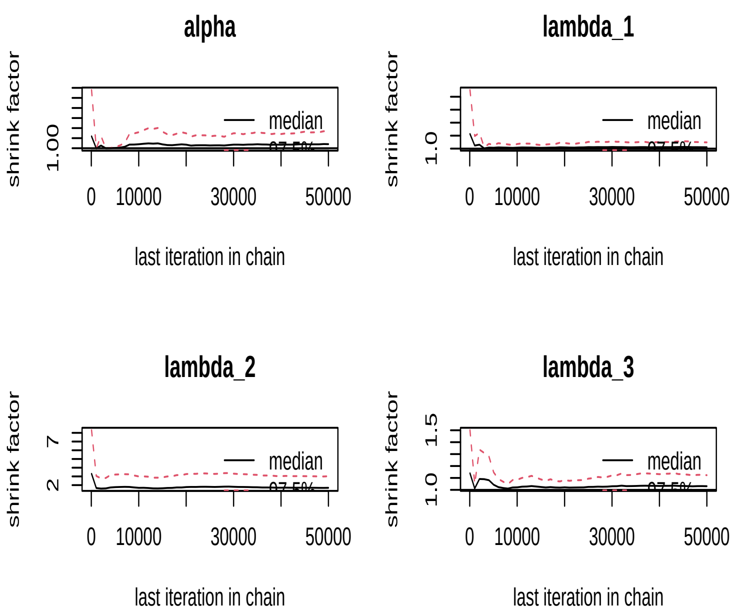

Figure 5: Gelman-Rubin plots for the parameters $\alpha, \lambda_1, \lambda_2, \lambda_3$ in case of $n = 15$ for chain 1

6. Data Analysis

To demonstrate the applicability of the *Weighted Marshall–Olkin Bivariate Exponential (WMOBE)* distribution, a dataset from the *National Football League (NFL)* was analyzed. The dataset consists of game times recorded over three consecutive weekends in 1986, representing the moment when the first points were scored in each match.

6.1 Dataset Description

The dataset includes two key variables:

- X_1: Time (in minutes) until the first points are scored via a **field goal**.
- X_2: Time until the first points are scored via a **touchdown**.

The relationship between these variables distinguishes different scoring patterns:

- If $X_1 < X_2$, the first score is a *field goal*.
- If $X_1 = X_2$, the first score is a *converted touchdown*.
- If $X_1 > X_2$, the first score is an *unconverted touchdown or safety*.

Since a touchdown and its extra-point attempt occur in quick succession, exact ties naturally arise in the data and cannot be ignored.

6.2 Dataset Table

Table 5 presents the recorded scoring times (X_1, X_2).

6.3 Model Selection and Parameter Estimation

Earlier research by *Csorgo and Welsch (1989)* analyzed this dataset using the *Marshall–Olkin Bivariate Exponential (MOBE) model*. Their findings indicated that while X_1 followed an *exponential distribution*, X_2 did not. This led to the rejection of the MOBE model as an adequate fit.

Using **maximum likelihood estimation (MLE)**, the parameters of the

Table 5: NFL First Scoring Time Data (in minutes)

X_1	X_2	X_1	X_2	X_1	X_2
2.05	3.98	5.78	25.98	10.40	14.25
9.05	9.05	13.80	49.75	2.98	2.98
0.85	0.85	7.25	7.25	3.88	6.43
3.43	3.43	4.25	4.25	0.75	0.75
7.78	7.78	1.65	1.65	11.63	17.37
10.57	14.28	6.42	15.08	1.38	1.38
7.05	7.05	4.22	9.48	10.35	10.35
2.58	2.58	15.53	15.53	12.13	12.13
7.23	9.68	2.90	2.90	14.58	14.58
6.85	34.58	7.02	7.02	11.82	11.82
32.45	42.35	6.42	6.42	5.52	11.27
8.53	14.57	8.98	8.98	19.65	10.70
31.13	49.88	10.15	10.15	17.83	17.83
14.58	20.57	8.87	8.87	10.85	38.07

WMOBE model were estimated as follows, refer to Jamalizadeh and Kundu (2013):

$$\hat{\lambda}_1 = 0.5996,$$
$$\hat{\lambda}_2 = 0.0346,$$
$$\hat{\lambda}_3 = 0.8639,$$
$$\hat{\alpha} = 2.5302.$$

The computed **log-likelihood value** for this model was -85.4447, and the corresponding **95% confidence intervals** for the estimated parameters were also calculated.

6.4 Model Comparison and Hypothesis Testing

To evaluate model fit, the *Kolmogorov–Smirnov (KS) test* was conducted. The resulting **p-values** were:

- X_1: 0.8351
- X_2: 0.7536
- $\min(X_1, X_2)$: 0.7671

Since all p-values were relatively high, the null hypothesis that the marginal distributions follow a **weighted exponential (WE) distribution** was not rejected, confirming that the WMOBE model is a good fit.

For additional comparison, the *Marshall–Olkin Bivariate Weibull (MOBW) model* was also tested, yielding a lower log-likelihood value (-90.4169) compared to WMOBE (-85.4447).

A likelihood ratio test was performed to assess whether the MOBE model could adequately fit the data. The log-likelihood value for the MOBE model was -93.3058, and the **p-value** for the test was less than 0.0001, leading to strong rejection of the MOBE model in favor of WMOBE.

6.5 Model Estimation and Goodness-of-Fit

We conducted data analysis using the **Weighted Marshall–Olkin Bivariate Exponential (WMOBE) model** in a Bayesian framework. The results are summarized in Table 6, which presents Bayesian parameter estimates under different loss functions.

Table 6: Bayesian Estimates and 95% HPD Intervals for WMOBE Model Parameters.

Loss Function	$\hat{\alpha}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\lambda}_3$
SEL	1.8952	0.6584	0.0423	0.9217
LINEX(C=1)	1.8321	0.6258	0.0384	0.8826
LINEX(C=-1)	1.9365	0.6895	0.0451	0.9642
HPD Interval	(1.4802, 2.2204)	(0.3902, 0.9005)	(0.0283, 0.0598)	(0.5201, 1.3208)

To assess the model's goodness of fit, we performed the **Kolmogorov–Smirnov (K-S) test** on the marginal distributions. The **K-S distances** and their corresponding **p-values** are reported in Table 7. The results suggest that the **WMOBE model provides a good fit** to the dataset.

Table 7: K-S Distances and p-values for WMOBE Model.

	SEL		LINEX(C=1)		LINEX(C=-1)	
Distribution	K-S Statistic D	p-value	K-S Statistic D	p-value	K-S Statistic D	p-value
X_1	0.1182	0.6821	0.1236	0.6217	0.1104	0.7218
X_2	0.1694	0.2486	0.1651	0.2714	0.1673	0.2603
$\min(X_1, X_2)$	0.1298	0.5583	0.1257	0.5921	0.1325	0.5312

7. Conclusion

In this study, we explored the WMOBE Model using the Metropolis-Hastings Markov Chain Monte Carlo (MCMC) method for parameter estimation. This model provides a flexible framework for modeling bivariate data, capturing dependencies between two random variables using a combination of exponential distributions and weight parameters. By applying the MCMC algorithm, we were able to estimate the parameters, such as the alpha (α) and (λ) values, which govern the dependence structure and the marginal distributions of the bivariate random variables. Through careful implementation, we used a combination of Gibbs sampling and Metropolis-Hastings updates to sample from the posterior distributions of the model's parameters. We validated the convergence of the MCMC chains using diagnostic tools like Gelman-Rubin convergence tests, trace plots, and autocorrelation functions. These diagnostics confirmed that the chains had successfully converged and the samples accurately represented the posterior distributions. The findings demonstrate the utility of the WMOBE Model in capturing the complex relationships between bivariate data. The estimated parameters provide insights into the dependence structure between the two variables, with implications for fields such as survival analysis, risk analysis, and insurance modeling, where understanding the joint behavior of two related events is crucial.

The analysis of the rejection rates and sampling efficiency also highlighted areas for potential improvement. Future work could focus on enhancing the proposal distribution to achieve better sampling efficiency, as well as extending the model to accommodate more complex dependencies or higher-dimensional datasets. Sensitivity analyses could further explore the robustness of the parameter estimates under different prior specifications.

Overall, this study successfully demonstrated the application of the WMOBE Model and provided a comprehensive framework for parameter estimation using MCMC methods. The results offer a strong foundation for future research and applications in modeling dependent bivariate events.

Compliance with Ethical Standards

Conflict of interest: The authors declare that they have no conflict of interest.

Acknowledgements

The authors sincerely thank the anonymous reviewers and the Editorial Office for their valuable comments and suggestions, which have greatly improved the quality of this manuscript.

References

Al-Mutairi, D., Ghitany, M., and Kundu, D. (2011). A new bivariate distribution with weighted exponential marginals and its multivariate generalization. *Statistical Papers*, **52**, 921–936.

Arnold, B., Beaver, R., Azzalini, A., Balakrishnan, N., Bhaumik, A., Dey, D., Cuadras, C., and Sarabia, J. (2002). Skewed multivariate models related to hidden truncation and/or selective reporting. *TEST: An Official Journal of the Spanish Society of Statistics and Operations Research*, **11**, 7–54.

Arnold, B. C. and Beaver, R. J. (2000). Hidden truncation models. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, **62(1)**, 23–35.

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

Basikhasteh, M. and Makhdoom, I. (2022). Bayesian inference of bivariate weibull geometric model based on linex and quadratic loss functions. *International Journal of System Assurance Engineering and Management*, **13(2)**, 867–880.

Calabria, R. and Pulcini, G. (1996). Point estimation under asymmetric loss functions for left-truncated exponential samples. *Communications in Statistics-Theory and Methods*, **25(3)**, 585–600.

Chang, Y.-c. and Hung, W.-l. (2007). Linex loss functions with applications to determining the optimum process parameters. *Quality & Quantity*, **41**, 291–301.

Dey, S. (2010). Bayesian estimation of the shape parameter of the generalised exponential distribution under different loss functions. *Pakistan Journal of Statistics and Operation Research*, **VI(2)**, 163–174.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, **7(4)**, 457–472.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculations of posterior moments. *Bayesian statistics*, **4**, 641–649.

Gupta, R. D. and Kundu, D. (2009). A new class of weighted exponential distributions. *Statistics*, **43(6)**, 621–634.

Hanagal, D. D. and Ahmadi, K. A. (2009). Bayesian estimation of the parameters of bivariate exponential distributions. *Communications in Statistics - Simulation and Computation*, **38(7)**, 1391–1413.

Jamalizadeh, A. and Kundu, D. (2013). Weighted marshall–olkin bivariate exponential distribution. *Statistics*, **47(5)**, 917–928.

Kotz, S., Balakrishnan, N., and Johnson, N. L. (2000). *Continuous Multivariate Distributions, Models and Applications.* John Wiley & Sons, New York.

Lindley, D. V. (1980). Approximate bayesian methods. *Trabajos de Estadistica Y de Investigacion Operativa*, **31(1)**, 223–245.

Makhdoom, I., Nasiri, P., and Pak, A. (2016). Bayesian approach for the reliability parameter of power lindley distribution. *International journal of system assurance engineering and management*, **7(3)**, 341–355.

Makhdoom, I. and Sakhaei, A. (2023). On dependence of the weighted marshall-olkin bivariate exponential model in the presence of the copula function. *Journal of Statistical Modelling: Theory and Applications*, **4(1)**, 157–172.

Nassar, M., Alotaibi, R., Okasha, H., and Wang, L. (2022). Bayesian estimation using expected linex loss function: A novel approach with applications. *Mathematics*, **10(3)**, 436.

Robert, C. P., Casella, G., and Casella, G. (1999). *Monte Carlo statistical methods*, volume 2. Springer.

Shahbaz, S., Shahbaz, M., and Butt, N. (2010). A Class of weighted weibull distribution. *Pakistan Journal of Statistics and Operation Research*, **6(1)**, 53 – 59.

Terven, J., Cordova-Esparza, D. M., Ramirez-Pedraza, A., and Chavez-Urbiola, E. A. (2023). Loss functions and metrics in deep learning. a review. *arXiv preprint arXiv:2307.02694.*

Tierney, L. and Kadane, J. B. (1986). Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association*, **81(393)**, 82–86.

Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss functions. *Journal of the American Statistical Association*, **81(394)**, 446–451.