

The Effectiveness of the Squared Error and Higgins-Tsokos Loss Functions on the Bayesian Reliability Analysis of Software Failure Times under the Power Law Process

Freeh N. Alenezi^{1,2}, Christ P. Tsokos¹

¹University of South Florida, Tampa, FL, USA

²Majmaah University, Al-Zulfi, Saudi Arabia

Email: fnalenezi@usf.edu, ctsokos@usf.edu

How to cite this paper: Alenezi, F.N. and Tsokos, C.P. (2019) The Effectiveness of the Squared Error and Higgins-Tsokos Loss Functions on the Bayesian Reliability Analysis of Software Failure Times under the Power Law Process. *Engineering*, 11, 272-299. <https://doi.org/10.4236/eng.2019.115020>

Received: April 24, 2019

Accepted: May 20, 2019

Published: May 23, 2019

Copyright © 2019 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Reliability analysis is the key to evaluate software's quality. Since the early 1970s, the Power Law Process, among others, has been used to assess the rate of change of software reliability as time-varying function by using its intensity function. The Bayesian analysis applicability to the Power Law Process is justified using real software failure times. The choice of a loss function is an important entity of the Bayesian settings. The analytical estimate of likelihood-based Bayesian reliability estimates of the Power Law Process under the squared error and Higgins-Tsokos loss functions were obtained for different prior knowledge of its key parameter. As a result of a simulation analysis and using real data, the Bayesian reliability estimate under the Higgins-Tsokos loss function not only is robust as the Bayesian reliability estimate under the squared error loss function but also performed better, where both are superior to the maximum likelihood reliability estimate. A sensitivity analysis resulted in the Bayesian estimate of the reliability function being sensitive to the prior, whether parametric or non-parametric, and to the loss function. An interactive user interface application was additionally developed using Wolfram language to compute and visualize the Bayesian and maximum likelihood estimates of the intensity and reliability functions of the Power Law Process for a given data.

Keywords

Power Law Process, Bayesian Reliability, Intensity Function, Kernel Density, Loss Function, Robustness

1. Introduction

Reliability analysis of a software under development is a key to assess whether a desired level of a quality product is achieved. Specially, when a software package is considered, and is tested after each failure detection, and then corrected until a new failure is observed. Over the past few decades, the reliability analysis of a software package has been studied, where graphical and numerical metrics have been introduced. One of the earliest, Duane (1964) [1], who introduced a graph to assess the reliability of a software over time using its failure times. It has the cumulative failure rate and the time on the y-axis and x-axis, respectively. In this graph, one can conclude a software reliability improvement if a negative curve is observed whereas a positive curve means the software reliability is deteriorating. On the other hand, a horizontal line indicates that the software reliability is stable. The failure numbers $N(t)$ in time interval $(0, t]$ is considered a Poisson counting process after satisfying the following conditions:

- 1) $N(t=0) = 0$.
- 2) Independent increment (counts of disjoint time intervals are independent).
- 3) It has an intensity function

$$V(t) = \lim_{\Delta t \rightarrow 0} \frac{P(N(t, t + \Delta t) = 1)}{\Delta t}.$$

- 4) Simultaneous failures do not exist

$$\lim_{\Delta t \rightarrow 0} \frac{P(N(t, t + \Delta t) = 2)}{\Delta t} = 0.$$

The probability of random value $N(t) = n$ is given by:

$$P(N(t) = n) = \frac{\exp\left\{-\int_0^t V(t) dt\right\} \left\{\int_0^t V(t) dt\right\}^n}{n!}, t > 0. \quad (1)$$

Crow (1974) proposed a Non-Homogeneous Poisson Process (NHPP), which is a Poisson Process with a time varying intensity function, given by:

$$V(t) = V(t; \beta, \theta) = \frac{\beta}{\theta} \left(\frac{t}{\theta}\right)^{\beta-1}, t > 0, \beta > 0, \theta > 0, \quad (2)$$

with β and θ are the shape and scale parameters, respectively. This Non-Homogeneous Poisson Process is also known as the Power Law Process (PLP).

The joint probability density function (PDF) of the ordered failure times T_1, T_2, \dots, T_n from a NHPP with intensity function $V(t; \beta, \theta)$ is given by:

$$f(t_1, \dots, t_n) = \prod_{i=1}^n V(t_i; \beta, \theta) \exp\left\{-\int_0^w V(t; \beta, \theta) dt\right\}, \quad (3)$$

where w is the so-called stopping time; $w = t_n$ for the failure truncated case. Considering the failure truncation case, the conditional reliability function of the failure time T_n given $T_1 = t_1, T_2 = t_2, T_3 = t_3, \dots, T_{n-2} = t_{n-2}, T_{n-1} = t_{n-1}$ is a function of $V(t; \beta, \theta)$.

As a numerical assessment, the estimate of the key parameter β in the

$V(t; \beta, \theta)$ has an important role in evaluating the reliability of a software package. When the estimates of β are less and larger than 1, they indicate that the software reliability is improving and decreasing, respectively. The PLP is reduced to a homogeneous Poisson process when the estimate of β equals to 1.

The NHPP has been used for analyzing software's failure times, and prediction of the next failure time. The subject model has been shown to be effective and useful not only in software reliability assessment [2]-[11], but also in cybersecurity; the attack detection in cloud systems [12] [13], breast and skin cancer treatments' effectiveness, [14] [15] [16], respectively, finance; modeling of financial markets at the ultra-high frequency level [17], transportation; modeling passengers' arrivals [18] [19] [20] [21] [22], and in the formulation of a software cost model [23].

Since the conditional reliability function of the PLP is a function of the $V(t; \beta, \theta)$, which includes the key parameter β . That being said updating the estimation methods for the key parameter will affect positively the $V(t; \beta, \theta)$ and the software reliability estimation, and therefore help the structuring of maintenance strategies. The authors [24] and [25] obtained the Bayesian estimates of the parameter β under the the squared-error and Higgins-Tsokos loss functions, respectively, and compared them to their approximate maximum likelihood estimate (MLE). They also showed the superiority of the Bayesian estimates to the MLE of the key parameter β , and the improvement in the reliability assessment under the PLP.

To perform Bayesian analysis on a real world problem, one needs to justify the applicability of such analysis. Then, the analysis process starts by identifying the probability distribution of the failure times of a software under development, the prior PDF of the key parameter β , and a loss function. The analytical tractability have made the squared-error loss function commonly used, where it places more weight on the estimates that are far from the true value than the estimates close to true value. Higgins and Tsokos [26] proposed a new loss function that maintains the analytical tractability feature and places exponentially more weight on extreme estimates of the true value.

In the present study, we investigate the effectiveness, in Bayesian Analysis, of using the commonly used squared-error (S-E) loss function versus the Higgins-Tsokos (H-T) loss function that puts the loss at the end of the process, for modeling software failure times. To accomplish this, we used the underline failure distribution to be the Power Law Process subject to using Burr PDF as a prior of the key parameter β . In addition, we utilize both loss functions to perform sensitive analysis of the prior selections. We perform parametric and non-parametric priors, namely Burr, Inverted Gamma, Jeffery, and two Kernel PDFs. Therefore, the primary objective of the study is to answer the following questions within a Bayesian framework:

- 1) How robust is the assumption of the squared-error loss function being

challenged by the Higgins-Tsokos loss function in estimating the key parameter β of PLP for modeling software failure times?

2) Is the Bayesian estimate of the intensity function, $V(t; \beta, \theta)$, of the PLP sensitive to the selections of the prior (parametric and non-parametric) and loss function (Higgins-Tsokos and S-E loss functions)?

The paper is organized as follows, Section 2 describes the theory and development of the Bayesian reliability model. Section 3 presents the results and discussion. Section 4 are the conclusions.

2. Theory and Bayesian Estimates

2.1. Review of the Analytical Power Law Process

The probability of achieving n failures of a given system in the time interval $(0, t]$ can be written as

$$P(x = n; t) = \frac{\exp\left\{-\int_0^t V(x) dx\right\} \left\{\int_0^t V(x) dx\right\}^n}{n!}, t > 0, \quad (4)$$

where $V(t)$ is the intensity function given by (2). The reduced expression is given by:

$$P(x = n; t) = \frac{1}{n!} \exp\left\{-\frac{t^\beta}{\theta}\right\} \frac{t^{n\beta}}{\theta}, \quad (5)$$

is the PLP that is commonly known as Weibull or Non-Homogeneous Poisson Process.

When the PLP is the underlying failure model of the failure times $t_1, t_2, t_3, \dots, t_{n-1}$ and t_n , the conditional reliability function of t_n given $t_1, t_2, t_3, \dots, t_{n-1}$ can be written mathematically as a function of the intensity function, given by:

$$R(t_n | t_1, t_2, \dots, t_{n-1}) = \exp\left\{\int_{t_{n-1}}^{t_n} -V(t; \beta, \theta) dt\right\}, t_n > t_{n-1} > 0, \quad (6)$$

since it is independent of $t_1, t_2, t_3, \dots, t_{n-2}$.

Note that the improvement in estimating the key parameter β in the $R(t_n | t_1, t_2, \dots, t_{n-1})$ of the PLP, Equation (6), will improve the reliability estimation.

The maximum likelihood estimation (MLE) of β is a function of the largest failure time and the MLE of θ is also a function of the MLE of β . Let T_1, T_2, \dots, T_n denote the first n failure times of the PLP, where $T_1 < T_2 < \dots < T_n$ are measured in global time; that is, the times are recorded since the initial startup of the system. Thus, the truncated conditional probability distribution function, $f_i(t | t_1, \dots, t_{i-1})$, in the Weibull process is given by

$$f_i(t | t_1, \dots, t_{i-1}) = \frac{\beta}{\theta} \left(\frac{t}{\theta}\right)^{\beta-1} \exp\left\{-\frac{t^\beta}{\theta} + \frac{t_{i-1}^\beta}{\theta}\right\}, t_{i-1} < t. \quad (7)$$

With $t = (t_1, t_2, \dots, t_n)$, the Likelihood function for the first n failure times of the PLP $T_1 = t_1, T_2 = t_2, \dots, T_n = t_n$ can be written as

$$L(t, \beta) = \exp\left(-\left(\frac{t_n}{\theta}\right)^\beta\right) \left(\frac{\beta}{\theta}\right)^n \prod_{i=1}^n \left(\frac{t_i}{\theta}\right)^{\beta-1}. \quad (8)$$

The MLE for the shape parameter is given by

$$\hat{\beta}_n = \frac{n}{\sum_{i=1}^n \log\left(\frac{t_n}{t_i}\right)}, \quad (9)$$

and for the scale parameter is

$$\hat{\theta}_n = \frac{t_n}{n^{1/\hat{\beta}_n}}. \quad (10)$$

Note that the MLE of θ depends on the MLE of β .

2.2. Development of the Bayesian Estimates

The authors [24] and [25] justified the applicability of Bayesian analysis to the PLP based on the Crow, [2] [27], failure data from a system undergoing developmental testing (Table 1), by showing that the MLE of the key parameter β varies depending on the last failure time (largest time). Moreover, the authors used the Crow data (40 successive failure times) to compute the MLE of β ($\hat{\beta}_{40} = 0.49$), then computed the estimate considering the $t_{39} = 3181$ is the largest failure time ($\hat{\beta}_{39} = 0.48$) and so on. After computing all MLEs of the key parameter β , they found that the MLEs of β follows a four-parameter Burr probability distribution, $g(\beta; \alpha, \gamma, \delta, \kappa)$, known as the four-parameter Burr type XII probability distribution, with a PDF given by:

$$g_B(\beta) = g(\beta; \alpha, \gamma, \delta, \kappa) = \begin{cases} \frac{\alpha \kappa \left(\frac{\beta - \gamma}{\delta}\right)^{\alpha-1}}{\delta \left(1 + \left(\frac{\beta - \gamma}{\delta}\right)^\alpha\right)^{\kappa+1}} & \gamma \leq \beta < \infty \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

where the hyperparameters α , γ , δ and κ are being estimated using MLEs in the Goodness of Fit (GOF) test applied to the β estimates. The MLE

Table 1. Crow's failure times of a system under development.

Failure times						
0.7	3.7	13.2	17.6	54.5	99.2	112.2
120.9	151	163	174.5	191.6	282.8	355.2
486.3	490.5	513.3	558.4	678.1	688	785.9
887	1010.7	1029.1	1034.4	1136.1	1178.9	1259.7
1297.9	1419.7	1571.7	1629.8	1702.4	1928.9	2072.3
2525.2	2928.5	3016.4	3181	3256.3	-	-

of the key parameter β is always affected by the largest failure, and therefore it is recommended not to consider it unknown constant. This recommendation provides the opportunity to study Bayesian analysis in the PLP with respect to various selections of loss functions and priors.

The Bayesian estimates of β will be derived using the squared-error and Higgins-Tsokos loss functions.

2.2.1. Bayesian Estimates Using Squared Error (S-E) Loss Function

The S-E loss function is given by:

$$L(\hat{\xi}, \xi) = (\hat{\xi} - \xi)^2. \quad (12)$$

The risk using the S-E loss function, where $\xi = \beta$ represents the estimate of $\hat{\xi} = \hat{\beta}$, is given by:

$$E[L(\hat{\beta}, \beta)] = \int_{-\infty}^{\infty} (\hat{\beta} - \beta)^2 h(\beta|t) d\beta, \quad (13)$$

By differentiating $E[L(\hat{\beta}, \beta)]$ with respect to β and setting it equal to zero we solve for $\hat{\beta}$, the Bayesian estimate of β with respect to the S-E loss function and Burr probability distribution, Equation (11), given by:

$$\hat{\beta}_{B.SE} = \int_{-\infty}^{\infty} \beta \cdot h(\beta|t) d\beta, \quad (14)$$

where the posterior PDF of β given data (t) , $h(\beta|t)$, using the Bayes' theorem, is given by:

$$h(\beta|t) = \frac{L(t|\beta)g_B(\beta)}{\int_{-\infty}^{\infty} L(t|\beta)g_B(\beta)d\beta}. \quad (15)$$

Then, the Bayesian estimate of β , under the squared-error loss, is given by

$$\hat{\beta}_{B.SE} = \frac{\int_{\gamma}^{\infty} \frac{\beta^{n+1}}{\theta^n} \exp\left\{-\left(\frac{t_n}{\theta}\right)^{\beta}\right\} \prod_{i=1}^n \left(\frac{t_i}{\theta}\right)^{\beta-1} \frac{\left(\frac{\beta-\gamma}{\delta}\right)^{\alpha-1}}{\left(1+\left(\frac{\beta-\gamma}{\delta}\right)^{\alpha}\right)^{\kappa+1}} d\beta}{\int_{\gamma}^{\infty} \frac{\beta^n}{\theta^n} \exp\left\{-\left(\frac{t_n}{\theta}\right)^{\beta}\right\} \prod_{i=1}^n \left(\frac{t_i}{\theta}\right)^{\beta-1} \frac{\left(\frac{\beta-\gamma}{\delta}\right)^{\alpha-1}}{\left(1+\left(\frac{\beta-\gamma}{\delta}\right)^{\alpha}\right)^{\kappa+1}} d\beta}. \quad (16)$$

2.2.2. Bayesian Estimates Using the Higgins-Tsokos Loss Function

The H-T loss function (1976) is given by

$$L(\hat{\xi}, \xi) = \frac{f_1 \exp\{f_2(\hat{\xi} - \xi)\} + f_2 \exp\{-f_1(\hat{\xi} - \xi)\}}{f_1 + f_2} - 1, f_1, f_2 > 0. \quad (17)$$

Higgins and Tsokos [26] showed that it places more weight on the extreme underestimation and overestimation when $f_1 > f_2$ and $f_1 < f_2$, respectively. The risk using the H-T loss function, where $\xi = \beta$ represents the estimate of

$\hat{\xi} = \hat{\beta}$, is given by:

$$E[L(\hat{\beta}, \beta)] = \int_{-\infty}^{\infty} \left[\frac{f_1 \exp\{f_2(\hat{\beta} - \beta)\} + f_2 \exp\{-f_1(\hat{\beta} - \beta)\}}{f_1 + f_2} - 1 \right] h(\beta | t) d\beta \quad (18)$$

By differentiating $E[L(\hat{\beta}, \beta)]$ with respect to β and setting it equal to zero we solve for $\hat{\beta}$, the Bayesian estimate of β with respect to the H-T loss function, given by:

$$\hat{\beta}_{B.TH} = \frac{1}{f_1 + f_2} \ln \left[\frac{\int_{-\infty}^{\infty} \exp\{f_1 \beta\} h(\beta | t) d\beta}{\int_{-\infty}^{\infty} \exp\{-f_2 \beta\} h(\beta | t) d\beta} \right]. \quad (19)$$

The Bayesian estimate of β with respect to the Higgins-Tsokos loss function and Burr probability distribution, as the prior, has $h(\beta | t)$ given by

$$h(\beta | t) = \frac{\left(\frac{\beta}{\theta}\right)^n \exp\left\{-\left(\frac{t_n}{\theta}\right)^\beta\right\} \prod_{i=1}^n \left(\frac{t_i}{\theta}\right)^{\beta-1} \frac{\left(\frac{\beta-\gamma}{\delta}\right)^{\alpha-1}}{\left(1 + \left(\frac{\beta-\gamma}{\delta}\right)^\alpha\right)^{\kappa+1}} d\beta}{\int_{\gamma}^{\infty} \left(\frac{\beta}{\theta}\right)^n \exp\left\{-\left(\frac{t_n}{\theta}\right)^\beta\right\} \prod_{i=1}^n \left(\frac{t_i}{\theta}\right)^{\beta-1} \frac{\left(\frac{\beta-\gamma}{\delta}\right)^{\alpha-1}}{\left(1 + \left(\frac{\beta-\gamma}{\delta}\right)^\alpha\right)^{\kappa+1}} d\beta}. \quad (20)$$

With the use of Equation (6), the conditional reliability of t_i , the analytical structure of the conditional Bayesian reliability estimate for the PLP that is subject to the above information is given by:

$$\hat{R}_B(t_i | t_1, t_2, \dots, t_{i-1}) = \exp\left\{-\int_{t_{i-1}}^{t_i} \hat{V}'_B(t; \beta, \theta) dt\right\}, t_i > t_{i-1} > 0, \quad (21)$$

where

$$\hat{V}'_B(t; \hat{\beta}_{B^*}, \theta) = \frac{\hat{\beta}_{B^*}}{\theta} \left(\frac{t}{\theta}\right)^{\hat{\beta}_{B^*}-1}, \theta > 0, t > 0, \quad (22)$$

where $\hat{\beta}_{B^*}$ is the Bayesian estimate using $\hat{\beta}_{B.SE}$ or $\hat{\beta}_{B.TH}$ for the squared error or Higgins-Tsokos loss functions, respectively. We are also interested in comparing the Bayesian estimates, using Higgins-Tsokos loss function, of the subject parameter for different parametric and non-parametric priors, and with respect to its MLE given by Equation (9), assuming β has a random behavior and θ as known; as well as, comparing Equation (10) with an adjusted MLE considered as a function of β .

2.3. Sensitivity Analysis: Prior and Loss Function

In this section, we seek the answer to the following question: Is the Bayesian MLE estimate of the intensity function, $V(t; \beta, \theta)$, of the PLP sensitive to the selections of the prior (parametric and non-parametric) and loss function (Higgins-Tsokos and S-E loss functions)? Assuming β is a random variable,

using simulated data, sensitive analysis was done for the following parametric and non-parametric priors ([25]):

1) Jeffreys' prior ([28]): Jeffreys' prior is proportional to the square root of the determinant of the Fisher information matrix ($I(\beta)$). It is a non-informative prior, where the Jeffreys' prior for the key parameter of the PLP $I(\beta)$ is scalar in this case, is given by:

$$g_J(\beta) \propto \sqrt{I(\beta)} = \sqrt{-E\left(\frac{\partial^2 \log L(t; \beta)}{\partial \beta^2}\right)} \propto \frac{1}{\beta}, \beta > 0. \quad (23)$$

2) The inverted gamma: The PLP and inverted gamma probability distributions belong to the exponential family of probability distributions, which makes the latter a logical choice for an informative parametric prior for β . The inverted gamma probability distribution is given by:

$$g_{IG}(\beta) \propto \left(\frac{\mu}{\beta}\right)^{v+1} \frac{1}{\mu \Gamma(v)} \exp\left\{\frac{-\mu}{\beta}\right\}, \beta > 0, \mu > 0, v > 0, \quad (24)$$

where v and μ are the shape and scale parameters.

3) Kernel' prior:

The kernel probability density estimation is a non-parametric method to approximately estimate the PDF of β using a finite data set. It is given by:

$$g_K(\beta) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{\beta - \beta_i}{h}\right), \quad (25)$$

where K is the kernel function and h is a positive number called the bandwidth.

2.3.1. The Jeffreys' Prior

Assuming Jeffreys' PDF (23) as the prior of β and using the likelihood (8) and (15), the posterior density of β is:

$$h_J(\bar{t} | \beta) = \frac{\exp\left\{\left(\frac{t_n}{\theta}\right)^\beta\right\} \frac{\beta^{n-1}}{\theta^{n\beta}} \prod_{i=1}^n (t_i)^{\beta-1}}{\int_0^\infty \exp\left\{\left(\frac{t_n}{\theta}\right)^\beta\right\} \frac{\beta^{n-1}}{\theta^{n\beta}} \prod_{i=1}^n (t_i)^{\beta-1} d\beta}. \quad (26)$$

Thus, the Jeffreys' Bayesian estimate of the key parameter β under the S-E and H-T loss functions, using (14) and (19), are given by:

$$\hat{\beta}_{B,SE}^J = \int_0^\infty \beta \cdot h_J(\bar{t} | \beta) d\beta, \quad (27)$$

and

$$\hat{\beta}_{B,HT}^J = \frac{1}{f_1 + f_2} \ln \left[\frac{\int_0^\infty \exp\{f_1 \beta\} h_J(\bar{t} | \beta) d\beta}{\int_0^\infty \exp\{-f_2 \beta\} h_J(\bar{t} | \beta) d\beta} \right]. \quad (28)$$

We must rely on a numerical estimation because we cannot obtain close solutions for both $\hat{\beta}_{B,SE}^J$ and $\hat{\beta}_{B,HT}^J$. Also note that it depends on knowing or being able to estimate the scale parameter θ .

2.3.2. The Inverted Gamma Prior

The following is an examination of the problem when the prior density of β is given by the inverted gamma (24). Using the likelihood (8), the posterior density of β is given by:

$$h_{IG}(t|\beta) = \frac{\frac{\beta^{n-v-1}}{\theta^{n\beta}} \exp\left\{-\left(\frac{t_n}{\theta}\right)^\beta - \frac{\mu}{\beta}\right\} \prod_{i=1}^n (t_i)^{\beta-1}}{\int_0^\infty \frac{\beta^{n-v-1}}{\theta^{n\beta}} \exp\left\{-\left(\frac{t_n}{\theta}\right)^\beta - \frac{\mu}{\beta}\right\} \prod_{i=1}^n (t_i)^{\beta-1} d\beta}. \quad (29)$$

Thus, the Bayesian estimates of β under the inverted gamma with respect to the S-E and H-T loss functions, using (14) and (19), are given by:

$$\hat{\beta}_{B,SE}^{IG} = \int_0^\infty \beta \cdot h_{IG}(\bar{t}|\beta) d\beta, \quad (30)$$

and

$$\hat{\beta}_{B,HT}^{IG} = \frac{1}{f_1 + f_2} \ln \left[\frac{\int_0^\infty \exp\{f_1\beta\} h_{IG}(t|\beta) d\beta}{\int_0^\infty \exp\{-f_2\beta\} h_{IG}(t|\beta) d\beta} \right]. \quad (31)$$

Here as well, we must rely on a numerical estimation because we cannot obtain close solutions for $\hat{\beta}_{B,SE}^{IG}$ and $\hat{\beta}_{B,HT}^{IG}$. Also note that it depends on knowing or being able to estimate the scale parameter θ .

2.3.3. The Kernel Prior

Assuming Kernel density (25) as the prior of β and using the likelihood (8), the posterior density of β is:

$$h_k(\bar{t}|\beta) = \frac{\exp\left\{\left(\frac{t_n}{\theta}\right)^\beta\right\} \frac{\beta^n}{\theta^{n\beta}} \prod_{i=1}^n (t_i)^{\beta-1} \frac{1}{nh} \sum_{i=1}^n K\left(\frac{\beta - \beta_i}{h}\right)}{\int_0^\infty \exp\left\{\left(\frac{t_n}{\theta}\right)^\beta\right\} \frac{\beta^n}{\theta^{n\beta}} \prod_{i=1}^n (t_i)^{\beta-1} \frac{1}{nh} \sum_{i=1}^n K\left(\frac{\beta - \beta_i}{h}\right) d\beta}. \quad (32)$$

Thus, the kernel Bayesian estimates of the key parameter β under the S-E and H-T loss functions, (14) and (19), are given by:

$$\hat{\beta}_{B,SE}^K = \int_0^\infty \beta \cdot h_k(\bar{t}|\beta) d\beta, \quad (33)$$

and

$$\hat{\beta}_{B,HT}^K = \frac{1}{f_1 + f_2} \ln \left[\frac{\int_\gamma^\infty \exp\{f_1\beta\} h_k(\bar{t}|\beta) d\beta}{\int_\gamma^\infty \exp\{-f_2\beta\} h_k(\bar{t}|\beta) d\beta} \right]. \quad (34)$$

We must rely on a numerical estimation because we cannot obtain close solutions for $\hat{\beta}_{B,SE}^K$ and $\hat{\beta}_{B,HT}^K$. Also note that it depends on knowing or being able to estimate the scale parameter θ . In addition, the kernel function, $K(u)$, and bandwidth, h , will be chosen to minimize the asymptotic mean integrated squared error (AMISE) given by:

$$\text{AMISE}(\hat{f}(\beta)) = \int E\left[\left(\hat{f}(\beta) - f(\beta)\right)^2\right] d\beta, \quad (35)$$

where $\hat{f}(\beta)$ and $f(\beta)$ are the estimated probability density of β and the true probability density of β respectively.

Table 2 shows the acronyms and notations used in this study.

3. Results and Discussion

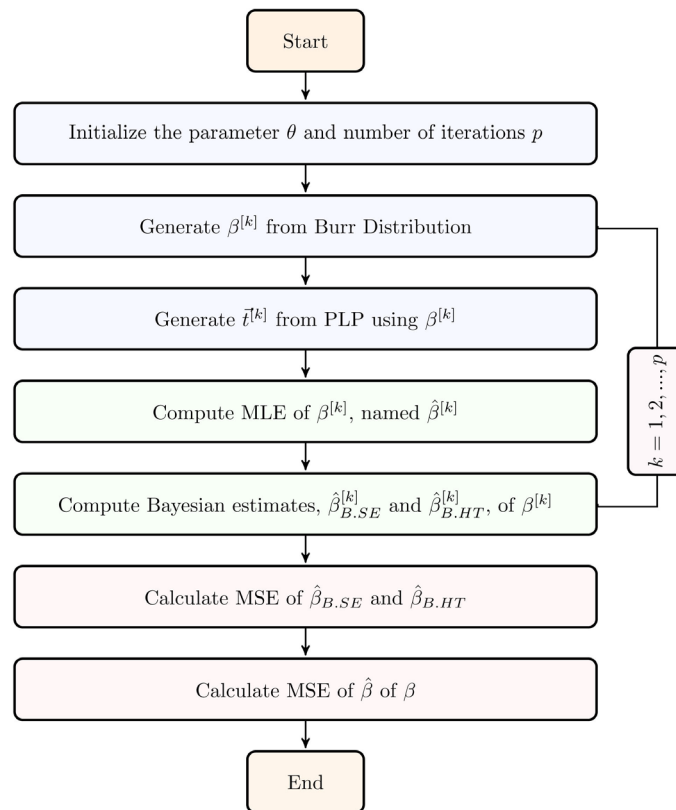
3.1. Numerical Simulation

A Monte Carlo simulation was used to compare the Bayesian, under the S-E and H-T loss functions, and the MLE approaches. The parameter β of the intensity function for the PLP was calculated using numerical integration techniques in conjunction with a Monte Carlo simulation to obtain its Bayesian estimates. Substituting these estimates in the intensity function we obtained the Bayesian intensity function estimates, from which the reliability function can be estimated.

For a given value of the parameter θ , a stochastic value for the parameter β was generated from a prior probability density. For a pair of values of θ and β , 400 samples of 40 failure times that follow a PLP were generated. This procedure was repeated 250 times and for three distinct values of θ . The procedure is based on the schematic diagram given by **Algorithm 1**.

Table 2. Acronyms and notations used in this study.

Acronyms	
HPP	Homogeneous Poisson Process
NHPP	Non-Homogeneous Poisson Process
PLP	Power Law Process
MLE	Maximum likelihood estimate
PDF	Probability density function
CDF	Cumulative density function
Notations	
β and θ	Shape and Scale parameters of PLP
T_1, T_2, \dots, T_n	First n successive failure times of the PLP
$V(t; \beta, \theta)$	Intensity function of the PLP
RE	Relative efficiency
AMISE	Asymptotic mean integrated squared error
H-T	Higgin-Tsokos
S-E	Squared-Error
$\hat{\beta}_{HT}$	Bayesian estimate of β under H-T loss function
$\hat{\beta}_{SE}$	Bayesian estimate of β under S-E loss function
\hat{V}_{HT}	Bayesian MLE estimate of $V(t; \beta, \theta)$ under H-T loss function
\hat{V}_{SE}	Bayesian MLE estimate of $V(t; \beta, \theta)$ under S-E loss function
$B.HT$	Bayesian estimate under Burr PDF and H-T loss function
$B.SE$	Bayesian estimate under Burr PDF and S-E loss function



Algorithm 1. Simulation to analyze Bayesian estimates of β for a given θ .

For each sample of size 40, the Bayesian estimates and MLEs of the parameter were calculated when $\theta \in \{0.5, 1.7441, 4\}$. The comparison is based on the mean squared error (MSE) averaged over the 100000 repetitions. The results are given in **Table 3**. It is observed that $\hat{\beta}_{B,SE}$ and $\hat{\beta}_{B,HT}$ maintain similar accuracy, where both are superior to $\hat{\beta}$ in estimating β .

For different sample sizes, the Bayesian estimates under S-E and H-T loss functions and the MLEs of the parameter β were calculated and averaged over 10,000 repetitions. **Table 4** displays the simulated result of comparing a true value of β with respect to its MLE and Bayesian estimates for $n = 20, 30, \dots, 160$.

It can be observed that the Bayesian estimates of β are closer to the true value than the MLE of β , where the Bayesian estimate under the H-T loss function is slightly performing better even for a very small sample size of $n = 20$. A graphical comparison of the true estimate of β along with the Bayesian estimates (under both S-E and H-T loss functions) and MLE as a function of sample size is given in **Figure 1**.

Figure 1 shows the excellent performance of the Bayesian estimates compared to the MLE of the key parameter β . The Bayesian estimates tend to underestimate while the MLE estimate tends to overestimate the true value, especially for small sample sizes. The MSEs of the MLE and Bayesian estimates of β for each sample size are given below by **Figure 2**.

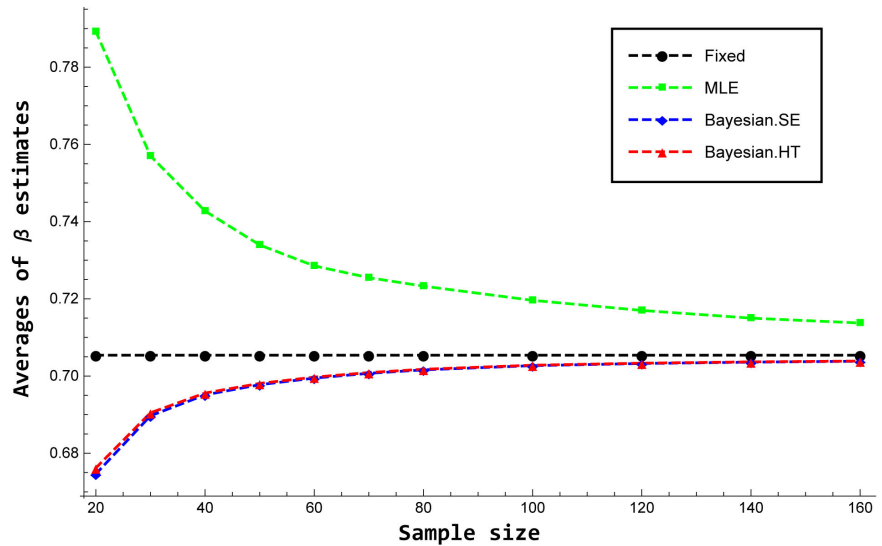


Figure 1. β estimates versus sample size.

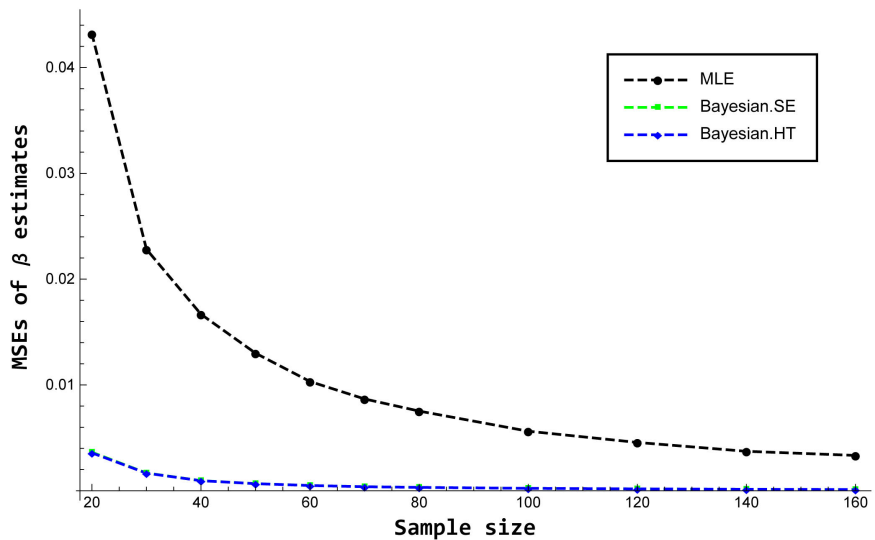


Figure 2. MSE of β Bayesian estimates versus sample size.

Table 3. MSE for Bayesian estimates, under squared error and Higgin-Tsokos loss functions, and MLEs of β .

θ	MSE of $\hat{\beta}$	MSE of $\hat{\beta}_{B,SE}$	MSE of $\hat{\beta}_{B,HT}$
0.50000	0.01124360	0.0005077610	0.000507356
1.74410	0.01105730	0.0005163560	0.000516057
400000	0.01096100	0.0005190550	0.000518632

For the considered sample sizes, the MSEs of the Bayesian estimates of β are sufficiently smaller than the MSEs for the MLE of β . The Bayesian estimate under the H-T loss function performed slightly better than the Bayesian estimate under the S-E loss function.

Table 4. Bayesian estimates, under squared error and Higgin-Tsokos loss functions, and MLEs for the parameter $\beta = 0.7054$ averaged over 10,000 repetitions.

n	β_{Fixed}	$\hat{\beta}$	$\hat{\beta}_{B,SE}$	$\hat{\beta}_{B,HT}$
20	0.7054	0.784026	0.673706	0.675263
30	0.7054	0.756617	0.689413	0.690189
40	0.7054	0.743982	0.695989	0.696467
50	0.7054	0.735310	0.698826	0.699158
60	0.7054	0.729563	0.700393	0.700642
70	0.7054	0.725977	0.701493	0.701690
80	0.7054	0.723338	0.702220	0.702382
100	0.7054	0.719117	0.703049	0.703165
120	0.7054	0.716315	0.703496	0.703585
140	0.7054	0.714821	0.703909	0.703980
160	0.7054	0.713641	0.704185	0.704244

Since the Bayesian estimates under both loss functions for β are superior to its MLE, Molinares and Tsokos [24] showed the improvement in the scale parameter (θ) when its estimate (10) is adjusted by using the Bayesian estimate of β instead of the corresponding MLE. Therefore, we calculated the adjusted estimate of θ using MLE and Bayesian estimates under S-E and H-T loss functions of β , shown in **Table 5**.

This proposed adjusted estimates, $\hat{\theta}_{B,SE}$ and $\hat{\theta}_{B,HT}$, were averaged over the 10,000 repetitions. It can be appreciated that, based on the Bayesian influence on β , $\hat{\theta}_{B,SE}$ and $\hat{\theta}_{B,HT}$ are better estimates than the MLE of θ ($\hat{\theta}$). This also can be seen in **Figure 3**, which visualize the performance of $\hat{\theta}_{B,SE}$ and $\hat{\theta}_{B,HT}$ compared to the corresponding MLE.

Figure 3 shows the excellent performance of the adjusted estimates of θ , where the adjusted estimate under the H-T was slightly closer to the true value. The MSEs of these estimates of θ are displayed in **Figure 4** given below.

The MSEs of the adjusted estimates of the shape parameter (θ) are significantly smaller than the MSEs of the MLE estimate. The MSEs of the adjusted estimates are then displayed alone in **Figure 5** to look closer at their performance.

It can be noticed that the adjusted estimate of θ under the influence of the Bayesian estimate with the H-T loss function, is better, particularly when considering small sample sizes.

We computed the adjusted estimate for the parameter θ and its MSE over 10000 repetitions for different values of θ and sample size $n = 40$. The results are given in **Table 6**.

The adjusted estimate of θ are more accurate when considering small true values of θ than the larger values.

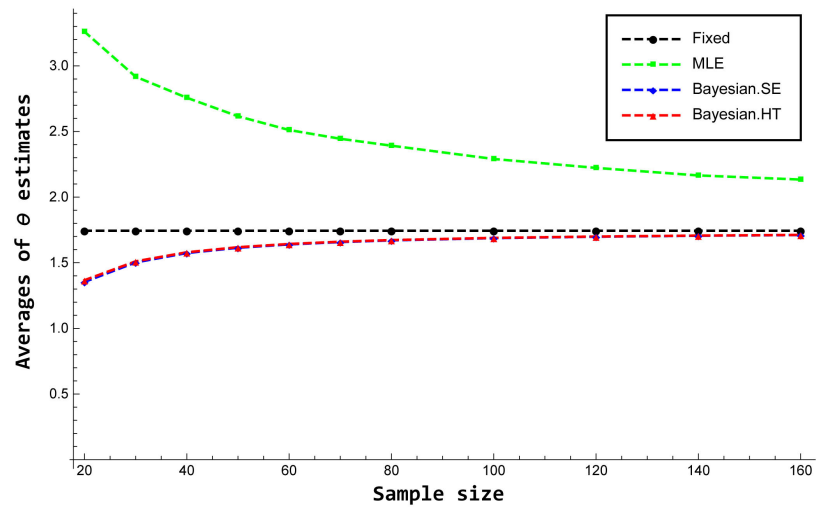


Figure 3. θ estimates versus sample size.

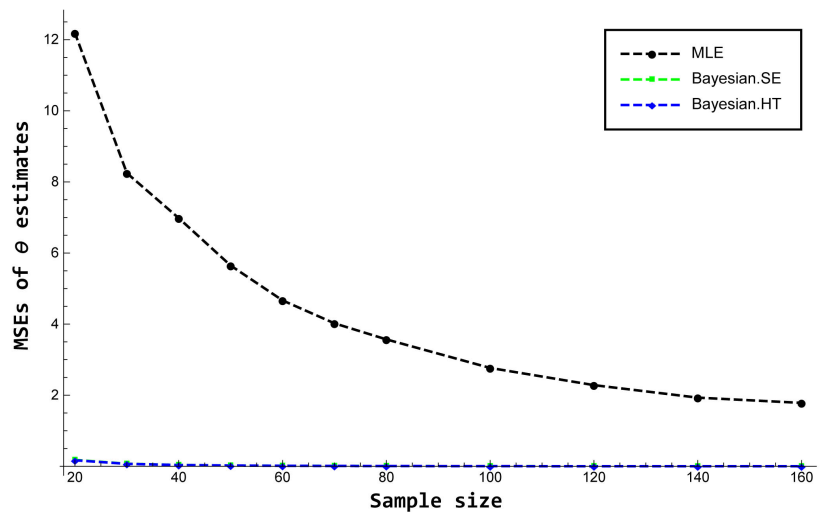


Figure 4. MSE of θ Bayesian and MLE estimates versus sample size.

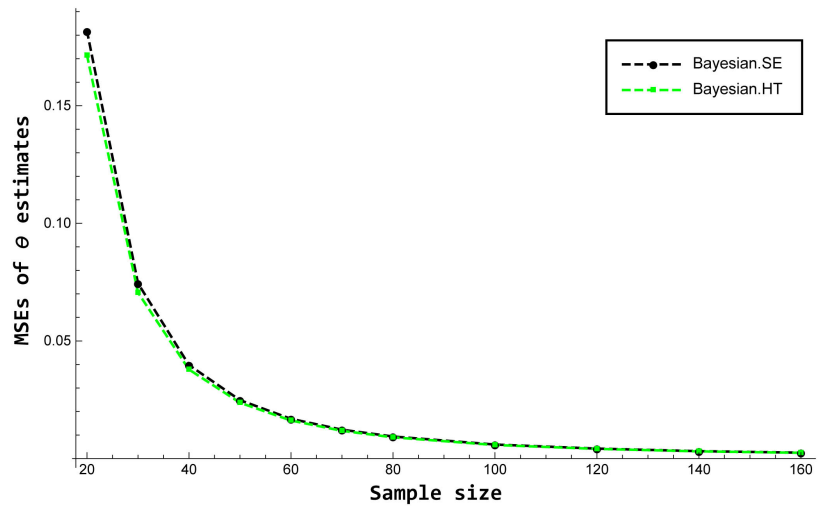


Figure 5. MSE of θ Bayesian estimates versus sample size.

Table 5. MLE Bayesian estimates, under squared error and Higgin-Tsokos loss functions, and MLEs for the parameter $\theta = 1.7441$ averaged over 10,000 repetitions.

n	θ	$\hat{\theta}_{MLE}$	$\hat{\theta}_{B,SE}$	$\hat{\theta}_{B,HT}$
20	1.7441	3.17139	1.3507	1.36422
30	1.7441	2.908	1.50142	1.5097
40	1.7441	2.73107	1.57545	1.58115
50	1.7441	2.59245	1.61556	1.61985
60	1.7441	2.48865	1.64065	1.64406
70	1.7441	2.41782	1.65803	1.66084
80	1.7441	2.36522	1.67055	1.67294
100	1.7441	2.26774	1.68719	1.68902
120	1.7441	2.20117	1.69776	1.69923
140	1.7441	2.15539	1.70537	1.70659
160	1.7441	2.11872	1.71089	1.71193

Table 6. MSE of θ estimates using Bayesian estimates, under squared error and Higgin-Tsokos loss functions, and MLE of β .

θ	$\hat{\theta}_{B,SE}$	$\hat{\theta}_{B,HT}$	MSE of $\hat{\theta}_{B,SE}$	MSE of $\hat{\theta}_{B,HT}$
0.50	0.5020220	0.5033140	0.00692919	0.00691164
1.74410	1.746390	1.75090	0.0830342	0.0827802
40	3.999920	4.010250	0.440474	0.439035

The slight improvements in the estimation of the shape and scale parameters of the PLP is expected to jointly improve the estimate of the intensity function and therefore the reliability estimation of a software. For a fixed value of $\theta = 1.7441$ and a sample size similar to the size of the collected data, $n = 40$, the estimates of the intensity function $\hat{V}_{MLE}(t)$, $\hat{V}_{B,SE}(t)$, and $\hat{V}_{B,HT}(t)$ were obtained when we use $\hat{\beta}$, $\hat{\beta}_{B,SE}$, and $\hat{\beta}_{B,HT}$, respectively, in (2). That is,

$$\hat{V}'_{MLE}(t) = \frac{\hat{\beta}}{\theta} \left(\frac{t}{\theta} \right)^{\hat{\beta}-1}, \theta > 0, t > 0. \quad (36)$$

$$\hat{V}'_{B,SE}(t) = \frac{\hat{\beta}_{B,SE}}{\theta} \left(\frac{t}{\theta} \right)^{\hat{\beta}_{B,SE}-1}, \theta > 0, t > 0. \quad (37)$$

$$\hat{V}'_{B,HT}(t) = \frac{\hat{\beta}_{B,HT}}{\theta} \left(\frac{t}{\theta} \right)^{\hat{\beta}_{B,HT}-1}, \theta > 0, t > 0. \quad (38)$$

Their graphs (Figure 6) reveal the superior performance of $\hat{V}'_{B,SE}(t)$ and $\hat{V}'_{B,HT}(t)$.

In order to obtain Bayesian estimates of the intensity function, $\hat{V}_{B,SE}^*$ and $\hat{V}_{B,HT}^*$, we substituted the Bayesian estimates of β and its corresponding θ MLE in (2):

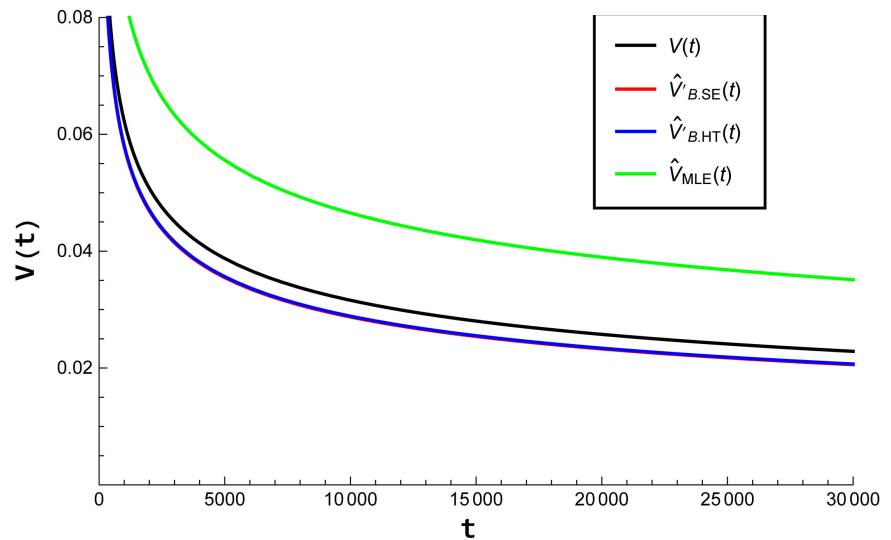


Figure 6. Graph for $\theta = 1.7441$ and the corresponding β Bayesian estimates and MLE's used in \hat{V}_{MLE}^* , $\hat{V}_{B,SE}^*$, and $\hat{V}_{B,HT}^*$ (of time t), $n = 40$.

$$\hat{V}_{B,SE}^*(t) = \frac{\hat{\beta}_{B,SE}}{\hat{\theta}} \left(\frac{t}{\hat{\theta}} \right)^{\hat{\beta}_{B,SE}-1}, t > 0. \quad (39)$$

$$\hat{V}_{B,HT}^*(t) = \frac{\hat{\beta}_{B,HT}}{\hat{\theta}} \left(\frac{t}{\hat{\theta}} \right)^{\hat{\beta}_{B,HT}-1}, t > 0. \quad (40)$$

The MLE of the intensity function, \hat{V}_{MLE} , is obtained using the MLEs of β and θ . That is,

$$\hat{V}_{MLE}(t) = \frac{\hat{\beta}}{\hat{\theta}} \left(\frac{t}{\hat{\theta}} \right)^{\hat{\beta}-1}, t > 0. \quad (41)$$

The Bayesian MLE of the intensity function under the influence of the Bayesian estimates of β , denoted by $\hat{V}_{B,SE}$ and $\hat{V}_{B,HT}$, are obtained by substituting $\hat{\beta}_{B,HT}$ and $\hat{\beta}_{B,SE}$ with $\hat{\theta}_{B,HT}$ and $\hat{\theta}_{B,SE}$, respectively, in (2):

$$\hat{V}_{B,SE}(t) = \frac{\hat{\beta}_{B,SE}}{\hat{\theta}_{B,SE}} \left(\frac{t}{\hat{\theta}_{B,SE}} \right)^{\hat{\beta}_{B,SE}-1}, t > 0, \quad (42)$$

and

$$\hat{V}_{B,HT}(t) = \frac{\hat{\beta}_{B,HT}}{\hat{\theta}_{B,HT}} \left(\frac{t}{\hat{\theta}_{B,HT}} \right)^{\hat{\beta}_{B,HT}-1}, t > 0. \quad (43)$$

To measure the robustness of $\hat{V}_{B,HT}$ with respect to $\hat{V}_{B,SE}$ and \hat{V}_{MLE} , we calculated the relative efficiency (RE) of the estimate $\hat{V}_{B,HT}$ compared to the estimate $\hat{V}_{B,SE}$ defined by:

$$RE(\hat{V}_{B,HT}, \hat{V}_{B,SE}) = \frac{\int_{-\infty}^{\infty} [\hat{V}_{B,HT}(t) - V(t)]^2 dt}{\int_{-\infty}^{\infty} [\hat{V}_{B,SE}(t) - V(t)]^2 dt}. \quad (44)$$

If $RE = 1$, $\hat{V}_{B,HT}$ and $\hat{V}_{B,SE}$ will be interpreted as equally efficient. If $RE < 1$, $\hat{V}_{B,HT}$ is more efficient than $\hat{V}_{B,SE}$. To the contrary, if $RE > 1$, $\hat{V}_{B,HT}$ is less efficient than $\hat{V}_{B,SE}$. Similarly, we compared $\hat{V}_{B,HT}$ and \hat{V}_{MLE} . Bayesian estimates and MLEs for the parameter $\beta = 0.7054$ and $\theta = 1.7441$ (Table 7), averaged over 10000 repetitions, are used, for $n = 40$, to compare $\hat{V}_{B,HT}$, $\hat{V}_{B,SE}$ and \hat{V}_{MLE} using (44). The results are given in Table 8 and Table 9.

For the comparison of $\hat{V}_{B,HT}$ and $\hat{V}_{B,SE}$, the $RE(\hat{V}_{B,HT}, \hat{V}_{B,SE})$ is less than 1, which implies that the intensity function using $\hat{\beta}_{B,HT}$ and $\hat{\theta}_{B,HT}$ is more efficient than the intensity function under $\hat{\beta}_{B,SE}$ and $\hat{\theta}_{B,SE}$. Comparing $\hat{V}_{B,HT}$ and $\hat{V}_{B,SE}$ to \hat{V}_{MLE} , we obtained a similar result, establishing the superior relative efficiency of Bayesian estimates over MLE estimates. The corresponding graphs for the intensity functions are given in Figure 7.

In addition, $\hat{V}_{B,HT}^*$ and $\hat{V}_{B,SE}^*$ are computed using Bayesian estimates for β and MLE estimates θ , which were less efficient compare to \hat{V}_{MLE} , $\hat{V}_{B,SE}$, and $\hat{V}_{B,HT}$. Based on the results, the Bayesian estimates under the H-T loss function will be used to analyze the real data.

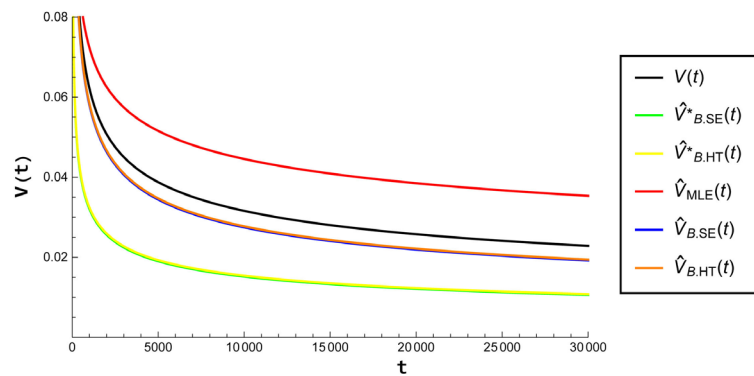


Figure 7. Estimates of the intensity function (of time t) using values in Table 7, $n = 40$.

Table 7. Averages of the Bayesian (under the under squared error and Higgin-Tsokos loss functions) and MLE estimates of β and θ .

β	$\hat{\beta}$	$\hat{\beta}_{B,SE}$	$\hat{\beta}_{B,HT}$	θ	$\hat{\theta}$	$\hat{\theta}_{B,SE}$	$\hat{\theta}_{B,HT}$
0.7054	0.743982	0.695989	0.696467	1.7441	2.73107	1.57545	1.58115

Table 8. Intensity functions with Bayesian and MLE estimates for β and θ .

$V(t)$	\hat{V}_{MLE}	$\hat{V}_{B,SE}$	$\hat{V}_{B,HT}$
$0.476465 \cdot t^{-0.2946}$	$0.352321 \cdot t^{-0.256018}$	$0.507238 \cdot t^{-0.304011}$	$0.5062 \cdot t^{-0.303533}$

Table 9. Relative efficiency of $\hat{V}_{B,HT}$ to \hat{V}_{MLE} and $\hat{V}_{B,SE}$.

$RE(\hat{V}_{B,SE}, \hat{V}_{MLE})$	$RE(\hat{V}_{B,HT}, \hat{V}_{MLE})$	$RE(\hat{V}_{B,HT}, \hat{V}_{B,SE})$
0.087746 0	0.07619190	0.868324 0

3.2. Using Real Data

Using the reliability growth data from **Table 1**, we computed $\hat{\beta}_{B,HT}$ and the adjusted estimate $\hat{\theta}_{B,HT}$ in order to obtain a Bayesian intensity function under H-T loss function. We followed **Algorithm 2** to obtain the Bayesian intensity function for the given real data.

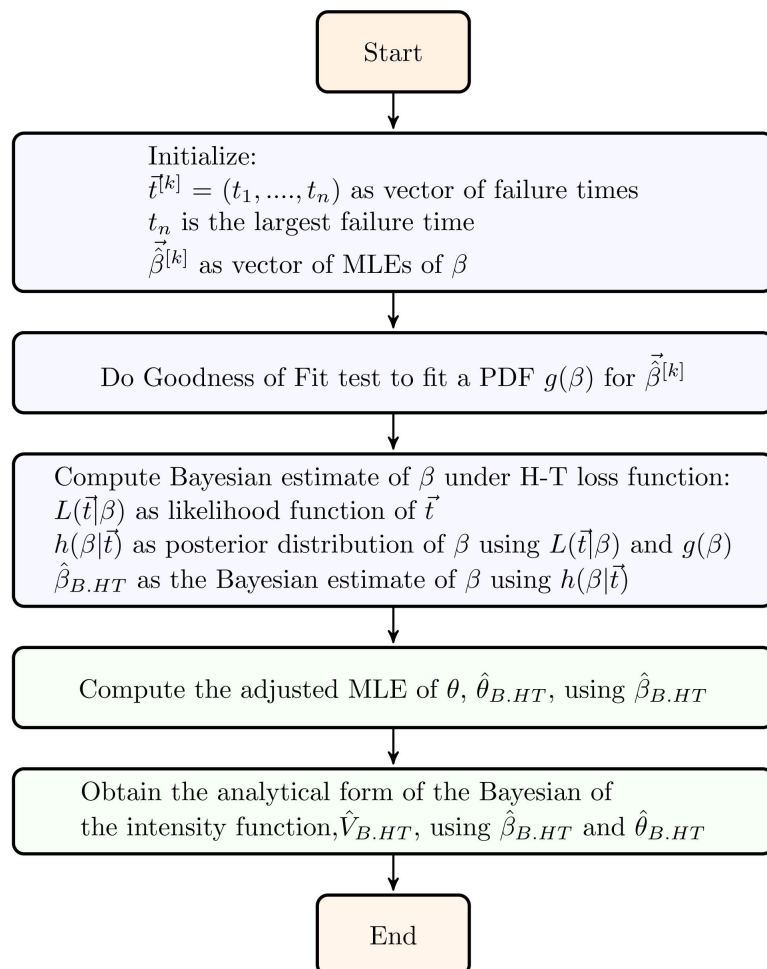
For the failure data of Crow, provided in **Table 1**, $\hat{\beta}_{B,HT}$ is approximately 0.501199 and $\hat{\theta}_{B,HT}$ is approximately 2.07144. Therefore, with the use of $\hat{\theta}_{B,HT}$, the Bayesian MLE of the intensity function ($\hat{V}_{B,HT}(t)$) for the data is given by:

$$\hat{V}_{B,HT}(t) = 0.347933 \cdot t^{-0.498801}, t > 0. \quad (45)$$

To obtain a Bayesian MLE for the reliability function under H-T loss function, we use this Bayesian estimate for the intensity function. The analytical form for the corresponding Bayesian reliability estimate, based on the data, is given by:

$$\hat{R}_{B,HT}(t_i | t_1, \dots, t_{i-1}) = \exp \left\{ -0.347933 \int_{t_{i-1}}^{t_i} x^{-0.498801} dx \right\}, t_i > t_{i-1} > 0. \quad (46)$$

Thus, the conditional reliability of the software given that the last two failure times were $t_{39} = 3181$ and $t_{40} = 3256.3$ is approximately 63%.

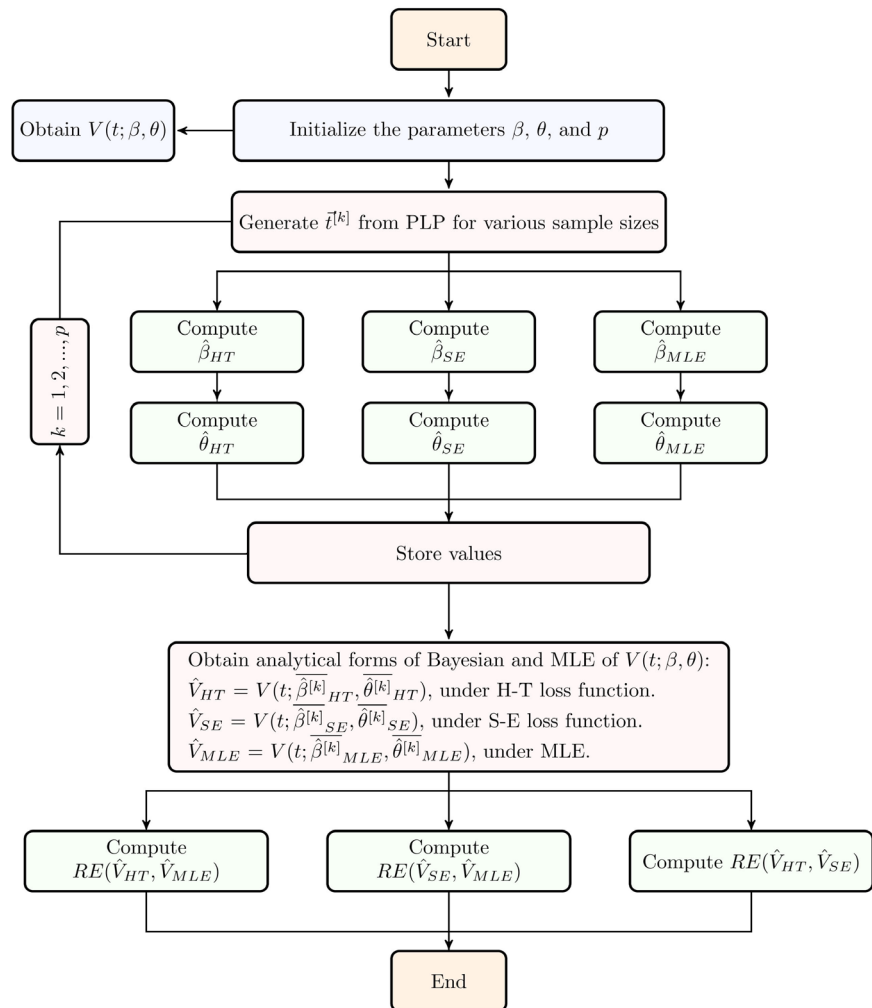


Algorithm 2. Estimate of the intensity function using Crow data in **Table 1**.

3.3. Sensitivity Analysis: Prior and Loss Function

To answer the second research question, “Is the Bayesian estimate of the intensity function, $V(t; \beta, \theta)$, of the PLP sensitive to the selections of the prior (both parametric and non-parametric priors) and loss function?”, we developed a simulation procedure, **Algorithm 3**, given below.

The algorithm compares the Bayesian and MLE estimates of the intensity function, $V(t; \beta, \theta)$, under different prior PDFs, for various sample sizes, with the H-T and S-E loss functions. The relative efficiency is used to compare these estimates of the $V(t; \beta, \theta)$. The relative efficiency with a value less than 1, larger than 1, and approximately equal to 1 indicate that the Bayesian estimates under the H-T loss function are more, less, equally efficient to the Bayesian estimate under the S-E loss function and the same analysis is applied when we compared to the MLE of $V(t; \beta, \theta)$, respectively. The algorithm starts by initializing the shape and scale parameters of the PLP, β and θ , respectively, and the number of iterations p .



Algorithm 3. Simulation to compare Bayesian and MLE estimates of the intensity function. Notations found in **Table 2**.

For various sample sizes ($n = 20, 40, 80, 140$), random failure times (time to failures) distributed according to the PLP are simulated using the initialized values of the PLP parameters. Then, the Bayesian and MLE estimates of the key parameter β are computed and used to compute the Bayesian estimates of θ , respectively. After a predetermined number of iterations, the average values of the Bayesian and MLE estimates of β and θ were used to obtain the analytical forms of the $V(t; \beta, \theta)$ under Bayesian, for both H-T and S-E loss functions and MLE, namely \hat{V}_{HT} , \hat{V}_{SE} , and \hat{V}_{MLE} , respectively. Informative parametric priors were considered such as the inverted gamma and the Burr PDFs, whereas the Jeffery prior was chosen as non-informative prior. In addition, probability kernel density function is selected as a non-parametric prior PDF. Probability kernel density estimation depends on the sample size, bandwidth, and the choice of the kernel function ($K(u)$). In this study, the optimal bandwidth (h^*) and kernel function were chosen to minimize the asymptotic mean integrated squared error (AMISE). The simplified form of the AMISE is reduced to:

$$\text{AMISE}(\hat{f}(\beta)) = \frac{C(K)}{n \cdot h} + \left(\frac{1}{4} \cdot h^4 \cdot k_2^2 \cdot R(f^{(2)}(\beta)) \right) \quad (47)$$

where:

- $C(K) = \int (K(u))^2 du$.
- n : sample size.
- h : bandwidth.
- $k_2 = \int_{-\infty}^{+\infty} u^2 \cdot K(u) du$.
- $f^{(2)}(\beta)$ is the second derivative of Burr PDF.
- $R(f^{(2)}(\beta)) = \int (f^{(2)}(\beta))^2 d\beta$.
- $h^* = \left[\frac{C(K)}{k_2^2 \cdot R(f^{(2)}(\beta))} \right]^{1/5} \cdot n^{-1/5}$.

AMISE was numerically calculated using the optimal bandwidth, with respect to different samples sizes for each kernel function considered in this study, namely Epanechnikov, Cosine, Biweight, Triweight, Gaussian, Triangle, Uniform, Tricube, and Logistic kernel functions. The results is given by **Table 10**.

The minimum AMISE corresponds to the Epanechnikov kernel function

($K(u) = \frac{3}{4}(1-u^2)I_{|u| \leq 1}$). In addition to the Epanechnikov kernel function, the

Gaussian kernel function ($K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) I_{\mathbb{R}}$) was also used in the calculation since it is commonly used for its analytical tractability.

Numerical integration techniques were used to compute the Bayesian estimates of the intensity function, $V(t; \beta, \theta)$, parameters under both H-T and S-E loss functions according to the equations defined in Section 2.3, for each of the parametric and non-parametric prior PDFs. Samples of size 20, 40, 80, and

Table 10. Calculations of the AMISE with respect to different sample size, optimal bandwidth, and kernel function.

Kernel function	Sample size				
	50	100	150	200	500
Epanechnikov	0.362827	0.208389	0.150662	0.119688	0.0575042
Cosine	0.362986	0.208481	0.150728	0.119741	0.0575295
Biweight	0.364607	0.209412	0.151401	0.120275	0.0577863
Triweight	0.366740	0.210637	0.152286	0.120979	0.0581243
Gaussian	0.377644	0.216900	0.156814	0.124576	0.0598525
Triangle	0.366972	0.210770	0.152383	0.121056	0.0581612
Uniform	0.384675	0.220938	0.159734	0.126895	0.0609669
Tricube	0.363433	0.208737	0.150913	0.119888	0.0576002
Logistic	0.399132	0.229241	0.165737	0.131665	0.0632582

140 were generated where the parameters β and θ were initialized to be 0.7054 and 1.7441, respectively. In the analytical form (17), f_1 and f_2 are conditioned to be positive numbers and play a big role in assigning the weight of loss depending on the estimator's behavior, whether underestimating or overestimating. Therefore, the simulation procedure was repeated three times according to the following cases:

- 1) $f_1 > f_2$
- 2) $f_1 < f_2$
- 3) $f_1 = f_2$

The results for 1000 repetitions, $f_1 > f_2$, and $n = 20, 40, 80, 140$, are shown in **Table 11**.

It can be observed that the Bayesian estimate of the $V(t; \beta, \theta)$ under the H-T loss function (\hat{V}_{HT}) and S-E loss function (\hat{V}_{SE}) had an outstanding efficiency compared to the MLE of the $V(t; \beta, \theta)$ (\hat{V}_{MLE}) for all sample sizes and prior PDFs, with the exception of the sample sizes 20 and 40 when inverted gamma PDF was the selected prior. The \hat{V}_{HT} was more efficient (6% - 11% estimation improvement) compared to the \hat{V}_{SE} when Burr PDF is selected to be the prior. The \hat{V}_{HT} had similar efficiency compared to the \hat{V}_{SE} when Jeffrey prior is selected and for large sample sizes, whereas unsurprisingly \hat{V}_{SE} was more efficient for small sample sizes since Jeffrey Bayesian estimate of the key parameter β tends to overestimate and for the H-T loss function gives more exponential weight on the extreme overestimate loss than the extreme under-estimate loss when $f_1 > f_2$. For Bayesian Gaussian and Epanechnikov kernel estimates, the \hat{V}_{HT} was more efficient compared to the \hat{V}_{SE} for sample sizes $n = 20, 40$ and 80 with 11% - 13% of estimation improvement even though they tend to underestimate and the H-T loss function puts more exponential weight on the extreme underestimation, but tend to have similar efficiency for sample size $n = 140$.

Table 11. The relative efficiency (RE) of the Bayesian estimate under H-T loss function, \hat{V}_{HT} when $f_1 > f_2$, compared to the Bayesian estimate under S-E loss function, \hat{V}_{SE} , and the MLE, \hat{V}_{MLE} , of $V(t; \beta, \theta)$.

Prior PDF		$RE(\hat{V}_{HT}, \hat{V}_{MLE})$	$RE(\hat{V}_{SE}, \hat{V}_{MLE})$	$RE(\hat{V}_{HT}, \hat{V}_{SE})$
Burr		0.1356	0.1519	0.8923
Inverted gamma		4.2461	4.1632	1.0199
Jeffrey	$n = 20$	0.0365	0.0289	1.2616
Gaussian kernel		0.1187	0.1346	0.8818
Epanechnikov kernel		0.1187	0.1346	0.8818
Burr		0.3047	0.3345	0.9107
Inverted gamma		6.3934	6.2832	1.0175
Jeffrey	$n = 40$	0.0166	0.0119	1.3947
Gaussian kernel		0.1234	0.1424	0.8663
Epanechnikov kernel		0.1221	0.1411	0.8659
Burr		0.0136	0.0151	0.9007
Inverted gamma		0.8058	0.7934	1.0156
Jeffrey	$n = 80$	0.0159	0.0144	1.1065
Gaussian kernel		0.0105	0.0117	0.8988
Epanechnikov kernel		0.0114	0.0127	0.8999
Burr		0.0035	0.0037	0.9367
Inverted gamma		0.1421	0.1399	1.0155
Jeffrey	$n = 140$	0.0040	0.0037	1.0680
Gaussian kernel		0.0019	0.0018	1.0119
Epanechnikov kernel		0.0021	0.0022	0.9670

The results for 1000 repetitions, $f_1 > f_2$, and $n = 20, 40, 80, 140$, are shown in **Table 12**.

Again, the Bayesian MLE estimate of the $V(t; \beta, \theta)$ under the H-T loss function (\hat{V}_{HT}) and S-E loss function (\hat{V}_{SE}) had an outstanding efficiency compared to the MLE of the $V(t; \beta, \theta)$ (\hat{V}_{MLE}) for all sample sizes and prior PDFs. When the inverted gamma was selected as prior, the \hat{V}_{HT} was more efficient compared to the \hat{V}_{SE} for all sample sizes with an approximately 2% of estimation improvement. As expected, the \hat{V}_{HT} was less efficient compared to the \hat{V}_{SE} when Burr PDF, and Gaussian and Epanechnikov kernel densities are selected as priors for sample sizes 20 and 40, since they tend to underestimate the $V(t; \beta, \theta)$ parameters, and the H-T loss function tends to put more weight on the extreme overestimation than on the extreme underestimation when $f_1 > f_2$. But the \hat{V}_{HT} and \hat{V}_{SE} had approximately similar efficiency for sample size $n = 80$, and the \hat{V}_{HT} tends to be slightly more efficient for large sample size ($n = 140$). The \hat{V}_{HT} was more efficient (4% - 24% estimation

Table 12. The relative efficiency (RE) of the Bayesian estimate under H-T loss function, \hat{V}_{HT} when $f_1 < f_2$, compared to the Bayesian estimate under S-E loss function, \hat{V}_{SE} , and the MLE, \hat{V}_{MLE} , of $V(t; \beta, \theta)$.

Prior PDF		$RE(\hat{V}_{HT}, \hat{V}_{MLE})$	$RE(\hat{V}_{SE}, \hat{V}_{MLE})$	$RE(\hat{V}_{HT}, \hat{V}_{SE})$
Burr		0.2068	0.1860	1.1116
Inverted gamma		4.7351	4.8309	0.9802
Jeffrey	$n = 20$	0.0232	0.0306	0.7589
Gaussian kernel		0.1948	0.1735	1.1226
Epanechnikov kernel		0.1949	0.1736	1.1227
Burr		0.1500	0.1327	1.1305
Inverted gamma		5.9173	6.0152	0.9837
Jeffrey	$n = 40$	0.0673	0.0785	0.8581
Gaussian kernel		0.0516	0.0431	1.1980
Epanechnikov kernel		0.051	0.0425	1.1985
Burr		0.0126	0.0121	1.0406
Inverted gamma		0.8155	0.8274	0.9856
Jeffrey	$n = 80$	0.0326	0.0349	0.9365
Gaussian kernel		0.0111	0.0108	1.0307
Epanechnikov kernel		0.0116	0.0112	1.0356
Burr		0.0180	0.0183	0.9814
Inverted gamma		0.2545	0.2576	0.9880
Jeffrey	$n = 140$	0.0329	0.0338	0.9733
Gaussian kernel		0.0222	0.0227	0.9762
Epanechnikov kernel		0.0204	0.0209	0.9772

improvement) compared to the \hat{V}_{SE} when Burr Jeffrey is chosen to be the prior PDF. The \hat{V}_{HT} had similar efficiency compared to the \hat{V}_{SE} for large sample sizes and when Jeffrey prior is selected, whereas unsurprisingly \hat{V}_{SE} was more efficient for small sample sizes since Jeffrey Bayesian estimate of the key parameter β tends to overestimate and for the H-T loss function gives more exponential weight on the extreme overestimate loss than the extreme under-estimate loss when $f_1 > f_2$. For Bayesian Gaussian and Epanechnikov kernel estimates, the \hat{V}_{HT} was more efficient compared to the \hat{V}_{SE} for sample sizes $n = 20, 40$ and 80 with 11% - 13% of estimation improvement even though they tend to underestimate and the H-T loss function puts more exponential weight on the extreme underestimation, but tend to have similar efficiency for sample size $n = 140$.

The results for 1000 repetitions, $f_1 > f_2$, and $n = 20, 40, 80, 140$, are shown in **Table 13**.

Table 13. The relative efficiency (RE) of the Bayesian estimate under H-T loss function, \hat{V}_{HT} when $f_1 = f_2$, compared to the Bayesian estimate under S-E loss function, \hat{V}_{SE} , and the MLE, \hat{V}_{MLE} , of $V(t; \beta, \theta)$.

Prior PDF		$RE(\hat{V}_{HT}, \hat{V}_{MLE})$	$RE(\hat{V}_{SE}, \hat{V}_{MLE})$	$RE(\hat{V}_{HT}, \hat{V}_{SE})$
Burr		0.0703	0.0702	1.0011
Inverted gamma		3.7132	3.7135	0.9999
Jeffrey	$n = 20$	0.0612	0.0613	0.9981
Gaussian kernel		0.0585	0.0583	1.0037
Epanechnikov kernel		0.0585	0.0583	1.0037
Burr		0.1195	0.1194	1.0008
Inverted gamma		7.3018	7.3022	0.9999
Jeffrey	$n = 40$	0.1351	0.1352	0.9993
Gaussian kernel		0.0384	0.0384	1.0008
Epanechnikov kernel		0.0381	0.0381	1.0008
Burr		0.0144	0.0144	1.0002
Inverted gamma		0.8626	0.8734	0.9876
Jeffrey	$n = 80$	0.0250	0.0250	0.9998
Gaussian kernel		0.0122	0.0122	1.0003
Epanechnikov kernel		0.0131	0.0131	1.0002
Burr		0.0065	0.0065	100000
Inverted gamma		0.1863	0.1863	100000
Jeffrey	$n = 140$	0.0117	0.0117	0.9999
Gaussian kernel		0.0070	0.0070	0.9999
Epanechnikov kernel		0.0064	0.0064	0.9999

Again, the Bayesian MLE estimate of the $V(t; \beta, \theta)$ under the H-T loss function (\hat{V}_{HT}) and S-E loss function (\hat{V}_{SE}) had an outstanding efficiency compared to the MLE of the $V(t; \beta, \theta)$ (\hat{V}_{MLE}) for all sample sizes and prior PDFs, with the exception of the sample sizes 20 and 40 when inverted gamma PDF was the selected prior. It is observed that both \hat{V}_{HT} and \hat{V}_{SE} had similar efficiency in estimation of the $V(t; \beta, \theta)$ for all sample sizes and priors considered in this study.

The sensitivity analysis shows that the Bayesian estimates of the intensity function of the PLP is sensitive to the prior and loss function selections. **Tables 11-13** indicate the efficiency of the Bayesian estimates under the H-T loss function when compared to the Bayesian estimate under S-E loss function and to the MLE, given that the engineer should choose the values of f_1 and f_2 based on his/her estimator's behaviour (underestimating and over estimating). Moreover, $f_1 > f_2$ is the recommended choice when the engineer selects Burr or kernel PDFs as their prior knowledge of the behavior of the key parameter β .

On the other hand, if the engineer does not have a prior knowledge of the key parameter β , it is still recommended to use H-T loss function in the Bayesian calculations with $f_1 < f_2$.

Thus far, we showed the more accuracy in estimating a software reliability when applying the Bayesian analysis under the H-T loss function compared to the Bayesian analysis under the S-E loss function and the MLE of the subject analysis. The performed extensive analysis requires efficiency in utilizing the existing programming languages, which therefore requires some programming experience, we developed an interactive user interface application using Wolfram language to compute and visualize the Bayesian and maximum likelihood estimates of the intensity and reliability functions of the Power Law Process for a given data.

4. Conclusions

In the present study, we developed the analytical Bayesian estimates of the key parameter β , under Higgin-Tsokos and squared-error loss functions, in the intensity function where the underlying failure distribution is the Power Law Process, that is used for software reliability assessment, among others. The reliability function of the subject model is written analytically as a function of the intensity function.

The behavior of the key parameter β is characterized by the Burr type XII probability distribution. Real data and numerical simulation were used to illustrate not only the robustness of the squared-error loss function being challenged by the assumption of the Higgins-Tsokos loss function, but also the efficiency improvement in the estimation of the intensity function of PLP under Higgins-Tsokos loss function ($\hat{V}_{B,HT}(t)$). For 100,000 samples of software failure times, based on Monte Carlo simulations and sample size of 40, the Bayesian estimate of β under Higgins-Tsokos loss function ($\hat{\beta}_{B,HT}$) performed slightly better than the Bayesian estimate of β under squared-error loss function ($\hat{\beta}_{B,SE}$) with respect to three different values of θ (0.5, 1.7441, 4). Even for different sample sizes (20, 30, 40, 50, 60, 70, 80, 100, 120, 140, and 160), similar results were achieved using $\beta = 0.7054$, $\theta = 1.7441$, and averaged over 10,000 samples of software failure times.

As the MLE of the second parameter in the intensity function (θ) depends on the estimate of β , the adjusted estimate of θ $\hat{\beta}_{B,HT}$ provided better performance compared to the adjusted estimate of θ using the $\hat{\beta}_{B,SE}(t)$. Moreover, the Relative Efficiency was used to compare the intensity function estimations, mainly using MLEs for both β and θ ($\hat{V}_{MLE}(t)$), using Bayesian estimate of β under the squared-error loss function and Bayesian of θ ($\hat{V}_{B,SE}(t)$), and using Bayesian estimate of β under the Higgins-Tsokos loss function and Bayesian of θ ($\hat{V}_{B,HT}(t)$), showing that $\hat{V}_{B,HT}(t)$ is more efficient in estimating the intensity function $V(t)$ with about 12% estimation improvement.

With respect to the question: Is the Bayesian estimate of the intensity function,

$V(t; \beta, \theta)$, of the PLP sensitive to the selections of the prior, both parametric and non-parametric priors, and loss function? The parametric prior PDFs were Burr, Jeffrey, and inverted gamma probability distributions whereas the non-parametric priors were Gaussian and Epanechnikov kernel densities. The priors' parameters were estimated using Crow failure times. Additionally, the optimal bandwidth and kernel functions were selected to minimize the asymptotic mean integrated squared error.

Using the developed algorithm, 1000 samples of software failure times with respect to four sample sizes of n (20, 40, 80, and 140) were generated from the PLP to compare the Bayesian estimates of $V(t; \beta, \theta)$ under the subject priors and loss functions using the Relative Efficiency among them. The simulation procedure was repeated three times for the cases when $f_1 > f_2$, $f_1 < f_2$, and $f_1 = f_2$. The results showed the efficacy of the Bayesian estimates of H-T loss function, and the choice of the f_1 and f_2 values depends on the prior knowledge of the key parameter β . It is recommended to choose values where $f_1 > f_2$ when the engineer thinks the prior knowledge of β is best characterized by Burr or Kernel based probability distributions with a proper justification, whereas a choice of $f_1 < f_2$ and Jeffery's prior is suggested when the engineer does not have a prior knowledge of β .

Thus, based on this aspect of our analysis, we can conclude that the Bayesian analysis approach under Higgins-Tsokos loss function not only as robust as the Bayesian analysis approach under squared error loss function but also performed better, where both are superior to the maximum likelihood approach in estimating the reliability function of the Power Law Process. The interactive user interface application can be used without any prior coding knowledge to compute and visualize the Bayesian and maximum likelihood estimates of the intensity and reliability functions of the Power Law Process for a given data.

Acknowledgements

We thank Majmaah University for funding the research, along with the support provided by the University of South Florida.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Duane, J.T. (1964) Learning Curve Approach to Reliability Monitoring. *IEEE Transactions on Aerospace*, **2**, 563-566. <https://doi.org/10.1109/TA.1964.4319640>
- [2] Crow, L.H. (1975) Tracking Reliability Growth. *Proceedings of the 20th Conference on Design of Experiments*, Report 75-2, US Army Research Office, Research Triangle Park, NC, 741-754.
- [3] Yamada, S., Ohba, M. and Osaki, S. (1983) S-Shaped Reliability Growth Modeling for Software Error Detection. *IEEE Transactions on Reliability*, **R-32**, 475-484.

- <https://doi.org/10.1109/TR.1983.5221735>
- [4] Goel, A.L. and Okumoto, K. (1979) Time-Dependent Error-Detection Rate Model for Software Reliability and Other Performance Measures. *IEEE Transactions on Reliability*, **R-28**, 206-211. <https://doi.org/10.1109/TR.1979.5220566>
 - [5] Calabria, R., Guida, M. and Pulcini, G. (1992) A Bayes Procedure for Estimation of Current System Reliability. *IEEE Transactions on Reliability*, **41**, 616-620. <https://doi.org/10.1109/24.249599>
 - [6] Bain, L. and Engelhardt, M. (1991) Statistical Analysis of Reliability and Life Testing Models. Marcel-Dekker, New York, NY, USA,.
 - [7] Goel, A.L. and Okumoto, K. (1984) Bayesian Inference for the Weibull Process with Applications to Assessing Software Reliability Growth and Predicting Software Failures. *Proc. Sixteenth Symp. Interface*, **R-28**, 206-211.
 - [8] Tsokos, C.P. and Rao, A.N.V. (1994) Estimation of Failure Intensity for the Weibull Process. *Reliability Engineering & System Safety*, **45**, 271-275. [https://doi.org/10.1016/0951-8320\(94\)90143-0](https://doi.org/10.1016/0951-8320(94)90143-0)
 - [9] Rigdon, S.E. and Basu, A.P. (1989) The Effect of Assuming a Homogeneous Poisson Process when the True Process Is a Power Law Process. *Journal of Quality Technology*, **22**, 111-117. <https://doi.org/10.1080/00224065.1990.11979222>
 - [10] Rigdon, S.E. and Basu, A.P. (1989) The Power Law Process: A Model for the Reliability of Repairable Systems. *Journal of Quality Technology*, **21**, 251-260. <https://doi.org/10.1080/00224065.1989.11979183>
 - [11] Rigdon, S.E. and Basu, A.P. (1990) Estimating the Intensity Function of a Power Law Process at the Current Time: Time Truncated Case. *Communications in Statistics—Simulation and Computation*, **19**, 1079-1104. <https://doi.org/10.1080/03610919008812906>
 - [12] Luo, L., Xing, L. and Levitin, G. (2018) Optimizing Dynamic Survivability and Security of Replicated Data in Cloud Systems under Co-Residence Attacks. *Reliability Engineering & System Safety*, in press. <https://doi.org/10.1016/j.res.2018.09.014>
 - [13] Movahedi, Y., Cukier, M., Andongabo, A. and Gashi, I. (2018) Cluster-Based Vulnerability Assessment of Operating Systems and Web Browsers. 2017 13th European Dependable Computing Conference, Geneva, 4-8 September 2017, 18-25. <https://doi.org/10.1109/EDCC.2017.27>
 - [14] Tsokos, C.P. and Xu, Y. (2011) Non-Homogenous Poisson Process for Evaluating Stage I & II Ductal Breast Cancer Treatment. *Journal of Modern Applied Statistical Methods*, **10**, 646-655. <https://doi.org/10.22237/jmasm/1320121320>
 - [15] But, A., Härkänen, T. and Haukka, J. (2017) Non-Parametric Bayesian Intensity Model: Exploring Time-to-Event Data on Two Time Scales. *Scandinavian Journal of Statistics*, **44**, 798-814, June. <https://doi.org/10.1111/sjos.12280>
 - [16] Turnbull, B.W., Abu-Libdeh, H. and Clark, L.C. (1990) Estimation of Failure Intensity for the Weibull Process. *Biometrics*, **46**, 1017-1034. <https://doi.org/10.2307/2532445>
 - [17] Raberto, M., Scalas, E., Ponta, L., Trinh, M. and Cincotti, S. (2019) Modeling Non-Stationarities in High-Frequency Financial Time Series. *Physica A: Statistical Mechanics and Its Applications*, **521**, 173-196. <https://doi.org/10.1016/j.physa.2019.01.069>
 - [18] Kieu, L.M. (2018) Analytical Modelling of Point Process and Application to Transportation. In: Zhou, J. and Chen, F., Eds., *Human and Machine Learning. Human-Computer Interaction Series*, Springer, Cham.

https://doi.org/10.1007/978-3-319-90403-0_19

- [19] Sayarshad, H.R. and Chow, J.Y.J. (2016) Survey and Empirical Evaluation of Non-homogeneous Arrival Process Models with Taxi Data. *Journal of Advanced Transportation*, **50**, 1275-1294. <https://doi.org/10.1002/atr.1401>
- [20] Qi, G., Pan, G., Li, S., Wu, Z., Zhang, D., Sun, L. and Yang, L.T. (2013) How Long a Passenger Waits for a Vacant Taxi—Large-Scale Taxi Trace Mining for Smart Cities. 2013 *IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing*, Beijing, 20-23 August 2013, 1029-1036. <https://doi.org/10.1109/GreenCom-iThings-CPSCoM.2013.175>
- [21] Menon, A.K. and Lee, Y. (2017) Predicting Short-Term Public Transport Demand via Inhomogeneous Poisson Processes. In: *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, ACM, New York, 2207-2210.
- [22] Yue, D., Zhao, G. and Yue, W. (2016) Analysis of a Multi-Server Queueing-Inventory System with Non-Homogeneous Poisson Arrivals. In: *Proceedings of the 11th International Conference on Queueing Theory and Network Applications*, ACM, New York.
- [23] Kimura, M., Toyota, T. and Yamada, S. (1999) Economic Analysis of Software Release Problems with Warranty Cost and Reliability Requirement. *Reliability Engineering & System Safety*, **66**, 49-55. [https://doi.org/10.1016/S0951-8320\(99\)00020-4](https://doi.org/10.1016/S0951-8320(99)00020-4)
- [24] Molinares, C.A. and Tsokos, C.P. (2013) Bayesian Reliability Approach to the Power Law Process with Sensitive Analysis to Prior Selection. *International Journal of Reliability, Quality and Safety Engineering*, **20**, No. 1. <https://doi.org/10.1142/S0218539313500046>
- [25] Alenezi, F.N. and Tsokos, C.P. (2018) Bayesian Reliability Analysis of the Power Law Process with Respect to the Higgins-Tsokos Loss Function for Modeling Software Failure Times. Submitted for Publication.
- [26] Higgins, J.J. and Tsokos, C.P. (1980) A Study of the Effect of the Loss Function on Bayes Estimates of Failure Intensity, MTBF, and Reliability. *Applied Mathematics and Computation*, **6**, 145-166. [https://doi.org/10.1016/0096-3003\(80\)90039-9](https://doi.org/10.1016/0096-3003(80)90039-9)
- [27] Crow, L.H. (1975) Reliability Analysis for Complex, Repairable Systems. In: Proschan, F. and Serfling, D.J., Eds., *Reliability and Biometry*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 379-410.
- [28] Jeffreys, H. (1946) An Invariant Form for the Prior Probability in Estimation Problems. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, **186**, 453-461. <https://doi.org/10.1098/rspa.1946.0056>