Article

A Three-Component Additive Weibull Distribution and Its Reliability Implications

Tien Thanh Thach (ID)

Faculty of Mathematics and Statistics, Ton Duc Thang University, Ho Chi Minh City 700000, Vietnam; thachthanhtien@tdtu.edu.vn

Abstract: We consider a new failure time distribution by combining three Weibull distributions. The proposed distribution exhibits decreasing, increasing, and bathtub-shaped failure rate functions. Some statistical properties of the proposed distribution are presented. The maximum likelihood method is used to estimate the parameters, reliability, and failure rate functions of the proposed distribution, and the uncertainties of the estimates are obtained by using asymptotic confidence intervals. The cross-entropy method, a stochastic optimization procedure, is used to find the maximum of the log-likelihood function. The superiority of the proposed distribution is demonstrated on two benchmark datasets.

Keywords: asymptotic confidence interval; bathtub-shaped failure rate function; cross-entropy method; maximum likelihood estimator; Weibull distribution

1. Introduction

The additive Weibull (AddW) distribution [1] is a distribution in reliability proposed for fitting bathtub-shaped failure rate data. The purpose of the AddW distribution is to provide a single model that can be used to model simultaneously all three phases of a bathtub curve. It combines two Weibull distributions [2], and its cumulative distribution function (CDF) is given by

$$F(x) = 1 - \exp\left[-(\theta x)^{\alpha} - (\lambda x)^{\beta}\right], \quad x \geq 0,$$

where θ and λ are non-negative and $0 < \beta < 1 < \alpha$. The shape parameters are restricted by $\beta < 1 < \alpha$ so that its failure rate function is bathtub-shaped only. Later, Lemonte et al. [3] has provided a detailed study on the AddW and has modified the conditions for the parameters as $\theta, \lambda > 0, 0 < \beta < \alpha$ which allows the AddW to have an increasing failure rate if $1 < \beta < \alpha$, a decreasing failure rate if $\beta < \alpha < 1$, and a bathtub-shaped failure rate if $\beta < 1 < \alpha$. Although the AddW distribution has bathtub-shaped failure rate functions, it is not flexible enough for fitting datasets with complex bathtub-shaped failure rates.

Many proposed distributions which result from combining different distributions, rather than purely from two Weibull distributions, have outperformed the AddW distribution. The new modified Weibull (NMW) distribution [4] combines the Weibull distribution and the modified Weibull distribution [5]. It has five parameters and its CDF is given by:

$$F(x) = 1 - e^{-\left(\alpha x^{\theta} + \beta x^{\gamma} e^{\lambda x}\right)}, \quad x \geq 0,$$

where $\alpha, \beta, \gamma, \theta > 0$, and $\lambda \geq 0$. A Bayes study of the improved NMW by using Hamiltonian Monte Carlo simulation is given in [6]. The additive modified Weibull (AMW)

distribution [7] combines the modified Weibull distribution and the Gompertz distribution [8]. This distribution has five parameters and its CDF is given by:

$$F(x) = 1 - e^{-\left(\alpha x^\theta e^{\gamma x} + e^{\lambda x - \beta} - e^{-\beta}\right)}, \quad x \geq 0,$$

where $\alpha, \beta, \theta > 0$, and $\gamma \lambda \geq 0$. The very flexible Weibull distributions [9] results from a linear combination of two logarithms of cumulative hazard functions. Its failure rate function can be monotone, bathtub-shaped, modified bathtub-shaped, or even upside-down bathtub-shaped. The log-normal modified Weibull distribution [10] results from combining the log-normal distribution and the modified Weibull distribution, and this distribution also has five parameters. The additive Chen–Weibull distribution [11] is a four-parameter distribution that combines the Weibull distribution and the Chen distribution. Recently, a generalization of the AddW distribution, called the generalized additive Weibull (GAW) distribution [12], has been proposed. Its CDF is given by:

$$F(x) = \left[1 - \exp\left(-\alpha x^\beta - \gamma x^\lambda\right)\right]^\theta, \quad x \geq 0,$$

where $\alpha, \theta, \gamma, \beta, \lambda > 0$. This distribution comprises a set of distributions, such as modified generalized linear failure rate, exponentiated exponential, exponentiated Weibull, modified Weibull distributions, and additive Weibull, among others.

In addition to the combined distributions that provide a flexible bathtub-shaped failure rate, other distributions also have a bathtub-shaped failure rate function. Some of them can be mentioned as the modified Weibull distribution [5], exponentiated Weibull distribution [13], Hjorth distribution [14], generalized modified Weibull distribution [15], beta modified Weibull distribution [16], upper truncated Weibull distribution [17], new three-parameter exponential-type distribution [18], beta Generalized Weibull distribution [19], Alpha logarithmic transformed Weibull distribution [20], logistic exponential distribution [21], and five-parameter spline distribution [22]. However, these distributions cannot provide bathtub-shaped failure rate functions with a long flat region.

More recently, the logarithmic transformed Weibull distribution was introduced by using a logarithm transformation method [23]. It has constant, increasing, decreasing, uni-modal, and unimodal then bathtub-shaped hazard rates. Shakhatreh et al. [24] introduced the generalized extended exponential-Weibull distribution for fitting non-monotonic failure rate data. This model includes the Weibull, generalized Weibull, exponentiated generalized linear exponential, and exponential-Weibull distributions. A new bounded distribution, called bounded weighted exponential distribution is introduced by Mallick et al. [25]. The proposed distribution exhibits increasing and bathtub-shaped failure rate functions.

In this study, we go one step further beyond the AddW distribution by introducing a distribution which combines three Weibull distributions. This distribution will be referred to as the three-component additive Weibull distribution (3CAW). The 3CAW distribution exhibits increasing, decreasing, and bathtub-shaped failure rate functions. It provides a flexible failure rate function, especially a bathtub-shaped failure rate with a long flat region, and it will be demonstrated to provide good fits to given well-known datasets. The 3CAW is useful for modeling either a series system with three independent components where each component follows a Weibull distribution or a component that is affected by three major failure modes, each following a Weibull distribution.

In fact, the combination of three Weibull components has already been proposed by Khalil et al. [26]. The model was named the flexible additive Weibull distribution and its CDF is given by:

$$F(x) = 1 - e^{-ax^b - cx^d - gx^k}, \quad x > 0,$$

where a, b, c, d, g, and $k > 0$. However, it is clear that this model cannot be identifiable, meaning that the same model might be produced by different sets of parameters' values. The identifiability of a distribution is a very important property of statistical models. It allows us to obtain precise estimates of the parameters' values. Without identifiability, we

cannot estimate the true parameters' values, even with an infinite number of observations. Although the idea of combining three Weibull distributions was already proposed by Khalil et al. [26], our independent study will differ from [26] in many aspects. Firstly, our proposed model will have a different parameterized form. Secondly, and most importantly, our proposed model will be proved to be identifiable. Thirdly, our study will use a stochastic optimization method that is more likely to find the global maximum since the log-likelihood function of the 6-parameter model is usually highly multimodal. Lastly, in the Applications section, we will carry out the assessments of all the uncertainties of the parameter estimates.

The rest of the paper is organized as follows. Section 2 introduces the 3CAW distribution and its reliability characteristics. Section 3 studies some statistical properties of the 3CAW distribution. Section 4 discusses the methods of maximum likelihood and asymptotic confidence interval estimations for parameters, reliability, and failure rate functions. Section 5 offers the simulation study. Section 6 brings applications of the 3CAW to two well-known data sets. Finally, Section 7 concludes the paper.

2. The 3CAW Distribution

In our previous studies [6,11] we realized that if we combine a model which has a bathtub-shaped failure rate function with a model which has an increasing failure rate function, we might obtain a model which has a very flexible bathtub-shaped failure rate function with a long flat region. In fact, we encountered such distribution in the literature, such as [4,7,10,11]. In this study, the 3CAW is formed by combining three Weibull components, where the first two components form a model which possibly exhibits a bathtub-shaped failure rate function and the last component can provide an increasing failure rate function. The CDF of the 3CAW which combines three Weibull distributions is given by:

$$F(x) = 1 - \exp\left[-(\theta x)^\alpha - (\lambda x)^\beta - (\eta x)^\gamma\right], \quad x > 0, \tag{1}$$

where $\theta, \lambda, \eta > 0$, and $0 < \alpha < \beta < \gamma$. The restriction of the shape parameters, i.e., $\alpha < \beta < \gamma$, leads to a very important property of the proposed model which will be stated in the following theorem.

Theorem 1 (Identifiability theorem). *The 3CAW distribution with the CDF given by Equation (1) is identifiable.*

Proof. The proof of Theorem 1 is given in Appendix A. □

The corresponding probability density function (PDF) is given by:

$$f(x) = \left[\alpha\theta(\theta x)^{\alpha-1} + \beta\lambda(\lambda x)^{\beta-1} + \gamma\eta(\eta x)^{\gamma-1}\right] \exp\left[-(\theta x)^\alpha - (\lambda x)^\beta - (\eta x)^\gamma\right].$$

It has the cumulative failure rate function given in the following form:

$$H(x) = (\theta x)^\alpha + (\lambda x)^\beta + (\eta x)^\gamma.$$

Its reliability and failure rate functions are, respectively, given by:

$$R(x) = \exp\left[-(\theta x)^\alpha - (\lambda x)^\beta - (\eta x)^\gamma\right],$$

and

$$h(x) = \alpha\theta(\theta x)^{\alpha-1} + \beta\lambda(\lambda x)^{\beta-1} + \gamma\eta(\eta x)^{\gamma-1}. \tag{2}$$

From the CDF of the 3CAW distribution, it is clear that it includes many other distributions as its sub-models. Table 1 shows a list of distributions that can be derived from the 3CAW distribution.

Table 1. The sub-models of the 3CAW.

Distribution	α	β	γ	θ	λ	η	$F(x)$	Reference
Flexible additive Weibull	b	d	h	$\sqrt[b]{a}$	$\sqrt[d]{c}$	$\sqrt[k]{g}$	$1 - e^{-ax^b - cx^d - gx^k}$	Khalil et al. [26]
Additive Weibull	-	-	0	-	-	0	$1 - e^{-(\theta x)^\alpha - (\lambda x)^\beta}$	Xie and Lai [1]
Non-linear failure rate	1	-	0	-	-	0	$1 - e^{-\theta x - (\lambda x)^\beta}$	Salem [27]
Linear failure rate	1	2	0	-	-	0	$1 - e^{-\theta x - (\lambda x)^2}$	Bain [28]
Weibull	-	0	0	-	0	0	$1 - e^{-(\theta x)^\alpha}$	Weibull [2]
Rayleigh	2	0	0	-	0	0	$1 - e^{-(\theta x)^2}$	Bain [28]
Exponential	1	0	0	-	0	0	$1 - e^{-\theta x}$	Bain [28]

The expression of the failure rate function in Equation (2) is the sum of three Weibull failure rate functions. It is decreasing if $\alpha < \beta < \gamma < 1$, increasing if $1 < \alpha < \beta < \gamma$, and bathtub-shaped if $\alpha < 1 < \beta < \gamma$ or $\alpha < \beta < 1 < \gamma$. Figure 1 shows the plots of the PDF and failure rate function of the 3CAW at selected parameter values. The 3CAW distribution has the PDF and failure rate function with flexible curves, and more importantly it has a long flat region of the bathtub-shaped failure rate function which is very important in reliability and biological studies.

(a)   (b)
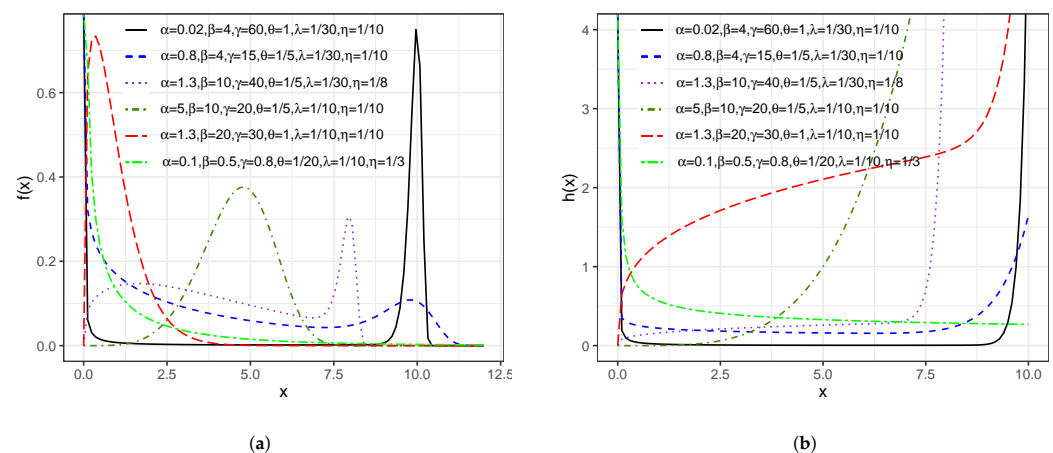
Figure 1. Plots of (**a**) PDF and (**b**) corresponding failure rate functions of the 3CAW.

3. Properties of the 3CAW Model

In this section, some statistical properties of the 3CAW distribution including the behavior of the failure rate function, the moments, order statistics, quantiles, and mean residual life are discussed.

3.1. The Failure Rate Function

We study the limiting behavior of $h(x)$ as x tend to 0 or $+\infty$.

- If $\gamma < 1$, then $\lim_{x \to 0} h(x) = +\infty$ and $\lim_{x \to +\infty} h(x) = 0$. In this case, the failure rate function is decreasing;
- If $\alpha > 1$, then $\lim_{x \to 0} h(x) = 0$ and $\lim_{x \to +\infty} h(x) = +\infty$. In this case, the failure rate function is increasing;
- If $\alpha < 1 < \beta$ or $\beta < 1 < \gamma$ then $\lim_{x \to 0} h(x) = +\infty$ and $\lim_{x \to +\infty} h(x) = +\infty$. In this case, the failure rate function is bathtub-shaped.

3.2. The Moments

In mathematical analysis we know that:

$$\int_0^{+\infty} x^r e^{-Ax^s} dx < +\infty, \quad \text{for all } A, s > 0 \text{ and } r \in \mathbb{Z}^+.$$

Therefore, from the expression of the CDF of the 3CAW, we see that the rth non-central moment μ_r (the rth moment about the origin) of the 3CAW always exists for all $r \in \mathbb{Z}^+$. It can be derived as follows:

$$
\begin{aligned}
\mu_r &= \int_0^{+\infty} x^r dF(x) \\
&= -\int_0^{+\infty} x^r de^{-(\theta x)^\alpha - (\lambda x)^\beta - (\eta x)^\gamma} \\
&= r \int_0^{+\infty} x^{r-1} e^{-(\theta x)^\alpha - (\lambda x)^\beta - (\eta x)^\gamma} dx,
\end{aligned}
\tag{3}
$$

for $r = 1, 2, \ldots$. Since (3) can not be obtained in closed form, some suitable numerical method, such as Gauss–Kronrod quadrature method, can be used to obtain this integral.

3.3. Order Statistics

Let a random sample $X_1, X_2, \ldots, X_n$ from the 3CAW distribution and $X_{k:n}$ is the kth order statistic of the sample, then the PDF of $X_{k:n}$ is given by:

$$
f_{k:n}(x) = \frac{1}{B(k, n-k+1)} F(x)^{k-1} (1 - F(x))^{n-k} f(x),
$$

where $B(\cdot, \cdot)$ is the beta function. Since:

$$
(1 - F(x))^{n-k} = e^{-(n-k)H(x)},
$$

and

$$
F(x)^{k-1} = (1 - e^{-H(x)})^{k-1} = \sum_{l=0}^{k-1} \binom{k-1}{l} (-1)^l e^{-lH(x)}.
$$

The PDF of $X_{k:n}$ is derived as:

$$
\begin{aligned}
f_{k:n}(x) &= \frac{1}{B(k, n-k+1)} \sum_{l=0}^{k-1} \binom{k-1}{l} (-1)^l h(x) e^{-(n+l+1-k)H(x)} \\
&= n \binom{n-1}{k-1} \sum_{l=0}^{k-1} \binom{k-1}{l} \frac{(-1)^l}{n+l+1-k} f(x; \alpha, \beta, \gamma, \theta', \lambda', \eta'),
\end{aligned}
$$

where $\theta' = \theta \sqrt[\alpha]{n+l+1-k}$, $\lambda' = \lambda \sqrt[\beta]{n+l+1-k}$ and $\eta' = \eta \sqrt[\gamma]{n+l+1-k}$.

3.4. Quantiles, Median, and Mode

The quantile x_q is the solution of the following non-linear equation

$$
(\theta x_q)^\alpha + (\lambda x_q)^\beta + (\eta x_q)^\gamma + \log(1 - q) = 0.
\tag{4}
$$

If $q = 1/2$, the quantile x_q is the median.

The mode is the value of x for which the PDF $f(x)$ attains its maximum. Therefore, the mode is the solution of the following non-linear equation:

$$
h'(x)R(x) + h(x)R'(x) = 0
\tag{5}
$$

where $R(x)$ and $h(x)$ are the reliability failure rate functions, respectively, and $R'(x)$ and $h'(x)$ are their corresponding derivatives with respect to x. Solving (4) and (5) requires some suitable numerical methods.

Data $x_1, \ldots, x_n$ can be simulated from the 3CAW distribution by using a simple algorithm as follows:

1:   Generate $u_i \sim U(0, 1)$, for $i = 1, \ldots, n$;
2:   Solve $(\theta x_i)^\alpha + (\lambda x_i)^\beta + (\eta x_i)^\gamma = \log \frac{1}{1 - u_i}$.

3.5. Mean Residual Life

The mean residual life at time x of a random variable $X \sim 3\text{CAW}$ distribution is the expectation of the conditional random variable $X - x | X > x$. It is the mean time to failure of an object that has survived to time x, and it is defined as:

$$M(x) = \mathbb{E}[X - x | X > x] = \frac{1}{R(x)} \int_x^{+\infty} R(u)du = \frac{1}{R(x)} \int_0^{+\infty} R(u + x)du$$

$$= e^{(\theta x)^\alpha + (\lambda x)^\beta + (\eta x)^\gamma} \int_0^{+\infty} e^{-(\theta(u+x))^\alpha - (\lambda(u+x))^\beta - (\eta(u+x))^\gamma} du. \tag{6}$$

In order to obtain the integral in (6), suitable numerical methods need to be used.

4. Parameter Estimation

4.1. Maximum Likelihood Estimation

Let $\mathcal{D}: x_1, \ldots, x_n$ be an observed sample from the 3CAW distribution with unknown parameter vector $\boldsymbol{\theta} = (\alpha, \beta, \gamma, \theta, \lambda, \eta)^T$. The likelihood function of the 3CAW is defined as:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n f(x_i; \boldsymbol{\theta})$$

$$= \prod_{i=1}^n \left[\left(\alpha\theta(\theta x_i)^{\alpha-1} + \beta\lambda(\lambda x_i)^{\beta-1} + \gamma\eta(\eta x_i)^{\gamma-1} \right) e^{-(\theta x_i)^\alpha - (\lambda x_i)^\beta - (\eta x_i)^\gamma} \right].$$

Then, the log-likelihood function is derived as:

$$l = \sum_{i=1}^n \log \left[\alpha\theta(\theta x_i)^{\alpha-1} + \beta\lambda(\lambda x_i)^{\beta-1} + \gamma\eta(\eta x_i)^{\gamma-1} \right] - \sum_{i=1}^n \left[(\theta x_i)^\alpha + (\lambda x_i)^\beta + (\eta x_i)^\gamma \right]. \tag{7}$$

We maximize this function to obtain the maximum likelihood estimates (MLEs) $\hat{\boldsymbol{\theta}}_{\text{ML}} = (\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\theta}, \hat{\lambda}, \hat{\eta})^T$ of the parameter vector $\boldsymbol{\theta} = (\alpha, \beta, \gamma, \theta, \lambda, \eta)^T$. This can be written in mathematical form as:

$$\hat{\boldsymbol{\theta}}_{\text{ML}} = \arg \max_{\boldsymbol{\theta} \in \Theta} l(\boldsymbol{\theta}),$$

where $\Theta = \{(\alpha, \beta, \gamma, \theta, \lambda, \eta)^T : \theta, \lambda, \eta > 0 \text{ and } 0 < \alpha < \beta < \gamma\}$.

One way to do so is to first calculate the first-order partial derivatives of the log-likelihood function with respect to the parameters:

$$\frac{\partial l}{\partial \alpha} = \sum_{i=1}^n \frac{\theta(\theta x_i)^{\alpha-1}(1 + \alpha \log(\theta x_i))}{\alpha\theta(\theta x_i)^{\alpha-1} + \beta\lambda(\lambda x_i)^{\beta-1} + \gamma\eta(\eta x_i)^{\gamma-1}} - \sum_{i=1}^n (\theta x_i)^\alpha \log(\theta x_i),$$

$$\frac{\partial l}{\partial \beta} = \sum_{i=1}^n \frac{\lambda(\lambda x_i)^{\beta-1}(1 + \beta \log(\lambda x_i))}{\alpha\theta(\theta x_i)^{\alpha-1} + \beta\lambda(\lambda x_i)^{\beta-1} + \gamma\eta(\eta x_i)^{\gamma-1}} - \sum_{i=1}^n (\lambda x_i)^\beta \log(\lambda x_i),$$

$$\frac{\partial l}{\partial \gamma} = \sum_{i=1}^n \frac{\eta(\eta x_i)^{\gamma-1}(1 + \gamma \log(\eta x_i))}{\alpha\theta(\theta x_i)^{\alpha-1} + \beta\lambda(\lambda x_i)^{\beta-1} + \gamma\eta(\eta x_i)^{\gamma-1}} - \sum_{i=1}^n (\eta x_i)^\gamma \log(\eta x_i),$$

$$\frac{\partial l}{\partial \theta} = \sum_{i=1}^n \frac{-(\alpha\theta)^2(\theta x_i)^{\alpha-1}}{\alpha\theta(\theta x_i)^{\alpha-1} + \beta\lambda(\lambda x_i)^{\beta-1} + \gamma\eta(\eta x_i)^{\gamma-1}} + \alpha\theta \sum_{i=1}^n (\theta x_i)^\alpha,$$

$$\frac{\partial l}{\partial \lambda} = \sum_{i=1}^n \frac{-(\beta\lambda)^2(\lambda x_i)^{\beta-1}}{\alpha\theta(\theta x_i)^{\alpha-1} + \beta\lambda(\lambda x_i)^{\beta-1} + \gamma\eta(\eta x_i)^{\gamma-1}} + \beta\lambda \sum_{i=1}^n (\lambda x_i)^\beta,$$

$$\frac{\partial l}{\partial \eta} = \sum_{i=1}^n \frac{-(\gamma\eta)^2(\eta x_i)^{\gamma-1}}{\alpha\theta(\theta x_i)^{\alpha-1} + \beta\lambda(\lambda x_i)^{\beta-1} + \gamma\eta(\eta x_i)^{\gamma-1}} + \gamma\eta \sum_{i=1}^n (\eta x_i)^\gamma,$$

and then set these partial derivatives equals zero and solve the system of non-linear equations by using some suitable numerical methods, such as Newton–Raphson or quasi-Newton methods.

Another better way to maximize the log-likelihood function in Equation (7) is to use the stochastic optimization procedure, called cross-entropy (CE) method [29], to maximize the log-likelihood function. The reason why CE works better than gradient-type methods is that the likelihood function is usually highly multimodal (that is, there are many local maxima and minima). For such multi-extremal optimization problems, a gradient-type method would typically find a local maximum rather than a global maximum. CE, on the other hand, is a global optimization technique that does not require local (i.e., gradient) information, and so is more likely to find the global maximum.

We prefer the latter due to its advantages in searching for a global maximum, and it is easy to implement by using the CE package in R or Matlab. All users need to do is to define the log-likelihood function, define the parameter space, specified algorithm's initial values, and run the algorithm through trial and error until we find the best optimizer which has the largest log-likelihood. The CE algorithm and R code for continuous optimization are, respectively, given in Appendices E and F.

After obtaining the MLEs of the model parameters, using the invariant property of MLEs, we obtain the MLEs of $R(x)$ and $h(x)$ at a certain time x as:

$$\hat{R}(x) = \exp\left[-(\hat{\theta}x)^{\hat{\alpha}} - (\hat{\lambda}x)^{\hat{\beta}} - (\hat{\eta}x)^{\hat{\gamma}}\right],$$

$$\hat{h}(x) = \hat{\alpha}\hat{\theta}(\hat{\theta}x)^{\hat{\alpha}-1} + \hat{\beta}\hat{\lambda}(\hat{\lambda}x)^{\hat{\beta}-1} + \hat{\gamma}\hat{\eta}(\hat{\eta}x)^{\hat{\gamma}-1}.$$

4.2. Asymptotic Confidence Interval

To obtain the asymptotic confidence intervals (CIs) for the parameters, and the reliability and failure rate functions, we first calculate the symmetric matrix of second-order partial derivatives.

$$\nabla^2 l = \begin{bmatrix} \frac{\partial^2 l}{\partial \alpha^2} & \frac{\partial^2 l}{\partial \alpha \partial \beta} & \frac{\partial^2 l}{\partial \alpha \partial \gamma} & \frac{\partial^2 l}{\partial \alpha \partial \theta} & \frac{\partial^2 l}{\partial \alpha \partial \lambda} & \frac{\partial^2 l}{\partial \alpha \partial \eta} \\ \cdot & \frac{\partial^2 l}{\partial \beta^2} & \frac{\partial^2 l}{\partial \beta \partial \gamma} & \frac{\partial^2 l}{\partial \beta \partial \theta} & \frac{\partial^2 l}{\partial \beta \partial \lambda} & \frac{\partial^2 l}{\partial \beta \partial \eta} \\ \cdot & \cdot & \frac{\partial^2 l}{\partial \gamma^2} & \frac{\partial^2 l}{\partial \gamma \partial \theta} & \frac{\partial^2 l}{\partial \gamma \partial \lambda} & \frac{\partial^2 l}{\partial \gamma \partial \eta} \\ \cdot & \cdot & \cdot & \frac{\partial^2 l}{\partial \theta^2} & \frac{\partial^2 l}{\partial \theta \partial \lambda} & \frac{\partial^2 l}{\partial \theta \partial \eta} \\ \cdot & \cdot & \cdot & \cdot & \frac{\partial^2 l}{\partial \lambda^2} & \frac{\partial^2 l}{\partial \lambda \partial \eta} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \frac{\partial^2 l}{\partial \eta^2} \end{bmatrix},$$

where the expressions for second-order partial derivatives are provided in Appendix C. Then, the observed (Fisher) information matrix is defined as

$$\mathbf{I}(\boldsymbol{\theta}) = -\nabla^2 l.$$

The matrix $\mathbf{I}$ can be evaluated at $\hat{\boldsymbol{\theta}}_{\text{ML}}$, the MLE of $\boldsymbol{\theta}$, by using the plug-in method

$$\mathbf{I}(\hat{\boldsymbol{\theta}}_{\text{ML}}) = -\nabla^2 l\Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{\text{ML}}}.$$

The approximate variance matrix is obtained by

$$\hat{V} = \begin{bmatrix} Var(\hat{\alpha}) & Cov(\hat{\alpha},\hat{\beta}) & Cov(\hat{\alpha},\hat{\gamma}) & Cov(\hat{\alpha},\hat{\theta}) & Cov(\hat{\alpha},\hat{\lambda}) & Cov(\hat{\alpha},\hat{\eta}) \\ \cdot & Var(\hat{\beta}) & Cov(\hat{\beta},\hat{\gamma}) & Cov(\hat{\alpha},\hat{\theta}) & Cov(\hat{\beta},\hat{\lambda}) & Cov(\beta,\hat{\eta}) \\ \cdot & \cdot & Var(\hat{\gamma}) & Cov(\hat{\gamma},\hat{\theta}) & Cov(\hat{\gamma},\hat{\lambda}) & Cov(\hat{\gamma},\hat{\eta}) \\ \cdot & \cdot & \cdot & Var(\hat{\theta}) & Cov(\hat{\theta},\hat{\lambda}) & Cov(\hat{\theta},\hat{\eta}) \\ \cdot & \cdot & \cdot & \cdot & Var(\hat{\lambda}) & Cov(\hat{\lambda},\hat{\eta}) \\ \cdot & \cdot & \cdot & \cdot & \cdot & Var(\hat{\eta}) \end{bmatrix} \approx \mathbf{I}^{-1}(\hat{\boldsymbol{\theta}}_{\text{ML}}).$$

Then, the CIs for the MLEs can be constructed by using the following approximation [30].

$$\hat{\boldsymbol{\theta}}_{\text{ML}} \approx \mathcal{N}(\boldsymbol{\theta}, \hat{V}.)$$

The $100(1 - \xi)\%$ CIs for $\alpha, \beta, \gamma, \theta, \lambda$, and η are, respectively, defined as:

$$\hat{\alpha} \pm z_{1-\frac{\xi}{2}}\sqrt{Var(\hat{\alpha})}, \quad \hat{\beta} \pm z_{1-\frac{\xi}{2}}\sqrt{Var(\hat{\beta})},$$

$$\hat{\gamma} \pm z_{1-\frac{\xi}{2}}\sqrt{Var(\hat{\gamma})}, \quad \hat{\theta} \pm z_{1-\frac{\xi}{2}}\sqrt{Var(\hat{\theta})},$$

$$\hat{\lambda} \pm z_{1-\frac{\xi}{2}}\sqrt{Var(\hat{\lambda})}, \quad \hat{\eta} \pm z_{1-\frac{\xi}{2}}\sqrt{Var(\hat{\eta})},$$

where $z_{1-\xi/2}$ is the $100(1 - \xi/2)$ percentage point of $N(0,1)$.

The delta method which was discussed in [31] and used recently by [32] is used to approximate the variance of $\hat{R}(x)$ and $\hat{h}(x)$.

$$\hat{\sigma}^2_{\hat{R}(x)} = \left[\nabla\hat{R}(x)\right]^T \left[\hat{V}\right]\left[\nabla\hat{R}(x)\right], \quad \hat{\sigma}^2_{\hat{h}(x)} = \left[\nabla\hat{h}(x)\right]^T \left[\hat{V}\right]\left[\nabla\hat{h}(x)\right],$$

where $\nabla\hat{R}(x)$ and $\nabla\hat{h}(x)$ are, respectively, the gradient of $R(x)$ and $h(x)$ with respect to $\alpha, \beta, \gamma, \theta, \lambda$, and η evaluated at $\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\theta}, \hat{\lambda}$, and $\hat{\eta}$. The gradient of $h(x)$ and $R(x)$ with respect to $\alpha, \beta, \gamma, \theta, \lambda$, and η are, respectively, given in Appendices B and D. Then, the $100(1 - \xi)\%$ CIs for $R(x)$ and $h(x)$ at time x are, respectively, defined as:

$$\hat{R}(x) \pm z_{1-\frac{\xi}{2}}\sqrt{\hat{\sigma}^2_{\hat{R}(x)}}, \quad \hat{h}(x) \pm z_{1-\frac{\xi}{2}}\sqrt{\hat{\sigma}^2_{\hat{h}(x)}}.$$

Although the parameters are positive, the CIs may result in a negative lower bound. In this regard, we can replace the negative values by zero. Another way is to use the normal approximation of log-transformed MLE [33]. Then, the $100(1 - \xi)\%$ CIs for $\alpha, \beta, \gamma, \theta, \lambda$, and η, are, respectively, defined as:

$$\hat{\alpha}\exp\left[\pm\frac{z_{1-\frac{\xi}{2}}\sqrt{Var(\hat{\alpha})}}{\hat{\alpha}}\right], \quad \hat{\beta}\exp\left[\pm\frac{z_{1-\frac{\xi}{2}}\sqrt{Var(\hat{\beta})}}{\hat{\beta}}\right],$$

$$\hat{\gamma}\exp\left[\pm\frac{z_{1-\frac{\xi}{2}}\sqrt{Var(\hat{\gamma})}}{\hat{\gamma}}\right], \quad \hat{\theta}\exp\left[\pm\frac{z_{1-\frac{\xi}{2}}\sqrt{Var(\hat{\theta})}}{\hat{\theta}}\right],$$

$$\hat{\lambda}\exp\left[\pm\frac{z_{1-\frac{\xi}{2}}\sqrt{Var(\hat{\lambda})}}{\hat{\lambda}}\right], \quad \hat{\eta}\exp\left[\pm\frac{z_{1-\frac{\xi}{2}}\sqrt{Var(\hat{\eta})}}{\hat{\eta}}\right].$$

Likewise, the $100(1 - \xi)\%$ CIs for $\hat{R}(x)$ and $\hat{h}(x)$ are, respectively, defined as:

$$\hat{R}(x)\exp\left[\pm\frac{z_{1-\frac{\xi}{2}}\sqrt{\hat{\sigma}^2_{\hat{R}(x)}}}{\hat{R}(x)}\right], \quad \hat{h}(x)\exp\left[\pm\frac{z_{1-\frac{\xi}{2}}\sqrt{\hat{\sigma}^2_{\hat{h}(x)}}}{\hat{h}(x)}\right].$$

5. Simulation Results

We conduct a Monte Carlo simulation study to assess the performances of the proposed estimation methods for the proposed model. We first select five different sets of parameters' values. For simplicity, we set all scale parameters to unity, i.e., $\theta = \lambda = \eta = 1$, in all cases, and vary the values of the shape parameters as given in Table 2. For each set of parameters' values, we simulate 1000 datasets with sample sizes $n = 25, 50, 100$, and 200, respectively, from the 3CAW distribution by using the algorithm given in Section 3.4. Then, we find the

MLEs of the parameters for fitting the 3CAW to each dataset, and calculate the biases and standard deviations (SDs) for the parameter estimates.

The results of the simulation study are listed in Table 2. The outcomes show that the MLEs of the parameters (except β and γ) have low bias and SD in virtually all the cases, demonstrating that the estimated values are close to the parameters' true values. In the results, we see that $\hat{\gamma}$ has the largest biases and SDs compared with the other estimators. The biases and SDs of $\hat{\beta}$ and $\hat{\gamma}$ increase as we increase the values of β and γ. This can be understandable because, with relatively small sample sizes, it is hard to estimate exactly the value of the Weibull shape parameter when it becomes larger and larger. In that case, the shape of the Weibull distribution does not change much when slightly varying the value of the shape parameter. The biases of $\hat{\lambda}$ are always negative in all cases which means that the MLE underestimates the parameter λ. In almost all instances, the biases and SDs decrease as the sample size increases, indicating the consistency of the proposed estimators.

Table 2. Biases and standard deviations of the 3CAW parameter estimates for some selected parameters' values at different sample sizes.

n	Bias						Standard Deviation					
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\eta}$
	$\alpha = 0.1, \beta = 2, \gamma = 10, \theta = 1, \lambda = 1, \eta = 1$											
25	0.0042	0.0877	1.3684	0.3379	−0.0195	0.0236	0.0288	1.1173	1.7485	1.3094	0.4039	0.1128
50	−0.0004	0.0645	1.0510	0.2783	−0.0094	0.0058	0.0185	0.9055	1.7630	1.2177	0.2560	0.0588
100	0.0001	0.0832	0.7688	0.3036	−0.0038	−0.0006	0.0126	0.7241	1.5956	1.0784	0.1810	0.0495
200	0.0000	0.0328	0.5359	0.1946	−0.0085	−0.0009	0.0087	0.5193	1.4143	0.9082	0.1399	0.0320
	$\alpha = 0.2, \beta = 4, \gamma = 20, \theta = 1, \lambda = 1, \eta = 1$											
25	0.0065	0.1413	1.3990	0.2191	−0.0448	0.0138	0.0548	1.6876	2.1558	1.0436	0.2270	0.0420
50	0.0046	0.0337	0.8171	0.2309	−0.0276	0.0042	0.0376	1.4517	2.2321	0.9281	0.1360	0.0291
100	0.0029	0.0137	0.5407	0.1184	−0.0119	0.0015	0.0251	1.1596	1.9485	0.6962	0.0888	0.0204
200	0.0004	0.0100	0.4517	0.0693	−0.0034	0.0001	0.0181	0.9421	1.7966	0.5369	0.0585	0.0141
	$\alpha = 0.4, \beta = 6, \gamma = 30, \theta = 1, \lambda = 1, \eta = 1$											
25	0.0263	0.1721	1.3080	0.1419	−0.0714	0.0075	0.1143	2.0221	2.3351	0.7280	0.2273	0.0274
50	0.0122	0.0848	0.8988	0.0682	−0.0208	0.0030	0.0770	1.6977	2.3738	0.5192	0.1039	0.0174
100	0.0048	0.0987	0.6592	0.0582	−0.0096	0.0002	0.0532	1.4810	2.2169	0.3873	0.0608	0.0126
200	−0.0001	0.0268	0.4164	−0.0049	−0.0028	0.0003	0.0352	1.2605	2.0607	0.2644	0.0395	0.0087
	$\alpha = 0.6, \beta = 8, \gamma = 40, \theta = 1, \lambda = 1, \eta = 1$											
25	0.0332	0.2336	1.4777	0.0530	−0.0541	0.0054	0.1727	2.1670	2.4230	0.5059	0.1903	0.0191
50	0.0130	0.1410	0.8352	0.0288	−0.0171	0.0021	0.1190	1.8803	2.5656	0.3610	0.0906	0.0129
100	0.0084	0.1571	0.6749	0.0145	−0.0060	0.0012	0.0817	1.6897	2.4045	0.2641	0.0449	0.0091
200	0.0019	0.0820	0.4934	0.0069	−0.0022	0.0001	0.0546	1.4847	2.1561	0.1843	0.0295	0.0067
	$\alpha = 0.8, \beta = 10, \gamma = 50, \theta = 1, \lambda = 1, \eta = 1$											
25	0.0536	0.1830	1.2681	0.0337	−0.0608	0.0049	0.2244	2.2711	2.5538	0.3692	0.1741	0.0167
50	0.0213	0.1323	1.0219	0.0051	−0.0144	0.0017	0.1527	2.0815	2.6315	0.2800	0.0776	0.0107
100	0.0030	0.0653	0.7603	0.0074	−0.0071	0.0008	0.1022	1.8486	2.5559	0.2047	0.0370	0.0077
200	0.0084	0.0476	0.4831	0.0038	−0.0021	0.0004	0.0766	1.6348	2.3628	0.1417	0.0247	0.0053

6. Applications

We use two real-life datasets known as the Aarset data [34] and Meeker and Escobar data [33] to investigate the performance of the 3CAW distribution. We also compare this distribution to some distributions, such as the AddW, NMW, AMW, and GAW. The log-likelihood ($\log(L)$) value, Kolmogorov–Smirnov (K-S) test statistic, Akaike information criterion (AIC), bias-corrected Akaike information criterion (AICc) (the adjusted version of AIC for small dataset), and Bayesian information criterion (BIC) are used for model selection. Among a set of alternative models, the model with the largest $\log(L)$ value and smallest K-S, AIC, BIC, and AICc values can be considered to be the best approximation model. Notice that K-S and AIC (as well as the AICc and BIC) values do not necessarily follow the same trend (i.e., the model with the smallest K-S may not have the smallest AIC).

We calculate the CIs for the MLEs of the parameters, reliability, and failure rate functions of the 3CAW by using the log-transformed MLE method mentioned above.

6.1. Aarset Data

Data in Table 3 represent the times to failure of 50 devices (in weeks) [34]. The empirical scaled TTT-transform plot [35] indicated that this dataset has a bathtub-shaped failure rate because it is first convex and then concave, as shown in Figure 2b, which matches the blue-dashed line in Figure 2a.

Table 3. Times to failure of $n = 50$ devices put on life test (in weeks) [34].

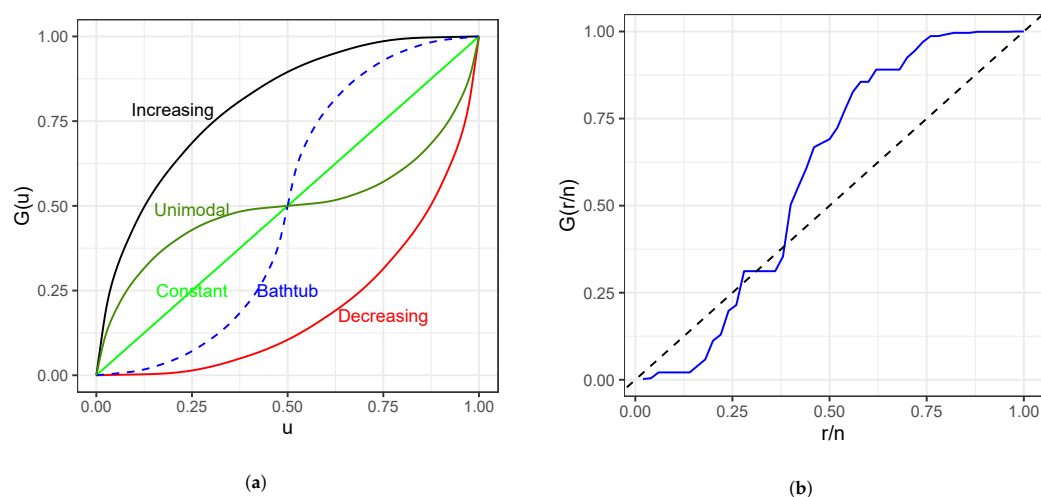|									
0.1	0.2	1.0	1.0	1.0	1.0	1.0	2.0	3.0	6.0
7.0	11.0	12.0	18.0	18.0	18.0	18.0	18.0	21.0	32.0
36.0	40.0	45.0	46.0	47.0	50.0	55.0	60.0	63.0	63.0
67.0	67.0	67.0	67.0	72.0	75.0	79.0	82.0	82.0	83.0
84.0	84.0	84.0	85.0	85.0	85.0	85.0	85.0	86.0	86.0

(a)

(b)

Figure 2. (**a**) Types of failure rate identify by shapes of the TTT-plot and (**b**) the empirical TTT-plot (blue) for Aarset data, where $G(u) = K^{-1}(u)/K^{-1}(1)$ for $0 < u < 1$, where $K^{-1}(u) = \int_0^{F^{-1}(u)} R(t)dt$, and its empirical version is $G_n(r/n) = \left(\sum_{i=1}^{r} x_{i:n} + (n-r)x_{r:n}\right)/\sum_{i=1}^{n} x_{i:n}$ for $r = 1, \ldots, n$, where $x_{i:n}$ denote the i-th ordered observation.

The MLEs of the parameters of the 3CAW, GAW, AMW, NMW, and AddW, for fitting to Aarset data, are given in Table 4, and Table 5 provides a numeric comparison between 3CAW, GAW, AMW, NMW, and AddW in terms of $\log(L)$ value, Kolmogorov–Smirnov (K-S) statistic, AIC, BIC, and AICc.

Table 4. The MLEs of parameters for Aarset data.

Model	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\eta}$
3CAW	0.49874	4.06058	97.88702	0.00748	0.01088	0.01175
GAW	5.53451×10^{-15}	7.53436	4.89809	0.05455	324.645	-
AMW	0.0763	90.1357	0.0104	0.4579	1.0604	-
NMW	0.0710	7.0150×10^{-8}	0.0160	0.5950	0.1970	-
AddW	0.7025	82.3350	-	0.0162	0.0118	-

Table 5. Log-likelihood, K-S statistic, AIC, BIC, and AICc for Aarset data.

Model	$\log(L)$	K-S (*p*-Value)	AIC	BIC	AICc
3CAW	-202.53	0.057 (0.997)	417.06	428.53	419.02
GAW	-215.93	0.121 (0.457)	441.86	451.42	443.22
AMW	-203.57	0.068 (0.976)	417.14	426.70	418.50
NMW	-212.90	0.088 (0.838)	435.80	445.36	437.16
AddW	-206.10	0.111 (0.568)	420.19	427.84	421.08

From this table, we can see that the 3CAW has largest $\log(L)$ value and smallest K-S and AIC values. The AMW model performs similarly to the 3CAW and even has smaller BIC and AICc value. This is because it has one less parameter than the 3CAW. However, its model form has a more complicated structure than the 3CAW.

We also provide a visual comparison by plotting the estimated reliability and failure rate functions of the 3CAW, GAW, AMW, NMW, and AddW along with the empirical estimates of reliability and failure rate functions in Figure 3. Once again this figure supports our conclusion that the 3CAW distribution provides the best fit for this dataset because the estimated reliability and failure rate functions of the 3CAW are closer to the corresponding empirical curves. In Figure 3b, the step function is a non-parametric estimation of the failure rate function performed by dividing the time domain into bins of equal width, and then estimating the failure rate in each bin as the number of events in that bin divided by the number of items at risk. This estimate is completed by using the "pehaz" function in the "muhaz" R package [36].
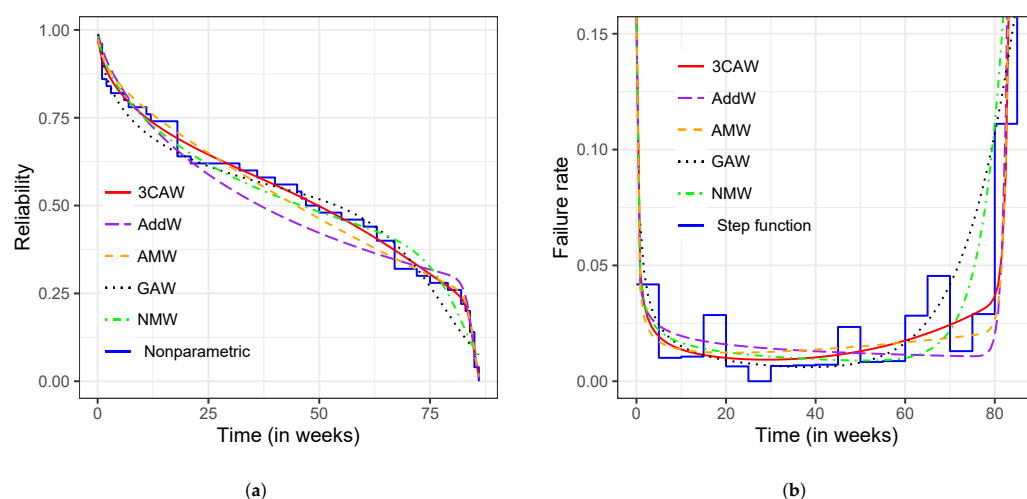

(a)

(b)

Figure 3. The estimated (**a**) reliability and (**b**) failure rate functions of 3CAW, GAW, AMW, NMW, and AddW for fitting to Aarset data.

Table 6 provides the MLEs of the parameters and their corresponding CIs for fitting 3CAW to Aarset data. The estimates of the shape parameters show that the first Weibull component has a decreasing failure rate which models the decreasing part of the bathtub curve; the second component has an increasing failure rate with slight change which models the middle part of the bathtub curve; the third component has an increasing failure rate with sharped change which models the sharped wear out of the bathtub curve. The 90% and 95% confidence intervals also support these claims. Table 7 provides the MLEs of the reliability and failure functions and their corresponding 95% confidence intervals at each given failure time. Figure 4 displays the results in Table 7 visually.

Table 6. MLEs and confidence intervals for the parameters of the 3CAW for fitting to Aarset data.

Parameter	MLE	90% CI	95% CI
α (shape)	0.49874	[0.34723, 0.71637]	[0.32396, 0.76783]
β (shape)	4.06058	[1.82322, 9.04351]	[1.56394, 10.5428]
γ (shape)	97.8870	[59.1332, 162.039]	[53.6904, 178.465]
θ (scale)	0.00748	[0.00271, 0.02064]	[0.00223, 0.02507]
λ (scale)	0.01088	[0.00911, 0.01300]	[0.00880, 0.01345]
η (scale)	0.01175	[0.01168, 0.01182]	[0.01166, 0.01184]

Table 7. MLEs and 95% confidence intervals for the estimated reliability and failure rate functions of the 3CAW for fitting to Aarset data.

| x	$\hat{R}(x)$	95% CI for $\hat{R}(x)$	$\hat{h}(x)$	95% CI for $\hat{h}(x)$
0.1	0.97278	[0.93988, 1.00683]	0.13765	[0.05822, 0.32544]
0.2	0.96175	[0.92103, 1.00428]	0.09725	[0.04646, 0.20356]
1	0.91666	[0.85448, 0.98335]	0.04340	[0.02593, 0.07264]
2	0.88430	[0.81239, 0.96257]	0.03066	[0.01922, 0.04893]
3	0.86026	[0.78273, 0.94547]	0.02502	[0.01580, 0.03963]
6	0.80840	[0.72113, 0.90624]	0.01769	[0.01092, 0.02866]
7	0.79477	[0.70518, 0.89574]	0.01638	[0.01001, 0.02680]
11	0.74981	[0.65271, 0.86136]	0.01311	[0.00778, 0.02211]
12	0.74025	[0.64155, 0.85414]	0.01258	[0.00743, 0.02130]
18	0.69129	[0.58457, 0.81749]	0.01049	[0.00618, 0.01782]
21	0.67049	[0.56071, 0.80176]	0.00992	[0.00589, 0.01669]
32	0.60416	[0.48815, 0.74775]	0.00939	[0.00549, 0.01606]
36	0.58157	[0.46482, 0.72765]	0.00971	[0.00549, 0.01716]
40	0.55882	[0.44180, 0.70683]	0.01029	[0.00566, 0.01872]
45	0.52941	[0.41247, 0.67949]	0.01140	[0.00621, 0.02096]
46	0.52333	[0.40646, 0.67381]	0.01168	[0.00637, 0.02141]
47	0.51718	[0.40039, 0.66805]	0.01197	[0.00655, 0.02187]
50	0.49822	[0.38178, 0.65017]	0.01296	[0.00722, 0.02327]
55	0.46467	[0.34944, 0.61791]	0.01500	[0.00871, 0.02583]
60	0.42845	[0.31564, 0.58159]	0.01755	[0.01047, 0.02942]
63	0.40541	[0.29477, 0.55757]	0.01935	[0.01146, 0.03264]
67	0.37323	[0.26626, 0.52319]	0.02206	[0.01250, 0.03895]
72	0.33105	[0.22912, 0.47834]	0.02601	[0.01319, 0.05131]
75	0.30499	[0.20566, 0.45230]	0.02870	[0.01333, 0.06177]
79	0.26963	[0.17264, 0.42111]	0.03350	[0.01437, 0.07810]
82	0.23714	[0.14461, 0.38889]	0.06727	[0.02709, 0.16702]
83	0.21538	[0.12866, 0.36056]	0.13852	[0.05370, 0.35734]
84	0.17123	[0.09605, 0.30525]	0.36197	[0.17828, 0.73493]
85	0.08976	[0.04106, 0.19623]	1.05836	[0.56081, 1.99736]
86	0.01296	[0.00176, 0.09529]	3.20497	[1.11010, 9.25304]

| (a) | (b) |
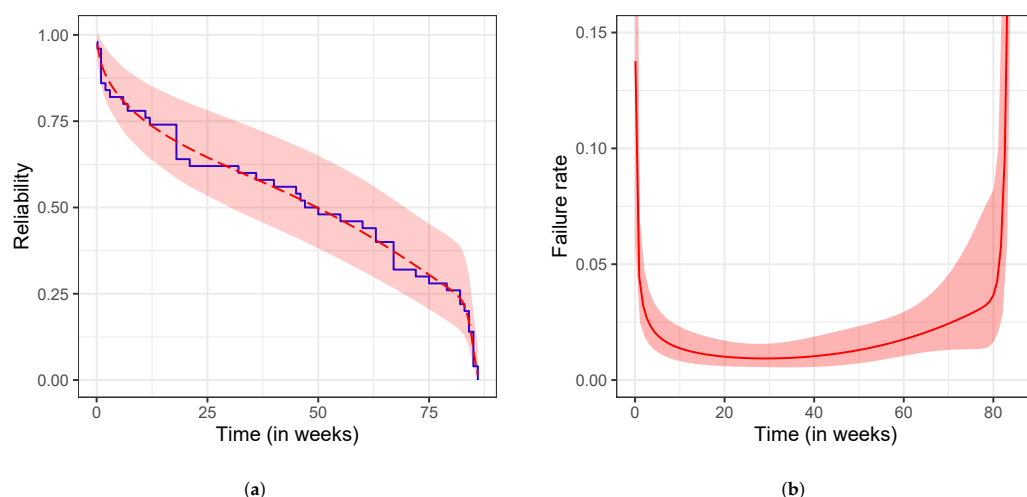
Figure 4. The 95% confidence interval for the estimated (**a**) reliability and (**b**) failure rate functions (red lines) of 3CAW for fitting to Aarset data. The blue line is the nonparametric estimate of the reliability function.

6.2. Meeker and Escobar Data

Data in Table 8 represent failure and running times of 30 devices (in thousands of cycles) ([33], p. 383). Two failure modes were observed for this dataset. These data are shown also to have a bathtub-shaped failure rate as indicated by the scaled TTT-transform plot given in Figure 5.

Table 8. Failure and running times of a sample of $n = 30$ devices (in thousands of cycles) [33].

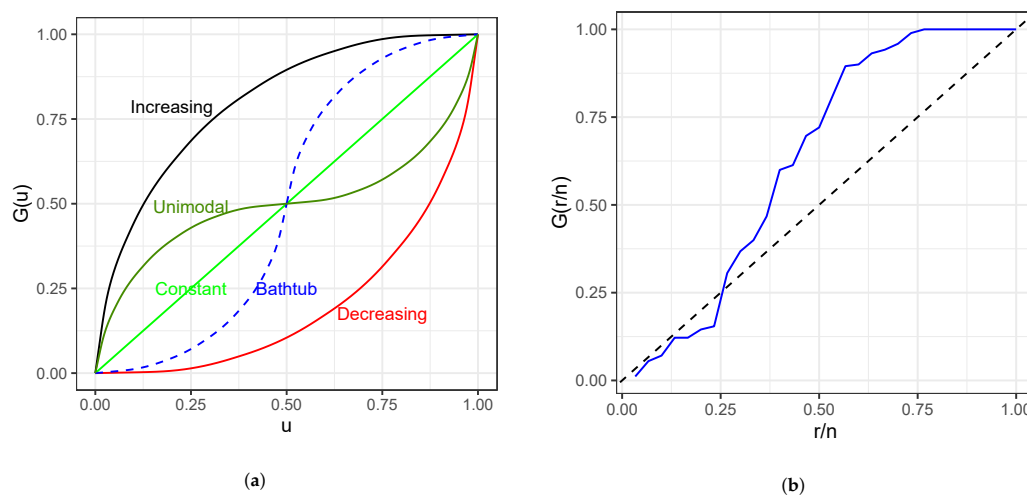| 2	10	13	23	23	28	30	65	80	88
106	143	147	173	181	212	245	247	261	266
275	293	300	300	300	300	300	300	300	300

| (a) | (b) |

Figure 5. (**a**) Types of failure rate identify by shapes of the TTT-plot and (**b**) the empirical TTT-plot (blue) for Meeker and Escobar data, where u and $G(u)$ are explained in Figure 2.

We provide the MLEs of the parameters of the given distributions in Table 9, and the numeric comparison between the given distributions, for fitting to Meeker and Escobar data, in Table 10. Figure 6 displays the visual comparison of the mention distributions by plotting the estimated reliability and failure rate functions. From these results, we see that the 3CAW provides a good fit to Meeker and Escobar data. Again, we can see that

AMW performs similar to 3CAW. Even though it has a smaller $\log(L)$ value and a larger K-S value, its other values are still slightly less than that of 3CAW.

Table 9. The MLEs of parameters for Meeker and Escobar data.

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\eta}$
3CAW	0.55141	4.07734	133.13084	0.00178	0.00287	0.00333
GAW	4.72286×10^{-8}	3.06873	3.48132	0.05456	120.502	-
AMW	0.0142	116.9665	0.0019	0.6788	0.3902	-
NMW	0.024	5.991×10^{-8}	0.012	0.629	0.056	-
AddW	0.84960	41.25756	-	0.00368	0.00335	-

Table 10. Log-likelihood, K-S statistic, AIC, BIC, and AICc for Meeker and Escobar data.

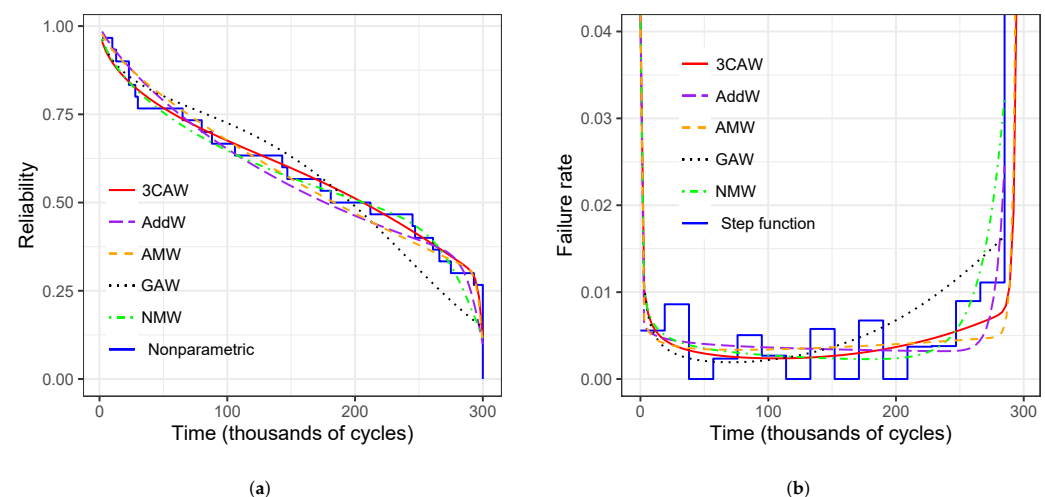| Model	$\log(L)$	K-S (p-Value)	AIC	BIC	AICc
3CAW	-154.67	0.147 (0.533)	321.33	329.74	324.98
GAW	-175.93	0.146 (0.545)	361.86	368.86	364.36
AMW	-155.58	0.167 (0.374)	321.16	328.17	323.66
NMW	-166.18	0.149 (0.522)	342.36	349.37	344.86
AddW	-162.58	0.168 (0.365)	333.17	340.82	334.05

(a)

(b)

Figure 6. The estimated (**a**) reliability and (**b**) failure rate functions of 3CAW, GAW, AMW, NMW, and AddW for fitting to Meeker and Escobar data.

Table 11 provides the MLEs of the parameters and their corresponding CIs for fitting 3CAW to Meeker and Escobar data. We also calculate the MLEs of the reliability and failure functions and their corresponding 95% confidence intervals at each given failure time in Table 12. The results in Table 12 are displayed visually in Figure 7.

Table 11. MLEs and confidence intervals for the parameters of the 3CAW for fitting to Meeker and Escobar data.

Parameter	MLE	90% CI	95% CI
α (shape)	0.55141	[0.31596, 0.96231]	[0.28399, 1.07064]
β (shape)	4.07734	[0.81092, 20.5010]	[0.59513, 27.9348]
γ (shape)	133.130	[71.4311, 248.122]	[63.4005, 279.553]
θ (scale)	0.00178	[0.00036, 0.00888]	[0.00026, 0.01208]
λ (scale)	0.00287	[0.00221, 0.00372]	[0.00211, 0.00391]
η (scale)	0.00333	[0.00331, 0.00335]	[0.00331, 0.00335]

Table 12. MLEs and 95% confidence intervals for the estimated reliability and failure rate functions of the 3CAW for fitting to Meeker and Escobar data.

x	$\hat{R}(x)$	95% CI for $\hat{R}(x)$	$\hat{h}(x)$	95% CI for $\hat{h}(x)$
2	0.95633	[0.89780, 1.01868]	0.01231	[0.00507, 0.02989]
10	0.89722	[0.81043, 0.99329]	0.00598	[0.00306, 0.01167]
13	0.88219	[0.79072, 0.98425]	0.00532	[0.00271, 0.01045]
23	0.84223	[0.73971, 0.95897]	0.00412	[0.00199, 0.00854]
28	0.82581	[0.71885, 0.94870]	0.00377	[0.00177, 0.00803]
30	0.81970	[0.71105, 0.94495]	0.00366	[0.00170, 0.00786]
65	0.73675	[0.60423, 0.89832]	0.00265	[0.00122, 0.00573]
80	0.70903	[0.56979, 0.88229]	0.00248	[0.00122, 0.00506]
88	0.69527	[0.55344, 0.87345]	0.00242	[0.00122, 0.00481]
106	0.66598	[0.52075, 0.85171]	0.00237	[0.00120, 0.00470]
143	0.60852	[0.46353, 0.79885]	0.00257	[0.00102, 0.00645]
147	0.60224	[0.45748, 0.79281]	0.00261	[0.00102, 0.00672]
173	0.55996	[0.41565, 0.75437]	0.00302	[0.00110, 0.00829]
181	0.54623	[0.40157, 0.74299]	0.00319	[0.00118, 0.00860]
212	0.48863	[0.34233, 0.69744]	0.00406	[0.00181, 0.00910]
245	0.41867	[0.27729, 0.63213]	0.00538	[0.00259, 0.01120]
247	0.41414	[0.27341, 0.62733]	0.00548	[0.00260, 0.01153]
261	0.38169	[0.24618, 0.59180]	0.00619	[0.00257, 0.01494]
266	0.36980	[0.23628, 0.57877]	0.00647	[0.00251, 0.01672]
275	0.34805	[0.21784, 0.55609]	0.00700	[0.00235, 0.02085]
293	0.29242	[0.16971, 0.50387]	0.02534	[0.00796, 0.08069]
300	0.11930	[0.04987, 0.28540]	0.39711	[0.12714, 1.24030]

(a)　　　　　(b)
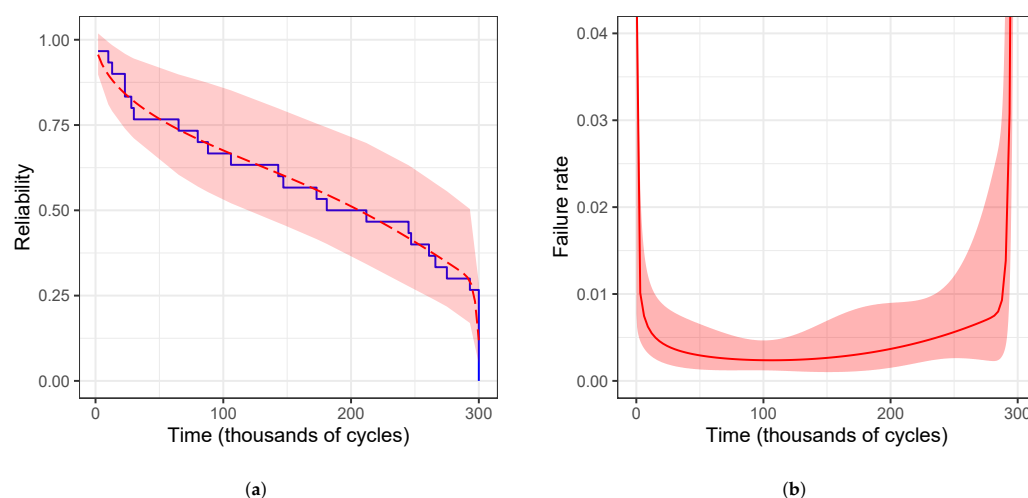
Figure 7. The 95% confidence interval for the estimated (**a**) reliability and (**b**) failure rate functions (red lines) of the 3CAW for fitting to Meeker and Escobar data. The blue line is the nonparametric estimate of the reliability function.

7. Conclusions

In this study, we have introduced the 3CAW distribution by combining three Weibull distributions. The properties of the model have been studied. The MLE method has been exploited for estimating the parameters and the reliability characteristics of the 3CAW, and the uncertainties of the estimates have been obtained by using asymptotic confidence intervals. The cross-entropy method has been deployed to optimize the log-likelihood function. Two sets of failure data have been used for testing the superiority of the 3CAW. It is interesting to see a Bayesian analysis of this six-parameter distribution by using Hamiltonian Monte Carlo simulation method, and to see more applications of the 3CAW to real datasets arising from practice.

Conflicts of Interest: The author declares no conflicts of interest.

Appendix A. Proof of Theorem 1

Let $\boldsymbol{\theta}_1 = (\alpha_1, \beta_1, \gamma_1, \theta_1, \lambda_1, \eta_1)$ where $0 < \alpha_1 < \beta_1 < \gamma_1$ and $\theta_1, \lambda_1, \eta_1 > 0$, and $\boldsymbol{\theta}_2 = (\alpha_2, \beta_2, \gamma_2, \theta_2, \lambda_2, \eta_2)$ where $0 < \alpha_2 < \beta_2 < \gamma_2$ and $\theta_2, \lambda_2, \eta_2 > 0$. To prove the model is identifiable, we show that: If $F_{\boldsymbol{\theta}_1}(x) = F_{\boldsymbol{\theta}_2}(x)$ for all $x \in \mathbb{R}^+$, then $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$. It means the mapping $\boldsymbol{\theta} \mapsto F_{\boldsymbol{\theta}}$ is one-to-one.

Indeed, for every $x \in \mathbb{R}^+$, we have $F_{\boldsymbol{\theta}_1}(x) = F_{\boldsymbol{\theta}_2}(x)$, so:

$$e^{-(\theta_1 x)^{\alpha_1} - (\lambda_1 x)^{\beta_1} - (\eta_1 x)^{\gamma_1}} = e^{-(\theta_2 x)^{\alpha_2} - (\lambda_2 x)^{\beta_2} - (\eta_2 x)^{\gamma_2}}. \tag{A1}$$

Taking logarithm both side of (A1), we obtain:

$$(\theta_1 x)^{\alpha_1} + (\lambda_1 x)^{\beta_1} + (\eta_1 x)^{\gamma_1} = (\theta_2 x)^{\alpha_2} + (\lambda_2 x)^{\beta_2} + (\eta_2 x)^{\gamma_2}. \tag{A2}$$

If $\alpha_1 \neq \alpha_2$, i.e., $\alpha_1 < \alpha_2$ or $\alpha_1 > \alpha_2$, without lost of generality we assume that $\alpha_1 < \alpha_2$, then dividing both side of (A2) by x^{α_1} gives:

$$\theta_1^{\alpha_1} + \lambda_1^{\beta_1} x^{\beta_1 - \alpha_1} + \eta_1^{\gamma_1} x^{\gamma_1 - \alpha_1} = \theta_2^{\alpha_2} x^{\alpha_2 - \alpha_1} + \lambda_2^{\beta_2} x^{\beta_2 - \alpha_1} + \eta_2^{\gamma_2} x^{\gamma_2 - \alpha_1}.$$

Letting $x \to 0$, we have $\theta_1^{\alpha_1} = 0$, which is a contradiction since $\theta_1, \alpha_1 > 0$. Therefore, $\alpha_1 = \alpha_2$. Now we have:

$$\theta_1^{\alpha_1} + \lambda_1^{\beta_1} x^{\beta_1 - \alpha_1} + \eta_1^{\gamma_1} x^{\gamma_1 - \alpha_1} = \theta_2^{\alpha_1} + \lambda_2^{\beta_2} x^{\beta_2 - \alpha_1} + \eta_2^{\gamma_2} x^{\gamma_2 - \alpha_1}.$$

Letting $x \to 0$, we obtain $\theta_1^{\alpha_1} = \theta_2^{\alpha_1}$, and hence $\theta_1 = \theta_2$. Then (A2) reduces to:

$$(\lambda_1 x)^{\beta_1} + (\eta_1 x)^{\gamma_1} = (\lambda_2 x)^{\beta_2} + (\eta_2 x)^{\gamma_2}.$$

Using the same reasoning as above, we finally obtain $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$.

Appendix B. The log-Likelihood Function and Its First-Order Partial Derivatives

We rewrite the log-likelihood in a compact form:

$$l = \sum_{i=1}^{n} (\log h(x_i; \boldsymbol{\theta}) - H(x_i; \boldsymbol{\theta})),$$

where $h(x_i; \boldsymbol{\theta})$ is the failure rate function and $H(x_i; \boldsymbol{\theta})$ is the cumulative failure rate function. Then its first-order partial derivatives can be derived as follows:

$$\frac{\partial l}{\partial \theta_k} = \sum_{i=1}^{n} \left(\frac{h_{\theta_k}(x_i; \boldsymbol{\theta})}{h(x_i; \boldsymbol{\theta})} - H_{\theta_k}(x_i; \boldsymbol{\theta}) \right), \quad k = \overline{1, 6},$$

where:

$$h(x_i; \boldsymbol{\theta}) = \alpha\theta(\theta x_i)^{\alpha-1} + \beta\lambda(\lambda x_i)^{\beta-1} + \gamma\eta(\eta x_i)^{\gamma-1},$$
$$h_\alpha(x_i; \boldsymbol{\theta}) = \theta(\theta x_i)^{\alpha-1}(1 + \alpha\log(\theta x_i)),$$
$$h_\beta(x_i; \boldsymbol{\theta}) = \lambda(\lambda x_i)^{\beta-1}(1 + \beta\log(\lambda x_i)),$$
$$h_\gamma(x_i; \boldsymbol{\theta}) = \eta(\eta x_i)^{\gamma-1}(1 + \gamma\log(\eta x_i)),$$
$$h_\theta(x_i; \boldsymbol{\theta}) = \alpha^2(\theta x_i)^{\alpha-1},$$
$$h_\lambda(x_i; \boldsymbol{\theta}) = \beta^2(\lambda x_i)^{\beta-1},$$
$$h_\eta(x_i; \boldsymbol{\theta}) = \gamma^2(\eta x_i)^{\gamma-1},$$
$$H(x_i; \boldsymbol{\theta}) = (\theta x_i)^\alpha + (\lambda x_i)^\beta + (\eta x_i)^\gamma,$$
$$H_\alpha(x_i; \boldsymbol{\theta}) = (\theta x_i)^\alpha \log(\theta x_i),$$
$$H_\beta(x_i; \boldsymbol{\theta}) = (\lambda x_i)^\beta \log(\lambda x_i),$$
$$H_\gamma(x_i; \boldsymbol{\theta}) = (\eta x_i)^\gamma \log(\eta x_i),$$
$$H_\theta(x_i; \boldsymbol{\theta}) = \alpha x_i(\theta x_i)^{\alpha-1},$$
$$H_\lambda(x_i; \boldsymbol{\theta}) = \beta x_i(\lambda x_i)^{\beta-1},$$
$$H_\eta(x_i; \boldsymbol{\theta}) = \gamma x_i(\eta x_i)^{\gamma-1}.$$

Appendix C. The log-Likelihood Function and Its Second-Order Partial Derivatives

The second-order partial derivatives of the log-likelihood function are given as:

$$\frac{\partial^2 l}{\partial \theta_k \partial \theta_l} = \sum_{i=1}^{n} \left(\frac{h_{\theta_k \theta_l}(x_i; \boldsymbol{\theta})h(x_i; \boldsymbol{\theta}) - h_{\theta_k}(x_i; \boldsymbol{\theta})h_{\theta_l}(x_i; \boldsymbol{\theta})}{h^2(x_i; \boldsymbol{\theta})} - H_{\theta_k \theta_l}(x_i; \boldsymbol{\theta}) \right), \quad k = \overline{1,6}, l = \overline{k,6},$$

where

$$h_{\alpha\alpha}(x_i; \boldsymbol{\theta}) = \theta(\theta x_i)^{\alpha-1} \log(\theta x_i)(2 + \alpha\log(\theta x_i)),$$
$$h_{\alpha\theta}(x_i; \boldsymbol{\theta}) = \alpha(\theta x_i)^{\alpha-1}(2 + \alpha\log(\theta x_i)),$$
$$h_{\theta\theta}(x_i; \boldsymbol{\theta}) = \alpha^2(\alpha - 1)x_i(\theta x_i)^{\alpha-2},$$
$$h_{\beta\beta}(x_i; \boldsymbol{\theta}) = \lambda(\lambda x_i)^{\beta-1} \log(\lambda x_i)(2 + \beta\log(\lambda x_i)),$$
$$h_{\beta\lambda}(x_i; \boldsymbol{\theta}) = \beta(\lambda x_i)^{\beta-1}(2 + \beta\log(\lambda x_i)),$$
$$h_{\lambda\lambda}(x_i; \boldsymbol{\theta}) = \beta^2(\beta - 1)x_i(\lambda x_i)^{\beta-2},$$
$$h_{\gamma\gamma}(x_i; \boldsymbol{\theta}) = \eta(\eta x_i)^{\gamma-1} \log(\eta x_i)(2 + \gamma\log(\eta x_i)),$$
$$h_{\gamma\eta}(x_i; \boldsymbol{\theta}) = \gamma(\eta x_i)^{\gamma-1}(2 + \gamma\log(\eta x_i)),$$
$$h_{\eta\eta}(x_i; \boldsymbol{\theta}) = \gamma^2(\gamma - 1)x_i(\eta x_i)^{\gamma-2},$$
$$h_{\alpha\beta}(x_i; \boldsymbol{\theta}) = h_{\alpha\gamma}(x_i; \boldsymbol{\theta}) = h_{\alpha\lambda}(x_i; \boldsymbol{\theta}) = 0,$$
$$h_{\alpha\eta}(x_i; \boldsymbol{\theta}) = h_{\beta\gamma}(x_i; \boldsymbol{\theta}) = h_{\beta\theta}(x_i; \boldsymbol{\theta}) = 0,$$
$$h_{\beta\eta}(x_i; \boldsymbol{\theta}) = h_{\gamma\theta}(x_i; \boldsymbol{\theta}) = h_{\gamma\lambda}(x_i; \boldsymbol{\theta}) = 0,$$
$$h_{\theta\lambda}(x_i; \boldsymbol{\theta}) = h_{\theta\eta}(x_i; \boldsymbol{\theta}) = h_{\lambda\eta}(x_i; \boldsymbol{\theta}) = 0,$$
$$H_{\alpha\alpha}(x_i; \boldsymbol{\theta}) = (\theta x_i)^\alpha \log^2(\theta x_i),$$
$$H_{\alpha\theta}(x_i; \boldsymbol{\theta}) = x_i(\theta x_i)^{\alpha-1}(1 + \alpha\log(\theta x_i)),$$
$$H_{\theta\theta}(x_i; \boldsymbol{\theta}) = \alpha(\alpha - 1)x_i^2(\theta x_i)^{\alpha-2},$$
$$H_{\beta\beta}(x_i; \boldsymbol{\theta}) = (\lambda x_i)^\beta \log^2(\lambda x_i),$$
$$H_{\beta\lambda}(x_i; \boldsymbol{\theta}) = x_i(\lambda x_i)^{\beta-1}(1 + \beta\log(\lambda x_i)),$$

$$H_{\lambda\lambda}(x_i; \boldsymbol{\theta}) = \beta(\beta - 1)x_i^2(\lambda x_i)^{\beta-2},$$
$$H_{\gamma\gamma}(x_i; \boldsymbol{\theta}) = (\eta x_i)^\gamma \log^2(\eta x_i),$$
$$H_{\gamma\eta}(x_i; \boldsymbol{\theta}) = x_i(\eta x_i)^{\gamma-1}(1 + \gamma \log(\eta x_i)),$$
$$H_{\eta\eta}(x_i; \boldsymbol{\theta}) = \gamma(\gamma - 1)x_i^2(\eta x_i)^{\gamma-2},$$
$$H_{\alpha\beta}(x_i; \boldsymbol{\theta}) = H_{\alpha\gamma}(x_i; \boldsymbol{\theta}) = H_{\alpha\lambda}(x_i; \boldsymbol{\theta}) = 0,$$
$$H_{\alpha\eta}(x_i; \boldsymbol{\theta}) = H_{\beta\gamma}(x_i; \boldsymbol{\theta}) = H_{\beta\theta}(x_i; \boldsymbol{\theta}) = 0,$$
$$H_{\beta\eta}(x_i; \boldsymbol{\theta}) = H_{\gamma\theta}(x_i; \boldsymbol{\theta}) = H_{\gamma\lambda}(x_i; \boldsymbol{\theta}) = 0,$$
$$H_{\theta\lambda}(x_i; \boldsymbol{\theta}) = H_{\theta\eta}(x_i; \boldsymbol{\theta}) = H_{\lambda\eta}(x_i; \boldsymbol{\theta}) = 0.$$

Appendix D. First-Order Partial Derivatives of the Reliability Function

Given the reliability function,

$$R(x; \boldsymbol{\theta}) = \exp\left[-(\theta x)^\alpha - (\lambda x)^\beta - (\eta x)^\gamma\right],$$

its first-order partial derivatives with respect to the parameters $\alpha, \beta, \gamma, \theta, \lambda$, and η are, respectively, given by:

$$R_\alpha(x; \boldsymbol{\theta}) = -(\theta x)^\alpha \log(\theta x)R(x; \boldsymbol{\theta}),$$
$$R_\beta(x; \boldsymbol{\theta}) = -(\lambda x)^\beta \log(\lambda x)R(x; \boldsymbol{\theta}),$$
$$R_\gamma(x; \boldsymbol{\theta}) = -(\eta x)^\gamma \log(\eta x_i)R(x; \boldsymbol{\theta}),$$
$$R_\theta(x; \boldsymbol{\theta}) = -\alpha x(\theta x)^{\alpha-1}R(x; \boldsymbol{\theta}),$$
$$R_\lambda(x; \boldsymbol{\theta}) = -\beta x(\lambda x)^{\beta-1}R(x; \boldsymbol{\theta}),$$
$$R_\eta(x; \boldsymbol{\theta}) = -\gamma x(\eta x)^{\gamma-1}R(x; \boldsymbol{\theta}).$$

Appendix E. Cross-Entropy (CE) Algorithm for Continuous Optimization

In our previous study [37], we have described the CE algorithm in a generic case on how it works. Here, we focus more on the specific case of this study. The objective is to find $\hat{\boldsymbol{\theta}}_{\text{ML}}$ that maximizes $l(\boldsymbol{\theta})$,

$$\hat{\boldsymbol{\theta}}_{\text{ML}} = \arg \max_{\boldsymbol{\theta} \in \Theta} l(\boldsymbol{\theta}), \tag{A3}$$

where $l(\boldsymbol{\theta})$ is the log-likelihood function, and $\Theta = \{(\alpha, \beta, \gamma, \theta, \lambda, \eta)^T : \theta, \lambda, \eta > 0 \text{ and } 0 < \alpha < \beta < \gamma\}$. For this continuous optimization, let $N_6(\boldsymbol{\mu}, \boldsymbol{\sigma}^2)$ denotes the 6-dimensional normal distribution with independent components, mean vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_6)$ and variance vector $\boldsymbol{\sigma}^2 = (\sigma_1^2, \ldots, \sigma_6^2)$. Then, the CE algorithm is proceeded as follows (Algorithm A1):

Algorithm A1 CE algorithm

1: Choose $\hat{\mu}_0$ and $\hat{\sigma}_0$. Set $t := 0$.
2: **while** $\max_j \{\tilde{\sigma}_{tj}\} \geq \varepsilon$ **do**
3:      $t = t + 1$.
4:      Generate $\theta_1, \ldots, \theta_N$ from the $N_6(\hat{\mu}_{t-1}, \hat{\sigma}_{t-1}^2)$ distribution.
5:      Calculate the performances $l(\theta_k)$ for all k and order them from smallest to largest, $l_{(1)} \leq \ldots \leq l_{(N)}$.
6:      Find the sample $(1 - \varrho)$-quantile $\hat{\gamma}_t = l_{(\lceil(1-\varrho)N\rceil)}$, where $10^{-2} \leq \varrho \leq 10^{-1}$.
7:      Let $\mathcal{E}_t = \{\theta_k : l(\theta_k) \geq \hat{\gamma}_t\}$ be best performing (=elite) samples.
8:      **for** $j = 1, \ldots, 6$ **do**
9:          $\tilde{\mu}_{tj} := \frac{\sum_{\theta_k \in \mathcal{E}_t} \theta_{ij}}{|\mathcal{E}_t|}$,
10:         $\tilde{\sigma}_{tj} := \sqrt{\frac{\sum_{\theta_k \in \mathcal{E}_t}(\theta_{ij}-\tilde{\mu}_{tj})^2}{|\mathcal{E}_t|}}$.
11:      **end for**
12:      $\hat{\mu}_t := \alpha\tilde{\mu}_t + (1-\alpha)\hat{\mu}_{t-1}$,
13:      $\hat{\sigma}_t := \beta_t\tilde{\sigma}_t + (1-\beta_t)\hat{\sigma}_{t-1}$, where $0.5 \leq \alpha \leq 0.9$, $\beta_t = \beta - \beta(1-\frac{1}{t})^q$, q is an integer (typically between 5 and 10) and β is a smoothing constant (typically between 0.8 and 0.99).
14: **end while**

While applying the CE algorithm, the mean vector $\hat{\mu}_t$ should converge to $\hat{\theta}_{\text{ML}}$ and the standard deviation vector $\hat{\sigma}_t$ should converge to zero vector. We refer to [38] for more details about the method. To solve (A3) with constraints on the parameters, we can apply the acceptance-rejection (AR) approach. The AR method works as follows: Generate a random vector θ from a normal distribution, then accept or reject it depending on whether the sample falls or not in the constrained parameter space [39].

Appendix F. R Code for the Cross-Entropy Method

We provide R code for the cross-entropy method by using the "CEoptim" package [40]. Here, we show the R code for the Aarset data analyzed in Section 6.1. Since the log-likelihood function is highly multi-model and CE method is a stochastic search algorithm, we might need to run the CEoptim function several times in order to obtain the best result.

- Install and invoke the CEoptim package

```
install.packages("CEoptim")
library(CEoptim)
```

- Input the Aarset data into R

```
Aarsetdata <- c(
136, 246, 255, 376, 421, 565, 616, 617, 652, 655,
658, 660, 662, 675, 681, 734, 736, 737, 757, 769,
777, 800, 807, 825, 855, 857, 864, 868, 870, 870,
873, 882, 895, 910, 934, 943, 1015, 1019
)
```

- Define the log-likelihood function

```
LogLikelihood <- function(x,t) {
sum(log(x[1]*x[4]*(t*x[4])^(x[1]-1) + x[2]*x[5]*(t*x[5])^(x[2]-1) +
x[3]*x[6]*(t*x[6])^(x[3]-1))-((t*x[4])^x[1] + (t*x[5])^x[2]
+ (t*x[6])^x[3]))
}
```

- Choose initial values for mean and standard deviation

```
mu0 <- c(0,5,90,0,0,0)
sigma0 <- c(1,10,10,1,1,1)
```

- Restrict the parameter space

```
A=matrix(c(-1,1,0,0,0,0, 0,-1,1,0,0,0, 0,0,-1,0,0,0,
0,0,0,-1,0,0, 0,0,0,0,-1,0, 0,0,0,0,0,-1), nrow = 6)
B=c(0,0,0,0,0,0)
```

- Execute the CEoptim function

```
opt <- CEoptim(
LogLikelihood, f.arg = list(t=Aarsetdata), maximize = T,
continuous = list(mean=mu0,sd=sigma0, conMat=A, conVec=B,
smoothMean=0.7, smoothSd=0.7, sdThr=0.00001), rho = 0.01, N=1000
)
```

- The optimum, optimizer (MLE), and AIC values are respectively given by

```
opt$optimum
opt$optimizer$continuous
AIC <- -2*opt$optimum+2*6
```

References

1. Xie, M.; Lai, C.D. Reliability analysis using an additive Weibull model with bathtub-shaped failure rate function. *Reliab. Eng. Syst.* **1995**, *52*, 87–93. [CrossRef]
2. Weibull, W.A. Statistical distribution function of wide applicability. *J. Appl. Mech.* **1951**, *18*, 293–296. [CrossRef]
3. Lemonte, A.J.; Cordeiroa, G.M.; Ortegab, E.M.M. On the additive Weibull distribution. *Commun. Stat. Theory Methods* **2014**, *43*, 2066–2080. [CrossRef]
4. Almalki, S.J.; Yuan, J. A new modified Weibull distribution. *Reliab. Eng. Syst.* **2013**, *111*, 164–170. [CrossRef]
5. Lai, C.D.; Xie, M.; Murthy, D.N.P. A modified Weibull distribution. *IEEE Trans. Reliab.* **2003**, *52*, 33–37. [CrossRef]
6. Thach, T.T.; Bris, R. Improved new modified weibull distribution: A bayes study using hamiltonian monte carlo simulation. *Proc. Inst. Mech. Eng. O J. Risk Reliab.* **2020**, *234*, 496–511. [CrossRef]
7. He, B.; Cui, W.; Du, X. An additive modified Weibull distribution. *Reliab. Eng. Syst.* **2016**, *145*, 28–37. [CrossRef]
8. Gompertz, B. On the nature of the function expressive of the law of human mortality and on the new model of determining the value of life contingencies. *Phil. Trans. R. Soc. Lon.* **1825**, *115*, 513–585.
9. Park, S.; Park, J. A general class of flexible Weibull distributions. *Commun. Stat. Theory Methods* **2018**, *47*, 767–778. [CrossRef]
10. Shakhatreh, M.K.; Lemonte, A.J.; Moreno-Arenas, G. The log-normal modified Weibull distribution and its reliability implications. *Reliab. Eng. Syst.* **2019**, *188*, 6–12. [CrossRef]
11. Thach, T.T.; Bris, R. An additive Chen-Weibull distribution and its applications in reliability modeling. *Qual. Reliab. Eng. Int.* **2021**, *37*, 352–373. [CrossRef]
12. Kumar, C.S.; Nair, S.R. On some aspects of a flexible class of additive Weibull distribution. *Commun. Stat. Theory Methods* **2018**, *47*, 1028–1049. [CrossRef]
13. Mudholkar, G.S.; Srivastava, D.K. Exponentiated Weibull family for analysing bathtub failure rate data. *IEEE Trans. Reliab.* **1993**, *42*, 299–302. [CrossRef]
14. Hjorth, U. A reliability distribution with increasing, decreasing, constant and bathtub-shaped failure rates. *Technometrics* **1980**, *17*, 99–107. [CrossRef]
15. Carrasco, M.; Ortega, E.M.; Cordeiro, G.M. A generalized modified Weibull distribution for lifetime modeling. *Comput. Stat. Data Anal.* **2008**, *53*, 450–462. [CrossRef]
16. Silva, G.O.; Ortega, E.M.; Cordeiro, G.M. The beta modified Weibull distribution. *Lifetime Data Anal.* **2010**, *16*, 409–430. [CrossRef]
17. Zhang, T.; Xie, M. On the upper truncated Weibull distribution and its reliability implications. *Reliab. Eng. Syst.* **2011**, *96*, 194–200. [CrossRef]
18. Lemonte, A.J. A new exponential-type distribution with constant, decreasing, increasing, upside-down bathtub and bathtub-shaped failure rate function. *Comput. Stat. Data Anal.* **2013**, *62*, 149–170. [CrossRef]
19. Singla, N.; Jain, K.; Sharma, S.K. The beta Generalized Weibull distribution: Properties and applications. *Reliab. Eng. Syst.* **2012**, *102*, 5–15. [CrossRef]
20. Nassar, M.; Afify, A.Z.; Dey, S.; Kumar, D. A new extension of Weibull distribution: Properties and different methods of estimation. *J. Comput. Appl. Math.* **2018**, *336*, 439–457. [CrossRef]
21. Ali, S.; Dey, S.; Tahir, M.H.; Mansoor, M. Two-Parameter Logistic-Exponential Distribution: Some New Properties and Estimation Methods. *Am. J. Math. Manag. Sci.* **2020**, *39*, 270–298. [CrossRef]

22. Shih, J.H.; Emura, T. Penalized Cox regression with a five-parameter spline model. *Commun. Stat. Theory Methods* **2021**, *50*, 3749–3768. [CrossRef]
23. Nassar, M.; Afify, A.Z.; Shakhatreh, M.K.; Dey, S. On a new extension of Weibull distribution: Properties, estimation, and applications to one and two causes of failures. *Qual. Reliab. Eng. Int.* **2020**, *36*, 2019–2043. [CrossRef]
24. Shakhatreh, M.K.; Lemonte, A.J.; Cordeiro, G.M. On the generalized extended exponential-Weibull distribution: Properties and different methods of estimation. *Int. J. Comput. Math.* **2020**, *97*, 1029–1057. [CrossRef]
25. Mallick, A.; Ghosh, I.; Dey, S.; Kumar, D. Bounded Weighted Exponential Distribution with Applications. *Am. J. Math. Manag. Sci.* **2021**, *40*, 68–87. [CrossRef]
26. Khalil, A.; Ijaz, M.; Ali, K.; Mashwani, W.K.; Shafiq, M.; Kumam, P.; Kumam, W. A novel flexible additive Weibull distribution with real-life applications. *Commun. Stat. Theory Methods* **2021**, *50*, 1557–1572. [CrossRef]
27. Salem, S.A. Bayesian estimation of a non-linear failure rate from censored samples type II. *Microelectron. Reliab.* **1992**, *32*, 1385–1388. [CrossRef]
28. Bain, L.J. Analysis for the linear failure-rate life-testing distribution. *Technometrics* **1974**, *16*, 551–559. [CrossRef]
29. Rubinstein, R.Y.; Kroese, D.P. *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte Carlo Simulation and Machine Learning*; Springer: New York, NY, USA, 2004.
30. Geyer, C.J. Fisher Information and Confidence Intervals Using Maximum Likelihood. Available online: http://www.stat.umn.edu/geyer/old03/5102/notes/fish.pdf (accessed on 10 August 2021).
31. Greene, W. *Econometric Analysis*, 8th ed.; Prentice-Hall: New York, NY, USA, 2020; pp. 118–121.
32. EL-Sagheer, R. Estimation of parameters of Weibull Gamma distribution based on progressively censored data. *Stat. Pap.* **2018**, *59*, 725–757. [CrossRef]
33. Meeker, W.Q.; Escobar, L.A. *Statistical Methods for Reliability Data*, 2nd ed.; John Wiley: New York, NY, USA, 1998.
34. Aarset, M.V. How to identify bathtub hazard rate. *IEEE Trans. Reliab.* **1987**, *36*, 106–108. [CrossRef]
35. Bergman, B.; Klefsjou, B. Burn–in models and TTT-transforms. *Qual. Reliab. Eng. Int.* **1985**, *1*, 125–130. [CrossRef]
36. Gentleman, R. Muhaz: A Package for Producing a Smooth Estimate of the Hazard Function for Censored Data. Available online: https://cran.r-project.org/web/packages/muhaz/muhaz.pdf (accessed on 20 August 2021).
37. Thach, T.T.; Bris, R.; Volf, P.; Coolen, F.P.A. Non-linear failure rate: A Bayes study using Hamiltonian Monte Carlo simulation. *Int. J. Approx. Reason.* **2020**, *123*, 55–76. [CrossRef]
38. Rubinstein, R.Y.; Kroese, D.P. *Simulation and the Monte Carlo Method*, 3rd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2016; pp. 257–303.
39. Kroese, D.P.; Porotsky, S.; Rubinstein, R.Y. The Cross-Entropy Method for Continuous Multi-Extremal Optimization. *Methodol. Comput. Appl. Probab.* **2006**, *8*, 383–407. [CrossRef]
40. Benham, T.; Duan, Q.; Kroese, D.P.; Liquet, B. CEoptim: Cross-Entropy R Package for Optimization. *J. Stat. Softw.* **2017**, *76*, 1–29. [CrossRef]