

Bias-Corrected and Variance-Corrected MLE for the New Median Based Unit Weibull Distribution (MBUW)

Abstract: The maximum likelihood estimation (MLE) method stands as the premier approach for parameter estimation, showcasing its attributes of being unbiased, consistent, efficient, and asymptotically normal. In this study, the author leverages MLE to fit the innovative distribution known as MBUW. It is critical to recognize, however, that MLE estimators can exhibit bias in small to moderate sample sizes, a contrast to the behavior observed with large sample sizes.

This paper vigorously discusses a bias-corrected approach for a new distribution known as Median Biased Unit Weibull (MBUW) and applies it effectively to real data analysis. The MLE estimators for MBUW, derived using optimization techniques such as the derivative-free Nelder-Mead algorithm, demonstrate significant high correlations that lead to substantial covariance between the parameters. This correlation has a detrimental impact on the parameter variances, which can inflate to a level that approaches infinity, obstructing the ability to construct reliable confidence intervals for each parameter. Such issues can emerge with any optimization technique, highlighting the urgent necessity for effective remedies. The author presents a robust variance correction approach that fundamentally relies on re-parameterizing the negative log-likelihood function.

Keywords: Cox and Snell bias-correction, Median Based Unit Weibull (MBUW), Maximum Likelihood estimators, Monte Carlo simulation, variance-corrected MLE, Bias-corrected MLE.

Comment: The abstract must be in single paragraph. Here, the author should explain the background (previous research and gap analysis), objective, method, and the result and or discussion and recommendation. Keyword should be 5 words (maximum).

Introduction

The method most frequently employed for parameter estimation is the maximum likelihood estimation (MLE), a cornerstone in the realm of probability distribution analysis (Pawitan, Y, 2001),(Millar, 2011). The MLE estimator boasts an array of advantageous characteristics: it is asymptotically unbiased, consistent, efficient, and follows an asymptotic normal distribution. However, these commendable traits primarily manifest in large sample sizes. In cases involving small to moderate sample sizes, certain properties—most notably unbiasedness—may begin to falter. As the sample size increases, the discrepancy between the expected value of a parameter and its true value diminishes, leading to more reliable estimates.

To address the challenges posed by smaller samples, researchers have devised nearly unbiased estimators tailored for various distributions. A glimpse into this extensive body of work reveals contributions from notable scholars including: (Saha & Paul, 2005),(Cordeiro et al., 1997),

(Giles, 2012), (Cribari-Neto & Vasconcellos, 2002), (Lemonte et al., 2007), additional insights can be found in the studies by (Giles, D. E & H. Feng, 2009), (Giles, 2012), as well as (Schwartz et al., 2013) who continue to expand this field. The exploration doesn't stop there; the research efforts of (Giles et al., 2013), (Teimouri & Nadarajah, 2013), (Zhang & Liu, 2017), and (Singh et al., 2015) further enrich the discourse. Scholars like (Lagos-Álvarez, B. et al., 2011), (Schwartz & Giles, 2016), and (Wang & Wang, 2017) also lend their voices, while (Ling & Giles, 2014), along with (Lemonte, 2011), and (Mazucheli & Dey, 2018), as well as (Reath, J., 2016), and (Teimouri & Nadarajah, 2016) contribute a wealth of knowledge to this important area of study. Also (Cordeiro & Cribari-Neto, 2014), (Mazucheli et al., 2019), (Menezes et al., 2021), (Rasekhi et al., 2019), and (Ferrari & Cribari-Neto, 1998) contributed to this subject. Each of these references plays an integral role in advancing our understanding of nearly unbiased estimation in statistical practice.

Estimating the bias of the maximum likelihood estimator (MLE) for distributions with a single parameter can be approached to the $O(n^{-1})$ even when the estimated parameter lacks a closed-form expression. Notably, pivotal work by statisticians such as (Bartlett, 1953a) (Haldane & Smith, 1956), (Bartlett, 1953b) and (Haldane, 1953) established the groundwork for this estimation by deriving analytical approximations for two-parameter log-likelihood functions through the application of Taylor series expansions. These expansions allow for an approximation of the behavior of the log-likelihood functions around certain parameter values, making it feasible to analyze the bias associated with the MLE.

However, the complexity increases considerably when extending these methodologies to multi-parameter distributions. This is highlighted in the research conducted by (Shenton & Bowman, 1963), which illustrated the difficulties encountered in applying similar Taylor series techniques. In multi-parameter contexts, the interactions among the parameters can complicate the approximation process, often leading to more intricate relationships that hinder straightforward calculations. As a result, tailoring methods to effectively estimate the bias in such cases necessitates more sophisticated analytical tools and techniques.

Numerous strategies have emerged to address the bias inherent in Maximum Likelihood Estimators (MLE). The first of these strategies, known as the "corrective approach," is a methodical framework championed by (Cox & Snell, 1968). This approach offers a precise analytical formula that quantifies the bias associated with MLE estimators, enabling researchers to refine these estimators into unbiased forms. Essentially, it lays the groundwork for making MLEs more accurate by systematically correcting the inherent biases. It is an analytical expression for the bias to $O(n^{-1})$ of the MLE estimators, then using these expressions to bias-correct the MLE estimator yielding estimators that are unbiased to $O(n^{-2})$.

The second method is the parametric Bootstrap resampling technique, put forth by (Efron, 1982). This ingenious approach involves a second-order bias correction, executing the corrections through numerical simulations rather than relying on analytical formulas. Through this technique, estimators can be adjusted without the need for explicit expressions for the bias function, showcasing its flexibility and robustness in addressing bias.

Lastly, the "preventive approach," advocated by (Firth, 1993), takes a different angle. This analytical procedure alters the score function of the log-likelihood before one attempts to solve for the MLEs, effectively minimizing bias to a specified degree. It reduces the bias to the order $O(n^{-2})$.

The first two strategies stand out for their relatively straightforward mathematical expressions, making them both attractive and easy to implement in practical settings. Each

approach, with its unique methodology, contributes invaluable tools for enhancing the accuracy of statistical estimations.

The MBUW distribution has been thoroughly explored in previous research conducted by the author (Attia, 2025) focusing on its properties and MLE estimation method, which have practical applications in real data analysis. The distribution is characterized by its probability density function (PDF), cumulative distribution function (CDF), and quantile function, which are detailed in equations (1-3).

$$f(y) = \frac{6}{\alpha^\beta} \left[1 - y^{\frac{1}{\alpha^\beta}} \right] y^{\left(\frac{2}{\alpha^\beta}-1\right)}, \quad 0 < y < 1, \quad \alpha > 0, \beta > 0 \quad (1)$$

$$F(y) = 3y^{\frac{2}{\alpha^\beta}} - 2y^{\frac{3}{\alpha^\beta}}, \quad 0 < y < 1, \quad \alpha > 0, \beta > 0 \quad (2)$$

$$u = F(y) = 3y^{\frac{2}{\alpha^\beta}} - 2y^{\frac{3}{\alpha^\beta}} = -2 \left(y^{\frac{1}{\alpha^\beta}} \right)^3 + 3 \left(y^{\frac{1}{\alpha^\beta}} \right)^2 \quad (3)$$

This distribution is meticulously defined within the confines of the unit interval, making it adept at accommodating skewed data presented as proportions. Depending on the parameters at play, the distribution can assume fascinating shapes, manifesting either an increasing, decreasing, or a striking unimodal form.

In this paper, the author delves deep into the intricacies of the bias-corrected maximum likelihood estimation (MLE) approach, alongside the innovative variance-corrected MLE technique. The structure of the paper unfolds in a series of well-organized sections. Section 2 is dedicated to elucidating the corrective procedures designed to address the biases inherent in MLE estimators. Here, the author methodically walks the reader through the necessary adjustments needed to refine these estimators.

In Section 3, the narrative unfolds further as the author derives the analytical function for the bias of MLE estimators specific to the MBUW distribution. This section enriches the discussion with significant theoretical findings. Results are then laid bare in Section 3, where the author elaborates on the variance-corrected MLE procedure for the MBUW distribution, showcasing its practical application with real-world data, thus grounding the theoretical elements in tangible outcomes.

The discourse continues in Section 4, where the author critically discusses the application of bias-corrected MLE on real datasets. Here, a comparative analysis is presented, juxtaposing the estimated parameters generated from the variance-corrected MLE with those derived through the bias-corrected approach, revealing insightful contrasts and nuances.

Finally, in Section 5 and 6, the author encapsulates the discussion with conclusions and recommendations, distilling the findings into key takeaways and future directions for research. This comprehensive exploration offers both clarity and depth to the reader, illuminating the path of MLE estimation in the context of the uniquely accommodating distributions discussed.

2. Methods

2.1. Corrective approach of the Bias-corrected MLE

Let Θ be a p -dimensional unknown parameter vector and $l = l(\Theta|y)$ be the log-likelihood function for a sample of n observations. Assume this log-likelihood is regular with respect to all derivatives up to and including those of the third order. The joint cumulants of the log-likelihood derivatives are defined as follows in equations (4-6)

$$k_{ij} = E \left(\frac{\partial^2 l}{\partial \theta_i \partial \theta_j} \right); i, j = 1, 2, \dots, p \quad (4)$$

$$k_{ijl} = E \left(\frac{\partial^3 l}{\partial \theta_i \partial \theta_j \partial \theta_l} \right); i, j = 1, 2, \dots, p \quad (5)$$

$$k_{ij,l} = E \left(\frac{\partial^2 l}{\partial \theta_i \partial \theta_j} \frac{\partial^3 l}{\partial \theta_l} \right); i, j = 1, 2, \dots, p \quad (6)$$

The derivative of these cumulants are defined as in equation (7)

$$k_{ij}^{(l)} = E \left(\frac{\partial k_{ij}}{\partial \theta_l} \right); i, j, l = 1, 2, \dots, p \quad (7)$$

The expressions in equations (4-7) are assumed to be $O(n)$.

According to (Cox & Snell, 1968), it was demonstrated that if the sample data is independent, though not necessarily identically distributed, the bias of the s^{th} element of the MLE of $\hat{\Theta}_s$, can be calculated as in equation (8):

$$Bias(\hat{\Theta}_s) = \sum_{i=1}^p \sum_{j=1}^p \sum_{l=1}^p k^{si} k^{jl} [0.5 * k_{ijl} + k_{ij,l}] + O(n^{-2}) \quad (8)$$

where : $s = 1, \dots, p$ and k^{ij} is the (i, j) th element of the inverse of the expected information matrix $K = -k_{ij}$.

(Cordeiro & Klein, 1994) established that equation (8) still holds if the data are non-independent, and it can be expressed as

$$Bias(\hat{\Theta}_s) = \sum_{i=1}^p k^{si} \sum_{j=1}^p \sum_{l=1}^p [k_{ij}^{(l)} - 0.5 * k_{ijl}] k^{jl} + O(n^{-2}) \quad (9)$$

The bias equation in (9) is largely simpler to calculate than in (8), because it does not include terms of the form given in (6).

Defining the following terms in equations (10-11):

$$a_{ij}^{(l)} = k_{ij}^{(l)} - 0.5 k_{ijl}; \quad i, j, l = 1, 2, \dots, p \quad (10)$$

$$A^{(l)} = \{ a_{ij}^{(l)} \}; \quad i, j, l = 1, 2, \dots, p \quad (11)$$

and collecting terms up into matrices $A = [A^{(1)} | \dots | A^{(p)}]$

The $O(n^{-2})$ bias of the MLE of $\hat{\Theta}$ in (9) can be rephrased in the convenient form equation (12):

$$\text{Bias}(\hat{\Theta}_s) = \hat{K}^{-1} \hat{A} \text{vec}(\hat{K}^{-1}) \quad (12)$$

where $\hat{K} = K|_{\Theta=\hat{\Theta}}$ and $\hat{A} = A|_{\Theta=\hat{\Theta}}$, the value of $\hat{\Theta}_s$ is obtained by solving the roots of log-likelihood equations using the numerical methods. $\text{Vec}(\cdot)$ means vectorization operator, which stacks the columns of the matrix in question one above the other, forming one extended column vector. Hence the bias adjusted-MLE is defined in equation (13) as

$$\tilde{\Theta} = \hat{\Theta} - \hat{K}^{-1} \hat{A} \text{vec}(\hat{K}^{-1}) \quad (13)$$

One of the advantages of this methodology is that these expressions can be calculated when determining the roots of the log-likelihood equations, which does not yield an analytic closed-form solution. In such cases, the bias-corrected maximum likelihood estimator (MLE) can be readily derived through standard numerical techniques, ensuring that the resulting estimator remains unbiased, and $\tilde{\Theta}$ is unbiased $O(n^{-2})$

2.2. Corrective Approach for Bias Reduction and MBUW

The PDF of MBUW satisfies the regularity conditions. In this context, the first-order partial derivatives of the log-likelihood function, concerning the alpha and beta parameters, are articulated in equation (14-15) as follows:

$$\frac{\partial l}{\partial \alpha} = \frac{\beta}{\alpha} + \sum_{i=1}^n \frac{y^{\alpha-\beta} \beta (\ln y) \alpha^{(-\beta-1)}}{1 - y^{\alpha-\beta}} - 2\beta \alpha^{(-\beta-1)} \sum_{i=1}^n \ln y \quad (14)$$

$$\frac{\partial l}{\partial \beta} = -n \ln \alpha + \sum_{i=1}^n \frac{y^{\alpha-\beta} \alpha^{-\beta} (\ln y) (\ln \alpha)}{1 - y^{\alpha-\beta}} - 2\alpha^{-\beta} (\ln \alpha) \sum_{i=1}^n \ln y \quad (15)$$

Equations (16-24) are the higher order derivatives (for one observation):

$$\begin{aligned} \frac{\partial^2 l}{\partial \alpha^2} &= \frac{\beta}{\alpha^2} - \left(\frac{y^{\alpha-\beta} \beta^2 (\ln y)^2 \alpha^{(-2\beta-2)}}{1 - y^{\alpha-\beta}} \right) - \frac{y^{\alpha-\beta} \beta (\beta + 1) (\ln y) \alpha^{(-\beta-2)}}{1 - y^{\alpha-\beta}} \\ &- \left(\frac{y^{2\alpha-\beta} \beta^2 (\ln y)^2 \alpha^{(-2\beta-2)}}{(1 - y^{\alpha-\beta})^2} \right) + 2\beta (\beta + 1) \alpha^{(-\beta-2)} (\ln y) \end{aligned} \quad (16)$$

$$\begin{aligned} \frac{\partial^2 l}{\partial \beta^2} &= \left(\frac{-y^{\alpha-\beta} (\ln \alpha)^2 (\ln y)^2 \alpha^{(-2\beta)}}{1 - y^{\alpha-\beta}} \right) - \left(\frac{y^{\alpha-\beta} (\ln \alpha)^2 (\ln y) \alpha^{(-\beta)}}{1 - y^{\alpha-\beta}} \right) \\ &- \left(\frac{y^{2\alpha-\beta} (\ln \alpha)^2 (\ln y)^2 \alpha^{(-2\beta)}}{(1 - y^{\alpha-\beta})^2} \right) + 2\alpha^{(-\beta)} (\ln y) (\ln \alpha)^2 \end{aligned} \quad (17)$$

$$\begin{aligned} \frac{\partial^2 l}{\partial \alpha \partial \beta} &= \frac{-1}{\alpha} - \left(\frac{y^{\alpha-\beta} \beta (\ln \alpha) (\ln y)^2 \alpha^{(-2\beta-1)}}{1 - y^{\alpha-\beta}} \right) + \frac{y^{\alpha-\beta} (\ln y) \alpha^{(-\beta-1)}}{1 - y^{\alpha-\beta}} \\ &- \left(\frac{y^{\alpha-\beta} \beta (\ln \alpha) (\ln y) \alpha^{(-\beta-1)}}{1 - y^{\alpha-\beta}} \right) - \left(\frac{y^{2\alpha-\beta} \beta (\ln \alpha) (\ln y)^2 \alpha^{(-2\beta-1)}}{(1 - y^{\alpha-\beta})^2} \right) \end{aligned}$$

$$-2\alpha^{(-\beta-1)} (\ln y) + 2\beta \alpha^{(-\beta-1)} (\ln \alpha)(\ln y) \quad (18)$$

$$\begin{aligned} \frac{\partial^3 l}{\partial \alpha^2 \partial \beta} &= \frac{1}{\alpha^2} + \frac{y^{\alpha^{-\beta}} \beta^2 (\ln y)^3 (\ln \alpha) \alpha^{(-3\beta-2)}}{1 - y^{\alpha^{-\beta}}} - \frac{2 y^{\alpha^{-\beta}} \beta (\ln y)^2 \alpha^{(-2\beta-2)}}{1 - y^{\alpha^{-\beta}}} \\ &+ \frac{2 y^{\alpha^{-\beta}} \beta^2 (\ln y)^2 (\ln \alpha) \alpha^{(-2\beta-2)}}{1 - y^{\alpha^{-\beta}}} + \frac{y^{2\alpha^{-\beta}} \beta^2 (\ln y)^3 (\ln \alpha) \alpha^{(-3\beta-2)}}{(1 - y^{\alpha^{-\beta}})^2} \\ &- \frac{y^{\alpha^{-\beta}} (2\beta + 1) (\ln y) \alpha^{(-\beta-2)}}{1 - y^{\alpha^{-\beta}}} + \frac{y^{\alpha^{-\beta}} (\beta^2 + \beta) (\ln \alpha) (\ln y)^2 \alpha^{(-2\beta-2)}}{1 - y^{\alpha^{-\beta}}} \\ &+ \frac{y^{\alpha^{-\beta}} (\beta^2 + \beta) (\ln \alpha) (\ln y) \alpha^{(-\beta-2)}}{1 - y^{\alpha^{-\beta}}} + \frac{y^{2\alpha^{-\beta}} (\beta^2 + \beta) (\ln y)^2 (\ln \alpha) \alpha^{(-2\beta-2)}}{(1 - y^{\alpha^{-\beta}})^2} \\ &+ \frac{2 y^{2\alpha^{-\beta}} \beta^2 (\ln y)^3 (\ln \alpha) \alpha^{(-3\beta-2)}}{(1 - y^{\alpha^{-\beta}})^2} - \frac{2 y^{2\alpha^{-\beta}} \beta (\ln y)^2 \alpha^{(-2\beta-2)}}{(1 - y^{\alpha^{-\beta}})^2} \\ &+ \frac{2 y^{2\alpha^{-\beta}} \beta^2 (\ln y)^2 (\ln \alpha) \alpha^{(-2\beta-2)}}{(1 - y^{\alpha^{-\beta}})^2} + \frac{2 y^{3\alpha^{-\beta}} \beta^2 (\ln \alpha) (\ln y)^3 \alpha^{(-3\beta-2)}}{(1 - y^{\alpha^{-\beta}})^3} \\ &+ 4\beta \alpha^{(-\beta-2)} (\ln y) + 2\alpha^{(-\beta-2)} (\ln y) - 2\beta^2 \alpha^{(-\beta-2)} (\ln y)(\ln \alpha) \\ &- 2\beta \alpha^{(-\beta-2)} (\ln y)(\ln \alpha) \quad (19) \end{aligned}$$

$$\begin{aligned} \frac{\partial^3 l}{\partial \beta^2 \partial \alpha} &= \frac{y^{\alpha^{-\beta}} \beta (\ln y)^3 (\ln \alpha)^2 \alpha^{(-3\beta-1)}}{1 - y^{\alpha^{-\beta}}} + \frac{2 y^{\alpha^{-\beta}} \beta (\ln y)^2 (\ln \alpha)^2 \alpha^{(-2\beta-1)}}{1 - y^{\alpha^{-\beta}}} \\ &- \frac{2 y^{\alpha^{-\beta}} (\ln y)^2 (\ln \alpha) \alpha^{(-2\beta-1)}}{1 - y^{\alpha^{-\beta}}} + \frac{y^{2\alpha^{-\beta}} \beta (\ln y)^3 (\ln \alpha)^2 \alpha^{(-3\beta-1)}}{(1 - y^{\alpha^{-\beta}})^2} \\ &+ \frac{y^{\alpha^{-\beta}} \beta (\ln y)^2 (\ln \alpha)^2 \alpha^{(-2\beta-1)}}{1 - y^{\alpha^{-\beta}}} + \frac{y^{\alpha^{-\beta}} (\beta) (\ln y) (\ln \alpha)^2 \alpha^{(-\beta-1)}}{1 - y^{\alpha^{-\beta}}} \\ &- \frac{2 y^{\alpha^{-\beta}} (\ln \alpha) (\ln y) \alpha^{(-\beta-1)}}{1 - y^{\alpha^{-\beta}}} + \frac{y^{2\alpha^{-\beta}} (\beta) (\ln y)^2 (\ln \alpha)^2 \alpha^{(-2\beta-1)}}{(1 - y^{\alpha^{-\beta}})^2} \\ &+ \frac{2 y^{2\alpha^{-\beta}} \beta (\ln y)^3 (\ln \alpha)^2 \alpha^{(-3\beta-1)}}{(1 - y^{\alpha^{-\beta}})^2} + \frac{2 y^{2\alpha^{-\beta}} \beta (\ln y)^2 (\ln \alpha)^2 \alpha^{(-2\beta-1)}}{(1 - y^{\alpha^{-\beta}})^2} \\ &- \frac{2 y^{2\alpha^{-\beta}} (\ln y)^2 (\ln \alpha) \alpha^{(-2\beta-1)}}{(1 - y^{\alpha^{-\beta}})^2} + \frac{2 y^{3\alpha^{-\beta}} \beta (\ln \alpha)^2 (\ln y)^3 \alpha^{(-3\beta-1)}}{(1 - y^{\alpha^{-\beta}})^3} \\ &- 2\beta \alpha^{(-\beta-1)} (\ln y)(\ln \alpha)^2 + 4\alpha^{(-\beta-1)} (\ln y) (\ln \alpha) \quad (20) \end{aligned}$$

$$\begin{aligned}
\frac{\partial^3 l}{\partial \alpha^3} &= \frac{-2\beta}{\alpha^3} + \frac{y^{\alpha-\beta} \beta^3 (\ln y)^3 \alpha^{(-3\beta-3)}}{1-y^{\alpha-\beta}} + \frac{3y^{\alpha-\beta} \beta^2 (\beta+1) (\ln y)^2 \alpha^{(-2\beta-3)}}{1-y^{\alpha-\beta}} \\
&+ \frac{3\beta^3 y^{2\alpha-\beta} (\ln y)^3 \alpha^{(-3\beta-3)}}{(1-y^{\alpha-\beta})^2} + \frac{y^{\alpha-\beta} \beta (\beta+1)(\beta+2) (\ln y) \alpha^{(-\beta-3)}}{1-y^{\alpha-\beta}} \\
&+ \frac{3\beta^2 (\beta+1) y^{2\alpha-\beta} (\ln y)^2 \alpha^{(-2\beta-3)}}{(1-y^{\alpha-\beta})^2} + \frac{2y^{3\alpha-\beta} \beta^3 (\ln y)^3 \alpha^{(-3\beta-3)}}{(1-y^{\alpha-\beta})^3} \\
&- 2\beta (\beta+1)(\beta+2) \alpha^{(-\beta-3)} (\ln y)
\end{aligned} \tag{21}$$

$$\begin{aligned}
\frac{\partial^3 l}{\partial \beta^3} &= \frac{y^{\alpha-\beta} (\ln \alpha)^3 (\ln y) \alpha^{(-\beta)}}{1-y^{\alpha-\beta}} + \frac{3y^{\alpha-\beta} (\ln \alpha)^3 (\ln y)^2 \alpha^{(-2\beta)}}{1-y^{\alpha-\beta}} \\
&+ \frac{y^{\alpha-\beta} (\ln \alpha)^3 (\ln y)^3 \alpha^{(-3\beta)}}{1-y^{\alpha-\beta}} + \frac{3y^{2\alpha-\beta} (\ln \alpha)^3 (\ln y)^2 \alpha^{(-2\beta)}}{(1-y^{\alpha-\beta})^2} \\
&+ \frac{3y^{2\alpha-\beta} (\ln \alpha)^3 (\ln y)^3 \alpha^{(-3\beta)}}{(1-y^{\alpha-\beta})^2} + \frac{2y^{3\alpha-\beta} (\ln \alpha)^3 (\ln y)^3 \alpha^{(-3\beta)}}{(1-y^{\alpha-\beta})^3} \\
&- 2\alpha^{(-\beta)} (\ln y) (\ln \alpha)^3
\end{aligned} \tag{22}$$

$$\begin{aligned}
\frac{\partial^3 l}{\partial \beta \partial \alpha \partial \alpha} &= \frac{1}{\alpha^2} - 2\beta(\beta+1)\alpha^{(-\beta-2)} (\ln y) (\ln \alpha) + 2(2\beta+1)\alpha^{(-\beta-2)} (\ln y) \\
&+ \frac{y^{\alpha-\beta} \beta(2\beta+1) (\ln y)^2 (\ln \alpha) \alpha^{(-2\beta-2)}}{1-y^{\alpha-\beta}} - \frac{y^{\alpha-\beta} \beta (\ln y)^2 \alpha^{(-2\beta-2)}}{1-y^{\alpha-\beta}} \\
&+ \frac{y^{\alpha-\beta} \beta^2 (\ln y)^3 (\ln \alpha) \alpha^{(-3\beta-2)}}{1-y^{\alpha-\beta}} + \frac{3y^{2\alpha-\beta} \beta^2 (\ln y)^3 (\ln \alpha) \alpha^{(-3\beta-2)}}{(1-y^{\alpha-\beta})^2} \\
&- \frac{y^{\alpha-\beta} \beta (\ln y)^2 \alpha^{(-2\beta-2)}}{1-y^{\alpha-\beta}} - \frac{y^{\alpha-\beta} (\beta+1) (\ln y) \alpha^{(-\beta-2)}}{1-y^{\alpha-\beta}} \\
&- \frac{3y^{2\alpha-\beta} \beta (\ln y)^2 \alpha^{(-2\beta-2)}}{(1-y^{\alpha-\beta})^2} + \frac{y^{\alpha-\beta} \beta (1+\beta) (\ln \alpha) (\ln y) \alpha^{(-\beta-2)}}{1-y^{\alpha-\beta}} \\
&+ \frac{y^{\alpha-\beta} \beta^2 (\ln \alpha) (\ln y)^2 \alpha^{(-2\beta-2)}}{1-y^{\alpha-\beta}} - \frac{y^{\alpha-\beta} \beta (\ln y) \alpha^{(-\beta-2)}}{1-y^{\alpha-\beta}} \\
&+ \frac{y^{2\alpha-\beta} \beta^2 (\ln y)^2 (\ln \alpha) \alpha^{(-2\beta-2)}}{(1-y^{\alpha-\beta})^2} + \frac{\beta y^{2\alpha-\beta} (2\beta+1) (\ln y)^2 (\ln \alpha) \alpha^{(-2\beta-2)}}{(1-y^{\alpha-\beta})^2}
\end{aligned}$$

$$+ \frac{2y^{3\alpha-\beta} \beta^2 (\ln \alpha) (\ln y)^3 \alpha^{(-3\beta-2)}}{(1-y^{\alpha-\beta})^3} \quad (23)$$

$$\begin{aligned} \frac{\partial^3 l}{\partial \alpha \partial \beta \partial \beta} &= 4 \alpha^{(-\beta-1)} (\ln y) (\ln \alpha) - 2\beta \alpha^{(-\beta-1)} (\ln y) (\ln \alpha)^2 \\ &+ \frac{\beta y^{\alpha-\beta} (\ln \alpha)^2 (\ln y)^3 \alpha^{(-3\beta-1)}}{1-y^{\alpha-\beta}} - \frac{2y^{\alpha-\beta} (\ln y)^2 (\ln \alpha) \alpha^{(-2\beta-1)}}{1-y^{\alpha-\beta}} \\ &+ \frac{3y^{\alpha-\beta} \beta (\ln y)^2 (\ln \alpha)^2 \alpha^{(-2\beta-1)}}{1-y^{\alpha-\beta}} + \frac{3y^{2\alpha-\beta} \beta (\ln y)^3 (\ln \alpha)^2 \alpha^{(-3\beta-1)}}{(1-y^{\alpha-\beta})^2} \\ &- \frac{2y^{\alpha-\beta} (\ln \alpha) (\ln y) \alpha^{(-\beta-1)}}{1-y^{\alpha-\beta}} - \frac{2y^{2\alpha-\beta} \beta (\ln y)^2 (\ln \alpha) \alpha^{(-2\beta-1)}}{(1-y^{\alpha-\beta})^2} \\ &+ \frac{y^{\alpha-\beta} \beta (\ln y) (\ln \alpha)^2 \alpha^{(-\beta-1)}}{1-y^{\alpha-\beta}} + \frac{3y^{2\alpha-\beta} \beta (\ln y)^2 (\ln \alpha)^2 \alpha^{(-2\beta-1)}}{(1-y^{\alpha-\beta})^2} \\ &+ \frac{2\beta y^{3\alpha-\beta} (\ln \alpha)^2 (\ln y)^3 \alpha^{(-3\beta-1)}}{(1-y^{\alpha-\beta})^3} \quad (24) \end{aligned}$$

The expectation of the derivative discussed previously is determined through the sophisticated technique of Monte Carlo integration, as illustrated in equations (25-31). Within these equations, the integrals can be computed using suitable numerical integration methods, resulting in a constant value. The author employed the trapezoidal method for this integration, ensuring accuracy in the calculations. By substituting the estimated parameters, alpha and beta, for each unique data set, we can derive the corresponding integrals, which yield fixed values specific to each dataset.

$$E\left(\frac{y^{\alpha-\beta} (\ln y)^3}{1-y^{\alpha-\beta}}\right) = \int_0^1 \frac{y^{\alpha-\beta} (\ln y)^3}{1-y^{\alpha-\beta}} f(y) dy = f_1 \quad (25)$$

$$E\left(\frac{y^{\alpha-\beta} (\ln y)^2}{1-y^{\alpha-\beta}}\right) = \int_0^1 \frac{y^{\alpha-\beta} (\ln y)^2}{1-y^{\alpha-\beta}} f(y) dy = f_2 \quad (26)$$

$$E\left(\frac{y^{\alpha-\beta} \ln y}{1-y^{\alpha-\beta}}\right) = \int_0^1 \frac{y^{\alpha-\beta} (\ln y)}{1-y^{\alpha-\beta}} f(y) dy = f_3 \quad (27)$$

$$E\left(\frac{y^{2\alpha-\beta} (\ln y)^3}{[1-y^{\alpha-\beta}]^2}\right) = \int_0^1 \frac{y^{2\alpha-\beta} (\ln y)^3}{[1-y^{\alpha-\beta}]^2} f(y) dy = f_4 \quad (28)$$

$$E\left(\frac{y^{2\alpha-\beta} (\ln y)^2}{[1-y^{\alpha-\beta}]^2}\right) = \int_0^1 \frac{y^{2\alpha-\beta} (\ln y)^2}{[1-y^{\alpha-\beta}]^2} f(y) dy = f_5 \quad (29)$$

$$E\left(\frac{y^{3\alpha-\beta} (\ln y)^3}{[1-y^{\alpha-\beta}]^3}\right) = \int_0^1 \frac{y^{3\alpha-\beta} (\ln y)^3}{[1-y^{\alpha-\beta}]^3} f(y) dy = f_6 \quad (30)$$

$$E(\ln y) = \int_0^1 (\ln y) f(y) dy = f_7 \quad (31)$$

Define the following quantities:

$$k_{11} = E\left(\frac{\partial^2 l}{\partial \alpha^2}\right), k_{22} = E\left(\frac{\partial^2 l}{\partial \beta^2}\right), \quad k_{12} = E\left(\frac{\partial^2 l}{\partial \alpha \partial \beta}\right) = k_{21} = E\left(\frac{\partial^2 l}{\partial \beta \partial \alpha}\right)$$

$$k_{111} = E\left(\frac{\partial^3 l}{\partial \alpha^3}\right), \quad k_{222} = E\left(\frac{\partial^3 l}{\partial \beta^3}\right), \quad k_{121} = E\left(\frac{\partial^3 l}{\partial \alpha \partial \beta \partial \alpha}\right) = k_{211} = E\left(\frac{\partial^3 l}{\partial \beta \partial \alpha \partial \alpha}\right)$$

$$k_{122} = E\left(\frac{\partial^3 l}{\partial \alpha \partial \beta \partial \beta}\right) = k_{212} = E\left(\frac{\partial^3 l}{\partial \beta \partial \alpha \partial \beta}\right)$$

$$k_{112} = E\left(\frac{\partial^3 l}{\partial \alpha^2 \partial \beta}\right), \quad k_{221} = E\left(\frac{\partial^3 l}{\partial \beta^2 \partial \alpha}\right)$$

$$k_{11}^{(1)} = \frac{\partial k_{11}}{\partial \alpha}, \quad k_{12}^{(1)} = \frac{\partial k_{12}}{\partial \alpha} = k_{21}^{(1)} = \frac{\partial k_{21}}{\partial \alpha}, \quad k_{22}^{(1)} = \frac{\partial k_{22}}{\partial \alpha}$$

$$k_{11}^{(2)} = \frac{\partial k_{11}}{\partial \beta}, \quad k_{12}^{(2)} = \frac{\partial k_{12}}{\partial \beta} = k_{21}^{(2)} = \frac{\partial k_{21}}{\partial \beta}, \quad k_{22}^{(2)} = \frac{\partial k_{22}}{\partial \beta}$$

$$a_{11}^{(1)} = k_{11}^{(1)} - 0.5 k_{111}, \quad a_{12}^{(1)} = k_{12}^{(1)} - 0.5 k_{121}, \quad a_{22}^{(1)} = k_{22}^{(1)} - 0.5 k_{221}$$

$$a_{11}^{(2)} = k_{11}^{(2)} - 0.5 k_{112}, \quad a_{12}^{(2)} = k_{12}^{(2)} - 0.5 k_{122}, \quad a_{22}^{(2)} = k_{22}^{(2)} - 0.5 k_{222}$$

The information matrix is $K = \{-k_{ij}\} = -n \times \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix}$, where n is the number of observations.

Defining $A_{ij}^{(q)} = \{a_{ij}^{(q)}\}; q = 1, 2$ and $A = [A^{(1)} \mid A^{(2)}]$

$$A = n \times \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{11}^{(2)} & a_{12}^{(2)} \\ a_{21}^{(1)} & a_{22}^{(1)} & a_{21}^{(2)} & a_{22}^{(2)} \end{bmatrix}$$

Upon using Cordeiro and Klein (1994) modification of the Cox and Snell (1968) result; the $Bias\left(\begin{matrix} \hat{\alpha} \\ \hat{\beta} \end{matrix}\right) = K^{-1} A vec(K^{-1})$

The bias-adjusted estimators can be obtained as in equation (32)

$$\begin{pmatrix} \alpha^* \\ \beta^* \end{pmatrix} = \hat{K}^{-1} \hat{A} vec(\hat{K}^{-1}) \quad (32)$$

Where $\hat{K} = K|_{\alpha=\hat{\alpha} \& \beta=\hat{\beta}}$ and $\hat{A} = A|_{\alpha=\hat{\alpha} \& \beta=\hat{\beta}}$

I think better to put here Section is 3,

3. Result

3. Variance-Corrected MLE procedure and MBUW

Maximum likelihood estimators (MLEs) for the parameters of the Median Based Unit Weibull (MBUW) distribution can sometimes exhibit significant variances, and in certain cases, the variances may even be infinite. This behavior arises primarily due to the inherent correlations among the parameters, which further contribute to large covariance values. When examining the landscape of the log-likelihood function, one may observe a relatively flat topology. This flatness indicates the presence of multiple pairs of estimators that effectively fit the distribution, leading to potential instabilities in the estimation process.

To explore this issue further, an examination of the surface of the negative log-likelihood function (nLL) is essential. This analysis reveals the specific pairs of parameters that yield the minimum negative likelihood values. The methodology adopted by the author in this study involves identifying these parameter pairs and analyzing the relationships between them. A key step in this process is to express one parameter in terms of the other, allowing for a more streamlined analysis.

Following this, the author re-parameterizes the negative log-likelihood function, with the possibility of applying a logarithmic scale to one of the parameters if necessary. This careful re-parameterization facilitates the estimation of the chosen parameter. After estimating this specific parameter, the value is substituted back into the established relationship to derive the corresponding value of the other parameter.

To calculate the variance of the parameter estimated through maximum likelihood estimation, one can utilize the inverse of the Fisher information matrix. Meanwhile, for the parameter derived from the established relationship, the author applies the delta method to obtain its variance. This comprehensive approach ensures a more robust understanding of the parameter estimations and their associated uncertainties.

3.1. Real data analysis using variance-corrected MLE

The OECD (Organization for Economic Co-operation and Development) is used. This organization is an international organization that aims to promote policies to improve the economic and social well-being of people around the world. The OECD provides a platform for its member countries to compare policy experiences, seek answers to common problems, identify good practices and coordinate domestic and international policies. The data is available at: <https://stats.oecd.org/index.aspx?DataSetCode=BLI>

Steps of the technique used by the author are outlined below:

1. Begin by inspecting the surface of the negative log-likelihood (nLL) function.
2. Identify and extract the pairs of alpha and beta parameters that minimize the nLL. To achieve this, the ranges for (α) and (β) are divided into 500 equally spaced points, with the specific ranges varying based on the datasets. These ranges are selected based on findings from maximum likelihood estimation (MLE) employing the Nelder-Mead algorithm.
3. Establish a relationship between (α) and (β) by fitting an optimal curve that describes this connection.

4. Re-parameterize the nLL using the established relationship between (α) and (β) .
5. Estimate the value of the (α) parameter.
6. Once (α) is estimated, substitute this value back into the relationship to determine the corresponding (β) parameter.
7. Utilize the inverse Fisher information obtained from the MLE to calculate the variance of the (α) parameter, and apply the delta method to derive the variance of the (β) parameter.
8. Finally, assess the goodness of fit for the distribution concerning the data using statistical tests such as the Kolmogorov-Smirnov (KS) test, Anderson-Darling (AD) test, and Cramér-von Mises (CVM) test. Other statistical indices like AIC, CAIC, BIA, HQIC, and nLL are also reported.

3.1.1. first dataset: time between failures data set.

Time between Failures of Secondary Reactor Pumps) (Maya et al., 2024, 1999) (Suprawhardana and Prayoto). The values of the data are: 0.216, 0.015, 0.4082, 0.0746, 0.0358, 0.0199, 0.0402, 0.0101, 0.0605, 0.0954, 0.1359, 0.0273, 0.0491, 0.3465, 0.007, 0.656, 0.106, 0.0062, 0.4992, 0.0614, 0.532, 0.0347, 0.1921. In this analysis, a thorough examination of the surface of the negative log-likelihood (nLL) was conducted. Figure 1 depicts this surface, which contains a notably flat region, indicating that a range of parameter values can yield comparable nLL results. A 246 pairs of parameters, that minimize the nLL, were successfully extracted. The relationship between these parameters is illustrated in Figure 2, which provides a clear visual representation of their interconnections. To further explore this dynamic, the relationship between the parameters α and β was defined by fitting a curve that best represents this correlation. This process is detailed in equations (33-35), which elucidate the interdependence of these parameters within the framework of the analysis.

$$\text{For exponential decay model: } \beta = 7.1594 e^{-0.97948(\alpha)} + 0.6735 \quad (33)$$

$$\text{For polynomial model: } \beta = 0.1011 \alpha^2 - 1.0218 \alpha + 3.2653 \quad (34)$$

$$\text{For reciprocal model: } \beta = \frac{3.1318}{\alpha} + 0.0648 \quad (35)$$

Follow the steps from 4 to 8

The exponential decay model demonstrates a strong fit to the data, characterized by a (RSS) value of 0.1239, an R-squared value of 0.9957, and a (RMSE) of 0.0225. These metrics indicate that the model not only accurately represents the dataset but also explains approximately 99.57% of the variability in the data, reflecting high precision in its predictions. On the other hand, the reciprocal model yields an RSS of 0.5583, an R-squared of 0.9808, and an RMSE of 0.0477. While these values illustrate that the reciprocal model also fits the data well, it does so somewhat less effectively than the exponential decay model, capturing about 98.08% of the data's variability, which still indicates a robust fit. Conversely, the quadratic model presents an RSS of 0.9653, an R-squared of 0.9668, and an RMSE of 0.0628. Although it captures around 96.68% of the variability within the dataset, its higher RSS and RMSE values suggest that this model is less optimal compared to the preceding two models. These results indicate that while the quadratic model provides a reasonable fit, it does not match the superior performance of the exponential decay and reciprocal models. Table 1 showcases the compelling results achieved by re-parameterizing the nLL function with the three distinct

models. These findings highlight the effectiveness of this approach and underline the significance of this analysis.

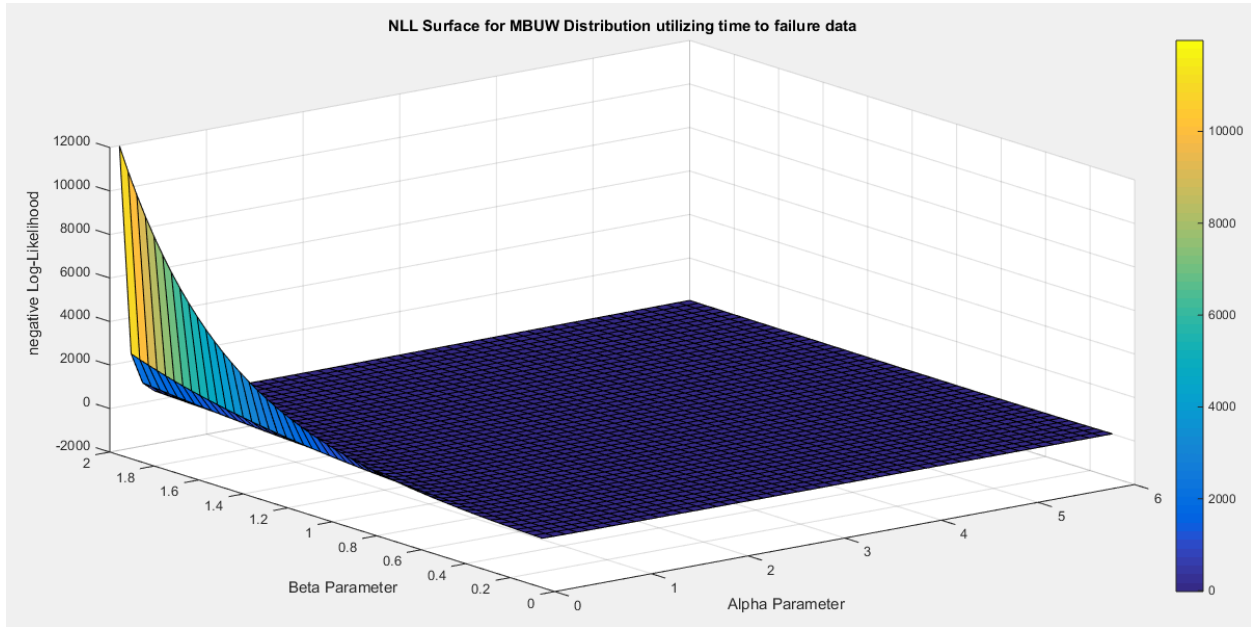


Fig. 1 Negative Log-Likelihood surface using the time between failures dataset.

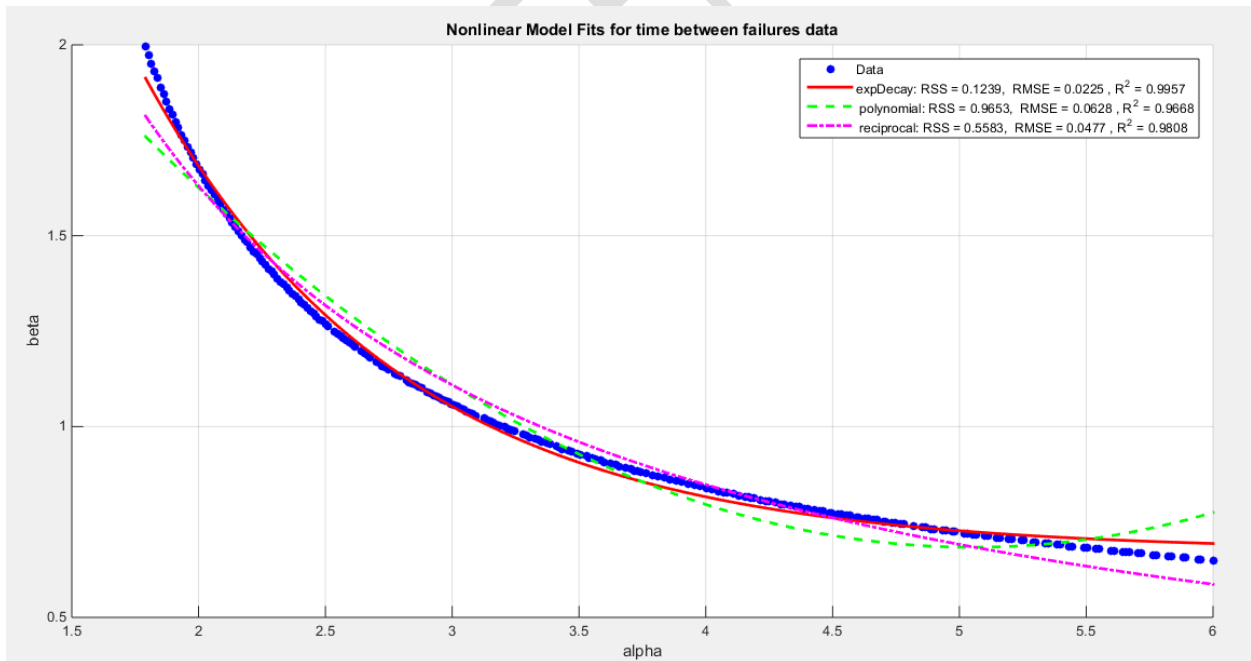


Fig. 2 shows a decreasing convex relationship between alpha and beta. The Nonlinear models for the time between failures dataset, with residual sum of squares (RSS), R² and root of mean square error(RMSE) are illustrated in the figure for each curve, high lightening that the exponential is the best model followed by the reciprocal model then the polynomial (quadratic).

Table 1: the results of the 3 nonlinear models using the time between failures data

metric	Exponential decay	Quadratic-polynomial	Reciprocal
α	1.977 (1.906, 2.048)	2.105 (2.0224, 2.1875)	2.141 (2.0551, 2.2269)
β	1.7062 (1.6343, 1.778)	1.5624 (1.5132, 1.6116)	1.5276 (1.4689, 1.5853)
Var (SE) α	0.0302 (0.0362)	0.0408 (0.0421)	0.0442 (0.0438)
Var (SE) β	0.0309 (0.0367)	0.0145 (0.0251)	0.0206 (0.0299)
nLL	-19.931	-19.931	-19.931
KS	0.1584	0.1584	0.1584
AD	0.6703	0.6703	0.6703
CVM	0.1253	0.1253	0.1253
AIC	-35.862	-35.862	-35.862
CAIC	-35.262	-35.262	-35.262
BIC	-33.591	-33.591	-33.591
HQIC	-35.2909	-35.2909	-35.2909
$H_0 = 0$	Fail to reject	Fail to reject	Fail to reject
P-value (KS-test)	0.5575	0.5575	0.5575

The evaluation of the models presented indicates that their variances are notably similar, suggesting a level of consistency in their performance. Furthermore, the estimators derived from these models show minimal differences, allowing for flexibility in model selection without significantly impacting the outcome. The (CIs) are quite narrow, reflecting a high degree of precision in the estimates. Interestingly, the stability of the estimators remains intact regardless of the initial assumptions or starting values used during the estimation process. This characteristic is particularly noteworthy, as it indicates robustness in the results; both the variance and the goodness-of-fit metrics exhibit little variability despite changes in these initial conditions. This pattern is an intriguing feature of the dataset under analysis.

3.1.2. second dataset: capacity factors data set.

data to evaluate the factors concerning the unit capacity, data compare factors between algorithms like SC 16 and P3)(Maya et al., 2024, 1999). The values of the data are: 0.853, 0.759, 0.866, 0.809, 0.717, 0.544, 0.492, 0.403, 0.344, 0.213, 0.116, 0.116, 0.092, 0.07, 0.059, 0.048, 0.036, 0.029, 0.021, 0.014, 0.011, 0.008, 0.006. Upon examining the surface of the negative log-likelihood (nLL), various features depicted in Figure 3 were observed, which notably displays a region that is

notably flat. This flat area suggests a range of parameter values that yield similar nLL outcomes, indicating potential stability in the parameter estimates. The author proceeded to extract pairs of parameters that effectively minimize the nLL, resulting in a total of 80 distinct pairs. The relationship between these parameter pairs is illustrated in Figure 3, which provides a visual representation of their interactions and dependencies. The exponential decay model demonstrates impressive performance with the following values: RSS of 0.0442, R squared of 0.9954, and RMSE of 0.0237. In comparison, the reciprocal model shows lesser effectiveness with an RSS of 0.2843, an R squared of 0.9701, and an RMSE of 0.06. Lastly, the quadratic model presents values of RSS at 0.3523, R squared at 0.963, and RMSE at 0.0668 as shown in Figure 4. Clearly, the exponential decay model stands out in terms of accuracy and reliability. To further understand the dynamics between the parameters, the author focuses on defining the relationship between α and β . This relationship is captured through a curve-fitting process and utilizing the best-fitting curve based on the provided equations (36-38). This analysis aims to uncover the underlying patterns in how alpha influences beta, enhancing our overall comprehension of the model's behavior.

$$\text{For exponential decay model: } \beta = 9.424 e^{-1.2276(\alpha)} + 0.6253 \quad (36)$$

$$\text{For polynomial model: } \beta = 0.1596 \alpha^2 - 1.3633 \alpha + 3.537 \quad (37)$$

$$\text{For reciprocal model: } \beta = \frac{3.0101}{\alpha} - 0.0811 \quad (38)$$

Follow the steps from 4 to 8

Table 2 presents the results of re-parameterizing the nLL function with three different models.

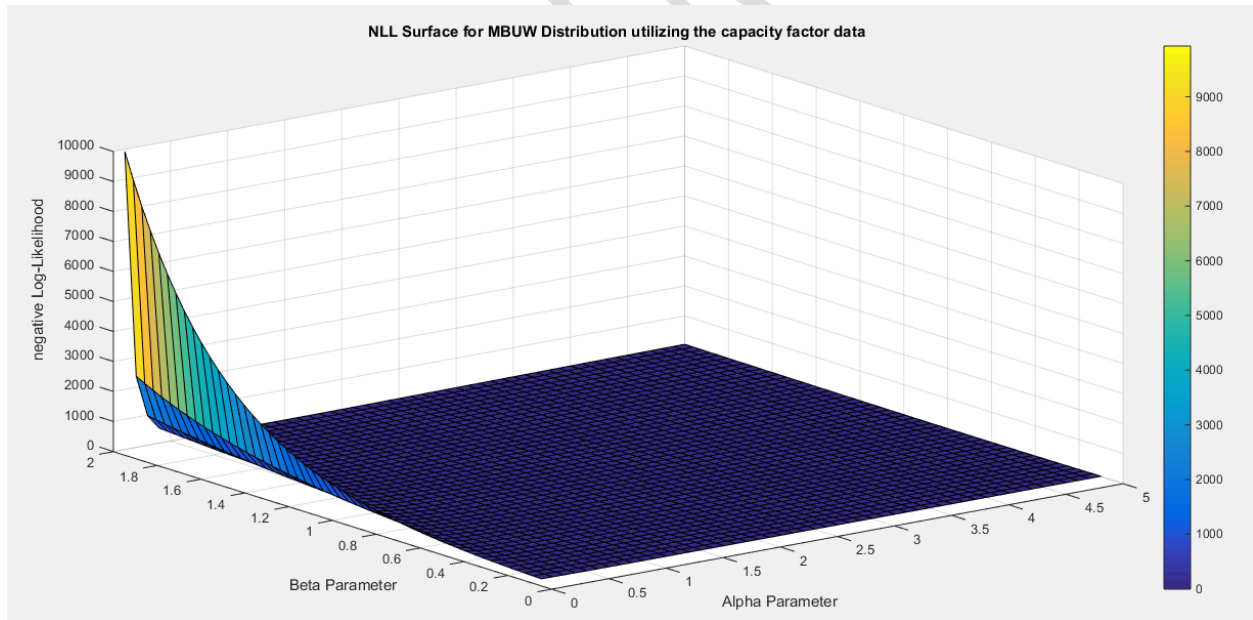


Fig. 3 Negative Log-likelihood surface using the factors affecting the unit capacity dataset.

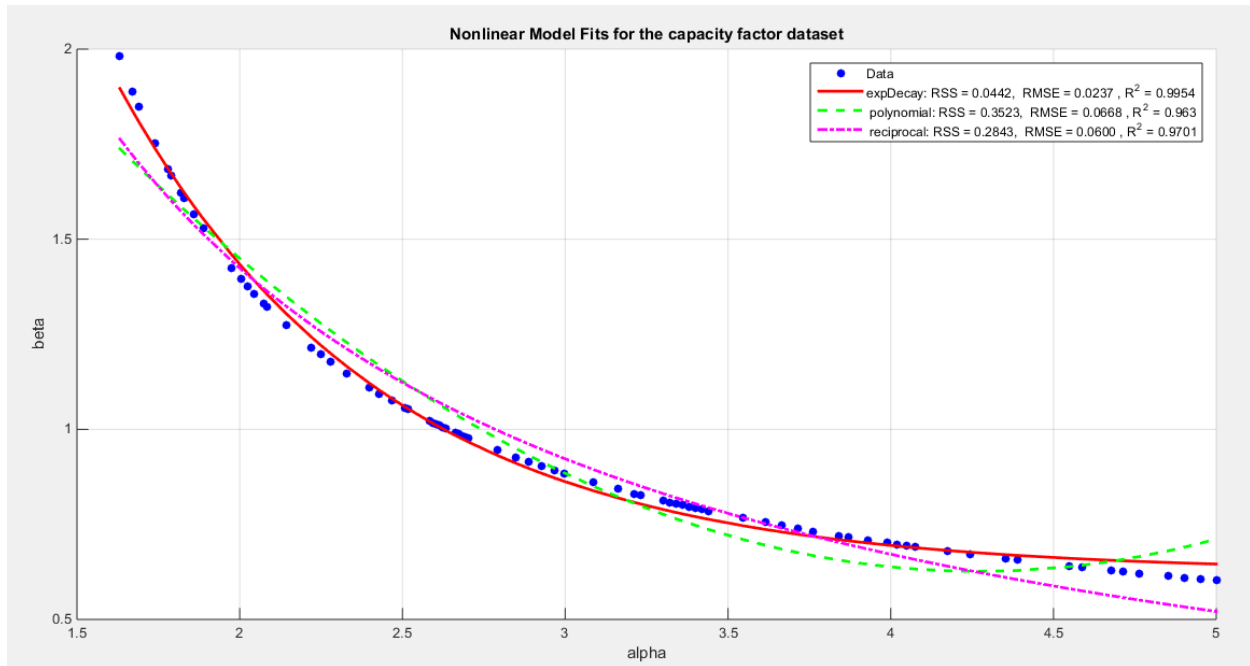


Fig. 4 shows a decreasing convex relationship between alpha and beta. The Nonlinear models for the capacity factor dataset, with residual sum of squares (RSS), R^2 and root of mean square error (RMSE) are illustrated in the figure for each curve, highlighting that the exponential is the best model followed by the reciprocal model then the polynomial (quadratic).

Table 3: the results of the 3 nonlinear models using the unit capacity data

metric	Exponential decay	Quadratic-polynomial	Reciprocal
α (CI)	1.7729(1.7086, 1.8373)	1.8767(1.8018, 1.9516)	1.9217(1.8421, 2.0012)
β (CI)	1.6943(1.6098, 1.7787)	1.5412(1.5984, 1.484)	1.4853(1.4205, 1.5502)
V&(SE) α	0.0248(0.0328)	0.0336(0.0382)	0.0379(0.0406)
V&(SE) β	0.0427(0.0431)	0.0196(0.0292)	0.0252(0.0331)
nLL	-7.6079	-7.6079	-7.6079
KS	0.1518	0.1518	0.1518
AD	1.9075	1.9075	1.9075
CVM	0.2033	0.2033	0.2033
AIC	-11.2158	-11.2158	-11.2158
CAIC	-10.6158	-10.6158	-10.6158
BIC	-8.9448	-8.9448	-8.9448
HQIC	-10.6446	-10.6446	-10.6446
$H_0 = 0$	Fail to reject	Fail to reject	Fail to reject
P-value (KS-test)	0.4074	0.4074	0.4074

The results generated by the three models exhibit a remarkable level of consistency and alignment and a high degree of comparability. The variances linked to the parameters across these models are strikingly low, underscoring their dependable and uniform performance in making

predictions. Furthermore, the confidence interval (CI) is notably narrow, indicating that the estimates are not only trustworthy but also robust, reliable and stable. Given these compelling attributes, any of the three models stands as a strong contender for implementation, as they all demonstrate an impressive degree and a commendable level of accuracy and precision in their outcomes.

4. Discussion

The bias-corrected methodology is implemented on the fourth and the fifth datasets to assess the estimated parameters derived from the variance-corrected approach utilizing the Quadratic polynomial model. This evaluation takes place before and after the application of the bias-corrected technique elaborated in section 3. In Table 3, readers will find a comprehensive display of the parameter values before and following the correction, accompanied by the corresponding bias values. Meanwhile, Table 4 reveals the updated metrics post-correction, as calculated using equation 32, offering a clearer view of the improvements made.

Table 4: the bias and values of parameters before and after correction

	Before correction	bias	After correction
Time between failure			
$\hat{\alpha}$	2.105	-0.000656	2.1057
$\hat{\beta}$	1.5624	0.0171	1.5453
Factors affecting unit capacity			
$\hat{\alpha}$	1.8767	0.000459	1.8762
$\hat{\beta}$	1.5412	0.0046	1.5366

The captivating illustrations displayed in Figures 5-6 present the cumulative distribution function (CDF) and the quantile-quantile (QQ) plot for the data both before and following the adjustments made for the specified parameter values. These visual representations vividly illustrate the transformations in the data's distribution, allowing for a clear comparison of the effects of the corrections applied.

Table 5: Metrics after applying bias-corrected MLE on the mentioned datasets

Metrics	Time between failures	Factors affecting unit capacity
nLL	-19.9277	-7.6076
AIC	-35.8553	-11.2153
CAIC	-35.2553	-10.6153
BIC	-33.5843	-8.9443
HQIC	-1.4131	0.5128
KS	0.1643	0.1536
AD	0.7135	1.9104
CVM	0.1371	0.2051
P(KS)	0.5072	0.4090
Inverse of K matrix= Var-cov matrix	$\begin{bmatrix} 0.0021 & -0.00064 \\ -0.00064 & 0.0032 \end{bmatrix}$	$\begin{bmatrix} 0.0027 & -0.00056 \\ -0.00056 & 0.007 \end{bmatrix}$

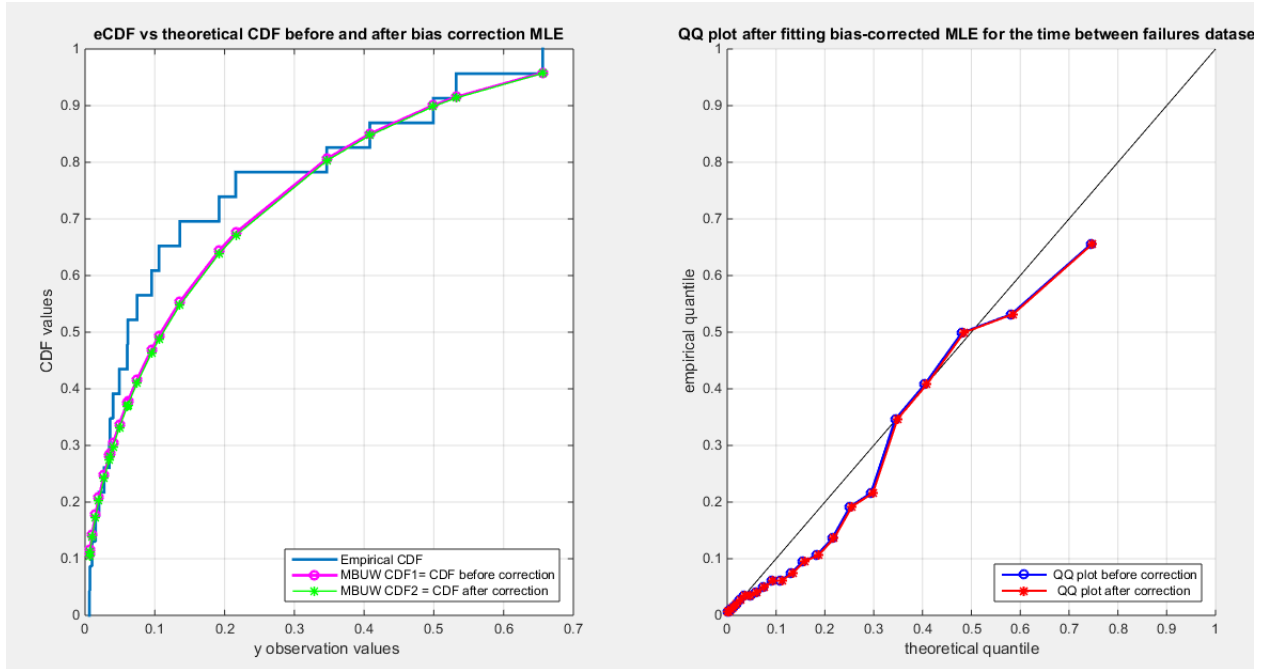


Fig. 5 shows the CDF on the right and QQ plot on the left for the time between failures dataset before and after correction

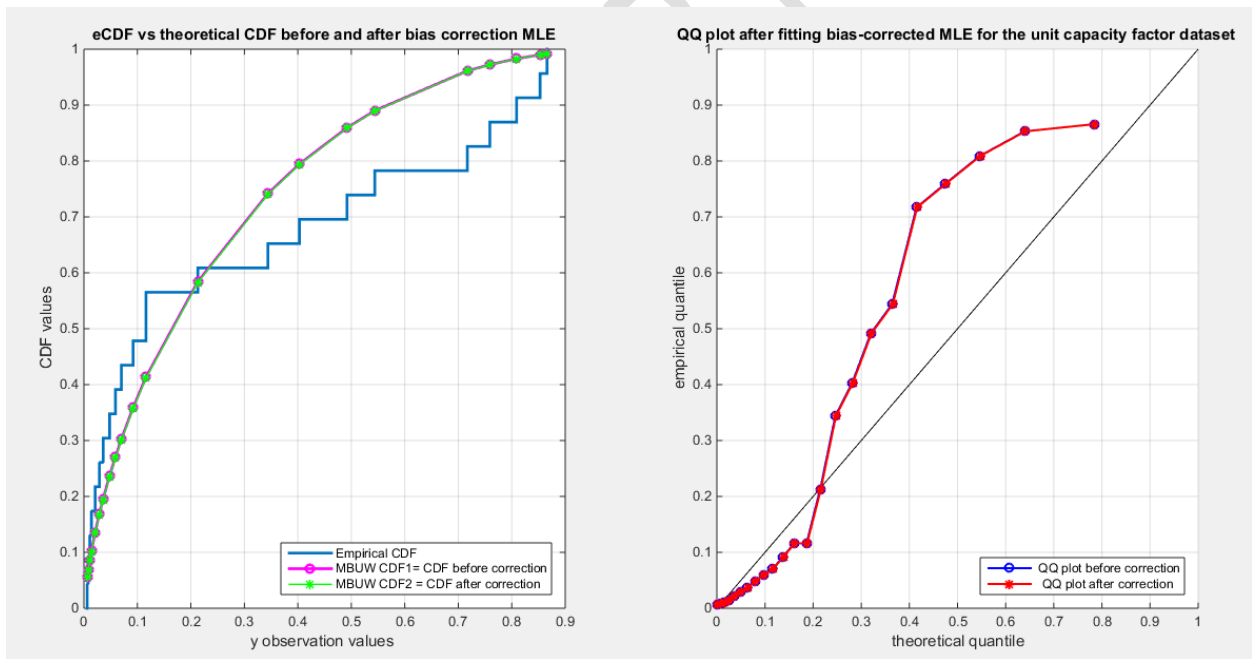


Fig. 6 shows the CDF on the right and QQ plot on the left for the factors affecting unit capacity dataset before and after correction

Table 4 presents a subtle yet noticeable decline in the metrics across all datasets following the correction. The variance-covariance matrix for each dataset reveals slight fluctuations in each parameter, accompanied by minimal negative co-variances. This indicates a delicate inverse relationship among the parameters, a relationship that was elucidated through the examination of the nLL surface in section 3.

Figures 5-6 illustrate the graphs for the time between failures dataset and the factors affecting unit capacity dataset exhibit almost no discernible differences. Intriguingly, the inverse of the information matrix, which serves as the variance-covariance matrix, is confirmed to be positive definite. In contrast, the variance-covariance matrices for the quality of support network, voter, and flood datasets are marked by non-positive definiteness. Nevertheless, the results of the KS test indicate a commendable fit of the distribution to these datasets, prompting the author to omit presenting further results.

The variance-corrected approach significantly diminishes both the variances and the correlation among the parameters, yielding a clearer understanding of the underlying data patterns. The author employed the finite difference method across all datasets, which astonishingly resulted in zero correlation among the parameters—particularly notable for those datasets exhibiting negative correlations, especially when applying reciprocal models (though these findings remain unpublished for now).

In contrast, the delta method proved to be even more effective, achieving a greater reduction in variance than that attained through the central finite difference method. This pursuit of uncorrelated estimators, paired with reduced variances, produced distributions adept at fitting various datasets. Conversely, the bias-corrected approach only achieved a positive definite variance-covariance matrix for three specific datasets, while the variance-corrected method consistently produced positive definite matrices across the board.

The bias-corrected technique is heavily reliant on the expectations of higher-order derivatives, introducing complexity in the integration process that can greatly influence the information matrix and, consequently, the variance-covariance matrix. Notably, the variance-corrected approach consistently outperformed traditional methods like the generalized method of moments (GMM) and percentile techniques previously employed by the author (Attia, 2025). While the alternative methods produced variances that were indeed smaller than those generated by the maximum likelihood estimation (MLE) using the Nelder-Mead optimization algorithm, it is noteworthy that these variances were still larger than the ones achieved through the variance-corrected MLE approach. This suggests that despite the improvements made by the latter methods, they do not quite match the precision of the variance-corrected MLE, which remains the most reliable option for minimizing variance in this context.

Remarkably, the implementation of the delta method led to a dramatic reduction in both variances and covariance, improving upon the figures achieved through GMMs and percentile methods. The variance-corrected procedure demonstrated a profound ability to diminish variance across the board. When the distribution aligns well with the dataset, the QQ plot reveals a discernible, consistent pattern, emphasizing the accuracy of the fitted values.

Furthermore, visualizing the negative log-likelihood (nLL) surface for each dataset proves invaluable, shedding light on the estimation process's outcomes, along with the uniqueness and identifiability of the parameters and their interrelations. The selection of estimated parameter values hinges on numerous pivotal factors. Foremost among them is the need for the chosen estimators to exhibit minimal variance. Simultaneously, a robust goodness-of-fit test should confirm a satisfactory alignment with the data. Metrics such as nLL, AIC, CAIC, BIC, and HQIC should ideally present the highest negative values. Stability is another critical characteristic; estimators should demonstrate resilience and minimal deviation from the initial estimates, particularly during iterative techniques, which is typically the case in practice.

5. Conclusion

The variance-corrected and bias-corrected approaches are essential techniques utilized after the initial estimation of MBUW parameters, which can be conducted using methods such as maximum likelihood estimation (MLE) with various optimizers tailored to the characteristics of the data, or by employing generalized methods of moments (GMMs) or percentile methods. These latter two methods specifically address the issue of parameter correlation, which is not considered in the initial estimations like (MLE).

In practical applications, a typical workflow begins with MLE, often using the Nelder-Mead optimization method or another suitable optimizer. However, if the characteristics of the data significantly influence the variance-covariance matrix, practitioners may transition to methods like the generalized method of moments or percentile methods that can further reduce the variance and the covariance though they may still exhibit relatively large variances.

Visualizing the negative log-likelihood (nLL) surface can be instrumental in identifying parameter values that lead to an optimal fit of the distribution to the data, aiding in the selection of initial parameter estimates. By leveraging this visualization, the variance-corrected and bias-corrected approaches can mitigate inflation in variance and decrease the correlation between parameters, resulting in more robust parameter estimates.

6. Future works

In the case of the datasets that demonstrate a variance-covariance matrix lacking positive definiteness, it is advisable to explore alternative numerical integration techniques or to employ Bayesian inference methods. These approaches can significantly enhance the accuracy and reliability of the estimation process when traditional methods fall short. Specifically, numerical integration techniques such as Monte Carlo integration or quadrature methods can effectively address issues arising from non-positive definiteness by providing a more robust framework for calculations. Meanwhile, Bayesian inference offers a powerful modeling perspective that allows the incorporation of prior information and uncertainty, which can lead to more informed and nuanced estimations despite the complications presented by the datasets in question.

Declarations:

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable

Availability of data and material

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Competing interests

The author declares no competing interests of any type.

References

- Attia, I. M. (2025). Median-Based Unit Weibull (MBUW): A New Variants of Generalized Method of Moments and Percentile Estimators. *Journal of Probability and Statistics*, 2025(1), 1503091. <https://doi.org/10.1155/jpas/1503091>
- Bartlett, M. S. (1953a). Approximate Confidence Intervals. *Biometrika*, 40(1/2), 12. <https://doi.org/10.2307/2333091>

- Bartlett, M. S. (1953b). Approximate Confidence Intervals.II. More than one Unknown Parameter. *Biometrika*, 40(3/4), 306. <https://doi.org/10.2307/2333349>
- Cordeiro, G. M., & Cribari-Neto, F. (2014). Analytical and Bootstrap Bias Corrections. In G. M. Cordeiro & F. Cribari-Neto, *An Introduction to Bartlett Correction and Bias Reduction* (pp. 69–99). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-55255-7_4
- Cordeiro, G. M., Da Rocha, E. C., Da Rocha, J. G. C., & Cribari-Neto, F. (1997). Bias-corrected maximum likelihood estimation for the beta distribution. *Journal of Statistical Computation and Simulation*, 58(1), 21–35. <https://doi.org/10.1080/00949659708811820>
- Cordeiro, G. M., & Klein, R. (1994). Bias correction in ARMA models. *Statistics & Probability Letters*, 19(3), 169–176. [https://doi.org/10.1016/0167-7152\(94\)90100-7](https://doi.org/10.1016/0167-7152(94)90100-7)
- Cox, D. R., & Snell, E. J. (1968). A General Definition of Residuals. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 30(2), 248–265. <https://doi.org/10.1111/j.2517-6161.1968.tb00724.x>
- Cribari-Neto, F., & Vasconcellos, K. L. P. (2002). Nearly Unbiased Maximum Likelihood Estimation for the Beta Distribution. *Journal of Statistical Computation and Simulation*, 72(2), 107–118. <https://doi.org/10.1080/00949650212144>
- Efron, B. (1982). *The Jackknife, the Bootstrap and other resampling plans* (Vol. 1–Vol. 38. Philadelphia, PA, USA: SIAM.). SIAM.
- Ferrari, S. L. P., & Cribari-Neto, F. (1998). On bootstrap and analytical bias corrections. *Economics Letters*, 58(1), 7–15. [https://doi.org/10.1016/S0165-1765\(97\)00276-0](https://doi.org/10.1016/S0165-1765(97)00276-0)
- Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, 80(1), 27–38. <https://doi.org/10.1093/biomet/80.1.27>
- Giles, D. E. (2012). A note on improved estimation for the Topp–Leone distribution. *Tech. Rep., Department of Economics, University of Victoria, Econometrics Working Papers*.
- Giles, D. E. (2012). Bias Reduction for the Maximum Likelihood Estimators of the Parameters in the Half-Logistic Distribution. *Communications in Statistics - Theory and Methods*, 41(2), 212–222. <https://doi.org/10.1080/03610926.2010.521278>
- Giles, D. E., Feng, H., & Godwin, R. T. (2013). On the Bias of the Maximum Likelihood Estimator for the Two-Parameter Lomax Distribution. *Communications in Statistics - Theory and Methods*, 42(11), 1934–1950. <https://doi.org/10.1080/03610926.2011.600506>
- Giles, D. E & H. Feng. (2009). Bias of the maximum likelihood estimators of the two-parameter Gamma distribution revisited. *Tech. Rep., Department of Economics, University of Victoria, Econometrics Working Papers*.
- Haldane, J. B. S., & Smith, S. M. (1956). The Sampling Distribution of a Maximum-Likelihood Estimate. *Biometrika*, 43(1/2), 96. <https://doi.org/10.2307/2333582>
- Haldane, J.B.S. (1953). The estimation of two parameters from a sample. *Sankhyā*, 12, 313-320.
- Lagos-Álvarez, B., M. D. Jiménez-Gamero, & V. Alba-Fernández. (2011). Bias correction in the Type I Generalized Logistic distribution. *Communications in Statistics - Simulation and Computation*, 40(4), 511–531.
- Lemonte, A. J. (2011). Improved point estimation for the Kumaraswamy distribution. *Journal of Statistical Computation and Simulation*, 81(12), 1971–1982. <https://doi.org/10.1080/00949655.2010.511621>
- Lemonte, A. J., Cribari-Neto, F., & Vasconcellos, K. L. P. (2007). Improved statistical inference for the two-parameter Birnbaum–Saunders distribution. *Computational Statistics & Data Analysis*, 51(9), 4656–4681. <https://doi.org/10.1016/j.csda.2006.08.016>
- Ling, X., & Giles, D. E. (2014). Bias Reduction for the Maximum Likelihood Estimator of the Parameters of the Generalized Rayleigh Family of Distributions. *Communications in Statistics - Theory and Methods*, 43(8), 1778–1792. <https://doi.org/10.1080/03610926.2012.675114>

- Maya, R., Jodrá, P., Irshad, M. R., & Krishna, A. (2024). The unit Muth distribution: Statistical properties and applications. *Ricerche Di Matematica*, 73(4), 1843–1866. <https://doi.org/10.1007/s11587-022-00703-7>
- Mazucheli, J., & Dey, S. (2018). Bias-corrected maximum likelihood estimation of the parameters of the generalized half-normal distribution. *Journal of Statistical Computation and Simulation*, 88(6), 1027–1038. <https://doi.org/10.1080/00949655.2017.1413649>
- Mazucheli, J., Menezes, A. F. B., & Dey, S. (2019). Bias-corrected maximum likelihood estimators of the parameters of the inverse Weibull distribution. *Communications in Statistics - Simulation and Computation*, 48(7), 2046–2055. <https://doi.org/10.1080/03610918.2018.1433838>
- Menezes, A., Mazucheli, J., Alqallaf, F., & Ghitany, M. E. (2021). Bias-Corrected Maximum Likelihood Estimators of the Parameters of the Unit-Weibull Distribution. *Austrian Journal of Statistics*, 50(3), 41–53. <https://doi.org/10.17713/ajs.v50i3.1023>
- Millar, R. (2011). *Maximum likelihood estimation and inference*. Chichester, West Sussex, United Kingdom: John Wiley & Sons, Ltd.
- Pawitan, Y. (2001). *In all likelihood: Statistical modelling and inference using likelihood*. Oxford: Oxford University Press.
- Rasekhi, M., دانشگاه ملایر, G. Hamedani, G., & مارکوت. (2019). Bias-corrected Maximum-Likelihood Estimator for the Parameter of the Logarithmic Series Distribution and its Characterizations. *Journal of Statistical Research of Iran*, 16(1), 59–72. <https://doi.org/10.52547/jsri.16.1.59>
- Reath, J. (2016). *Improved parameter estimation of the log-logistic distribution with applications* [PhD. thesis]. Michigan Technological University.
- Saha, K., & Paul, S. (2005). Bias-Corrected Maximum Likelihood Estimator of the Negative Binomial Dispersion Parameter. *Biometrics*, 61(1), 179–185. <https://doi.org/10.1111/j.0006-341X.2005.030833.x>
- Schwartz, J., & Giles, D. E. (2016). Bias-reduced maximum likelihood estimation of the zero-inflated Poisson distribution. *Communications in Statistics - Theory and Methods*, 45(2), 465–478. <https://doi.org/10.1080/03610926.2013.824590>
- Schwartz, J., Godwin, R. T., & Giles, D. E. (2013). Improved maximum-likelihood estimation of the shape parameter in the Nakagami distribution. *Journal of Statistical Computation and Simulation*, 83(3), 434–445. <https://doi.org/10.1080/00949655.2011.615316>
- Shenton, L. R., & Bowman, K. (1963). Higher Moments of a Maximum-Likelihood Estimate. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 25(2), 305–317. <https://doi.org/10.1111/j.2517-6161.1963.tb00511.x>
- Singh, A. K., Singh, A., & Murphy, D. J. (2015). On Bias Corrected Estimators of the Two Parameter Gamma Distribution. *2015 12th International Conference on Information Technology - New Generations*, 127–132. <https://doi.org/10.1109/ITNG.2015.151>
- suprawhardana. (1999). Suprawhardana, M.S., Prayoto, S.: Total time on test plot analysis for mechanical components of the RSG-GAS reactor. *At. Indones.* 25(2), 81–90 (1999). *25(2), 81-90(1999)*, 25(5), 81–90.
- Teimouri, M., & Nadarajah, S. (2013). Bias corrected MLEs for the Weibull distribution based on records. *Statistical Methodology*, 13, 12–24. <https://doi.org/10.1016/j.stamet.2013.01.001>
- Teimouri, M., & Nadarajah, S. (2016). Bias corrected MLEs under progressive type-II censoring scheme. *Journal of Statistical Computation and Simulation*, 86(14), 2714–2726. <https://doi.org/10.1080/00949655.2015.1123709>
- Wang, M., & Wang, W. (2017). Bias-Corrected maximum likelihood estimation of the parameters of the weighted Lindley distribution. *Communications in Statistics - Simulation and Computation*, 46(1), 530–545. <https://doi.org/10.1080/03610918.2014.970696>

Zhang, G., & Liu, R. (2017). Bias-corrected estimators of scalar skew normal. *Communications in Statistics - Simulation and Computation*, 46(2), 831–839.
<https://doi.org/10.1080/03610918.2014.980512>

Generally, the paper is good and well done, and a lot formula has been derived. So, the paper is eligible to be published

UNDER PEER REVIEW