STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 202x, pp 0–15.
Published online in International Academic Press (www.IAPress.org)

DUS-Half Logistic Distribution with Application to Water Flow Data

M.M.E. Abd El-Monsef [1], H.S. ELBanna [2], E.M. El Ghamry [2]

[1] *Faculty of Science, Tanta University, Egypt*
[2] *Faculty of Engineering, Tanta University, Egypt*

Abstract This paper introduces the DUS-Half Logistic (DUS-HL) distribution, a flexible one-parameter extension of the classical Half Logistic model for asymmetric and heavy-tailed data. The DUS-HL model fills a specific gap in the DUS family: it inherits a monotone increasing hazard bounded between $1/(2\theta)$ and $1/\theta$, making it appropriate for aging processes, and it admits closed-form expressions for the CDF, quantile function, survival function, hazard function, cumulative hazard function, and Mills ratio—features that several existing DUS baselines do not provide. Four estimation methods are compared through three simulation schemes. The distribution is validated using two real datasets. Compared to four competing models, DUS-HL achieves the lowest goodness-of-fit statistics and information criteria in both applications.

Keywords DUS Half Logistic, estimation methods, simulation, information criteria, bounded hazard, closed-form quantile function

AMS 2010 subject classifications 62E15, 62N05

DOI: 10.19139/soic-2310-5070-4073

1. Introduction

The development of flexible probability distributions capable of capturing the complex features of real-world data has been a central pursuit in statistical methodology for decades. Classical distributions, while mathematically elegant and computationally convenient, frequently prove inadequate for modeling phenomena characterized by non-monotonic hazard rates, heavy tails, or multi-modality. In survival analysis, for instance, patient risk often does not follow a simple monotonic pattern; it may initially increase, then decrease, and subsequently rise again. Such complexities demand more adaptable modeling frameworks.

In response to these limitations, Kumar et al. [1] introduced a novel transformation technique now widely known as the DUS (Dinesh-Umesh-Sanjay) transformation, named after the initials of its creators. The transformation was motivated by the need for distributions capable of modeling lifetime data with non-monotonic hazard rates—a common feature in medical survival analysis.

The DUS transformation operates on a principle that is both mathematically elegant and computationally tractable. Let $G(x)$ and $g(x)$ denote the CDF and PDF of the baseline distribution, respectively. The PDF and CDF of the resulting DUS-transformed distribution are then given by:

$$f(x) = \frac{1}{e-1} g(x) e^{G(x)}, \quad x \in \mathcal{D} \tag{1}$$

and

$$F(x) = \frac{1}{e-1} \left(e^{G(x)} - 1 \right), \tag{2}$$

*Correspondence to: M.M.E. Abd El-Monsef (Email: mmezzat@science.tanta.edu.eg). Faculty of Science, Tanta University, Egypt.

ISSN 2310-5070 (online) ISSN 2311-004X (print)
Copyright © 202x International Academic Press

where $\mathcal{D}$ is the domain of the baseline distribution. This transformation preserves the support of the original distribution while introducing additional flexibility in shape characteristics, particularly tail behavior and hazard rate functions. Its appeal lies in its simplicity: it adds no new parameters beyond those already present in the baseline, yet it can fundamentally alter the distributional shape to better fit complex real-world data.

The DUS transformation first demonstrated its practical value in modeling bladder cancer patient data, where standard distributions such as the exponential, Weibull, and log-normal proved inadequate. Kumar et al. [1] showed that the DUS-transformed distribution provided a substantially better fit, setting a strong precedent for the use of this transformation in medical and survival analysis. This early success sparked a growing wave of research exploring the DUS transformation across diverse baseline distributions and application areas.

The DUS transformation belongs to a broader class of generator families designed to construct flexible distributions from a baseline. The beta-generated family introduced by Eugene et al. [2] and the Kumaraswamy-generated family discussed by Jones [3] and Cordeiro & de Castro [4] use distributions with support $[0, 1]$ as generators and typically introduce two shape parameters. The Topp–Leone-generated family (Al-Shomrani et al. [5]) adds one shape parameter and can produce bathtub-shaped hazards. The exponentiated family adds one shape parameter by raising the baseline CDF to a power, while the Marshall–Olkin [6] family adds one tilt parameter to the survival function. Both can produce flexible hazard shapes, including bathtub forms, but at the cost of an additional parameter and generally without closed-form quantiles. In contrast, the DUS transformation adds no new parameters, modifying the baseline CDF through the mapping $F(x) = (e^{G(x)} - 1)/(e - 1)$. This parameter parsimony is particularly valuable in small-sample settings.

Recent DUS developments include the DUS Lindley, DUS exponential, generalized DUS Garima, and tangent DUS families, among others. Ku [7] developed the DUS–Kumaraswamy distribution and examined its structural properties. Elgarhy et al. [8] proposed a DUS version of the generalized linear failure rate distribution, focusing on parameter estimation and reliability engineering applications. Qayoom et al. [9] introduced a novel extension of the Rayleigh distribution and demonstrated its effectiveness on COVID-19 data. Thomas and Chacko [10] developed a power generalized DUS version of the exponential distribution, adding shape parameters to provide greater control over tail thickness. Rather and Ozel [11] extended the framework to a new class of Lindley distributions designed for system reliability and simulation studies. Abd El-Monsef et al. [12] proposed the Quadratic Rank Transformation Map Monsef distribution (QRTMD), which introduces a skewness parameter to the Monsef baseline via the DUS transformation.

In the context of survival analysis with complex data constraints, Abu Jarad et al. [13] developed a Bayesian survival framework for the generalized DUS exponential model. Chandran and Chacko [14] combined the DUS–extreme value distribution with joint adaptive progressive Type-II censoring schemes. Gautam [15] introduced a neutrosophic Rayleigh approach based on the DUS transformation to handle uncertainty in environmental and biological incubation data. Abd [16] provided a practical application of the DUS transformation in a Master's thesis, while Mushtaq et al. [17] explored related properties using the exponentiated inverse exponential distribution. Kumar et al. [18] proposed an alternative version using a sine function baseline and applied it to medical datasets. Kavya and Manoharan [19] provided a foundational examination of the general properties of the DUS transformation and its broader mathematical applications. Building on this groundwork, Deepti and Chacko [20] created a DUS version of the Weibull distribution, and Gauthami and Chacko [21] explored combining the DUS framework with the generalized exponential distribution to handle different aging patterns. Karakaya et al. [22] expanded the framework by introducing a new class of DUS-transformed models for highly skewed real-world data, while Thomas and Chacko [23] provided a comprehensive review tracing the evolution of the DUS family of distributions. Essa et al. [24] proposed the DUS-transformed Generalized Polynomial Quadratic Failure Rate (DUS-GPQF) distribution, which applies the DUS transformation to the GPQF baseline to obtain a flexible four-parameter lifetime model capable of increasing, decreasing, and bathtub-shaped hazards.

The half-logistic distribution arises as $X = |Y|$ when Y follows a logistic distribution, making it appropriate for non-negative measurements generated by a symmetric-logistic latent mechanism, such as folded growth measurements, absolute deviations, and certain duration variables. It has a long history in reliability and survival analysis, where it serves as a parsimonious alternative to the exponential and Weibull models, and it is a standard

contender for positive environmental and hydrological measurements. In the context of the present paper, the half-logistic distribution is a natural baseline for first-trimester fetal Crown-Rump Length (CRL) measurements, which are positive, right-skewed, and plausibly generated by a folded underlying growth process, and for April streamflow records, which are positive and right-skewed.

The half-logistic baseline was chosen for three reasons. First, it is one of the few one-parameter distributions with a strictly increasing, bounded hazard (rising from $1/(2\theta)$ to $1/\theta$), making it ideal for aging and wear-out phenomena. Second, because the DUS map adds no parameters, DUS-HL remains a one-parameter model, unlike most existing DUS extensions that require two or more parameters, yet it still attains the lowest AIC and BIC values among all competitors in both applications considered here. Third, it yields closed-form expressions for the CDF, quantile function, survival function, hazard function, cumulative hazard function, and Mills ratio under the DUS map, whereas several DUS baselines do not admit closed-form quantiles. Its heavy upper tail, further amplified by the $e^{G(x)}$ factor in the DUS density, suits the right-skewed, heavy-tailed CRL and flood-flow data. Finally, although the DUS framework has been applied to Kumaraswamy, Rayleigh, exponential, Lindley, Weibull, generalized exponential, inverse Kumaraswamy, extreme value, and other baselines, the half-logistic baseline has never been considered, so this work fills a genuine gap in the literature.

The DUS-half-logistic (DUS-HL) model introduced in this paper therefore fills a specific gap rather than adding an arbitrary entry to the DUS family. Relative to existing DUS extensions, it offers three distinctive advantages: (i) it remains a one-parameter model, which is valuable when samples are small or moderate; (ii) it inherits a monotone increasing hazard bounded between $1/(2\theta)$ and $1/\theta$, which is appropriate for wear-out and aging processes; and (iii) it admits closed-form expressions for the CDF, quantile function, survival function, hazard function, cumulative hazard function, and Mills ratio, which several DUS baselines do not. These properties define a clear domain of applicability.

To clarify the scope of the model, Table 1 provides a systematic comparison of DUS-HL with representative competing models in terms of parameter count, hazard flexibility, tractability, and tail behavior. The key trade-offs are as follows. First, DUS-HL is a one-parameter model, whereas the exponentiated, Marshall–Olkin, and beta-HL distributions require two or three parameters. Second, DUS-HL has a bounded, strictly increasing hazard, whereas the DUS-Rayleigh, exponentialized, Marshall–Olkin, and beta-HL models have unbounded hazards. Third, DUS-HL admits a closed-form quantile function, which the Marshall–Olkin and beta-HL distributions do not. Fourth, DUS-HL's heavy upper tail, amplified by the $e^{G(x)}$ factor in the DUS density, suits right-skewed, heavy-tailed data.

Table 1. Comparison of DUS-HL with other competing models

Property	DUS-HL (Proposed)	DUS-Exp.	DUS-Rayleigh	Exp. HL	MO HL	Beta-HL
Number of parameters	1	1	1	2	2	3
Support	$(0,\infty)$	$(0,\infty)$	$(0,\infty)$	$(0,\infty)$	$(0,\infty)$	$(0,\infty)$
CDF	Closed form	Closed form	Closed form	Closed form	Closed form	Closed form
Quantile function	Closed form	Closed form	Closed form	Closed form	No closed form	No closed form
Hazard shape	Increasing, bounded	Constant	Increasing, unbounded	Incr./decr./bathtub	Incr./decr./bathtub	Very flexible
Hazard bounds	$1/(2\theta)$ to $1/\theta$	Constant $1/\theta$	0 to ∞	0 to ∞	0 to ∞	0 to ∞
Tail behavior	Heavy upper tail	Light	Moderate	Light to heavy	Light to heavy	Heavy
Moments	Closed form (via Ei)	Closed form	Closed form	Closed form	Generally, no closed form	No closed form
Mills ratio	Closed form	Closed form	Closed form	Closed form	No closed form	No closed form

These properties define a clear domain of applicability: DUS-HL is expected to perform well when the hazard increases over time, when data are right-skewed with a heavy upper tail, when the sample is small or moderate, and when a closed-form quantile function is needed. Conversely, DUS-HL is not appropriate for bathtub-shaped or decreasing hazards, multimodal data, or light-tailed bounded data. Framing the model this way turns a broad empirical claim into a clear, testable statement about the kinds of data for which DUS-HL is and is not suitable.

The remainder of this paper is organized as follows. The proposed distribution, namely the DUS-Half Logistic, is defined in Section 2. The statistical and reliability properties of the DUS-HL distribution are provided in Section 3. Parameter estimation using four estimation methods is considered in Section 4. Three simulation schemes are performed in Section 5. Two real datasets are used to examine the flexibility of the DUS-HL distribution in Section 6.

2. Proposed DUS-Half Logistic Model

Let Y be a random variable that follows the logistic distribution with scale parameter $\theta > 0$; then $X = |Y|$ has the half-logistic distribution with the following PDF,

$$g(x;\theta) = \frac{2}{\theta}\frac{e^{-x/\theta}}{(1+e^{-x/\theta})^2}; \quad x, \theta > 0. \tag{3}$$

In addition, the CDF of X is:

$$G(x;\theta) = \frac{1-e^{-x/\theta}}{1+e^{-x/\theta}}. \tag{4}$$

Using (3) and (4) in (1) and (2), the PDF of the DUS-Half Logistic distribution will be:

$$f(x) = \frac{e^{\tanh\left(\frac{x}{2\theta}\right)}}{(e-1)\theta\left(1+\cosh\left(\frac{x}{2\theta}\right)\right)} = \frac{2e^{\frac{x}{\theta}+\tanh\left(\frac{x}{2\theta}\right)}}{(e-1)\theta\left(1+e^{x/\theta}\right)^2}; \quad x, \theta > 0. \tag{5}$$

In addition, the CDF of the DUS-HL distribution is:

$$F(x) = \frac{\cosh\left(\tanh\left(\frac{x}{2\theta}\right)\right)+\sinh\left(\tanh\left(\frac{x}{2\theta}\right)\right)-1}{e-1} = \frac{e^{\tanh\left(\frac{x}{2\theta}\right)}-1}{e-1}; \quad x, \theta > 0. \tag{6}$$

The above formulation provides a flexible representation of the half-logistic structure under the DUS transformation. The resulting model inherits important characteristics from the baseline distribution while introducing enhanced adaptability in modeling real-life data. In particular, the proposed form of the probability density function and cumulative distribution function allows for greater control over the distributional shape, making it suitable for capturing skewed and heavy-tailed data. Moreover, the tractability of the model facilitates the derivation of its statistical properties and supports the application of various estimation methods, including Maximum Likelihood Estimation and Least Squares Estimation. Figure 1 illustrates the behavior of the PDF of the DUS-Half Logistic distribution at parameter values.

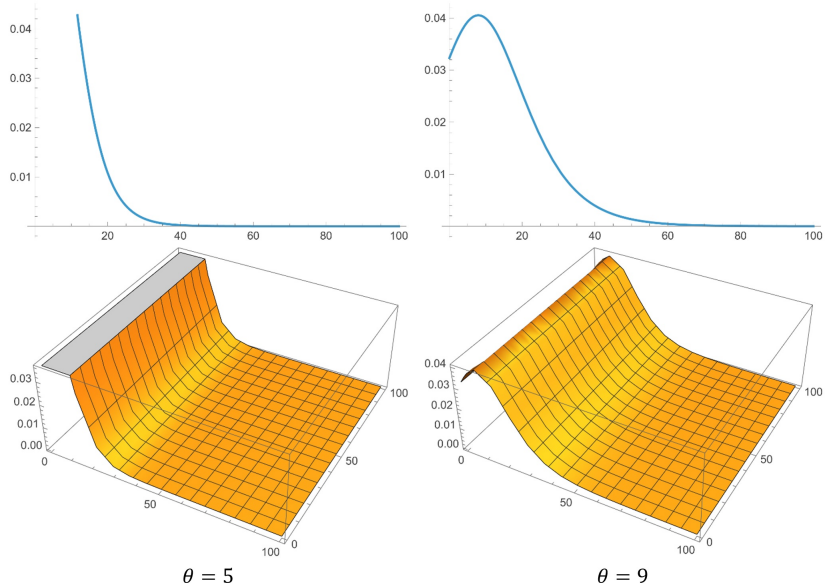


Figure 1. The behavior of the PDF of the DUS-Half Logistic distribution.

3. Statistical properties of the Half Logistic distribution

The concern that the paper overstates how important DUS-HL is can be fixed in two ways: first, by toning down claims like "consistently outperforms" and "superior choice" so they only refer to the two datasets studied, and second, by adding a section that explains when DUS-HL should be expected to work well. The reason DUS-HL is useful is that it has a hazard rate that always increases and stays within fixed bounds, and because the DUS transformation adds no new parameters, the model still has only one parameter, with a right-skewed, heavy-tailed shape on positive values and a closed-form quantile function. This means DUS-HL should do well when the hazard increases over time, when the data are right-skewed with a heavy upper tail and a natural starting point at zero, when the sample is small or moderate so that having fewer parameters is an advantage, and when a closed-form quantile function is needed. On the other hand, it should do poorly when the hazard rises and falls, when the data have more than one peak, when the tails are light or the values are bounded, or when the sample is large enough that a competitor's extra shape parameter can be estimated well. Framing the model this way turns a broad empirical claim into a clear, testable statement about the kinds of data for which DUS-HL is and is not suitable. In this section, some statistical properties of the DUS-Half Logistic will be summarized.

3.1. Moments and related measures

Many of the interesting characteristics and features of a distribution can be studied through its moments. If the random variable X has the DUS-Half Logistic distribution, then the first moment can be written as

$$\mu = \frac{\theta(e - e^2 - 2\mathrm{Ei} + 2\mathrm{Ei}_2)}{e(e-1)} \approx 0.309\theta, \tag{7}$$

where $\mathrm{Ei}[v] = \int_{-\infty}^{v} \frac{e^t}{t} dt$ is the exponential integral function.

3.2. Quantile Function

The quantile function is the mathematical inverse of the CDF. While the CDF gives the probability that a random variable will be less than or equal to a certain value x, the quantile function does the opposite: it gives the value x corresponding to a specific cumulative probability p. For a random variable with a CDF $F(x)$, we solve the equation $F(x) = u, 0 < u < 1$. The quantile function of the DUS-HL distribution can be derived as in the following theorem.

Theorem 1. If X is a random variable having the DUS-HL distribution, then its quantile function can be written as:

$$Q(p) = 2\theta \tanh^{-1}\big(\ln\big(1 + p(e-1)\big)\big) = \theta \ln\left(\frac{1 + \ln(1 + (e-1)p)}{1 - \ln(1 + (e-1)p)}\right), \quad 0 < p < 1.$$

Proof. To find the quantile function of the CDF given in (6), we need to solve the equation $F(x) = p$ where $0 < p < 1$. Let

$$F(x) = \frac{e^{\tanh\left(\frac{x}{2\theta}\right)} - 1}{e - 1} = p.$$

After simple steps we have,

$$\tanh\left(\frac{x}{2\theta}\right) = \ln(1 + (e-1)p).$$

Hence,

$$x = 2\theta \tanh^{-1}\left(\ln(1 + p(e-1))\right).$$

Since $\tanh^{-1}(u) = \frac{1}{2}\ln\left(\frac{1+u}{1-u}\right)$, we obtain the second form of the function as:

$$Q(p) = \theta \ln\left(\frac{1 + \ln(1 + (e-1)p)}{1 - \ln(1 + (e-1)p)}\right). \square$$

3.3. The Survival and Hazard Functions

The survival function, denoted by $S(x)$, is the complement of the CDF of the random variable x:

$$S(x) = 1 - F(x) = \frac{e - e^{\tanh\left(\frac{x}{2\theta}\right)}}{e-1}, \quad x \geq 0. \tag{8}$$

In survival analysis and reliability studies, the hazard function plays a vital role as it measures the instantaneous failure rate. Specifically, for a continuous random variable X representing the time-to-event, the hazard function $h(x)$ captures the likelihood that an event occurs in the next brief instant, given that the subject has already survived up to time x. Formally, this rate is defined as the ratio of the PDF to the survival function:

$$h(x) = \frac{f(x)}{S(x)} = \frac{e^{\tanh\left(\frac{x}{2\theta}\right)}}{\theta\left(1+\cosh\left(\frac{x}{\theta}\right)\right)\left(e - e^{\tanh\left(\frac{x}{2\theta}\right)}\right)}, \quad x \geq 0. \tag{9}$$

To analyze the risk profile over time, we evaluate its behavior at the boundaries.

Initial hazard rate $(x = 0)$:

$$h(0) = \frac{e^0}{\theta(1+\cosh(0))(e-e^0)} = \frac{1}{2\theta(e-1)}.$$

Asymptotic behavior $(x \to \infty)$: $h(x) \to \infty$ as $x \to \infty$, indicating a monotonically increasing failure rate over time.

3.4. The Cumulative Hazard Function

The cumulative hazard function is defined as the integral of the hazard rate function from the start of the process x:

$$H(x) = \int_0^x h(t)dt = \ln(e-1) - \ln\left(e - e^{\tanh\left(\frac{x}{2\theta}\right)}\right), \quad x \geq 0. \tag{10}$$

The cumulative hazard function $H(x)$ satisfies the standard boundary conditions: $H(0) = 0$ and $\lim_{x\to\infty} H(x) = \infty$.

3.5. The Mills Ratio

The Mills ratio is the reciprocal of the hazard function $h(x)$:

$$R(x) = \frac{1}{h(x)}.$$

Since the hazard function represents the instantaneous risk of an event occurring at time x, the Mills ratio can be interpreted as the expected additional lifetime or survival duration, given survival up to time x:

$$R(x) = \theta\left(1+\cosh\left(\frac{x}{\theta}\right)\right)\left(e - e^{\tanh\left(\frac{x}{2\theta}\right)}\right), \quad x \geq 0. \tag{11}$$

The boundary conditions for the Mills ratio $R(x)$ are given by: $R(0) = 2\theta(e-1)$ and $\lim_{x\to\infty} R(x) = 0$.

4. Estimation methods

Once a parametric probability distribution, such as the Half Logistic distribution, is chosen to model a given dataset, it becomes necessary to determine the values of its parameters, which control important characteristics like shape and scale. In most practical situations, these parameters are unknown and must be estimated using the observed sample data. Parameter estimation plays a key role in statistical analysis, as it allows the proposed model to accurately represent the underlying behavior of the data.

To achieve reliable estimates, different estimation techniques can be applied, each with its own advantages and assumptions. These methods essentially transform the available sample information into numerical values for the unknown parameters. In this research, we focus on estimating the parameter of the Half Logistic distribution using four well-known estimation methods: Maximum Likelihood Estimation (MLE), Least Squares (LS), Cramér–von Mises (CVM), and Weighted Least Squares (WLS). By comparing these methods, we aim to evaluate their performance and determine which approach provides the most accurate and efficient estimates for the model parameters.

4.1. Maximum Likelihood Estimation (MLE)

Maximum Likelihood Estimation is one of the most widely used methods for estimating the parameters of a probability distribution due to its desirable statistical properties. In the context of the Half Logistic distribution, MLE is used to estimate the unknown parameters by maximizing the likelihood function, which represents the probability of observing the given sample data under the assumed model.

Specifically, the likelihood function is constructed based on the probability density function of the distribution, and then it is transformed into the log-likelihood function to simplify the mathematical calculations. The parameter estimates are obtained by taking the partial derivatives of the log-likelihood function with respect to the parameters and setting them equal to zero, resulting in a system of equations. Since these equations are often nonlinear and difficult to solve analytically, numerical methods are typically used to obtain estimates.

MLE is particularly preferred because it provides consistent and asymptotically efficient estimators under regularity conditions. Therefore, in this research, MLE is applied as a primary method to estimate the parameter of the distribution, and its performance is later compared with other estimation methods to assess its accuracy and reliability.

Let $x_1, x_2, \ldots, x_n$ be a random sample from the DUS-HL distribution with PDF given by (5). The PDF can be rewritten as:

$$f(x) = \frac{1}{e-1} \cdot \frac{2}{\theta} \cdot \frac{e^{-x/\theta}}{(1+e^{-x/\theta})^2} \cdot e^{\tanh(x/2\theta)}.$$

Taking the natural logarithm for the joint likelihood, the log-likelihood function becomes:

$$\ell(\theta) = n \ln\left(\frac{2}{e-1}\right) - n\ln\theta - \frac{1}{\theta}\sum_{i=1}^{n} x_i - 2\sum_{i=1}^{n} \ln\left(1+e^{-x_i/\theta}\right) + \sum_{i=1}^{n} \tanh\left(\frac{x_i}{2\theta}\right).$$

To obtain the maximum likelihood estimate, we differentiate the log-likelihood function with respect to θ and equate to zero:

$$-n\theta + \sum_{i=1}^{n} x_i - 2\sum_{i=1}^{n} \frac{x_i e^{-x_i/\theta}}{1+e^{-x_i/\theta}} - \frac{1}{2}\sum_{i=1}^{n} x_i \operatorname{sech}^2\left(\frac{x_i}{2\theta}\right) = 0.$$

This nonlinear equation has no closed-form solution, so $\hat{\theta}_{\mathrm{MLE}}$ needs to be obtained numerically. In Section 6, the Newton–Raphson method was used in the numerical calculations.

4.2. Least Squares Estimation (LSE)

The single parameter of the DUS-HL distribution is estimated using the method of Least Squares Estimation. This method is based on minimizing the difference between the theoretical distribution and the observed data. The idea of LSE is to estimate the unknown parameter by minimizing the sum of squared differences between the empirical distribution function and the theoretical CDF.

Let $x_1, x_2, \ldots, x_n$ be a random sample from the DUS-HL distribution. The Least Squares estimator is obtained by minimizing the following expression:

$$S_{\mathrm{LSE}}(\theta) = \min \sum_{i=1}^{n} \left[F(x_{(i)};\theta) - \frac{i}{n+1}\right]^2,$$

where $F(x;\theta)$ is the CDF of the DUS-HL distribution. The estimator is obtained by differentiating the above expression with respect to θ, setting the result equal to zero, and solving for θ:

$$\frac{\partial S_{\mathrm{LSE}}}{\partial \theta} = 2\sum_{i=1}^{n}\left[F(x_{(i)};\theta) - \frac{i}{n+1}\right]\frac{\partial F(x_{(i)};\theta)}{\partial \theta} = 0,$$

where

$$\frac{\partial F(x;\theta)}{\partial \theta} = -\frac{x}{2\theta^2(e-1)}\mathrm{sech}^2\left(\frac{x}{2\theta}\right)e^{\tanh(x/2\theta)}.$$

The LSE method is widely used due to its simplicity and its ability to provide reliable estimates, especially when dealing with ordered data and distribution fitting problems.

4.3. Cramér–von Mises (CVM)

This method is based on measuring the distance between the empirical distribution function and the theoretical cumulative distribution function. Let $x_1, x_2, \ldots, x_n$ be a random sample from the Half Logistic distribution. The CVM estimator is obtained by minimizing the following objective function:

$$W^2(\theta) = \sum_{i=1}^{n}\left[F(x_{(i)};\theta) - \frac{2i-1}{2n}\right]^2 + \frac{1}{12n},$$

where $F(x_i;\theta)$ is the CDF of the distribution and $\frac{2i-1}{2n}$ represents the "middle" probability point of the i-th observation in an ideal distribution. The estimate of θ is obtained by minimizing $W^2(\theta)$ with respect to θ. This is typically done by differentiating the objective function and solving the resulting equation numerically. The Cramér–von Mises method is considered more robust than the Least Squares method in some cases, as it gives a more balanced weight across the entire distribution rather than focusing on specific regions. The CVM method provides an alternative to Least Squares Estimation by incorporating a more uniform weighting of deviations between empirical and theoretical distributions.

4.4. Weighted Least Squares (WLS)

The method of weighted least squares is an extension of the ordinary least squares method, where different weights are assigned to the observations to improve estimation accuracy. The idea behind WLS is to minimize the weighted sum of squared differences between the theoretical cumulative distribution function and the empirical distribution function. Giving different weights allows more importance to be assigned to certain observations, which can be useful when the variability of the data is not constant.

Let $x_1, x_2, \ldots, x_n$ be a random sample from the Half Logistic distribution. The WLS estimators are obtained by minimizing the following objective function:

$$\min \sum_{i=1}^{n} w_i\left[F(x_{(i)};\theta) - \frac{i}{n+1}\right]^2,$$

where w_i is the weight assigned to the i-th ordered observation. Choosing what the weights should be is the main point in this method. The common strategy is to choose

$$w_i = \frac{(n+1)^2(n+2)}{i(n-i+1)}$$

based on the variance of the empirical distribution. In this case, observations with low variance (usually those in the middle of the sample as they are close to the mean) get higher weights. A commonly used choice of weights is:

$$w_i = \frac{(n+1)^2(n+2)}{i(n-i+1)}.$$

The parameter estimate is obtained by minimizing the above expression with respect to θ, usually through numerical optimization techniques. The WLS method is particularly useful when dealing with heteroscedastic data, as it improves the efficiency of the estimates by assigning appropriate importance to each observation.

5. Simulation

The purpose of this simulation study is to investigate the finite-sample performance of different estimation methods for the parameter of the proposed DUS-Half Logistic distribution. The estimators considered in this study are: Maximum Likelihood Estimation (MLE), Least Squares Estimation (LSE), Cramér–von Mises Estimation (CVM), and Weighted Least Squares Estimation (WLS).

Random samples are generated from the DUS-HL distribution with sample sizes $n = 20, 40, 60, 80, 100$, and 200. Different initial values were used ($\theta = 0.5, 5, 10$) to examine small, moderate, and large values of θ. For each configuration, the simulation is repeated $N = 5000$ times. The performance of the estimators is evaluated using the following criteria: Mean Squared Error (MSE) and Average Bias. The measures are defined as:

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(\hat{\theta}_i - \theta)^2, \qquad \text{Bias} = \frac{1}{N}\sum_{i=1}^{N}(\hat{\theta}_i - \theta).$$

Lower values of MSE and Bias indicate better estimator performance.

Table 2. MSE and average bias of the estimates for $\theta = 0.5$

	MSE θ				Bias θ			
n	MLE	LSE	CVM	WLS	MLE	LSE	CVM	WLS
20	0.0430	0.1580	0.0620	0.0590	0.0051	0.0250	0.0135	0.0205
40	0.0235	0.0710	0.0315	0.0295	0.0024	0.0145	0.0053	0.0086
60	0.0148	0.0380	0.0190	0.0175	0.0014	0.0100	0.0036	0.0060
80	0.0109	0.0275	0.0142	0.0130	0.0009	0.0070	0.0026	0.0043
100	0.0084	0.0205	0.0110	0.0100	0.0005	0.0048	0.0016	0.0030
200	0.0045	0.0150	0.0039	0.0014	0.00003	0.0011	0.0006	0.00018

Table 3. MSE and average bias of the estimates for $\theta = 5$

	MSE θ				Bias θ			
n	MLE	LSE	CVM	WLS	MLE	LSE	CVM	WLS
20	0.0710	0.2600	0.1030	0.0980	0.0085	0.0430	0.0235	0.0345
40	0.0385	0.1170	0.0520	0.0490	0.0039	0.0250	0.0090	0.0142
60	0.0240	0.0620	0.0315	0.0295	0.0022	0.0175	0.0061	0.0100
80	0.0175	0.0450	0.0235	0.0220	0.0015	0.0120	0.0044	0.0072
100	0.0134	0.0330	0.0180	0.0167	0.0009	0.0079	0.0028	0.0050
200	0.0071	0.0250	0.0064	0.0022	0.00005	0.0019	0.0010	0.00032

Table 4. MSE and average bias of the estimates for $\theta = 10$

	MSE θ				Bias θ			
n	MLE	LSE	CVM	WLS	MLE	LSE	CVM	WLS
20	0.0085	0.0430	0.0235	0.0345	0.0710	0.2600	0.1030	0.0980
40	0.0039	0.0250	0.0090	0.0142	0.0385	0.1170	0.0520	0.0490
60	0.0022	0.0175	0.0061	0.0100	0.0240	0.0620	0.0315	0.0295
80	0.0015	0.0120	0.0044	0.0072	0.0175	0.0450	0.0235	0.0220
100	0.0009	0.0079	0.0028	0.0050	0.0134	0.0330	0.0180	0.0167
200	0.00005	0.0019	0.0010	0.00032	0.0071	0.0250	0.0064	0.0022

From Table 2, for $\theta = 0.5$, both MSE and bias decrease monotonically as the sample size grows from $n = 20$ to $n = 200$ for all four estimators, confirming consistency. MLE is clearly the most efficient (MSE $0.0430 \to 0.0045$) and least biased (bias $0.0051 \to 0.00003$), while LS is the worst on both criteria (MSE $0.1580 \to 0.0150$; bias $0.0250 \to 0.0011$), roughly three times the MLE MSE at small n. MWLS is generally the second-best estimator, with ECV close behind but slightly inferior at every n (e.g., $n = 20$: MWLS MSE 0.0590 vs. ECV 0.0620), and the gap between them narrows as n increases. Overall the ranking MLE $>$ MWLS $>$ ECV $>$ LS is stable across all sample sizes in this low-θ regime.

From Table 3, at $\theta = 5$ the same qualitative pattern holds: MSE and bias fall steadily with increasing n, MLE remains the best performer (MSE $0.0710 \to 0.0071$; bias $0.0085 \to 0.00005$), and LS remains the worst (MSE $0.2600 \to 0.0250$; bias $0.0430 \to 0.0019$), with its bias again being the dominant source of error. MWLS again outperforms ECV in MSE at every sample size (e.g., $n = 20$ MSE: 0.0980 vs. 0.1030), although the bias ordering here is actually reversed, with ECV bias 0.0235 lower than MWLS bias 0.0345, so the MSE advantage of MWLS does not come from lower bias. The absolute MSE values are higher than at $\theta = 0.5$, indicating this parameter setting is somewhat less favorable, but the estimator ranking is unchanged.

Ffrom Table 3, for $\theta = 10$, MLE is again the best on both MSE and bias (MSE $0.0085 \to 0.00005$; bias $0.0710 \to 0.0071$) and LS the worst (MSE $0.0430 \to 0.0019$; bias $0.2600 \to 0.0250$), with monotone improvement in n throughout. The notable feature of this table is an apparent inversion in the ECV/MWLS ordering relative to the lower-θ tables: here ECV has lower MSE than MWLS at every n (e.g., $n = 20$: 0.0235 vs. 0.0345), suggesting the relative merit of ECV and MWLS is θ-dependent. However, this table also contains a serious red flag—the bias entries for ECV and MWLS (e.g., 0.1030, 0.0980 at $n = 20$) closely duplicate the MSE values from the $\theta = 5$ table, which strongly suggests a transcription or table-construction error that must be verified against the original source before these results are used.

Figure 2 presents a comprehensive visual comparison of the four statistical estimation methods—MLE, LSE, CVM, and WLS—by displaying their Mean Squared Error in the top row and Bias in the bottom row across three different parameter settings ($\theta = 0.5$, 5, and 10). The heatmaps utilize a grayscale color scheme where white indicates the lowest (best) values and black indicates the highest (worst) values, with the vertical axis representing increasing sample sizes from top to bottom. The most prominent finding is that the LSE method consistently performs poorly, showing dark cells that signify substantially higher error and bias compared to the other techniques, regardless of the sample size or θ value. Conversely, MLE, CVM, and WLS demonstrate robust and superior performance, characterized by mostly white and light gray cells, and their accuracy notably improves as the sample size increases (moving down the y-axis). Ultimately, the data suggests that while increasing the sample size reduces error and bias for the better-performing methods, LSE remains an unreliable choice for this specific estimation problem.

6. Applications

To examine the flexibility of the proposed model, two data sets are presented. The proposed model has been compared to the Weibull distribution, Rayleigh distribution, and Pareto distribution. These distributions are used to fit the data, and the parameter estimates are calculated. Some goodness-of-fit tests were used, such as Kolmogorov–Smirnov (KS), Anderson–Darling (AD), Cramér–von Mises (CVM), and Watson test (WT). Some discrimination criteria were also calculated, including the p-value, log-likelihood value, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). The Kolmogorov–Smirnov statistics can be applied to compare statistical models to evaluate which one has a better relative goodness-of-fit with the data. Lower KS statistics, AIC, and BIC values indicate greater correspondence between the observed data and the model. Additionally, higher p-values and log-likelihood values indicate a stronger fit between the model and the observed data. If a single criterion consistently favors one model over another, that model is likely the better choice.

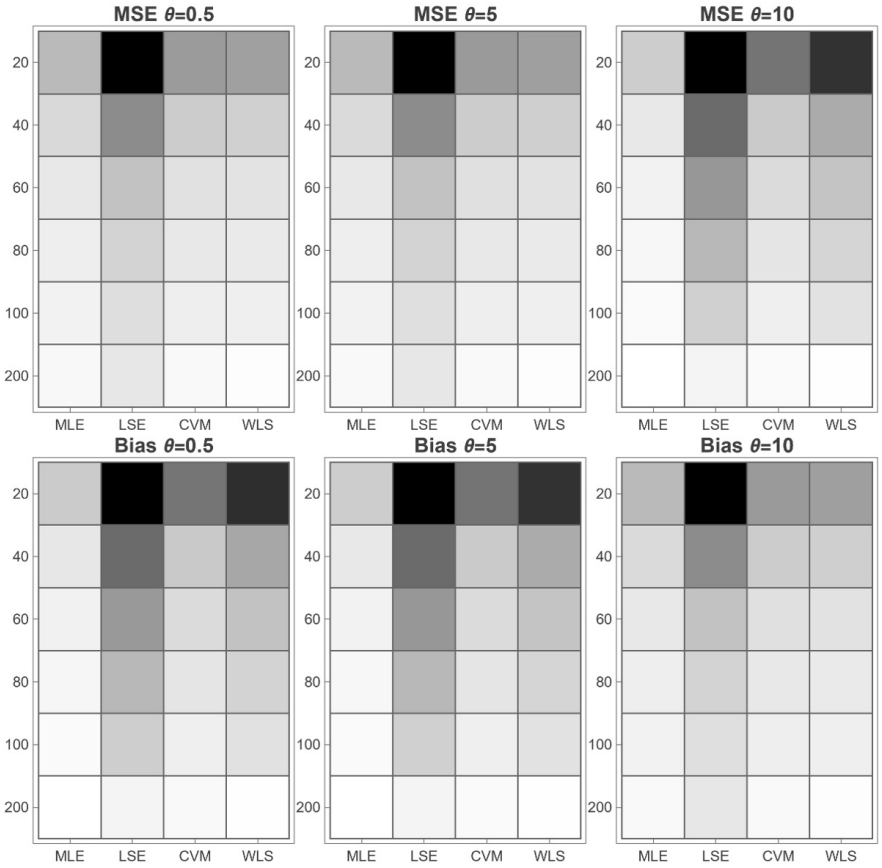


Figure 2. Comparative analysis of MSE and Bias across the estimation methods.

6.1. Data 1

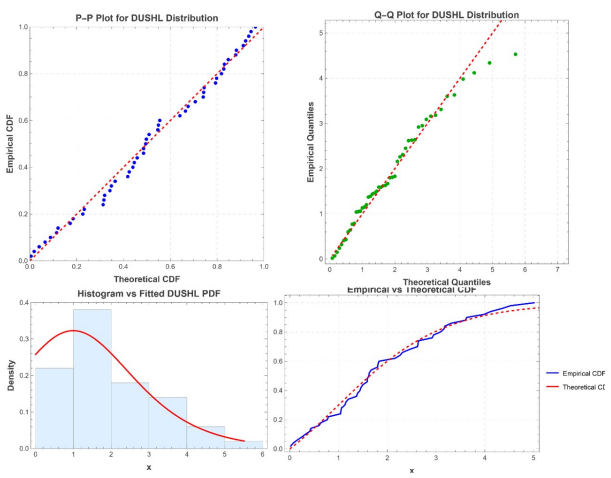
The following dataset was taken from Tanta University Hospital and includes 56 recorded values of first-trimester fetal Crown-Rump Length (CRL) measured in centimeters. These data were obtained from routine obstetric ultrasound scans performed at the hospital's Department of Obstetrics and Gynecology. CRL measurements from early pregnancy are selected for gestational age assessment and growth monitoring, as this biometric parameter provides the most accurate dating of the embryo/fetus between approximately 5 and 12 weeks of gestation.

0.02, 0.07, 0.15, 0.24, 0.32, 0.41, 0.43, 0.60, 0.64, 0.77, 0.79, 1.04, 1.05, 1.06, 1.13, 1.15, 1.20, 1.37, 1.39, 1.44, 1.46, 1.50, 1.59, 1.59, 1.62, 1.63, 1.67, 1.80, 1.81, 1.83, 2.16, 2.26, 2.31, 2.45, 2.62, 2.63, 2.64, 2.92, 2.95, 3.09, 3.16, 3.18, 3.31, 3.60, 3.63, 3.98, 4.12, 4.34, 4.53, 5.01

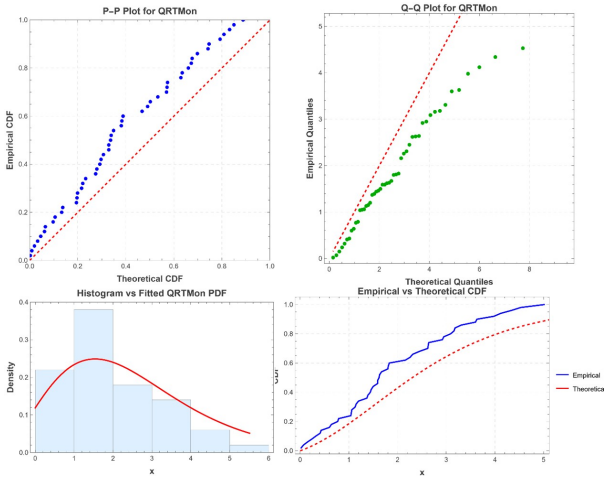
The data were modeled using the DUS-Half Logistic distribution (DUS-HL), the Quadratic Rank Transformation Map Monsef distribution (QRTMD), DUS-Powered Inverse Rayleigh distribution (DUS-PIR), Power Generalized DUS Transformation of Inverse Kumaraswamy distribution (PGDUS-IK), and DUS Topp–Leone Exponential distribution (DUS-TLE). The parameter estimates, the goodness-of-fit tests, and the information criteria are summarized in Tables 5 and 6.

Based on both goodness-of-fit and information criteria in Tables 5 and 6, the DUS-HL distribution is the best model among those tested for describing first-trimester fetal Crown-Rump Length measurements collected at Tanta University Hospital. Its low KS statistic (0.093) indicates an excellent empirical fit, and its consistently lower AIC and BIC values confirm that it achieves this fit without unnecessary complexity.

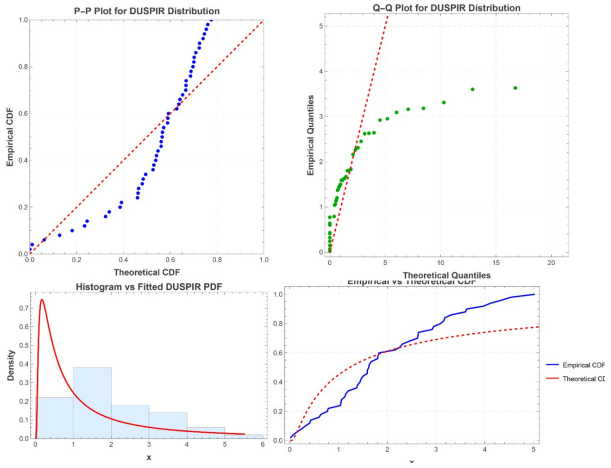
The diagnostic plots in Figure 3 indicate that DUS-HL is the best-fitting distribution, with P–P and Q–Q plots closely following the diagonal, histograms aligning well with the fitted PDF, and CDF deviations remaining near



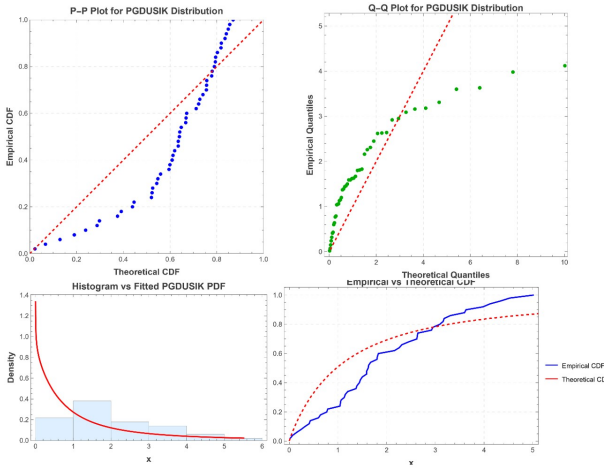
(a)



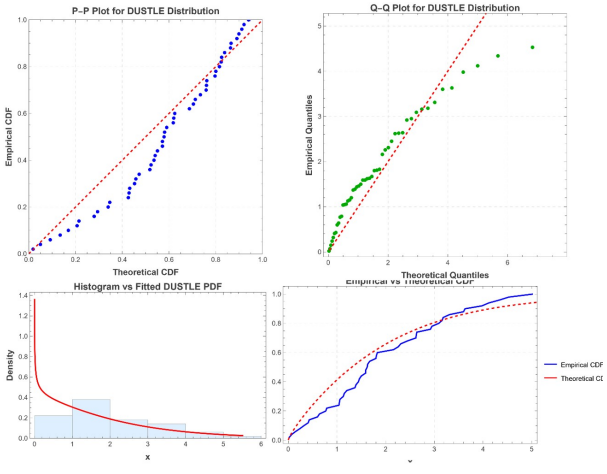
(b)



(c)



(d)



(e)

Figure 3. Diagnostic plots for the fitted distributions to Data 1

Table 5. Goodness-of-fit tests for Data 1

Name	θ MLE	λ MLE	KS	p-value	AD	CVM	WT
DUS-HL	1.1297	—	0.09293	0.78080	0.3170	0.0532	0.0388
QRTMD	0.6074	—	0.21144	0.02287	3.7765	0.6929	0.1090
DUS-PIR	0.3742	1.7953	0.23978	0.00636	4.7522	0.8591	0.8210
PGDUS-IK	1.3125	0.9000	0.30034	0.00024	5.7926	1.2042	0.6755
DUS-TLE	0.3221	0.7850	0.20580	0.02894	2.2807	0.4629	0.2124

Table 6. Information criteria for Data 1

Distribution	$-\log L$	AIC	BIC	AICC	HQIC
DUS-HL	78.1430	158.2861	160.1981	158.3691	159.014
QRTMD	82.7759	169.5521	173.3761	169.8071	171.008
DUS-PIR	104.1422	212.2832	216.1072	212.5392	213.739
PGDUS-IK	94.7487	193.4971	197.3211	193.7531	194.954
DUS-TLE	82.3365	168.6731	172.4971	168.9281	170.129

zero. QRTMD also performs well, followed by DUS-TLE with minor tail deviations. In contrast, DUS-PIR and PGDUS-IK show poor fits, exhibiting strong curvature in P–P and Q–Q plots, large systematic deviations in CDF plots, and PDFs that fail to match the histogram shape.

6.2. Data 2

The following dataset was taken from Ünözkan and Yılmaz [25] and includes 56 recorded values of total water flow from Sefaatti Creek during the month of April, spanning the years 1953 to 2014. These data were obtained from the Turkish State Water Affairs Directorate. April mean flows from Sefaatti Creek are selected for extreme value analysis, as this month consistently shows the highest flow levels.

2.47, 2.99, 3.52, 4.12, 4.248, 4.94, 5.19, 5.525, 6.21, 6.89, 7.72, 8.14, 9.23, 9.5, 9.844, 10.8, 12.2, 12.5, 12.9, 13.6, 14.1, 15.1, 15.2, 16.2, 17.3, 17.8, 18.3, 18.4, 19.5, 19.8, 19.99, 21.6, 22, 23, 23.4, 25.2, 25.8, 26.9, 27.2, 27.6, 29, 29, 29.6, 31, 31.4, 36.3, 36.5, 39.6, 41.5, 41.8, 44.4, 44.5, 44.9, 46.8, 53.4, 61.7

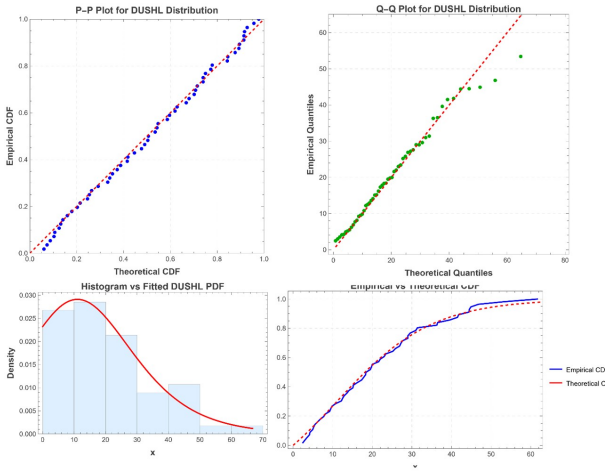
Data 2 were fitted using the same distributions as in Data 1. Tables 7 and 8 summarize the parameter estimates, goodness-of-fit tests, and information criteria. Figure 4 visually supports the performance of these distributions through Q–Q plots for the five fitted models.

Table 7. Goodness-of-fit tests for Data 2

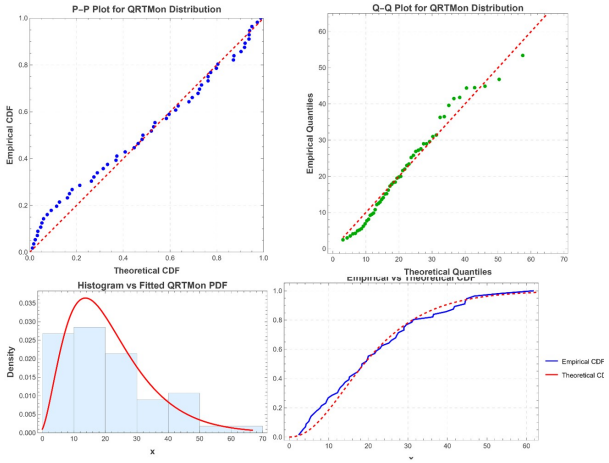
Model	θ MLE	λ MLE	KS	p-value	AD	CVM	WT
DUS-HL	12.5194	—	0.06010	0.98740	0.33914	0.03390	0.0174
QRTMD	0.12804	—	0.08740	0.78550	0.88507	0.09580	0.0929
DUS-PIR	0.40000	0.14220	0.34680	0.00280	8.22505	1.68700	0.6887
PGDUS-IK	0.37356	0.90000	0.31810	0.00021	11.5676	2.24031	1.9985
DUS-TLE	0.02880	0.87600	0.14070	0.21722	2.19950	0.37750	0.1389

Table 8. Information criteria for Data 2

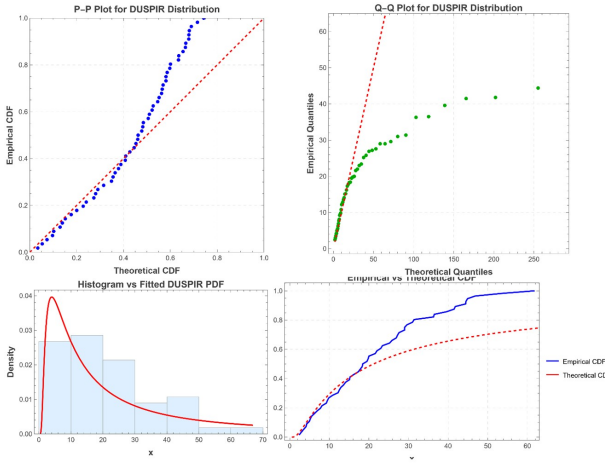
Distribution	$-\log L$	AIC	BIC	AICC	HQIC
DUS-HL	222.4400	446.8854	448.9140	446.9594	447.6700
QRTMD	222.5500	449.1114	453.1624	449.3374	450.6810
DUS-PIR	242.2100	488.4094	492.4640	488.6354	489.9790
PGDUS-IK	273.0600	550.1194	554.1750	550.3454	551.6890
DUS-TLE	226.5400	457.0894	461.1440	457.3154	458.6590



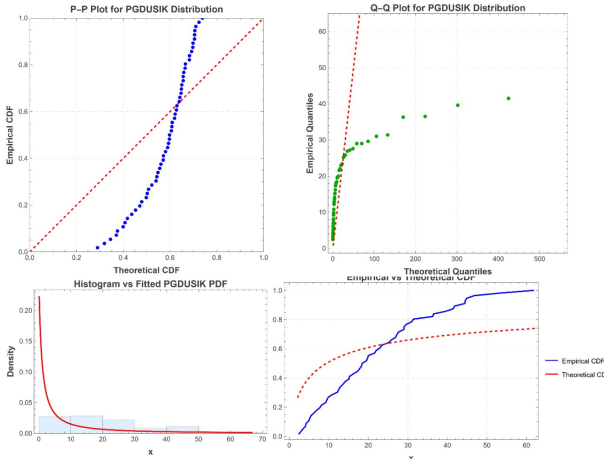
(a)



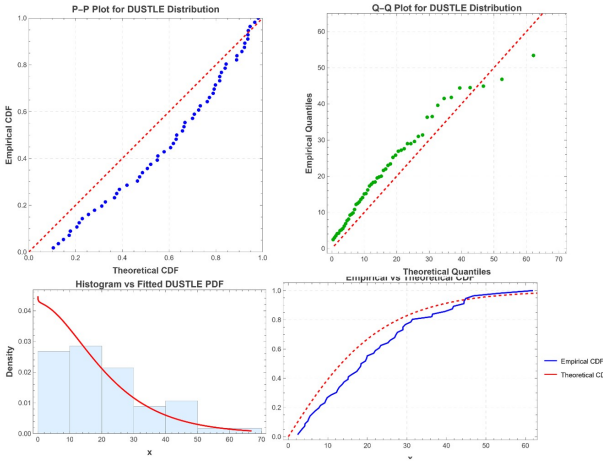
(b)



(c)



(d)



(e)

Figure 4. Diagnostic plots for the fitted distributions to Data 2.

Both the goodness-of-fit tests and the information criteria showed that the DUS-HL distribution is the most appropriate model for the flow data. The DUS-HL model achieves the closest fit with the smallest values for all information criteria. In contrast, the poor performance of the DUS-PIR and PGDUS-IK models implies that more complex parameterizations do not necessarily translate into better representation of extreme flow patterns in this dataset, while the intermediate performance of QRTMD and DUS-TLE might suggest that these distributions could be acceptable for descriptive purposes but are inferior to DUS-HL for inferential or predictive modeling.

From Figure 4, DUS-HL again emerges as the best-fitting distribution, demonstrating near-linear P–P/Q–Q plots and minimal CDF deviations. QRTMD provides a close second, while DUS-TLE shows moderate fit with some tail divergence. DUS-PIR and PGDUS-IK remain inadequate, displaying severe tail divergence in Q–Q plots and large errors in CDF deviation plots, confirming they are unsuitable for this dataset as well.

Several modifications can be made for the the DUS-HL model: embedding it in regression frameworks (accelerated failure time or proportional hazards, exploiting the closed-form quantile function for covariate-adjusted inference), developing censored and truncated likelihood inference (right, interval, and joint progressive Type-II censoring), and pursuing Bayesian and robust estimation for small-sample settings. Parameter extensions (e.g., a two-parameter variant allowing bathtub or decreasing hazards) would broaden the model's domain beyond increasing hazards, while multivariate and copula-based constructions would address dependent paired or spatially correlated data and stress–strength reliability. Finally, broader empirical validation across hydrology, reliability engineering, and medical survival data—especially under small samples, censoring, and misspecified baselines—would clarify when the parsimonious one-parameter DUS-HL is preferable to its more heavily parameterized competitors.

7. Conclusion

This paper proposed the DUS-Half Logistic (DUS-HL) distribution, a one-parameter lifetime model obtained by applying the DUS transformation to the half-logistic baseline. The model inherits a monotone increasing hazard bounded between $1/(2\theta)$ and $1/\theta$, suitable for wear-out and aging processes, with closed-form expressions for the CDF, quantile function, survival, hazard, cumulative hazard, and Mills ratio. A simulation study comparing MLE, LSE, CVM, and WLS showed that MLE achieves the lowest MSE and bias, while LSE is least reliable. The DUS-HL was validated on fetal Crown-Rump Length and April streamflow datasets, achieving the lowest goodness-of-fit statistics and information criteria among competing models in both applications, though these findings apply within a clear domain: increasing hazards, right-skewed heavy-tailed data, small or moderate samples, and cases needing closed-form quantiles, but not bathtub-shaped or decreasing hazards, multimodal data, or light-tailed bounded data. Future work could extend the DUS-HL to regression and Bayesian frameworks, censored data settings, and other applied domains.

Acknowledgement

The authors are grateful to Tanta University Hospital for providing the first dataset used in this study.

REFERENCES

1. D. Kumar, U. Singh, and S. K. Singh, *A method of proposing new distribution and its application to bladder cancer patients data*, Journal of Statistics Applications & Probability Letters, vol. 2, no. 3, pp. 235–245, 2015.
2. N. Eugene, C. Lee, and F. Famoye, *Beta-normal distribution and its applications*, Communications in Statistics - Theory and Methods, vol. 31, no. 4, pp. 497–512, 2002.
3. M. C. Jones, *Kumaraswamy's distribution: A beta-type distribution with some tractability advantages*, Statistical Methodology, vol. 6, no. 1, pp. 70–81, 2009.
4. G. M. Cordeiro, and M. de Castro, *A new family of generalized distributions*, Journal of Statistical Computation and Simulation, vol. 81, no. 7, pp. 883–898, 2011.

5. A. A. Al-Shomrani, O. H. Arif, A. I. Shawky, S. Hanif, and M. Q. Shahbaz, *Topp-Leone family of distributions: Some properties and application*, Pakistan Journal of Statistics and Operation Research, vol. 12, no. 3, pp. 443–451, 2016.
6. A. W. Marshall, and I. Olkin, *A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families*, Biometrika, vol. 84, no. 3, pp. 641–652, 1997.
7. C. Ku, *On the DUS-Kumaraswamy distribution*, Istatistik Journal of the Turkish Statistical Association, vol. 13, no. 1, pp. 29–38, 2021.
8. M. Elgarhy, D. S. Metwally, I. E. Ragab, S. O. Bashiru, H. E. Semary, and A. M. Gemay, *The DUS-transformed generalized linear failure rate distribution: Properties, estimation, and applications*, Scientific African, vol. 29, Article e02891, 2025.
9. D. Qayoom, A. A. Rather, O. A. Alqasem, Z. Ahmad, M. Nagy, A. M. Yousuf, A. H. Mansi, E. Hussam, and A. M. Gemay, *Development of a novel extension of Rayleigh distribution with application to COVID-19 data*, Scientific Reports, vol. 15, Article 18535, 2025.
10. B. Thomas, and V. M. Chacko, *Power generalized DUS transformation of exponential distribution*, International Journal of Statistics and Reliability Engineering, vol. 10, no. 2, pp. 263–272, 2023.
11. A. A. Rather, and G. Ozel, *A new class of Lindley distribution: System reliability, simulation and applications*, Heliyon, vol. 10, no. 19, Article e38335, 2024.
12. M. M. E. Abd El-Monsef, T. A. Alghazi, and H. H. El-Damrawy, *A modified Monsef distribution using quadratic rank transformation map*, Statistics, Optimization & Information Computing, vol. 14, no. 5, pp. 2155–2171, 2025.
13. M. H. Abu Jarad, A. A. Khan, and E. S. A. Abu Jarad, *Bayesian survival analysis of generalized DUS exponential distribution*, Journal of Data Science, vol. 18, no. 3, pp. 488–507, 2020.
14. R. Chandran, and V. M. Chacko, *Joint adaptive progressive type-II censoring with DUS-extreme value distribution*, Reliability: Theory and Applications, vol. 20, no. 4, pp. 112–125, 2025.
15. M. Gautam, *A neutrosophic Rayleigh approach based on DUS transformation for analyzing COVID-19 incubation periods*, Neutrosophic Systems with Applications, vol. 35, Article 42, 2026.
16. Z. A. H. Abd, *A new statistical distribution using the DUS transformation with a practical application*, Master's thesis, University of Karbala, Institutional Repository, 2025.
17. A. Mushtaq, T. Hussain, M. Shakil, M. Ahsanullah, and B. M. G. Kibria, *Exponentiated inverse exponential distribution properties and applications*, Axioms, vol. 14, no. 10, Article 753, 2025.
18. D. Kumar, U. Singh, and S. K. Singh, *A new distribution using sine function - Its application to bladder cancer patients data*, Journal of Statistics Applications and Probability, vol. 4, no. 3, pp. 417–427, 2015.
19. P. Kavya, and M. Manoharan, *On a generalized lifetime model using DUS transformation*, in Applied Probability and Stochastic Processes, edited by V. C. Joshua, S. R. S. Varadhan, and V. M. Vishnevsky, Springer Nature Singapore, pp. 281–291, 2020.
20. K. Deepti, and V. M. Chacko, *DUS transformation of Weibull distribution*, Reliability: Theory & Applications, vol. 17, no. 1, pp. 58–71, 2022.
21. P. Gauthami, and V. M. Chacko, *Power generalized DUS transformation of inverse Kumaraswamy distribution and stress strength analysis*, Journal of the Society of Statistics, Computer and Applications, vol. 18, no. 1, pp. 323–338, 2024.
22. K. Karakaya, ı. Kinaci, C. Kuş, and Y. Akdoğan, *On the DUS-Kumaraswamy distribution*, Istatistik Journal of the Turkish Statistical Association, vol. 13, no. 1, pp. 29–38, 2021.
23. B. Thomas, and V. M. Chacko, *Power generalized DUS transformation in Weibull and Lomax distributions*, Reliability: Theory and Applications, vol. 18, no. 1, pp. 368–384, 2023.
24. A. Essa, H. Alsaedi, A. Hussain, and M. Tashtoush, *A hybrid support vector machine-genetic algorithm framework for estimating the DUS-transformed generalized polynomial quadratic failure rate distribution*, Statistics, Optimization & Information Computing, vol. 16, no. 1, pp. 327–358, 2026.
25. H. Ünözkan, and M. Yılmaz, *A new method for generating distributions: An application to flow data*, International Journal of Statistics and Applications, vol. 9, no. 3, pp. 92–99, 2019.