J. Stat. Appl. Pro. Vol. **14**, No. 6, 701-716 (2025)

Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/140614

Estimation Methods for the Generalized Extreme Weibull Distribution: Theory, Simulation, and Applications to Cancer Survival Data

Silvana El Rabih and Noura Yassine[*]

Department of Mathematics & Computer Science, Faculty of Science, 1100, Beirut Arab University, Beirut, Lebanon

Received: 27 Aug. 2025, Revised: 28 Sep. 2025, Accepted: 22 Oct. 2025.
Published online: 1 Nov. 2025.

Abstract: In this study, motivated by the importance of the Weibull distribution and its widespread use in modeling real-life data, we propose a new distribution called the generalized extreme Weibull distribution (GEW), based on the maximum of multiple Weibull distributions. The statistical properties of the distribution were examined, and expressions for these measures were obtained. Furthermore, the distribution parameters were estimated using different estimation methods. The accuracy of these methods is also highlighted. Monte Carlo simulation is employed to validate the results. To demonstrate the practical relevance and effectiveness of our novel distribution in modeling empirical phenomena, the results are applied to two types of cancer datasets. The findings provide strong evidence of the GEW distribution's efficiency as a flexible and powerful tool for modeling on-hand data.

Keywords: Extreme Weibull Distribution, Estimations Methods, Monte Carlo Simulation, Survival Cancer Model.

1. Introduction

Data analysis plays a crucial role in various fields, including management science, reliability analysis, economics, and health sciences. The choice of an appropriate model or distribution ensures that the analysis can capture the key features of the data, such as skewness, variability, tail behavior and hazard rates. Many statistical distributions have been extensively applied to real-life situations, such as environmental studies for modeling wind speed and energy [1], optimization in inventory systems [2, 3,4,5,6], and biomedical research [7,8]. The analysis of survival data is no exception, as it has wide applicability across various domains, particularly in medical research [9,10]. This type of study requires careful statistical consideration and multilevel adjustments to establish a model that provides the best fit to the data [11,12]. Therefore, the proper selection of models and distributions is central to producing valid, generalizable, and practically useful results for data analysis. However, traditional distributions may exhibit features such as asymmetry, heavy tails, multimodality, and complex dependence structures, which limit their ability to accurately represent real-world problems. This has driven researchers to explore novel and more flexible distributions that can better capture the data characteristics and underlying patterns. The functions of random variables provide one approach to developing new distributions, often involving additional parameters that allow the distribution to relax its assumptions and more effectively represent the behavior of the data.

Prior studies have recognized mixtures of continuous distributions as a method for developing new distributions [13,14], and introduced new distributions based on the extremes of several existing ones, providing a practical tool for modeling stochastic inventory problems with random properties. Other researchers have proposed weighted sums, products, and ratios of distributions as approaches to create or extend new distributions [15,16]. Additionally, introducing extra parameters can refine or relax existing distributions, allowing them to better adapt to data characteristics [17,18].

This study focuses on the Weibull distribution [19], a widely recognized and extensively used model for mortality and failure data, owing to its flexibility in describing different failure rates. Despite its popularity. The original Weibull distribution, with its two parameters, is limited to modeling monotonically increasing or decreasing hazard function. Consequently, its application to biomedical data often requires extension [9,12,17,18,20]. To alleviate these challenges, we introduce the Generalized Extreme Weibull (GEW) distribution, which is derived as the maximum of several non- identical Weibull distributions. By incorporating additional parameters, the GEW enhances flexibility, enabling the model to capture non-monotonic hazard shapes and represent a wider variety of data patterns.

To highlight the significance and properties of our distribution, a comparison is established between the proposed Generalized Extreme Weibull (GEW) distribution and several existing generalizations of the Weibull distribution, including the Weibull [19], Beta Weibull (BW) [26], Kumaraswamy Weibull (KumW) [27], Alpha Power Weibull (APW) [29], Alpha Power Kumaraswamy Weibull (APKumW) [9], Exponentiated Generalized Weibull (EGW) [25], and Exponentiated Kumaraswamy

[*]Corresponding author e-mail: noura.yassein@bau.edu.lb

Weibull (EKumW) [28] distributions. Table 1 provides a quantitative comparison among the more prevalent Weibull generalizations, detailing the number of parameters that reflects the flexibility of the distribution and their key features, see Table 1.

Table 1: Comparative Analysis of Weibull Generalizations

Distribution	Ref	No of Parameters	Key Features
Weibull	[19]	2	Baseline model, models monotone hazard rates only
APW	[29]	3	Flexible in modeling skewness, better fit for non-monotonic data than standard Weibull.
EGW	[25]	4	Capable of modeling unimodal and bathtub-shaped hazard rates
BW	[26]	4	Classic generalization; widely used for heavy-tailed data and complex hazard shapes.
KumW	[27]	4	Highly flexible; handles bounded unit interval properties within a Weibull framework.
EKumW	[28]	5	Extended flexibility for extreme values; captures subtle tail behaviors.
APKumW	[9]	5	Superior flexibility and the capability to model non-monotone hazard rates, including bathtub and unimodal shapes.
GEW	New Model	flexible	Unimodal shape achieves high flexibility with fewer parameters, can adopt more parameters to better fit data when needed.

This article extends prior studies by presenting detailed properties of the GEW distribution, exploring advanced estimation methods, and demonstrating its relevance in several relevant scenarios. The results are expected to benefit both researchers and practitioners in survival analysis, extreme value theory, and related fields.

This article provides a structured overview in Section 1, along with supporting literature on the Generalized Extended Weibull (GEW) distribution that reinforces this research. In Section 2, expressions for the Cumulative Distribution Function (CDF) and Probability Density Function (PDF) are developed. Moreover, in Section 3, the fundamental statistics associated with the GEW distribution and order statistics are studied. In Section 4, four methods for estimating GEW distribution parameters: maximum likelihood estimation (MLE), least squares estimation (LSE), weighted least squares estimation (WLSE), and Cramér–von Mises estimation (CVME) are discussed. In Section 5, a comprehensive simulation study is conducted, and the parameter estimation methodologies are evaluated based on extensive numerical experiments to assess their accuracy and performance. Based on these results**,** Section 6 uses Monte Carlo Simulations to assess the statistical validity and robustness of the theoretical findings regarding the GEW Distribution. Finally, Section 7 illustrates real-world applications of the GEW distribution using two actual cancer survival datasets. While. Section 8 summarizes the overall contributions of this study and prospective future work.

2. Generalized Extreme Weibull Distribution

The cumulative probability of the Weibull distribution [19], $X \sim Weib(\alpha, k)$, with a scale parameter α and shape parameter k is expressed as

$$F(x) = 1 - \exp\left(-\left(\frac{x}{\alpha}\right)^k\right), \quad x \geq 0 \tag{1}$$

Differentiating (1), the probability density function(pdf) of the Weibull distribution becomes:

$$f(x) = \frac{k}{\alpha}\left(\frac{x}{\alpha}\right)^{k-1} \exp\left(-\left(\frac{x}{\alpha}\right)^k\right), \quad x \geq 0, \quad k, \alpha > 0 \tag{2}$$

where k follows the distribution's shape and α is the scale parameter.

We define the Generalized Extreme Weibull (GEW) distribution as a new family of distribution that arises as the maximum of independent Weibull random variables with distinct scale and shape parameters. Define $S = \{X_1, X_2, \dots, X_n\}$ as a set of mutually independent random variables that follow a Weibull distribution parameterized by scale α_i and shape k_i. The cumulative distribution (CDF) of X_i is expressed as

$$F_i(t) = P(X_i \leq t) = 1 - exp\left(-\left(\frac{t}{\alpha_i}\right)^{k_i}\right). \tag{3}$$

Describe the Generalized Extreme Weibull (GEW) random variable as $Y = \text{Max}(\{X_1, X_2, \dots, X_n)$. The CDF of Y is

$$F_Y(t) = P(Y \leq t) = P(max(X_1, X_2, \dots, X_n) \leq t) = P(X_1 \leq t, X_2 \leq t, \dots, X_n \leq t).$$

Based on the assumption that the random variables are independent, we have

$$F_Y(t) = P(X_1 \le t)P(X_2 \le t)\dots P(X_n \le t) = \prod_{i=1}^{n} F_i(t). \quad (4)$$

Using (3), the GEW distribution's cumulative function becomes:

$$F_Y(t) = P(Y \le t) = \prod_{i=1}^{n}\left(1 - exp\left(-\left(\frac{t}{\alpha_i}\right)^{k_i}\right)\right). \quad (5)$$

Using the inclusion-exclusion formula [21], the product term in (5) is expanded as follows:

$$\prod_{i=1}^{n}\left(1 - exp\left(-\left(\frac{t}{\alpha_i}\right)^{k_i}\right)\right) = \sum_{k=0}^{n}(-1)^k\left(\sum_{1\le i_1<\cdots<i_k\le n}\exp\left(-\sum_{l=1}^{k}\left(\frac{t}{\alpha_{i_l}}\right)^{k_{i_l}}\right)\right). \quad (6)$$

Hence, the CDF becomes:

$$F_Y(t) = \sum_{k=0}^{n}(-1)^k\left(\sum_{1\le i_1<\cdots<i_k\le n}\exp\left(-\sum_{l=1}^{k}\left(\frac{t}{\alpha_{i_l}}\right)^{k_{i_l}}\right)\right). \quad (7)$$

By differentiating equation (5), GEW's probability density function becomes:

$$f_Y(t) = \sum_{j=1}^{n}\left(\frac{k_j}{\alpha_j}\left(\frac{t}{\alpha_j}\right)^{k_j-1} exp\left(-\left(\frac{t}{\alpha_j}\right)^{k_j}\right)\prod_{\substack{i=1\\ i\ne j}}^{n}\left(1 - exp\left(-\left(\frac{t}{\alpha_i}\right)^{k_i}\right)\right)\right). \quad (8)$$

Using the inclusion-exclusion formula [21] again, the product term in (8) is expanded such that:

$$\prod_{\substack{i=1\\ i\ne j}}^{n}\left(1 - exp\left(-\left(\frac{t}{\alpha_i}\right)^{k_i}\right)\right) = \sum_{k=0}^{n-1}(-1)^k\left(\sum_{\substack{1\le i_1<\cdots<i_k\le n\\ i_l\ne j}}\exp\left(-\sum_{l=1}^{k}\left(\left(\frac{t}{\alpha_{i_l}}\right)^{k_{i_l}}\right)\right)\right). \quad (9)$$

Hence, the GEW's probability density function becomes

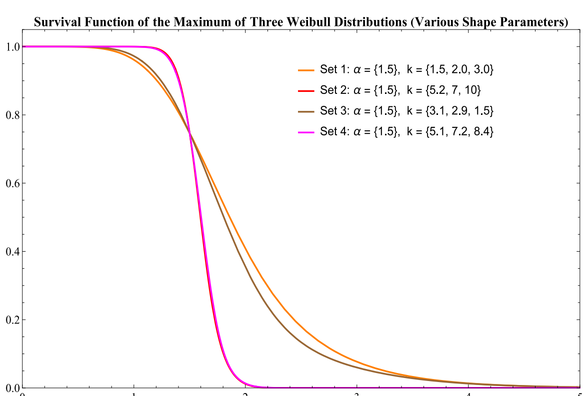
$$f_Y(t) = \sum_{j=1}^{n}\left(\frac{k_j}{\alpha_j}\left(\frac{t}{\alpha_j}\right)^{k_j-1} exp\left(-\left(\frac{t}{\alpha_j}\right)^{k_j}\right)\right)\sum_{k=0}^{n-1}(-1)^k\left(\sum_{\substack{1\le i_1<\cdots<i_k\le n\\ i_l\ne j}}\exp\left(-\sum_{l=1}^{k}\left(\left(\frac{t}{\alpha_{i_l}}\right)^{k_{i_l}}\right)\right)\right). \quad (10)$$

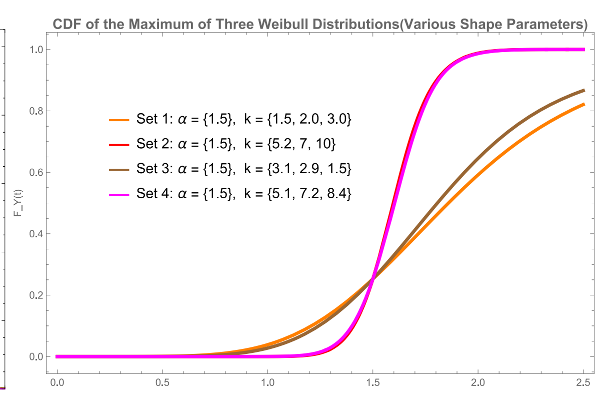
The survival and hazard functions are respectively:

$$S_Y(t) = 1 - F_Y(t) = 1 - \left(\sum_{k=0}^{n}(-1)^k\left(\sum_{1\le i_1<\cdots<i_k\le n} exp\left(-\sum_{l=1}^{k}\left(\left(\frac{t}{\alpha_{i_l}}\right)^{k_{i_l}}\right)\right)\right)\right), \quad (11)$$

and

$$h_Y(t) = \frac{f_Y(t)}{1-F_Y(t)} = \frac{\sum_{j=1}^{n}\left[\frac{k_i}{\alpha_j}\left(\frac{t}{\alpha_j}\right)^{k_j-1} exp\left(-\left(\frac{t}{\alpha_j}\right)^{k_j}\right)\right]\sum_{k=0}^{n-1}(-1)^k\left(\sum_{\substack{1\le i_1<\cdots<i_k\le n\\ i_l\ne j}}\exp\left(-\sum_{l=1}^{k}\left(\left(\frac{t}{\alpha_{i_l}}\right)^{k_{i_l}}\right)\right)\right)}{1-\left(\sum_{k=0}^{n}(-1)^k\left(\sum_{1\le i_1<\cdots<i_k\le n} exp\left(-\sum_{l=1}^{k}\left(\left(\frac{t}{\alpha_{i_l}}\right)^{k_{i_l}}\right)\right)\right)\right)}. \quad (12)$$





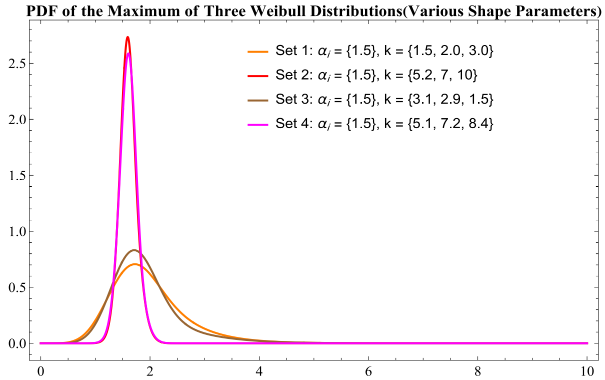


Fig. 1: PDF, CDF and Survival Function of GEW for n=3, different k_i, and identical scale parameters**.**

In the GEW distribution, the shape parameters k_i play a crucial role in controlling the concentration and spread of the distribution. As k_i increases, the PDF becomes more concentrated around the mode, the CDF has a steeper slope, and the survival function decays quickly with less variability and a greater probability of the maximum falling within a narrower interval of values (see Figure 1). However, with a small value of k_i, there is a flatter distribution's PDF, gradual CDF, and slower-decaying Survival Function (see Figure 1) resulting in a wider interval that is heavily populated with the maximum value due to the diffusion of maximum values.

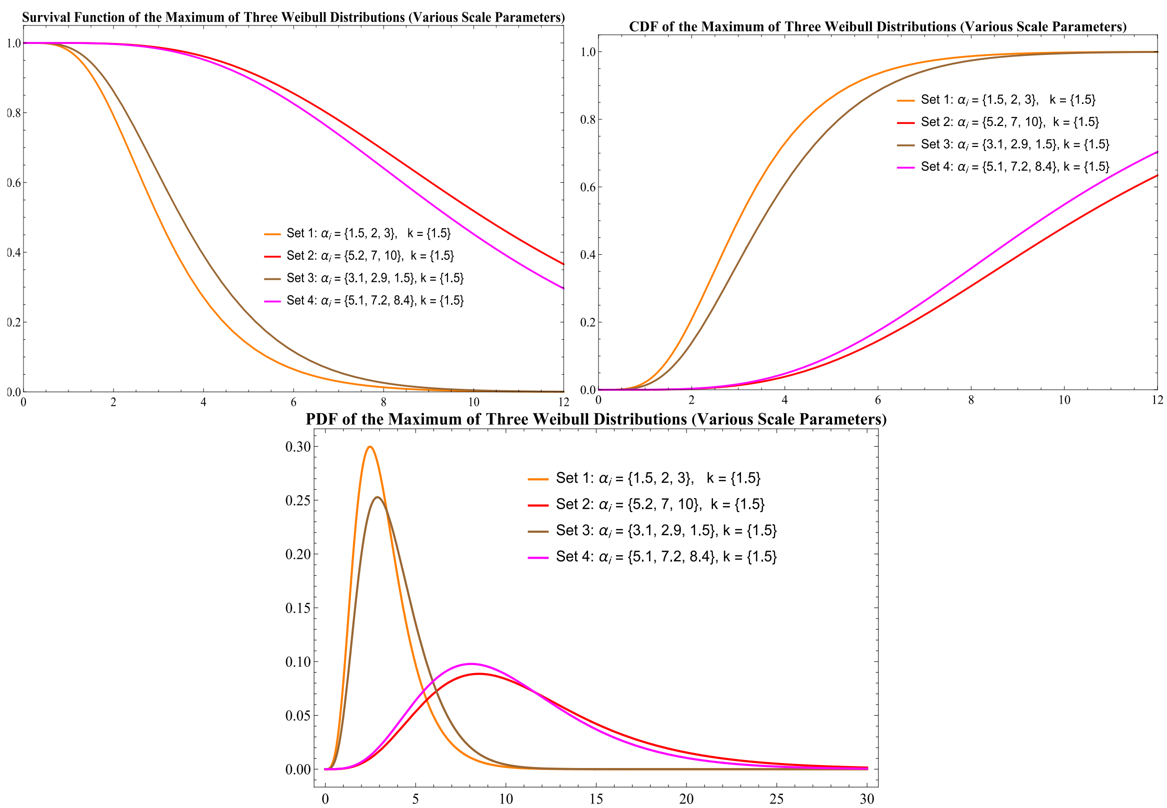


Fig. 2: PDF, CDF & Survival Function of GEW for n=3, different α_i and identical shape parameters**.**

As the scale parameters α_i increased, the PDF of GEW distribution shifted rightward, resulting in wider and more dispersed PDFs with larger maximum values. This causes a gradual increase in the flatness of the CDFs, and a less rapid decline in the survival functions (see Figure 2). Thus, the maximum values occur at a higher probability of taking extremes with heavier tail. However, decreasing the scale parameters will cause the PDFs to be closer in proximity to the origin, have a steeper peak, and have lower variance. The CDFs will rise quickly, whereas the survival functions decrease rapidly, indicating a shorter duration between maximum values as well as more concentrated clusters of maximum values. The scale parameters are therefore critical components for determining where and how widely distributed the GEW is across time.

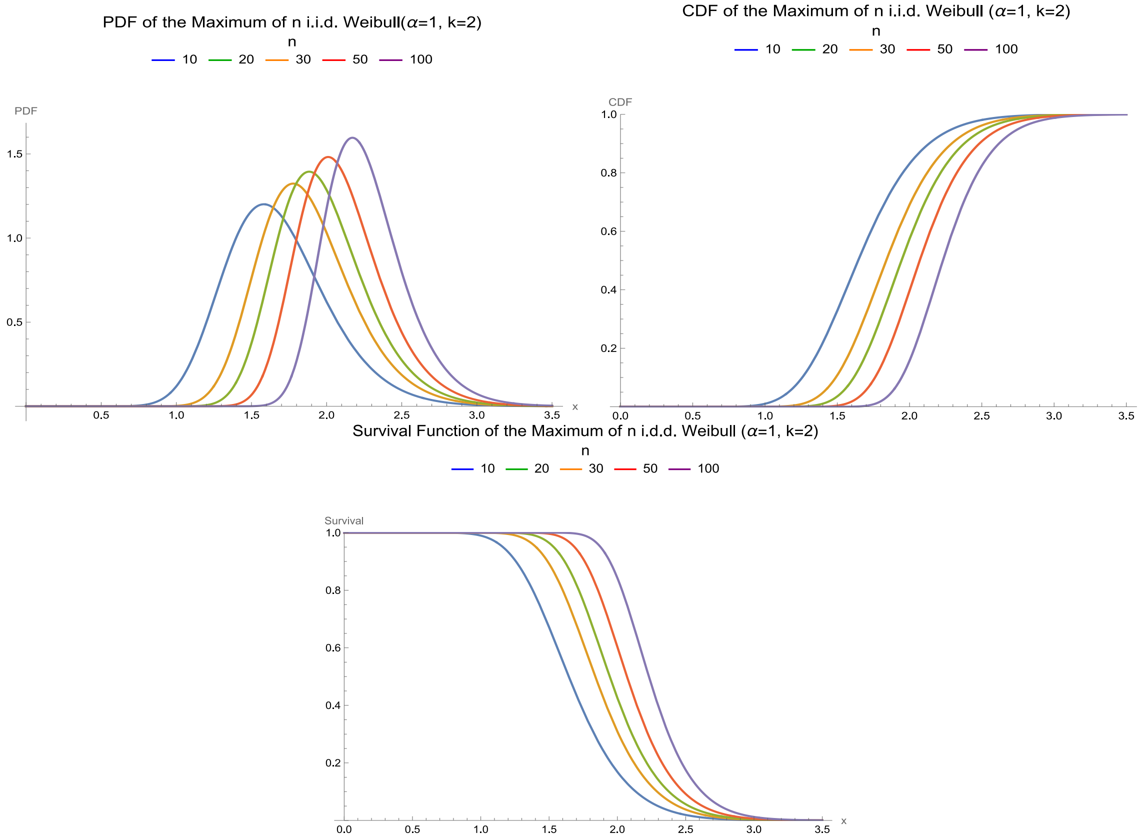


Fig. 3: PDF, CDF & Survival Function of the GEW distributions with different n, identical scale and shape parameters

As shown in Figure 3, increasing the parameter *n* shifts the PDF right, and makes it narrower. The CDF becomes steeper, and the survival function decays more rapidly. This indicates a higher concentration and lower variability of the maxima (see Figure 3). Conversely, when *n* decreases, the distribution's PDF becomes flatter and more spread out, the CDF rises more gradually, and the survival function decays more slowly (see Figure 3). Overall, n acts as a shape-controlling parameter, influencing the concentration of the maximum without changing the individual Weibull scale

3. Statistical Measures

This section provides an overview of some important statistical features of the GEW distribution including quantile, mode, skewness, kurtosis, r^{th} moments and order statistics.

3.1 Quantile and Mode

In this subsection, the quantile function and mode of the GEW distribution are provided.

Suppose $X_i \sim Weib(\alpha_i, k_i)$ where X_i are mutually independent. Define the Generalized Extreme Weibull (GEW) random variable as $Y = \text{Max}(\{X_1, X_2, \dots, X_n)$. The q^{th} quantile *x*, is defined such that $P(X \leq x) = q$ using the cumulative function assumed by (5), expressions for quantiles can be found by setting $F(x) = q$. Hence,

$$\mathrm{F}_Y(x) = \prod_{i=1}^{n} \left(1 - exp\left(-\left(\frac{x}{\alpha_i}\right)^{k_i}\right)\right) \,]$$

$$= \sum_{k=0}^{n} (-1)^k \left(\sum_{1 \leq i_1 < \cdots < i_k \leq n} \exp\left(-\sum_{l=1}^{k}\left(\left(\frac{x}{\alpha_l}\right)^{k_l}\right)\right)\right) = q. \tag{13}$$

Since (13) is not linear, the quantile of the GEW distribution is determined numerically. By setting $q = 0.5$ and solving for *x*, the median can be determined. On the other hand, the mode is the value that maximizes the probability density function. Hence, by solving $f_Y'(x) = 0$ one can obtain the modal value.

3.2 Measures of Skewness and Kurtosis

The skewness and kurtosis of the GEW distribution are evaluated using its quantile-based definitions.

The coefficient of skewness of the GEW distribution is expressed in terms of quartiles as follows:

$$Sk = \frac{Q_3+Q_1-2Q_2}{Q_3-Q_1}, \tag{14}$$

where Q_1, Q_2 and Q_3 denote the first, second, and third quartiles, respectively. As for the kurtosis measure, it is given in terms of the octiles. That is,

$$Ku = \frac{Q_7-Q_5+Q_3-Q_1}{Q_6-Q_2}, \tag{15}$$

where $Q_i = \frac{i}{8}$

3.3 The Moments

In this section, expressions for the r^{th} moments of the GEW distribution are developed.

Let Y be a random variable following the Generalized Extreme Weibull (GEW) distribution. The r^{th}-moments of Y are expressed as

$$\mu_Y(t) = E(Y^r) = \int_0^{+\infty} t^r f_Y(t)dt. \tag{16}$$

Integrating (16) we obtain

$$\mu_Y(t) = r\int_0^{+\infty} t^{r-1}P(Y > t)dt = r\int_0^{+\infty} t^{r-1}\left(1-\left(\prod_{i=1}^{n}\left(1-exp\left(-\left(\frac{t}{\alpha_i}\right)^{k_i}\right)\right)\right)\right)dt. \tag{17}$$

Using the expansion of the product in (6), the GEW's moments are given by

$$\mu_Y(t) = \sum_{k=1}^{n}(-1)^{k+1}\left(\sum_{1\le i_1<\cdots<i_k\le n}\int_0^{+\infty} rt^{r-1}\, exp\left(-\sum_{l=1}^{k}\left(\left(\frac{t}{\alpha_{i_l}}\right)^{k_{i_l}}\right)\right)dt\right). \tag{18}$$

Table 2 summarizes both the estimated quantiles and the main statistical properties of the GEW distribution with parameter set $S = \{(\alpha_1 = 0.5, k_1 = 1.2), (\alpha_2 = 1, k_2 = 1.5), (\alpha_3 = 1.5, k_3 = 2)\}$. The quantiles illustrate how the data are distributed across different probability levels, while the accompanying measures (mean, variance, mode, skewness, and kurtosis) describe the overall shape, spread, and asymmetry of the distribution.

By calculating the first two moments, the expression for the variance of the GEW distribution can be obtained. Thus, using the moments in (18), we have

$$E(X) = \sum_{k=1}^{n}(-1)^{k+1}\left(\sum_{1\le i_1<\cdots<i_k\le n}\int_0^{+\infty} exp\left(-\sum_{l=1}^{k}\left(\left(\frac{t}{\alpha_{i_l}}\right)^{k_{i_l}}\right)\right)dt\right). \tag{19}$$

$$E(X^2) = \sum_{k=1}^{n}(-1)^{k+1}\left(\sum_{1\le i_1<\cdots<i_k\le n}\int_{0s}^{+\infty} 2t\, exp\left(-\sum_{l=1}^{k}\left(\left(\frac{t}{\alpha_{i_l}}\right)^{k_{i_l}}\right)\right)dt\right). \tag{20}$$

The variance is defined by:

$$Var(X) = E(X^2) - \left(E(X)\right)^2 \tag{21}$$

Using (19) and (20) in (21) the GEW's variance is determined.

To illustrate, Mathematica software is used to provide the estimated quantiles and some statistical properties of the GEW distribution. Considering the parameters listed in the set $S = \{(\alpha_1 = 0.5, k_1 = 1.2), (\alpha_2 = 1, k_2 = 1.5), (\alpha_3 = 1.5, k_3 = 2)\}$, the quantiles, mean, variance, mode, skewness, and kurtosis are computed and depicted in Table 2.

Table 2: Quantiles and Statistical Properties of GEW for the set *S*

Section	**Metric**	**Value**
Quantiles	Q(1/8)	0.8463
	Q(1/4)	1.0785
	Q(3/8)	1.2759
	Q(1/2)	1.4699
	Q(5/8)	1.6800
	Q(3/4)	1.9327
	Q(7/8)	2.2985

Statistical Properties	Mean	1.5458
	Variance	0.4047
	Mode	1.3032
	Skewness	0.6710
	Kurtosis	3.4979

3.4 Order Statistics

In this section, the order statistics of the GEW distribution are investigated.

Suppose an independent random sample $y_1, y_2, \dots, y_N$ where each y_j follow the GEW distribution. Let $Y_{1,N}, Y_{2,N}, \dots, Y_{N,N}$ denote the corresponding order statistics. The PDF of the r^{th} order statistics of the GEW's is expressed as follows:

$$f_{Y_{r,N}}(x) = \frac{N!}{(r-1)!(N-r)!}\left(1 - F_Y(x)\right)^{N-r}\left(F_Y(x)\right)^{r-1} f_Y(x). \text{ where } x > 0 \tag{22}$$

By substituting the CDF (6) and the PDF (10) in (22), the r^{th} order statistic of the GEW distribution is obtained as follows

$$f_{Y_{r,N}}(x) = \frac{N!}{(r-1)!(N-r)!}\left(1 - \left(\sum_{k=0}^{n}(-1)^k\left(\sum_{1\le i_1<\cdots<i_k\le n}\exp\left(-\sum_{l=1}^{k}\left(\left(\frac{x}{\alpha_l}\right)^{k_l}\right)\right)\right)\right)\right)^{N-r}$$

$$\left(\sum_{k=0}^{n}(-1)^k\left(\sum_{1\le i_1<\cdots<i_k\le n}\exp\left(-\sum_{l=1}^{k}\left(\left(\frac{x}{\alpha_l}\right)^{k_l}\right)\right)\right)\right)^{r-1}\sum_{j=1}^{n}\left(\frac{k_j}{\alpha_j}\left(\frac{t}{\alpha_j}\right)^{k_j-1} exp\left(-\left(\frac{t}{\alpha_j}\right)^{k_j}\right)\right)$$

$$\sum_{k=0}^{n-1}(-1)^k\left(\sum_{\substack{1\le i_1<\cdots<i_k\le n\\ i_l\ne j}}\exp\left(-\sum_{l=1}^{k}\left(\left(\frac{t}{\alpha_{i_l}}\right)^{k_{i_l}}\right)\right)\right). \tag{23}$$

By setting $r = 1$ and $N = 1$, expression (23) reduces to the PDF of the GEW distribution given in (6). The r^{th} order statistics of the GEW distribution for $r = 1$ and $r = N$ are respectively given by

$$f_{Y_{1,N}}(x) = N\left(1 - \left(\sum_{k=0}^{n}(-1)^k\left(\sum_{1\le i_1<\cdots<i_k\le n}\exp\left(-\sum_{l=1}^{k}\left(\left(\frac{x}{\alpha_l}\right)^{k_l}\right)\right)\right)\right)\right)^{N-1}$$

$$\sum_{j=1}^{n}\left(\frac{k_j}{\alpha_j}\left(\frac{t}{\alpha_j}\right)^{k_j-1} exp\left(-\left(\frac{t}{\alpha_j}\right)^{k_j}\right)\right)\sum_{k=0}^{n-1}(-1)^k\left(\sum_{\substack{1\le i_1<\cdots<i_k\le n\\ i_l\ne j}}\exp\left(-\sum_{l=1}^{k}\left(\left(\frac{t}{\alpha_{i_l}}\right)^{k_{i_l}}\right)\right)\right). \tag{24}$$

$$f_{Y_{N,N}}(x) = N\left(\sum_{k=0}^{n}(-1)^k\left(\sum_{1\le i_1<\cdots<i_k\le n}\exp\left(-\sum_{l=1}^{k}\left(\left(\frac{x}{\alpha_l}\right)^{k_l}\right)\right)\right)\right)^{N-1}\sum_{j=1}^{n}\left(\frac{k_j}{\alpha_j}\left(\frac{t}{\alpha_j}\right)^{k_j-1} exp\left(-\left(\frac{t}{\alpha_j}\right)^{k_j}\right)\right)$$

$$\sum_{k=0}^{n-1}(-1)^k\left(\sum_{\substack{1\le i_1<\cdots<i_k\le n\\ i_l\ne j}}\exp\left(-\sum_{l=1}^{k}\left(\left(\frac{t}{\alpha_{i_l}}\right)^{k_{i_l}}\right)\right)\right). \tag{25}$$

4. Parameter Estimation

The parameters of the GEW distribution are estimated using four methods: the maximum likelihood estimators (MLE), the least-squares estimators (LSE), the weighted least-squares estimators (WLSE), and the Cramér–von Mises estimators (CVME). A sensitivity analysis of the performance of these methods is conducted to examine their effectiveness and the influencing factors.

4.1 MLE

Let $y_1, y_2, \dots, y_N$ be *a* random sample of size N following the GEW distribution. Define the vector $\theta = (\alpha_1, k_1, \alpha_2, k_2, \dots, \alpha_n, k_n)$ be the parameters to be estimated. The GEW's likelihood function, denoted by $L(\theta)$, is given as follows

$$L(\theta) = \prod_{r=1}^{N} f_Y(y_r). \tag{26}$$

Moreover, using (8) in (26), the function of the log-likelihood GEW, represented as $l(\theta)$, can be expressed as the following:

$$l(\theta) = \sum_{r=1}^{N} ln\left(\sum_{j=1}^{n}\left(\frac{k_j}{\alpha_j}\left(\frac{y_r}{\alpha_j}\right)^{k_j-1} exp\left(-\left(\frac{y_r}{\alpha_j}\right)^{k_j}\right)\prod_{\substack{i=1\\ i\neq j}}^{n}\left(1-exp\left(-\left(\frac{y_r}{\alpha_i}\right)^{k_i}\right)\right)\right)\right). \quad (27)$$

The MLEs [22] $\widehat{\alpha_l}$ and $\widehat{k_l}$ can be determined by the maximization of (27) through the application of numerical methods. To obtain the normal equations of $l(\theta)$, the first partial derivatives of $l(\theta)$, with respect to α_l and k_l are determined. Thus, $\forall l = 1,..,n$, the following equations are obtained

$$\frac{\partial l}{\partial \alpha_l} = \sum_{r=1}^{N}\frac{1}{S_r}\left(E_{l,r}D_{l,r}\left(-\frac{k_l}{\alpha_l}+\frac{k_l\left(\frac{y_r}{\alpha_l}\right)^{k_l}}{\alpha_l}\right)-\sum_{\substack{j=1\\ j\neq l}}^{n}D_{j,r}E_{j,r}\frac{k_l\left(\frac{y_r}{\alpha_l}\right)^{k_l}exp\left(-\left(\frac{y_r}{\alpha_l}\right)^{k_l}\right)}{\alpha_l\left(1-exp\left(-\left(\frac{y_r}{\alpha_l}\right)^{k_l}\right)\right)}\right), \quad (28)$$

and

$$\frac{\partial l}{\partial k_l} = \sum_{r=1}^{N}\frac{1}{S_r}\left(E_{l,r}D_{l,r}\left(\frac{1}{k_l}+\left(1-\left(\frac{y_r}{\alpha_l}\right)^{k_l}\right)ln\left(\frac{y_r}{\alpha_l}\right)\right)-\sum_{\substack{j=1\\ j\neq l}}^{n}D_{j,r}E_{j,r}\frac{\left(\frac{y_r}{\alpha_l}\right)^{k_l}ln\left(\frac{y_r}{\alpha_l}\right)exp\left(-\left(\frac{y_r}{\alpha_l}\right)^{k_l}\right)}{1-exp\left(-\left(\frac{y_r}{\alpha_l}\right)^{k_l}\right)}\right), \quad (29)$$

Where $D_{j,r} = \frac{k_j}{\alpha_j}\left(\frac{y_r}{\alpha_j}\right)^{k_j-1} exp\left(-\left(\frac{y_r}{\alpha_j}\right)^{k_j}\right)$, $E_{j,r} = \prod_{\substack{i=1\\ j\neq i}}^{n}\left(1-exp\left(-\left(\frac{y_r}{\alpha_j}\right)^{k_j}\right)\right)$, and

$S_r = \sum_{j=1}^{n} D_{j,r}E_{j,r}.\ \forall j = 1,,,n.$

The parameter estimates under MLE are computed numerically using the Newton- Raphson method by solving equations (28) and (29) after equating to zero. To find the MLE $\hat{\theta}$, iterate θ using

$$\theta^{(t+1)} = \theta^{(t)} + J^{-1}\left(\theta^{(t)}\right).\nabla l\left(\theta^{(t)}\right) \quad (30)$$

Where $\nabla l(\theta)$ is the gradient vector from the partial derivatives (25) and (26) and $J(\theta)$ is the negative Hessian of the log-likelihood. That is,

$$J(\theta) = -\left(\frac{\partial^2 l}{\partial\theta_i\partial\theta_j}\right)_{i,j};\ \text{i,j =1,..n and } \frac{\partial^2 l}{\partial\theta_i\partial\theta_j} \approx \frac{\partial l(\theta_i+h,\theta_j)-\partial l(\theta_i-h,\theta_j)}{2h} \quad (31)$$

Thus, the 95% confidence interval for each parameter θ_i is expressed as

$$C.I = \hat{\theta}_i \pm 1.96\sqrt{J^{-1}(\theta)} \quad (32)$$

4.2 LSE

The least square objective of GEW denoted as $S(\theta)$, is defined by

$$S(\theta) = \sum_{j=1}^{N}\left(F_Y(y_j)-\frac{j}{N+1}\right)^2. \quad (33)$$

Substituting the CDF (5) in (33), the least square objective of GEW denoted as $S(\theta)$, is expressed as

$$S(\theta) = \sum_{j=1}^{N}\left(\prod_{i=1}^{n}\left(1-exp\left(-\left(\frac{y_j}{\alpha_i}\right)^{k_i}\right)\right)-\frac{j}{N+1}\right)^2. \quad (34)$$

The LSEs [23], $\widehat{\alpha_l}$ and $\widehat{k_l}$ can be determined by minimizing equation (34). Hence, by numerically solving the nonlinear equations (35) and (36) using the Newton-Raphson method, which are set equal to zero, we obtain the parameter estimates

$$\frac{\partial S}{\partial \alpha_l} = 2\sum_{j=1}^{N}\left(\prod_{i=1}^{n}\left(1-exp\left(-\left(\frac{y_j}{\alpha_i}\right)^{k_i}\right)\right)-\frac{j}{N+1}\right)E_{i,j}\left(-\frac{k_l}{\alpha_l}\right)\left(\frac{y_j}{\alpha_l}\right)^{k_l}exp\left(-\left(\frac{y_j}{\alpha_l}\right)^{k_l}\right). \quad (35)$$

$$\frac{\partial S}{\partial k_l} = 2\sum_{j=1}^{N}\left(\prod_{i=1}^{n}\left(1-exp\left(-\left(\frac{y_j}{\alpha_i}\right)^{k_i}\right)\right)-\frac{j}{N+1}\right)E_{i,j}\left(\frac{y_j}{\alpha_l}\right)^{k_l}ln\left(\frac{y_j}{\alpha_l}\right)^{k_l}exp\left(-\left(\frac{y_j}{\alpha_l}\right)^{k_l}\right). \quad (36)$$

where $E_{i,j} = \prod_{l\neq i}\left(1-exp\left(-\left(\frac{y_j}{\alpha_i}\right)^{k_i}\right)\right) = \sum_{k=0}^{n-1}(-1)^k\sum_{S\subseteq\{1,..,n\}/\{i\}}exp\left(-\sum_{l\in S}\left(\frac{y_j}{\alpha_l}\right)^{k_l}\right)$,, |S|=k,

$\forall l = 1,..,n.$

4.3 WLSE

The weighted least square objective function of GEW is denoted as $S_w(\theta)$ and is defined as

$$S_w(\theta) = \sum_{j=1}^{N} w_j \left(F_Y(y_j) - \frac{j}{N+1}\right)^2. \quad (37)$$

Substituting (5) in (37), the weighted least square objective function of GEW is denoted as $S_w(\theta)$ and is given by the following

$$S_w(\theta) = \sum_{j=1}^{N} w_j \left(\prod_{i=1}^{n}\left(1 - exp\left(-\left(\frac{y_j}{\alpha_i}\right)^{k_i}\right)\right) - \frac{j}{N+1}\right)^2, \quad (38)$$

where w_j is the weighted least squares fit one gives at each empirical point y_j given by

$$w_j = \frac{(N+1)^2}{j(N+1-j)}. \quad (39)$$

The WLSE [23] estimates $\widehat{\alpha_l}$ and $\widehat{k_l}$ can be determined by the minimizing (39), hence by solving numerically the nonlinear equations (40) & (41) using the Newton- Raphson method, which are set equal to zero we can obtain the parameter estimates

$$\frac{\partial S_w}{\partial \alpha_l} = 2\sum_{j=1}^{N} w_j \left(\prod_{i=1}^{n}\left(1 - exp\left(-\left(\frac{y_j}{\alpha_i}\right)^{k_i}\right)\right) - \frac{j}{N+1}\right) E_{i,j}\left(-\frac{k_l}{\alpha_l}\right)\left(\frac{y_j}{\alpha_l}\right)^{k_l} exp\left(-\left(\frac{y_j}{\alpha_l}\right)^{k_l}\right). \quad (40)$$

$$\frac{\partial S_w}{\partial k_l} = 2\sum_{j=1}^{N} w_j \left(\prod_{i=1}^{n}\left(1 - exp\left(-\left(\frac{y_j}{\alpha_i}\right)^{k_i}\right)\right) - \frac{j}{N+1}\right) E_{i,j}\left(\frac{y_j}{\alpha_l}\right)^{k_i} ln\left(\frac{y_j}{\alpha_l}\right)^{k_l} exp\left(-\left(\frac{y_j}{\alpha_l}\right)^{k_L}\right). \quad (41)$$

Where $E_{i,j} = \prod_{l \neq i}\left(1 - exp\left(-\left(\frac{y_j}{\alpha_I}\right)^{k_I}\right)\right) = \sum_{k=0}^{n-1}(-1)^k \sum_{S\subseteq\{1,..,n\}/\{i\}} exp\left(-\sum_{l\in S}\left(\frac{y_j}{\alpha_l}\right)^{k_l}\right),$, |S|=k,

$\forall l = 1,..,n$

4.4 CVME

The Cramer–von Mises objective of GEW is denoted as $C(\theta)$ is defined by

$$C(\theta) = \frac{1}{12N} + \sum_{j=1}^{N}\left(F_Y(y_j) - \frac{2j-1}{2N}\right)^2. \quad (42)$$

The Cramer–von Mises objective of GEW is denoted as $C(\theta)$ is given by

$$C(\theta) = \frac{1}{12N} + \sum_{j=1}^{N}\left(\prod_{i=1}^{n}\left(1 - exp\left(-\left(\frac{y_j}{\alpha_i}\right)^{k_i}\right)\right) - \frac{2j-1}{2N}\right)^2. \quad (43)$$

The CVME [24] estimators $\widehat{\alpha_l}$ and $\widehat{k_l}$ are determined by minimizing (43), this is achieved by solving numerically the nonlinear equations (44) and (45) using the Newton-Raphson method, which are set equal to zero

$$\frac{\partial C}{\partial \alpha_l} = 2\sum_{j=1}^{N}\left(\prod_{i=1}^{n}\left(1 - exp\left(-\left(\frac{y_j}{\alpha_i}\right)^{k_i}\right)\right) - \frac{2j-1}{2N}\right) E_{i,j}\left(-\frac{k_l}{\alpha_l}\right)\left(\frac{y_j}{\alpha_l}\right)^{k_l} exp\left(-\left(\frac{y_j}{\alpha_l}\right)^{k_l}\right). \quad (44)$$

$$\frac{\partial C}{\partial k_l} = 2\sum_{j=1}^{N}\left(\prod_{i=1}^{n}\left(1 - exp\left(-\left(\frac{y_j}{\alpha_i}\right)^{k_i}\right)\right) - \frac{2j-1}{2N}\right) E_{i,j}\left(\frac{y_j}{\alpha_l}\right)^{k_i} ln\left(\frac{y_j}{\alpha_l}\right)^{k_l} exp\left(-\left(\frac{y_j}{\alpha_l}\right)^{k_l}\right). \quad (45)$$

Where $E_{i,j} = \prod_{l \neq i}\left(1 - exp\left(-\left(\frac{y_j}{\alpha_i}\right)^{k_i}\right)\right) = \sum_{k=0}^{n-1}(-1)^k \sum_{S\subseteq\{1,..,n\}/\{i\}} exp\left(-\sum_{l\in S}\left(\frac{y_j}{\alpha_l}\right)^{k_l}\right)$, |S|=k,

$\forall l = 1,..,n.$

5. Simulation Analysis

The accuracy of the estimation methods is investigated using the parameters $(\alpha_1 = 0.5, k_1 = 1.2), (\alpha_2 = 1, k_2 = 1.5), (\alpha_3 = 1.5, k_3 = 2)$. Furthermore, the effect of changes in the sample size N on the GEW parameter estimates is examined. The estimated parameters for each estimator, $\hat{\theta}$, are assessed using measures of accuracy, namely, the bias, the variance and the root mean squared error (RMSE), which is calculated respectively by using the expressions:

$Bias(\hat{\theta}) = \frac{1}{R}\sum_{i=1}^{R}(\hat{\theta}_i - \theta)$

$Var(\hat{\theta}) = \frac{1}{R-1}\sum_{i=1}^{R}\left(\hat{\theta}_i - \bar{\hat{\theta}}\right)^2$ where $\bar{\hat{\theta}} = \frac{1}{R}\sum_{i=1}^{R}\hat{\theta}_i$

and

$RMSE(\hat{\theta}) = \sqrt{\frac{\sum_{i=1}^{R}(\hat{\theta}_i - \theta_i)^2}{R}} = \sqrt{Var(\hat{\theta}) + Bias(\hat{\theta})^2}$

Where R: Simulation Number, R=1000

θ_i: Actual value of the i^{th} observation

$\hat{\theta}_i$: Estimated Value of i^{th} observation

$\bar{\hat{\theta}}$: Mean of the estimated values

Once the bias and variance of each estimated parameter are obtained from the simulation results, the RMSE of each estimated parameter is computed for different sample sizes N and reported in table 3

Table 3: Parameter Estimates for Different *N*

Sample Size		**MLE (Est)**	**MLE (RMSE)**	**LSE (Est)**	**LSE (RMSE)**	**WLSE (Est)**	**WLSE (RMSE)**	**CVME (Est)**	**CVME (RMSE)**
N = 25	α_1	0.5683	0.3168	0.5462	0.4135	0.5529	0.4089	0.5488	0.3917
	α_2	1.0337	0.3394	1.0715	0.3752	1.0637	0.3500	1.0661	0.3287
	α_3	1.4279	0.1817	1.4638	0.1987	1.4571	0.1978	1.4619	0.2017
	k_1	1.0602	2.2724	1.1274	2.4797	1.1124	2.2715	1.1198	2.2336
	k_2	1.3465	1.2609	1.4226	1.4448	1.4362	1.2212	1.4301	1.5042
	k_3	2.0618	0.5195	2.0841	0.6141	2.0928	0.5472	2.0887	0.6162
N = 50	α_1	0.5294	0.2719	0.5368	0.3430	0.5416	0.3361	0.5392	0.3220
	α_2	0.9795	0.2571	1.0369	0.2469	1.0294	0.2322	1.0348	0.2335
	α_3	1.4418	0.1225	1.4742	0.1491	1.4695	0.1356	1.4711	0.1438
	k_1	1.1447	2.1541	1.1583	2.4306	1.1492	2.1581	1.1531	2.0628
	k_2	1.5578	0.8516	1.4691	1.0767	1.4638	0.8435	1.4657	1.2672
	k_3	2.0793	0.3357	2.0527	0.3825	2.0607	0.3649	2.0579	0.4198
N = 100	α_1	0.5126	0.2310	0.5247	0.2859	0.5281	0.2700	0.5312	0.2626
	α_2	1.0089	0.1801	1.0186	0.1668	1.0142	0.1520	1.0201	0.1602
	α_3	1.4735	0.0885	1.4839	0.0994	1.4862	0.0946	1.4827	0.1057
	k_1	1.1864	1.9299	1.1719	2.0582	1.1658	1.9934	1.1627	1.8820
	k_2	1.5132	0.5172	1.4923	0.7424	1.4951	0.5554	1.4939	0.6413
	k_3	2.0281	0.2264	2.0398	0.2731	2.0449	0.2387	2.0421	0.2630
N = 500	α_1	0.5038	0.1399	0.5074	0.1497	0.5049	0.1318	0.5041	0.1430
	α_2	1.0112	0.0660	1.0061	0.0712	1.0054	0.0707	1.0060	0.0672
	α_3	1.4926	0.0445	1.4948	0.0435	1.4959	0.0416	1.4949	0.0454
	k_1	1.1971	0.4920	1.1957	0.3976	1.1974	0.3864	1.1979	0.3978
	k_2	1.4956	0.1443	1.4979	0.2142	1.4988	0.1880	1.4985	0.2186
	k_3	2.0114	0.0924	2.0068	0.1177	2.0061	0.1050	2.0052	0.1194
N = 1000	α_1	0.5019	0.1016	0.5027	0.1053	0.5018	0.0955	0.5015	0.1043
	α_2	1.0047	0.0475	1.0031	0.0487	1.0042	0.0467	1.0049	0.0488
	α_3	1.4981	0.0297	1.4970	0.0307	1.4976	0.0306	1.4979	0.0314
	k_1	1.1985	0.3142	1.1989	0.3063	1.1991	0.2985	1.1990	0.2878
	k_2	1.5026	0.0995	1.5012	0.1492	1.5006	0.1271	1.5007	0.1515
	k_3	2.0043	0.0652	2.0037	0.0829	2.0047	0.0730	2.0041	0.0819

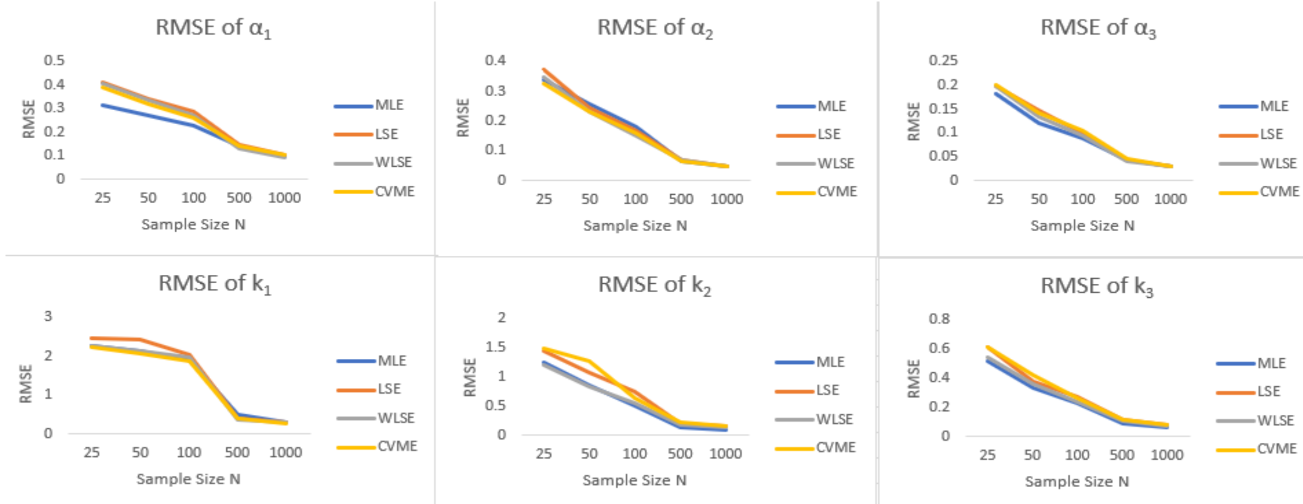


Fig. 4: RMSE of the Estimated Parameters of the GEW Distribution at Different Sample Sizes

Mathematica software was used to simulate parameter estimates using various estimation methods across different sample sizes. The accuracy of the methods was assessed using the root mean squared error (RMSE), as presented in Table 3. The results show that as N increases, RMSE consistently decreases, indicating that a larger sample size improves the accuracy of the estimates, regardless of the estimation method used. Overall, for the proposed model, the maximum likelihood estimation (MLE) method proved to be the most effective, yielding the most accurate parameter estimates. This trend is visually illustrated in Figure 4, which demonstrates the consistency of the estimators as N increases.

6. Monte Carlo Simulation

Simulation is applied to estimate the distribution of the GEW random variables using different parameter listed in the set $S = \{(\alpha_1 = 0.5, k_1 = 1.2), (\alpha_2 = 1, k_2 = 1.5), (\alpha_3 = 1.5, k_3 = 2)\}$. Moreover, the GEW's probability density function of this set is constructed using simulations and presented in Figure 5. This graph suggests that the GEW distribution follows a recognizable pattern and fits into theoretical assumptions reasonably well. The Monte Carlo method represents a reliable tool for finding extreme values in Weibull-distributed data because it approximates actual distribution rather well.

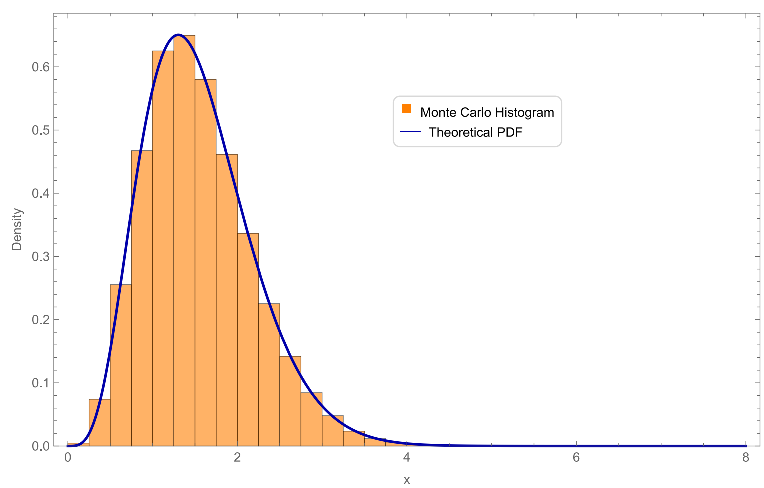


Fig. 5: Theoretical and Simulated PDFs of the GEW in the Set S

7. Applications

Two real datasets of cancer patients, adopted from [9], are utilized to evaluate the goodness of fit of our distribution. For simplicity and without loss of generality, we consider the GEW distribution with $n = 3$ for the non-identical parameters case. Moreover, the fit of our novel GEW distribution is compared against existing distributions, such as the Alpha Power Kumaraswamy Weibull distribution (APKumW) [9], the Weibull distribution [19], the Exponentiated Generalized Weibull (EGW) distribution [25], the Beta Weibull (BW) distribution [26], the KumW distribution [27], the Exponentiated

Kumaraswamy Weibull (EKumW) distribution [28], and the Alpha Power Weibull (APW) distribution [29]. The results show that the data are closely fitted by GEW distribution with n = 3, outperforming the competing distributions and highlighting its significance and applicability. The efficiency of-fit for each competing distribution is assessed using the AIC and the KS goodness-of-fit test.

Using (10) the pdf of GEW distribution for n = 3 is given by:

$$f_Y(x) = \sum_{j=1}^{3}\left(\frac{k_j}{\alpha_j}\left(\frac{t}{\alpha_j}\right)^{k_j-1} exp\left(-\left(\frac{t}{\alpha_j}\right)^{k_j}\right)\right)\sum_{k=0}^{2}(-1)^k\left(\sum_{\substack{1\le i_1<\cdots<i_k\le 3\\ i_l\neq j}} \exp\left(-\sum_{l=1}^{k}\left(\left(\frac{t}{\alpha_{i_l}}\right)^{k_{i_l}}\right)\right)\right). \quad (46)$$

The novel distribution is employed to model real-life data sourced from [9]. For the purpose of estimation, only records with fully observed (uncensored) data were included.

Based on Prior studies, [40, 41], the performance of these fitted models is assessed. Three important statistical measures are mainly, the Akaike Information Criterion (AIC) the Kolmogorov-Smirnov statistic, and the accompanying p-value. The AIC provides a means of model selection in which lower values indicate a better fit while penalizing model complexity. The KS statistic captures the largest difference between the empirical and theoretical distribution functions with smaller values the better agreement. The p-value of the KS test represents its statistical significance. The larger the p-value, usually greater than 0.05, the better the model fits to the data.

7.1 Acute Bone Cancer Data

The dataset comprises the survival times (in days) of 73 patients detected with acute bone cancer, obtained from [30]. The observed values are as follows:

“0.09, 0.76, 1.81, 1.10, 3.72, 0.72, 2.49, 1.00, 0.53, 0.66, 31.61, 0.60, 0.20, 1.61, 1.88, 0.70, 1.36, 0.43, 3.16, 1.57, 4.93, 11.07, 1.63, 1.39, 4.54, 3.12, 86.01, 1.92, 0.92, 4.04, 1.16, 2.26, 0.20, 0.94, 1.82, 3.99, 1.46, 2.75, 1.38, 2.76, 1.86, 2.68, 1.76, 0.67, 1.29, 1.56, 2.83, 0.71, 1.48, 2.41, 0.66, 0.65, 2.36, 1.29, 13.75, 0.67, 3.70, 0.76, 3.63, 0.68, 2.65, 0.95, 2.30, 2.57, 0.61, 3.93, 1.56, 1.29, 9.94, 1.67, 1.42, 4.18, 1.37.”

Table 4 displays the maximum likelihood estimates (MLEs) for both the GEW distribution and the competing models, which were used to model acute bone cancer data. The GEW distribution was found to fit the data best, as indicated by its AIC value (286.2710) and KS statistic (0.0647), both of which were the lowest among all tested models, and its P-value (0.9194), which was the highest.

Table 4: Parameters Estimation for the Acute Bone Cancer Data

Distribution	**$\hat{\theta}$ MLE Estimated Parameters**	**AIC**	**KS**	**P Value**
GEW	$\hat{\theta} = \begin{pmatrix}\alpha_1 = 0.1055;\ \alpha_2 = 0.1769;\ \alpha_3 = 1.789;\\ k_1 = 0.2440;\ k_2 = 0.6415;\ k_3 = 1.4766\end{pmatrix}$	286.2710	0.0647	0.9194
APKumW	$\hat{\theta} = \begin{pmatrix}\alpha = 0.0046;\ a = 5.0887;\ b = 0.4137;\\ c = 0.5358;\ \lambda = 1.3007\end{pmatrix}$	291.7005	0.0680	0.8888
Weibull	$\hat{\theta} = (c = 0.7656;\ \lambda = 2.9260)$	326.8033	0.1887	0.0111
EGW	$\hat{\theta} = \begin{pmatrix}a = 2.7262;\ b = 80.5514;\ c = 0.2353;\\ \lambda = 0.15070\end{pmatrix}$	294.0796	0.0924	0.5612
BW	$\hat{\theta} = \begin{pmatrix}a = 59.2646;\ b = 62.394;\ c = 0.1262;\\ \lambda = 41.4347\end{pmatrix}$	298.9643	0.0988	0.4747
KumW	$\hat{\theta} = \begin{pmatrix}a = 2.7498;\ b = 0.3506;\ c = 0.6483;\\ \lambda = 0.3447\end{pmatrix}$	311.4273	0.1470	0.0853
EKumW	$\hat{\theta} = \begin{pmatrix}a = 1.273;\ b = 1.9974;\ c = 0.4002;\\ \lambda = 1.0345;\ \alpha = 5.4855\end{pmatrix}$	302.2774	0.1168	0.2720
APW	$\hat{\theta} = (c = 0.9218;\ \lambda = 0.0791;\ \alpha = 0.0021)$	309.0348	0.1884	0.0112

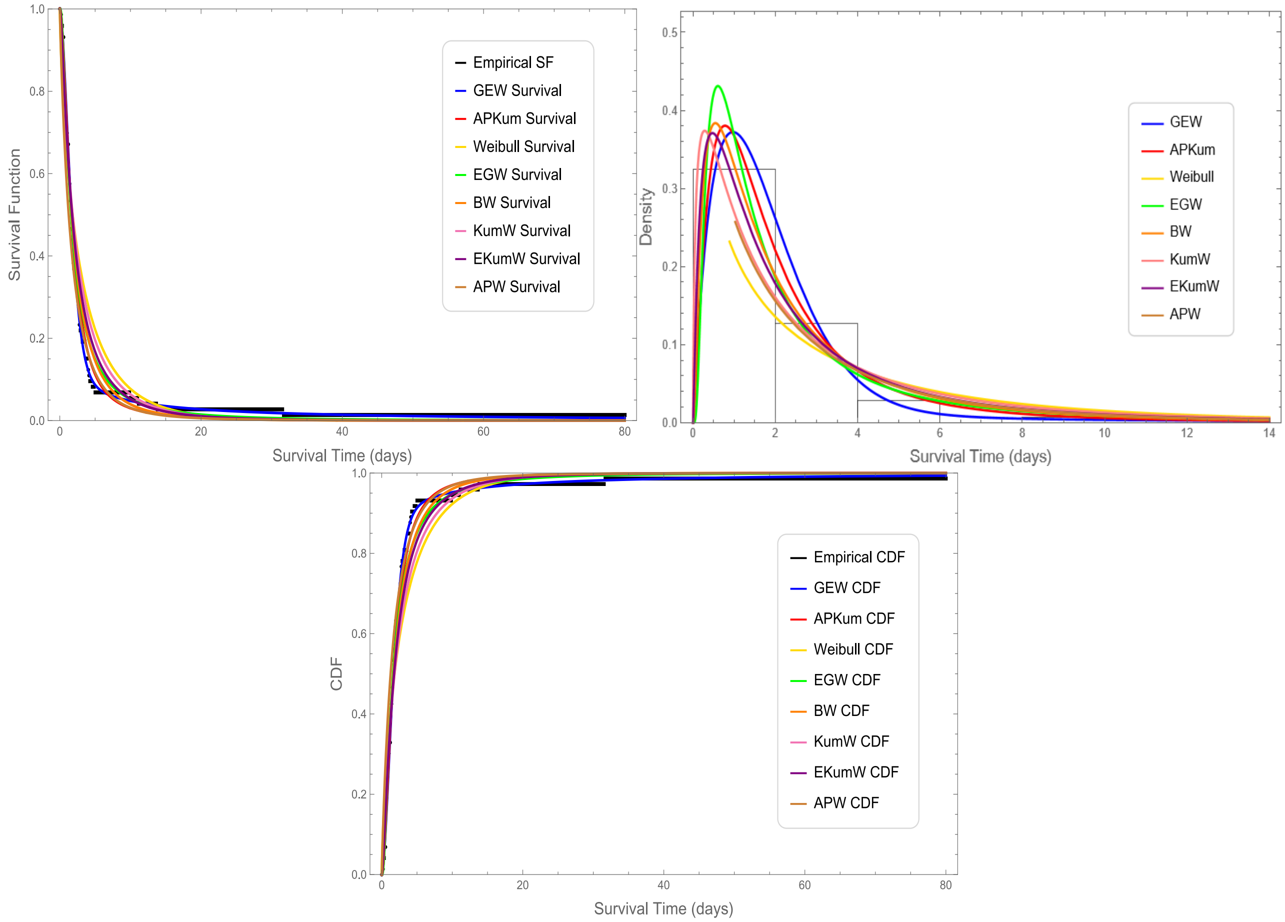


Fig. 6: CDFs, PDFs and Survival functions of Acute Bone Cancer Data

7.2 Head and Neck Cancer Data

Survival time of a sample of 44 patients having Head and Neck cancer disease patients recently examined by [31] is presumed. The information is: "12.20, 23.56, 23.74, 25.87, 31.98, 37, 41.35, 47.38, 55.46, 58.36, 63.47, 68.46, 78.26, 74.47, 81.43, 84, 92, 94, 110, 112, 119, 127, 130, 133, 140, 146, 155, 159, 173, 179, 194, 195, 209, 249, 281, 319, 339, 432, 469, 519, 633, 725, 817, 1776".

Table 5 displays the maximum likelihood estimates (MLEs) for both the GEW distribution and the competing models, which were used to model the cancer data for head and neck. The GEW distribution clearly provides the best fit with the lowest AIC (564.3977), lowest KS statistic (0.0476), and very high P-value (0.9999), demonstrating a high fit with the empirical data.

Table 5: Parameters Estimation for the Head and Neck Cancer Data

Distribution	$\hat{\theta}$ MLE Estimated Parameters	AIC	KS	P Value
GEW	$\hat{\theta} = \begin{pmatrix} \alpha_1 = 81.5693;\ \alpha_2 = 11.9971;\ \alpha_3 = 132.8173; \\ k_1 = 1.3325;\ k_2 = 1.2705;\ k_3 = 0.6132 \end{pmatrix}$	564.3977	0.0476	0.9999
APKumW	$\hat{\theta} = \begin{pmatrix} a = 4.8428;\ b = 0.6523;\ c = 0.5719; \\ \lambda = 0.0260;\ \alpha = 0.3173 \end{pmatrix}$	565.1112	0.0751	0.4435
Weibull	$\hat{\theta} = (c = 0.9386;\ \lambda = 213.6881)$	567.6877	0.1267	0.4435
EGW	$\hat{\theta} = \begin{pmatrix} a = 0.0784;\ b = 1.6582;\ c = 0.3437; \\ \lambda = 0.0513 \end{pmatrix}$	602.3591	0.2687	0.0027

BW	$\hat{\theta}=\begin{pmatrix} a=2.3728;\ b=0.0759;\ c=0.3999; \\ \lambda=0.2225 \end{pmatrix}$	602.5257	0.3075	0.0003
KumW	$\hat{\theta}=\begin{pmatrix} \mathrm{a}=0.3532;\ \mathrm{b}=0.0782;\ \mathrm{c}=0.4911; \\ \lambda=1.2577 \end{pmatrix}$	604.0207	0.3142	0.0002
EKumW	$\hat{\theta}=\begin{pmatrix} a=7.8983;\ b=6.8318;\ c=0.2203; \\ \lambda=0.0541;\ \alpha=1.8148 \end{pmatrix}$	566.0263	0.0973	0.7625
APW	$\hat{\theta}=(c=0.8779;\ \lambda=0.0105;\ \alpha=1.6918)$	570.2769	0.1277	0.4342

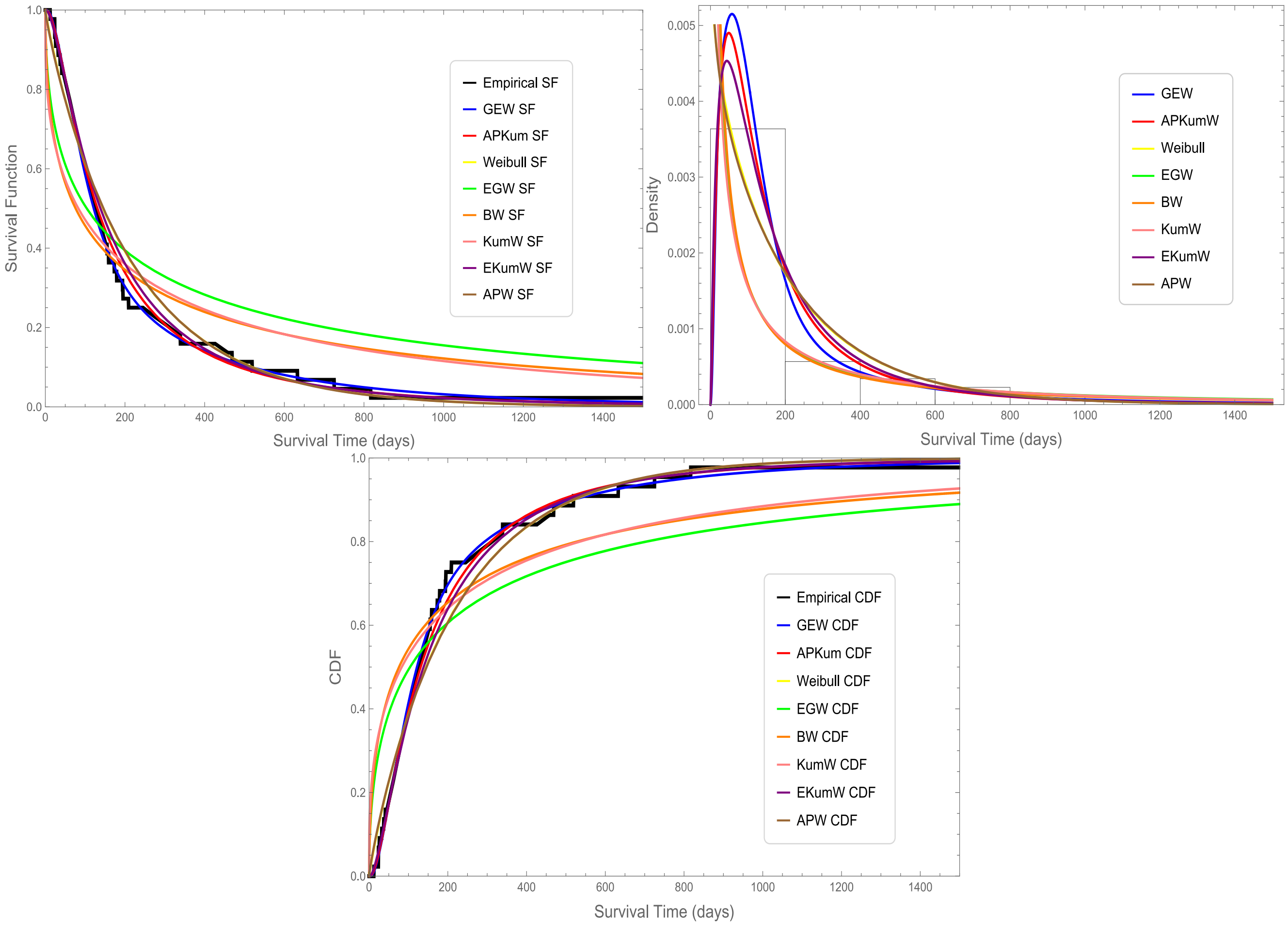


Fig. 7- CDFs, PDFs and Survival Functions for the Head and Neck Cancer Data

The Generalized Extreme Weibull (GEW) distribution demonstrated the best overall fit among all the competing models. For each dataset analyzed, the GEW distribution yielded the lowest AIC and KS values with the highest p-values, demonstrating a superior goodness of fit (Tables 4–5). Moreover, the graphical analyses presented in Figures 6–7 further support this conclusion.

8. Conclusion

A new generalization arising from Weibull distributions, the Generalized Extreme Weibull distribution (GEW), is formulated. The statistical properties of our novel distribution are investigated, including PDF, CDF, mean, mode, moment generating function, quartiles, and others. Furthermore, the order statistics are evaluated. The parameters are estimated using four estimation methods, namely, MLE, LSE, WLSE, and CVME. Numerical analysis is employed to evaluate the accuracy of the estimation methods. Two cancer data sets were considered to evaluate the suitability of this distribution in modeling such data. The findings underline the effectiveness of the GEW distribution as a practical and reliable option for survival analysis, particularly in medical and reliability studies characterized by complex hazard rate structures.

Acknowledgments:

The authors would like to express their sincere gratitude to the editor and anonymous reviewers for their constructive comments and careful review, which have significantly improved the quality of this manuscript.

Conflicts of Interest:

The authors state that there is no conflict of interest regarding the publication of this article.

References

[1] R. Kollu, S. R. Rayapudi, S. V. L. Narasimham, and K. M. Pakkurthi, Mixture probability distribution functions to model wind speed distributions, *Int. J. Energy Environ. Eng.,* **3**, 27 (2012).

[2] N. Yassine, Inventory planning under supplier uncertainty in a two-level supply chain, *Int. J. Logist. Manag.,* **34**, 497–516 (2023).

[3] N. Yassine, The EOQ model with items of imperfect quality and replenishment from different suppliers, *Int. J. Oper. Res.,* **40**, 117–136 (2021).

[4] N. Yassine and A. Al Sagheer, Optimizing a collaborative two-layer supply chain with probabilistic quality, *WSEAS Trans. Bus. Econ.,* **14**, 145–154 (2017).

[5] N. Yassine, A sustainable economic production model: Effects of quality and emissions tax from transportation, *Ann. Oper. Res.,* **290**, 73–94 (2020).

[6] N. Yassin and S. El Rabih, Assembling components with probabilistic lead times, *IOP Conf. Ser. Mater. Sci. Eng.,* **521**, 012025 (2019).

[7] M. N. Alshahrani, Half logistic Zeghdoudi distribution: Statistical properties, estimation, simulation and applications, *J. Stat. Appl. Probab.,* **14**, 165–182 (2025).

[8] J. Akan-nyaatemi, D. Jakperik, and M. Akamba, Bayesian survival analysis of recovery for tuberculosis patients in Bulsa Traditional Area, *J. Stat. Appl. Probab.,* **13**, 53–61 (2024).

[9] H. S. Klakattawi, Survival analysis of cancer patients using a new extended Weibull distribution, *PLoS One,* **17**, e0264229 (2022).

[10] T. Ghazal, B. Tashkandy, M. Balgoun, and M. Bakr, Exponentiated extended extreme value distribution: Properties, estimation, and applications in applied fields, *AIMS Math.,* **9**, 17634–17656 (2024).

[11] M. Kumar, S. K. Maurya, S. K. Singh, U. Singh, and A. Pathak, Model suitability analysis of survival time to ovarian cancer patients data, *J. Stat. Appl. Probab.,* **9**, 609–620 (2020).

[12] M. A. Selim and H. M. Salem, The exponentiated generalized power generalized Weibull distribution: Properties and applications, *J. Stat. Appl. Probab.,* **12**, 19–38 (2023).

[13] N. Yassine, Joint probability distribution and the minimum of a set of normalized random variables, *Procedia Soc. Behav. Sci.,* **230**, 235–239 (2016).

[14] N. Yassine, A. AlSagheer, and N. Azzam, A bundling strategy for items with different quality based on functions involving the minimum of two random variables, *Int. J. Eng. Bus. Manag.,* **10**, 1–12 (2018).

[15] Yassine N, Singh SK. Sustainable supply chains based on supplier selection and HRM practices. *J. Enterp. Inf. Manag.* **34**, 399–426 (2021).

[16] T. Hitchen and S. Nadarajah, Exact results for the distribution of randomly weighted sums, *Mathematics,* **12**, 149 (2024).

[17] A.M. Basheer, E. M. Almetwally, and H. M. Okasha, Marshall–Olkin alpha power inverse Weibull distribution: Non-Bayesian and Bayesian estimations, *J. Stat. Appl. Probab.,* **10**, 327–345 (2021).

[18] E. M. Almetwally, H. H. Ahmad, and H. M. Almongy, Extended odd Weibull Pareto distribution: Estimation and applications, *J. Stat. Appl. Probab.,* **11**, 795–809 (2022).

[19] W. Weibull, A statistical distribution function of wide applicability, *J. Appl. Mech.,* **18**, 293–297 (1951).

[20] M. M. Al Sobhi, The extended Weibull distribution with its properties, estimation and modeling skewed data, *J. King Saud Univ. Sci.,* **34**, 101801 (2022).

[21] G. Grimmett and D. Stirzaker, *Probability and Random Processes*, 3rd ed., Oxford University Press, UK, (2001).

[22] G. H. Lemon, Maximum likelihood estimation for the three-parameter Weibull distribution based on various left, right, and interval censored sampling, *Technometrics,* **17**, 247–254 (1975).

[23] J. J. Swain, S. Venkatraman, and J. R. Wilson, Least-squares estimation of distribution functions in Johnson's translation system, *J. Stat. Comput. Simul.,* **29**, 271–297 (1988).

[24] A.Luceo, Fitting the generalized Pareto distribution to data using maximum goodness-of-fit estimators, *Comput. Stat. Data Anal.,* **51**, 904–917 (2006).

[25] G. M. Cordeiro, E. M. Ortega, and D. C. da Cunha, The exponentiated generalized class of distributions, *J. Data Sci.,* **11**, 1–27 (2013).

[26] F. Famoye, C. Lee, and O. Olumolade, The beta-Weibull distribution, *J. Stat. Theory Appl.,* **4**, 121–136 (2005).

[27] G. M. Cordeiro, E. M. Ortega, and S. Nadarajah, The Kumaraswamy–Weibull distribution with application to failure data, *J. Franklin Inst.,* **347**, 1399–1429 (2010).

[28] F. H. Eissa, The exponentiated Kumaraswamy–Weibull distribution with application to real data, *Int. J. Stat. Probab.,* **6**, 167–182 (2017).

[29] M. Nassar, A. Alzaatreh, M. Mead, and O. Abo-Kasem, Alpha power Weibull distribution: Properties and applications, *Commun. Stat. Theory Methods,* **46**, 10236–10252 (2017).

[30] M. Mansour, H. M. Yousof, W. A. Shehata, and M. Ibrahim, A new two-parameter Burr XII distribution: Properties, copula, different estimation methods and modeling acute bone cancer data, *J. Nonlinear Sci. Appl.,* **13**, 223–238 (2020).

[31] I.Sule, S. I. Doguwa, A. Isah, and H. M. Jibril, Topp–Leone Kumaraswamy-G family of distributions with applications to cancer disease data, *J. Biostat. Epidemiol.,* **6**, 37–48 (2020).

[32] B. Efron, Logistic regression, survival analysis, and the Kaplan–Meier curve, *J. Am. Stat. Assoc.,* **83**, 414–425 (1988).

[33] E. T. Lee and J. W. Wang, *Statistical Methods for Survival Data Analysis*, 3rd ed., John Wiley & Sons, NJ, USA, (2003).

[34] R. Hibatullah, Y. Widyansingh, and S. Abdullah, Marshall–Olkin extended power Lindley distribution with application, *J. Ris. Apl. Mat.,* **2**, 84–92 (2018).

[35] Z. A. Al-Saiary and R. A. Bakoban, The Topp–Leone generalized inverted exponential distribution with real data applications, *Entropy,* **22**, 1144 (2020).

[36] N. Balakrishnan, E. Castilla, M. Jaenada, and L. Pardo, Robust inference for destructive one-shot device test data under Weibull lifetimes and competing risks, *J. Comput. Appl. Math.,* **439**, 115701 (2024).

[37] H. Fu, D. Nicolet, K. Mrózek, R. M. Stone, A.-K. Eisfeld, J. C. Byrd, and K. J. Archer, Controlled variable selection in Weibull mixture cure models for high-dimensional data, *Stat. Med.,* **41**, 4340–4366 (2022).

[38] X. Li, P. Lin, and H. Wang, Weibull parametric model for survival analysis in women with endometrial cancer using clinical and T2-weighted MRI radiomic features, *BMC Med. Res. Methodol.,* **24**, 2234 (2024).

[39] M. G. M. Ghazal, A new extension of the modified Weibull distribution with applications for engineering data, *Probab. Eng. Mech.,* **74**, 103523 (2023).

[40] K. P. Burnham and D. R. Anderson, *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd ed., Springer, New York, NY, USA, (2002).

[41] M. A. Stephens, EDF statistics for goodness of fit and some comparisons, *J. Am. Stat. Assoc.,* **69**, 730–737 (1974).