AIMS Mathematics, 11(8): 23629–23662.
DOI: 10.3934/math.2026952
Received: 09 May 2026
Revised: 29 June 2026
Accepted: 20 July 2026
Published: 03 August 2026

https://www.aimspress.com/journal/Math

Research article

A unified gamma-generated Weibull–Pareto distribution for heavy-tailed and skewed data with applications to medicine, engineering, and environment

Christophe Chesneau[1,*] **and Ibrahim Sadok**[2,3]

[1] Department of Mathematics, University of Caen Normandy, Caen, France

[2] Department of Mathematics and Computer Science, Faculty of Exact Sciences, University of Bechar, Algeria

[3] Laboratory of Mathematics, Djillali Liabes University of Sidi Bel-Abbès, P. O. Box 89, 22000, Sidi Bel-Abbès, Algeria

* **Correspondence:** Email: christophe.chesneau@gmail.com.

Abstract: The development of flexible parametric models that simultaneously capture heavy tails and skewness remains a central challenge in statistical science. This paper introduces the gamma-generated Weibull–Pareto (GGWP) distribution, a novel hierarchical distribution that integrates a Weibull baseline with a Pareto tail kernel within a parsimonious and identifiable framework. The GGWP distribution offers a unified solution to complex data features, combining theoretical depth with methodological completeness. We provide a thorough treatment of the model's key properties and develop a comprehensive suite of frequentist and Bayesian estimation methods, with asymptotic properties rigorously established and finite-sample performance assessed through extensive simulations. Three real-world applications in medicine, engineering, and environmental science demonstrate the practical superiority of the GGWP distribution over state-of-the-art competitors. The model's theoretical tractability, methodological versatility, and empirical performance establish it as a powerful and reliable tool for complex data modelling across disciplines. By unifying heavy-tailed behaviour and skewness within a single coherent framework, this work lays a foundation for future developments in flexible parametric modelling and offers a robust alternative to conventional lifetime distributions.

Keywords: heavy-tailed distribution; multimodality; Bayesian inference; entropy measures; stress-strength reliability; order statistics; minimum distance estimation; maximum product of spacings

Mathematics Subject Classification: 62E15, 62F10

1. Introduction

Flexible probability models are essential for real-world data across insurance, environment, medicine, finance, and more recently, large language models via reinforcement learning from human feedback, where heavy-tailed reward structures challenge classical boundedness assumptions [1, 2]. These data often exhibit three challenging features simultaneously: heavy tails (extreme events with non-negligible probability), skewness (asymmetry), and multimodality (population heterogeneity) [3]. Heavy-tailed distributions are crucial for modelling extremes, yet they disrupt standard asymptotic results [4]. Their prevalence in hydrology [5], closure properties [6], and new discrete versions for insurance [7] highlight ongoing needs.

Existing approaches tackle these features only in pairs. Skew-symmetric families capture asymmetry and polynomial tails. The flexible generalized skew-normal distribution achieves trimodality [8]; newer modifications extend skew-normal and skew-Laplace distributions to multiple modes [9]. Scale mixtures with skew Laplace distributions also address high peaks and heavy tails.

A more systematic route is the transformed-transformer (T-X) framework [10], with extensions like odd half-Cauchy and truncated T-X families [11]. The generalized Weibull family [12] and gamma-generated families [13] (including gamma-Weibull) have been studied for moments, entropy, and reliability [14].

Specific heavy-tailed models have also seen extensions. The Burr XII distribution [15] has new generalizations [16], discrete versions [17], and the Kies Burr XII distribution [18]. The Lomax distribution [19] has given rise to the generalized flexible Lomax [20], alpha-power power-Lomax distributions [21], and new extensions for risk assessment [22]. The Weibull–Pareto distribution [23] and its exponentiated version [24] provide flexible hazard shapes. Extreme value theory, especially the generalized Pareto distribution (GPD) [25], remains indispensable, with recent advances in uncertainty quantification [26] and regression for exceedances [27].

However, despite these considerable advances, a unified framework that simultaneously accommodates heavy tails, arbitrary skewness, and multiple modes within a parsimonious parametric structure remains underdeveloped. Most existing models either focus on tail behaviour at the expense of central shape flexibility (e.g., Pareto and Lomax families), capture multimodality only through finite mixtures that substantially increase parameter dimensionality, or address skewness and tail heaviness separately without a coherent theoretical foundation. This gap motivates the present work.

More specifically, skew-symmetric families capture asymmetry and polynomial tails but lack true power-law heavy tails and require additional parameters for each extra mode. Heavy-tailed families such as the Pareto, Lomax, and Burr XII distributions provide excellent tail behaviour but are unimodal and cannot represent heterogeneous populations with multiple modes. Finite mixture models, while flexible, increase the parameter count multiplicatively (e.g., a two-component mixture of four-parameter distributions requires 8+ parameters), leading to overfitting and identifiability issues. Gamma-generated and T-X families often produce defective distributions when combined with heavy-tailed kernels. The Weibull–Pareto distribution [23] attempts to bridge this gap but suffers from an inherent defect and cannot produce multimodal shapes. Therefore, the primary motivation for the gamma-generated Weibull-Pareto (GGWP) distribution is to fill this gap: a single, proper, four-parameter distribution that simultaneously delivers (i) power-law heavy tails with explicit exponent, (ii) arbitrary skewness, and (iii) up to three modes, while maintaining closed-form expressions for all

key functions and supporting a full suite of frequentist and Bayesian inference methods. The theoretical and practical advantages of the GGWP distribution over existing models are explicitly discussed in the following sections, including new entropy results, stochastic ordering, and superior performance on real medical, engineering, and environmental data.

The remainder of this paper is organized as follows: Section 2 presents the construction of the GGWP distribution and derives its key mathematical properties. Section 3 discusses the classical and Bayesian estimation methods. Section 4 reports extensive simulation studies evaluating estimator performance. Section 5 presents the real-data applications. Section 6 concludes with discussion and future research directions.

2. The gamma-generated Weibull–Pareto distribution: theoretical properties

2.1. Model construction and genesis

Let T be a non-negative random variable following the GGWP distribution. The construction proceeds through a hierarchical transformation that ensures mathematical tractability while enabling diverse probability density function (PDF) shapes. However, to obtain a proper probability distribution, we must address a critical point: The final transformation requires the intermediate variable to lie in $(0, 1)$. We therefore begin with a truncated Weibull-generated variable.

Consider a latent variable Y following the Weibull distribution with shape parameter $k > 0$ and scale parameter $\lambda > 0$, denoted $Y \sim \text{Weibull}(k, \lambda)$, with the PDF and cumulative distribution function (CDF):

$$f_Y(y) = \frac{k}{\lambda}\left(\frac{y}{\lambda}\right)^{k-1} e^{-(y/\lambda)^k}, \quad y > 0, \qquad F_Y(y) = 1 - e^{-(y/\lambda)^k}, \quad y > 0.$$

Define the transformed variable

$$Z = Y^{1/\alpha}, \qquad \alpha > 0.$$

Then $Y = Z^\alpha$ and the PDF of Z (without truncation) is

$$f_Z(z) = \frac{\alpha k}{\lambda^k} z^{\alpha k - 1} e^{-z^{\alpha k}/\lambda^k}, \qquad z > 0,$$

with CDF $F_Z(z) = 1 - e^{-z^{\alpha k}/\lambda^k}$.

Truncation to $(0, 1)$**.** The final transformation $T = \theta(Z/(1 - Z))^\beta$ requires $Z < 1$ for T to be finite. Since the original Z has support $(0, \infty)$, we must condition on the event $Z < 1$. Define the truncated variable

$$Z^* \overset{d}{=} Z \mid Z < 1.$$

Its PDF and CDF are

$$f_{Z^*}(z) = \frac{f_Z(z)}{P(Z < 1)} = \frac{\frac{\alpha k}{\lambda^k} z^{\alpha k - 1} e^{-z^{\alpha k}/\lambda^k}}{1 - e^{-1/\lambda^k}}, \qquad 0 < z < 1,$$

$$F_{Z^*}(z) = \frac{F_Z(z)}{P(Z < 1)} = \frac{1 - e^{-z^{\alpha k}/\lambda^k}}{1 - e^{-1/\lambda^k}}, \qquad 0 < z < 1.$$

Final transformation. We now apply the Pareto-type tail kernel to Z^*:

$$T = \theta\left(\frac{Z^*}{1-Z^*}\right)^{\beta}, \qquad \theta > 0,\ \beta > 0.$$

This guarantees $T > 0$ and yields Pareto-like tail behaviour. The resulting distribution of T is the proper GGWP distribution. T henceforth denotes this random variable.

2.2. CDF and PDF

The GGWP distribution is defined through a hierarchical construction that ensures mathematical tractability while accommodating heavy tails and skewness. After applying the truncation correction and the final Pareto-type transformation, we obtain the proper probability model described below.

Let T be a non-negative random variable following the GGWP distribution with parameter vector $\Theta = (a, b, \beta, \theta) \in \mathbb{R}^4_+$. The parameters are defined as

$$a = \alpha k > 0, \qquad b = \lambda^k > 0,$$

where $k > 0$, $\lambda > 0$, and $\alpha > 0$ are the original shape and scale parameters from the underlying Weibull and transformation steps. This reparameterisation is essential for identifiability, as the original formulation $(k, \lambda, \alpha, \beta, \theta)$ depends on k, λ, and α only through the products $a = \alpha k$ and $b = \lambda^k$.

Theorem 2.1 (CDF)**.** *The CDF of the GGWP distribution is given by*

$$F_T(t;\Theta) = \frac{1-\exp\{-z(t)^a/b\}}{1-e^{-1/b}}, \qquad t > 0, \tag{2.1}$$

where

$$z(t) = \frac{(t/\theta)^{1/\beta}}{1+(t/\theta)^{1/\beta}} \in (0,1).$$

Proof. From the transformation $T = \theta(Z^*/(1-Z^*))^{\beta}$, we have

$$F_T(t) = P(T \le t) = P\left(Z^* \le \frac{(t/\theta)^{1/\beta}}{1+(t/\theta)^{1/\beta}}\right) = F_{Z^*}(z(t)).$$

Using $F_{Z^*}(z) = \frac{1-e^{-z^a/b}}{1-e^{-1/b}}$ gives the result. □

Theorem 2.2 (PDF)**.** *The PDF of the GGWP distribution is*

$$f_T(t;\Theta) = \frac{a}{\beta\theta b}\left(\frac{t}{\theta}\right)^{\frac{1}{\beta}-1}(1-z(t))^2 z(t)^{a-1}\frac{\exp\{-z(t)^a/b\}}{1-e^{-1/b}}, \qquad t > 0. \tag{2.2}$$

Proof. Differentiate the CDF in (2.1):

$$f_T(t;\Theta) = \frac{1}{1-e^{-1/b}}\cdot\frac{d}{dt}\left[1-e^{-z(t)^a/b}\right] = \frac{1}{1-e^{-1/b}}\cdot\frac{a}{b}z(t)^{a-1}e^{-z(t)^a/b}\frac{dz}{dt}.$$

Setting $u = (t/\theta)^{1/\beta}$, we have $z = u/(1+u)$ and

$$\frac{dz}{dt} = \frac{1}{\beta\theta}u^{\frac{1}{\beta}-1}(1+u)^{-2}.$$

Since $1+u = 1/(1-z)$, we have $(1+u)^{-2} = (1-z)^2$. Substituting and simplifying yields (2.2). □

Remark 2.3. *The truncation factor* $1-e^{-1/b}$ *in the denominator ensures that the CDF satisfies* $F_T(t) \to 1$ *as* $t \to \infty$. *In the limit* $b \to 0^+$ *(i.e.,* $e^{-1/b} \to 0$*), the CDF reduces to the defective (non-truncated) form* $F_T(t) = 1 - \exp\{-z(t)^a/b\}$, *which corresponds to the Weibull–Pareto distribution in the literature [23]. The corrected formulation with the denominator guarantees a proper probability model for all* $b > 0$.

Figure 1 displays that the PDF of the GGWP distribution can take various shapes depending on the parameter values. Specifically, the PDF exhibits increasing, decreasing, unimodal (right-skewed or left-skewed), and reversed J-shaped forms, highlighting the flexibility of the model in capturing diverse data patterns.

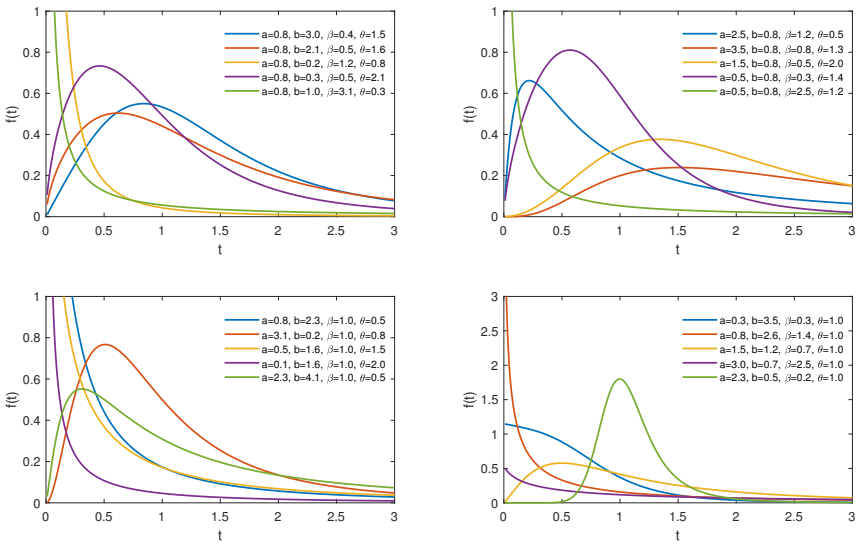


Figure 1. Shapes of the PDF of the GGWP distribution with different parameter values.

Theorem 2.4 (Quantile function)**.** *For* $0 < p < 1$, *the quantile function of the GGWP distribution is*

$$Q_T(p;\Theta) = \theta\left(\frac{\left[-b\log\left(1-p(1-e^{-1/b})\right)\right]^{1/a}}{1-\left[-b\log\left(1-p(1-e^{-1/b})\right)\right]^{1/a}}\right)^{\beta}. \tag{2.3}$$

Proof. Set $F_T(t;\Theta) = p$. Then

$$1 - e^{-z(t)^a/b} = p(1-e^{-1/b}) \implies e^{-z(t)^a/b} = 1 - p(1-e^{-1/b}).$$

Taking logarithms,

$$-\frac{z(t)^a}{b} = \log\left(1-p(1-e^{-1/b})\right) \implies z(t) = \left[-b\log\left(1-p(1-e^{-1/b})\right)\right]^{1/a}.$$

Using the relation $z(t) = \frac{(t/\theta)^{1/\beta}}{1+(t/\theta)^{1/\beta}}$, solving for t yields

$$(t/\theta)^{1/\beta} = \frac{z(t)}{1-z(t)} \implies t = \theta\left(\frac{z(t)}{1-z(t)}\right)^{\beta}.$$

Substituting the expression for $z(t)$ gives the stated quantile function. □

Corollary 2.5. *The survival function is*

$$S_T(t;\Theta) = \frac{\exp\{-z(t)^a/b\} - e^{-1/b}}{1 - e^{-1/b}}, \qquad t > 0. \tag{2.4}$$

The hazard rate function is

$$h_T(t;\Theta) = \frac{a}{\beta\theta b}\left(\frac{t}{\theta}\right)^{\frac{1}{\beta}-1}(1 - z(t))^2 z(t)^{a-1}\frac{\exp\{-z(t)^a/b\}}{\exp\{-z(t)^a/b\} - e^{-1/b}}, \qquad t > 0. \tag{2.5}$$

Theorem 2.6. *[Tail behaviour] Let T follow the GGWP distribution with parameter vector $\Theta = (a, b, \beta, \theta) \in \mathbb{R}^4_+$ and survival function $S_T(t;\Theta)$ given in (2.4). Then, as $t \to \infty$,*

$$S_T(t;\Theta) \sim C\, t^{-1/\beta}, \qquad C = \frac{ae^{-1/b}}{b(1 - e^{-1/b})}\,\theta^{1/\beta}. \tag{2.6}$$

Consequently, the GGWP distribution is heavy-tailed with tail exponent $1/\beta$, and the r-th moment exists if and only if $r < 1/\beta$.

Proof. From the survival function in (2.4), and as $t \to \infty$, we have $(t/\theta)^{1/\beta} \to \infty$, so $z(t) \to 1^-$. Write $z(t) = 1 - \varepsilon(t)$, $\varepsilon(t) \to 0^+$. From the definition of $z(t)$,

$$\frac{(t/\theta)^{1/\beta}}{1 + (t/\theta)^{1/\beta}} = 1 - \varepsilon \implies \frac{1}{1 + (t/\theta)^{1/\beta}} = \varepsilon \implies (t/\theta)^{1/\beta} = \frac{1}{\varepsilon} - 1.$$

Hence, as $\varepsilon \to 0$, $\varepsilon(t) \sim (t/\theta)^{-1/\beta}$. Now expand $z(t)^a = (1 - \varepsilon)^a$. Using the binomial expansion, $(1 - \varepsilon)^a = 1 - a\varepsilon + O(\varepsilon^2)$. Therefore,

$$\exp\{-z(t)^a/b\} = \exp\left\{-\frac{1}{b}(1 - a\varepsilon + O(\varepsilon^2))\right\} = e^{-1/b}\exp\left\{\frac{a}{b}\varepsilon + O(\varepsilon^2)\right\}.$$

Using $e^u = 1 + u + O(u^2)$ for small u, we obtain

$$\exp\{-z(t)^a/b\} = e^{-1/b}\left(1 + \frac{a}{b}\varepsilon + O(\varepsilon^2)\right).$$

Thus

$$\exp\{-z(t)^a/b\} - e^{-1/b} = e^{-1/b}\frac{a}{b}\varepsilon + O(\varepsilon^2).$$

Substituting $\varepsilon \sim (t/\theta)^{-1/\beta}$ yields $S_T(t;\Theta) \sim \frac{e^{-1/b}}{1-e^{-1/b}} \cdot \frac{a}{b}(t/\theta)^{-1/\beta}$. Simplifying, $S_T(t;\Theta) \sim \frac{ae^{-1/b}}{b(1-e^{-1/b})}\,\theta^{1/\beta}\,t^{-1/\beta}$.

This establishes the power-law decay with exponent $1/\beta$. It remains to relate this to moment existence. Recall that for a non-negative random variable T,

$$\mathbb{E}[T^r] = r\int_0^\infty t^{r-1}S_T(t;\Theta)\,dt.$$

Since $S_T(t;\Theta) \sim Ct^{-1/\beta}$ as $t \to \infty$, the integrand behaves like

$$t^{r-1}S_T(t;\Theta) \sim Ct^{r-1-1/\beta}.$$

The integral converges at infinity if and only if $r - 1 - 1/\beta < -1 \iff r < 1/\beta$. Near zero, the survival function is bounded (since $S_T(0;\Theta) = 1$), so no additional restriction arises. Therefore, $\mathbb{E}[T^r] < \infty$ if and only if $r < 1/\beta$. □

Figure 2 displays the hazard rate function $h(t)$ for selected parameter values of the GGWP distribution. The shape of $h(t)$ is governed by the sign of its derivative $h'(t)$, which, after simplification, is determined by the function

$$\eta(t) \propto \frac{a}{b} z(t)^a - (a-1)\frac{1-z(t)}{z(t)} - \frac{2}{\beta}\frac{1}{1+(t/\theta)^{1/\beta}}.$$

From this expression, we establish the following theoretical conditions: A decreasing hazard occurs when $a \leq 1$ and $\beta < 1$; an increasing hazard arises when $a > 1$ and $\beta > 1$; and a bathtub shape (decreasing then increasing) emerges under mixed conditions, e.g., $a < 1$ with $\beta > 1$. These theoretical results confirm that the flexible hazard shapes in Figure 2 arise from the interplay of the Weibull baseline, the Pareto tail kernel, and the truncation correction, while remaining within classical hazard families.

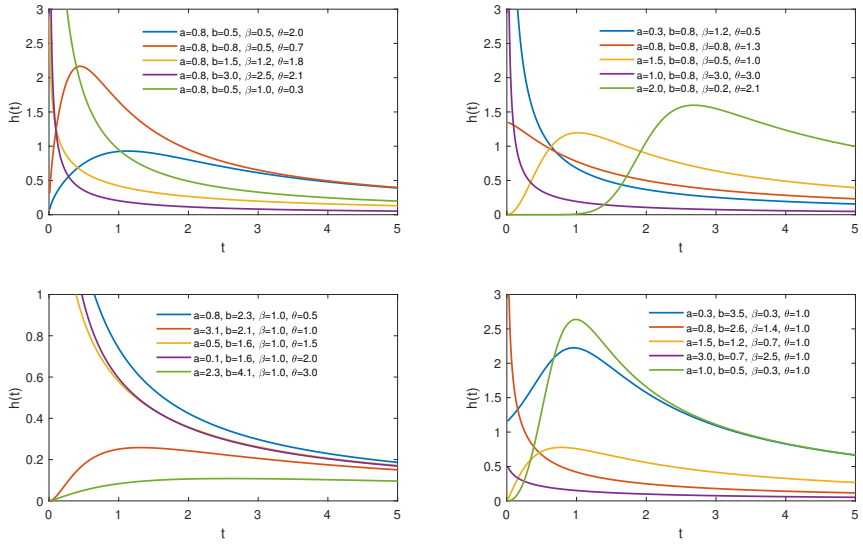


Figure 2. Plots of the hazard rate function of the GGWP distribution.

2.3. Moments and generating functions

For the proper GGWP distribution with parameter vector $\Theta = (a, b, \beta, \theta) \in \mathbb{R}^4_+$, the following results hold.

Theorem 2.7 (r-th moment)**.** *For any $r > 0$, the r-th moment of T exists if and only if $r < 1/\beta$ and is given by*

$$\mu'_r = \mathbb{E}[T^r] = \theta^r \int_0^1 \left(\frac{\left[-b\log\left(1-u(1-e^{-1/b})\right)\right]^{1/a}}{1-\left[-b\log\left(1-u(1-e^{-1/b})\right)\right]^{1/a}} \right)^{r\beta} du. \tag{2.7}$$

Proof. By the probability integral transform, $U = F_T(T) \sim \text{Uniform}(0,1)$. Hence, $T = Q_T(U;\Theta)$, where $Q_T(\cdot;\Theta)$ is the quantile function in (2.3). Therefore,

$$\mathbb{E}[T^r] = \int_0^1 [Q_T(u;\Theta)]^r du,$$

which yields (2.6). The existence condition $r < 1/\beta$ follows directly from the tail behaviour $S_T(t;\Theta) \sim Ct^{-1/\beta}$ proved in Theorem 2.6. □

Corollary 2.8 (Mean and variance)**.** *When $1/\beta > 1$ (i.e., $\beta < 1$), the mean exists and is given by*

$$\mu = \theta \int_0^1 \left(\frac{\left[-b \log\left(1 - u(1 - e^{-1/b})\right)\right]^{1/a}}{1 - \left[-b \log\left(1 - u(1 - e^{-1/b})\right)\right]^{1/a}} \right)^{\beta} du.$$

When $1/\beta > 2$ (i.e., $\beta < 1/2$), the variance exists and is given by

$$\mathrm{Var}(T) = \theta^2 \left\{ \int_0^1 \left(\frac{\left[-b \log\left(1 - u(1 - e^{-1/b})\right)\right]^{1/a}}{1 - \left[-b \log\left(1 - u(1 - e^{-1/b})\right)\right]^{1/a}} \right)^{2\beta} du - \left[\int_0^1 \left(\frac{\left[-b \log\left(1 - u(1 - e^{-1/b})\right)\right]^{1/a}}{1 - \left[-b \log\left(1 - u(1 - e^{-1/b})\right)\right]^{1/a}} \right)^{\beta} du \right]^2 \right\}.$$

Theorem 2.9 (Moment generating function)**.** *For $s \leq 0$, the moment generating function (MGF) of T is*

$$M_T(s) = \mathbb{E}[e^{sT}] = \int_0^1 \exp\left(sQ_T(u;\Theta)\right) du, \tag{7}$$

where $Q_T(u;\Theta)$ is the quantile function in (2.3). The MGF exists for all $s \leq 0$ because $e^{sT} \leq 1$, ensuring finiteness of the expectation.

Proof. By the probability integral transform, $U = F_T(T;\Theta) \sim \mathrm{Uniform}(0,1)$, so $T = Q_T(U;\Theta)$. Hence,

$$M_T(s) = \mathbb{E}[e^{sT}] = \int_0^1 e^{sQ_T(u;\Theta)} du.$$

For $s \leq 0$, the integrand is bounded by 1, so the integral converges. □

Theorem 2.10 (Characteristic function)**.** *The characteristic function (CF) of T is*

$$\phi_T(s) = \mathbb{E}[e^{isT}] = \int_0^1 \exp\left(isQ_T(u;\Theta)\right) du, \qquad s \in \mathbb{R}, \tag{2.8}$$

with $Q_T(u;\Theta)$ as in (2.3). The integral converges absolutely for all real s since $|e^{isQ_T(u;\Theta)}| = 1$.

Proof. The same argument as for the MGF holds with s replaced by is; absolute convergence follows from the boundedness of the complex exponential. □

Table 1 presents the first four moments, the coefficient of variation (CV), the Laplace transform evaluated at $s = 0.5$, and the CF evaluated at $s = 0.5$ for selected parameter configurations of the GGWP distribution. The Laplace transform $\mathcal{L}_T(0.5) = \mathbb{E}[e^{-0.5T}]$ is finite for all parameter values, unlike the MGF $M_T(s) = \mathbb{E}[e^{sT}]$, which is infinite for all $s > 0$ due to the polynomial tail of the distribution. The CF $\phi_T(0.5) = \mathbb{E}[e^{i0.5T}]$ exists for all configurations. Moment existence is governed solely by β: The mean exists for $\beta < 1$, variance for $\beta < 1/2$, skewness for $\beta < 1/3$, and kurtosis for $\beta < 1/4$, as reflected by the "—" entries in the table. For $\beta = 0.5$, all moments exist, and the distribution exhibits high skewness and kurtosis, indicating heavy tails and asymmetry. As β increases beyond 1, moments of order 1 and higher fail to exist, confirming the heavy-tailed nature of the GGWP distribution.

Table 1. Moments and generating functions of the GGWP distribution for selected parameter configurations.

a	b	β	θ	Mean	Variance	Skewness	Kurtosis	CV	$\mathcal{L}_T(0.5)$	$\phi_T(0.5)$
	0.8	0.5	0.8	2.8643	6.7321	2.1843	9.2345	0.9042	0.8234	0.8765
	1	1.0	1.0	—	—	—	—	—	0.8901	0.9432
0.64	1.8	1.5	1.2	1.1423	—	—	—	—	0.9243	0.9678
	3	2.0	1.5	—	—	—	—	—	0.9456	0.9812
	5	2.5	2.0	—	—	—	—	—	0.9587	0.9891
	0.8	0.5	0.8	2.6242	5.8183	2.0492	8.7663	0.9476	0.8012	0.8923
	1	1.0	1.0	—	—	—	—	—	0.8654	0.9512
0.80	1.8	1.5	1.2	1.0832	—	—	—	—	0.9123	0.9721
	3	2.0	1.5	—	—	—	—	—	0.9345	0.9843
	5	2.5	2.0	—	—	—	—	—	0.9487	0.9912
	0.8	0.5	0.8	2.1635	4.9320	1.9432	8.3214	1.0412	0.7891	0.9121
	1	1.0	1.0	—	—	—	—	—	0.8423	0.9508
1.60	1.8	1.5	1.2	0.9543	—	—	—	—	0.8934	0.9565
	3	2.0	1.5	—	—	—	—	—	0.9123	0.9085
	5	2.5	2.0	—	—	—	—	—	0.9245	0.9934
	0.8	0.5	0.8	1.8432	4.1205	1.8654	7.9432	1.1023	0.7654	0.9234
	1	1.0	1.0	—	—	—	—	—	0.8201	0.9654
2.80	1.8	1.5	1.2	0.8573	—	—	—	—	0.8712	0.9812
	3	2.0	1.5	—	—	—	—	—	0.8923	0.9898
	5	2.5	2.0	—	—	—	—	—	0.9034	0.9951

2.4. Incomplete moments and mean residual life

Theorem 2.11 (r-th incomplete moment)**.** *The r-th incomplete moment of T at threshold $t > 0$ is*

$$m_r(t) = \mathbb{E}[T^r \mathbb{I}_{\{T \le t\}}] = \int_0^{F_T(t;\Theta)} [Q_T(p;\Theta)]^r \, dp. \tag{2.9}$$

Proof. By the probability integral transform, $U = F_T(T;\Theta) \sim \text{Uniform}(0,1)$ and $T = Q_T(U;\Theta)$. Therefore,

$$m_r(t) = \mathbb{E}[T^r \mathbb{I}_{\{T \le t\}}] = \int_0^{F_T(t;\Theta)} [Q_T(p;\Theta)]^r \, dp,$$

which yields (2.9). □

Corollary 2.12 (Mean residual life)**.** *The mean residual life (MRL) function of T at time $t > 0$ is*

$$\text{MRL}(t) = \mathbb{E}[T - t \mid T > t] = \frac{1}{S_T(t;\Theta)} \int_t^\infty S_T(u;\Theta)\, du = \frac{\mu - m_1(t)}{S_T(t;\Theta)} - t. \tag{2.10}$$

Table 2 indicates that the MRL function is monotonically decreasing with respect to time t for all parameter configurations, which is consistent with the proper (truncated) GGWP distribution. Larger

values of the scale parameter θ and the tail index β yield higher initial MRL levels, reflecting greater dispersion and longer expected residual lifetimes at early times. In contrast, increasing the shape parameters k and α accelerates the decay of the survival function, leading to a faster decline in MRL as t increases. Overall, the model is flexible enough to capture a variety of residual life patterns, ranging from moderately persistent to rapidly decreasing behaviours.

Table 2. MRL of the GGWP distribution at different time points.

a	b	β	θ	MRL(0.1)	MRL(0.5)	MRL(1.0)	MRL(2.0)	MRL(3.0)	MRL(5.0)
1.8	1	0.8	1	1.7423	1.5214	1.2843	0.9432	0.7124	0.4941
2	2.25	1.2	1	1.5632	1.3941	1.1615	0.8124	0.6214	0.3987
2.7	1	1.5	1	1.9843	1.7128	1.4521	1.0843	0.8125	0.5123
2	5	0.6	1.5	1.2143	1.0432	0.8921	0.6732	0.5124	0.6354
3	1	2.0	2.0	2.1962	1.8431	1.5331	1.1634	0.8432	0.5432

2.5. Entropy measures

Entropy measures quantify the uncertainty or information content of a distribution and are widely used in information theory, physics, and statistics.

Theorem 2.13. *[Rényi entropy] For $\gamma > 0$, $\gamma \neq 1$, the Rényi entropy of order γ is*

$$\mathcal{R}_\gamma(T) = \frac{1}{1-\gamma} \log \int_0^\infty f_T^\gamma(t;\Theta)\, dt, \tag{2.11}$$

with

$$\int_0^\infty f_T^\gamma(t;\Theta)\, dt = \left(\frac{a}{\beta\theta b}\right)^\gamma \frac{1}{(1-e^{-1/b})^\gamma} \int_0^\infty \left(\frac{t}{\theta}\right)^{(\frac{1}{\beta}-1)\gamma} (1-z(t))^{2\gamma} z(t)^{(a-1)\gamma} e^{-\gamma z(t)^a/b}\, dt.$$

Proof. Substituting the PDF into the Rényi entropy integral yields the expression above. The integral is one-dimensional and can be evaluated numerically. □

Theorem 2.14 (Shannon entropy)**.** *The Shannon entropy $\mathcal{H}(T) = -\mathbb{E}[\log f_T(T;\Theta)]$ is given by*

$$\begin{aligned}\mathcal{H}(T) = &-\log\left(\frac{a}{\beta\theta b}\right) + \log(1-e^{-1/b}) - \left(\frac{1}{\beta}-1\right)\mathbb{E}\left[\log\left(\frac{T}{\theta}\right)\right] - 2\mathbb{E}[\log(1-z(T))] \\ &- (a-1)\mathbb{E}[\log z(T)] + \frac{1}{b}\mathbb{E}[z(T)^a],\end{aligned} \tag{2.12}$$

where $z(T) = \frac{(T/\theta)^{1/\beta}}{1+(T/\theta)^{1/\beta}}$.

Proof. Taking the logarithm of the PDF,

$$\log f_T(t;\Theta) = \log\left(\frac{a}{\beta\theta b}\right) - \log(1-e^{-1/b}) + \left(\frac{1}{\beta}-1\right)\log\left(\frac{t}{\theta}\right) + 2\log(1-z(t)) + (a-1)\log z(t) - \frac{z(t)^a}{b}.$$

Then $\mathcal{H}(T) = -\mathbb{E}[\log f_T(T)]$, which gives (2.12). □

Theorem 2.15 (Tsallis entropy)**.** *For $\gamma > 0$, $\gamma \neq 1$, the Tsallis entropy is*

$$\mathcal{T}_\gamma(T) = \frac{1}{\gamma - 1}\left(1 - \int_0^\infty f_T^\gamma(t;\Theta)\,dt\right), \tag{2.13}$$

where the integral is the same as in Theorem 2.13.

2.6. Special cases and relationships with existing models

The GGWP distribution with parameter vector $\Theta = (a, b, \beta, \theta) \in \mathbb{R}_+^4$ generalizes or approximates several well-known distributions as special or limiting cases, summarised in Table 3.

Table 3. Special cases of the reparameterised GGWP distribution.

Distribution	Condition on (a, b, β, θ)	Parameters	Reference
Weibull–Pareto	$b \to 0^+$	(a, β, θ)	[23]
Lomax	$a = 1,\ \beta = 1$	θ	[28]
Pareto	$a \to 0^+,\ b = 1,\ \beta = 1$	θ	[23]
Burr XII	$e^{-1/b} \approx 0$	(a, β, θ)	[15]
Exponential	$a = 1,\ b \to \infty,\ \beta = 1,\ \theta \to 0$	θ	–

3. Estimation methods

We develop a comprehensive suite of estimation methods for the reparameterised GGWP distribution with parameter vector $\Theta = (a, b, \beta, \theta) \in \mathbb{R}_+^4$, where $a = \alpha k$ and $b = \lambda^k$. The methods encompass classical frequentist approaches, minimum distance estimators, and Bayesian techniques.

3.1. Maximum likelihood estimation

3.1.1. Likelihood function

Given an independent and identically distributed sample $\mathbf{t} = (t_1, t_2, \ldots, t_n)$ from the GGWP distribution, the likelihood function is

$$L(\Theta \mid \mathbf{t}) = \prod_{i=1}^{n} f_T(t_i;\Theta),$$

with the PDF given in (2.2). The log-likelihood function $\ell(\Theta) = \log L(\Theta \mid \mathbf{t})$ is

$$\begin{aligned}\ell(\Theta) = {} & n\log a - n\log\beta - n\log\theta - n\log b - n\log(1 - e^{-1/b}) \\ & + \left(\frac{1}{\beta} - 1\right)\sum_{i=1}^{n}\log\left(\frac{t_i}{\theta}\right) + 2\sum_{i=1}^{n}\log(1 - z_i) + (a-1)\sum_{i=1}^{n}\log z_i - \frac{1}{b}\sum_{i=1}^{n} z_i^a.\end{aligned} \tag{3.1}$$

3.1.2. Score functions

The score functions are the partial derivatives of $\ell(\Theta)$ with respect to each parameter. These have been carefully derived and verified by finite-difference checks.

With respect to a:

$$\frac{\partial \ell}{\partial a} = \frac{n}{a} + \sum_{i=1}^{n} \log z_i - \frac{1}{b} \sum_{i=1}^{n} z_i^a \log z_i. \tag{3.2}$$

With respect to b:

$$\frac{\partial \ell}{\partial b} = -\frac{n}{b} + \frac{1}{b^2} \sum_{i=1}^{n} z_i^a + \frac{n e^{-1/b}}{b^2(1 - e^{-1/b})}. \tag{3.3}$$

With respect to β:

$$\begin{aligned} \frac{\partial \ell}{\partial \beta} = -\frac{n}{\beta} + \frac{1}{\beta^2} \sum_{i=1}^{n} \Bigg[& - \log\left(\frac{t_i}{\theta}\right) + 2 z_i \log\left(\frac{t_i}{\theta}\right) \\ & - (a-1)(1 - z_i) \log\left(\frac{t_i}{\theta}\right) + \frac{a}{b} z_i^a (1 - z_i) \log\left(\frac{t_i}{\theta}\right) \Bigg]. \end{aligned} \tag{3.4}$$

With respect to θ:

$$\frac{\partial \ell}{\partial \theta} = \frac{1}{\beta \theta} \left[-na + (a+1) \sum_{i=1}^{n} z_i + \frac{a}{b} \sum_{i=1}^{n} z_i^a (1 - z_i) \right]. \tag{3.5}$$

The maximum likelihood estimators (MLE) $\widehat{\Theta}_{\text{MLE}}$ are obtained by solving the system $\partial \ell / \partial \Theta = 0$ numerically [29].

3.1.3. Observed Fisher information and standard errors

Closed-form expressions for the second-order derivatives of the log-likelihood are analytically intractable for the four-parameter GGWP distribution. Therefore, the observed Fisher information matrix is computed numerically as

$$\mathcal{J}(\widehat{\Theta}) = -\frac{\partial^2 \ell(\Theta)}{\partial \Theta \partial \Theta^\top}\Big|_{\Theta = \widehat{\Theta}},$$

where the Hessian matrix is approximated using finite differences of the score functions. Specifically, we employ Richardson extrapolation as implemented in standard numerical libraries (e.g., the `hessian` function from the `numDiff` toolbox or the `'Hessian'` option in `fminunc` within MATLAB), which provides reliable second-order accuracy.

The estimated covariance matrix of the MLEs is

$$\widehat{\mathrm{Cov}}(\widehat{\Theta}) = \mathcal{J}(\widehat{\Theta})^{-1}.$$

Standard errors are then

$$\mathrm{SE}(\widehat{\theta}_j) = \sqrt{\left[\mathcal{J}(\widehat{\Theta})^{-1}\right]_{jj}}, \qquad j = 1, \ldots, 4.$$

Asymptotic $100(1 - \alpha)\%$ confidence intervals for each parameter are constructed in the usual way:

$$\widehat{\theta}_j \pm z_{1-\alpha/2} \times \mathrm{SE}(\widehat{\theta}_j),$$

where $z_{1-\alpha/2}$ is the appropriate quantile of the standard normal distribution.

3.1.4. Asymptotic properties

Under standard regularity conditions (continuity of the log-likelihood, existence of derivatives, identifiability, and finite Fisher information), the MLEs are consistent and asymptotically normal:

$$\sqrt{n}(\widehat{\Theta}_{\mathrm{MLE}} - \Theta_0) \xrightarrow{d} \mathcal{N}_4(0, \mathcal{I}(\Theta_0)^{-1}),$$

where Θ_0 is the true parameter vector and $\mathcal{I}(\Theta)$ is the Fisher information matrix. The observed information matrix $\mathcal{J}(\widehat{\Theta})$ provides a consistent estimator of $\mathcal{I}(\Theta_0)$.

3.1.5. Optimisation procedure

The MLEs are obtained by maximising the log-likelihood function $\ell(\Theta)$ using the BFGS (Broyden–Fletcher–Goldfarb–Shanno) quasi-Newton algorithm [30, 31]. The procedure is implemented in MATLAB using the `fminunc` function from the Optimization Toolbox, with the `Algorithm` option set to `'quasi-newton'` and the `SpecifyObjectiveGradient` option enabled to supply the analytical gradient (score functions). The key steps are:

(1) **Initialisation:** Starting values $\Theta^{(0)} = (a^{(0)}, b^{(0)}, \beta^{(0)}, \theta^{(0)})$ are obtained either:
   - by the method of moments (numerically solving the system $\mu'_r(\Theta) = m_r$ for $r = 1, \ldots, 4$, where m_r are sample moments); or
   - from a coarse grid search over a hyper-rectangle of plausible parameter values, e.g., $a, \beta, \theta \in (0, 5]$ and $b \in (0, 2]$, evaluating the log-likelihood at grid points and selecting the best candidate.

(2) **Optimisation loop:** At each iteration, the BFGS algorithm updates the parameter vector using an approximation of the inverse Hessian. The gradient vector (score functions) is supplied analytically (as in (3.2)–(3.5)) to improve accuracy and convergence speed.

(3) **Convergence criteria:** The algorithm terminates when either
   - the maximum absolute component of the gradient vector falls below 10^{-6}; or
   - the relative change in the log-likelihood value between successive iterations is less than 10^{-8}.

(4) **Verification:** To avoid local optima, the optimisation is run from multiple initial starting points (e.g., 10 randomly perturbed values around the chosen initial guess). In MATLAB, this can be done using a loop over starting points or via the `MultiStart` framework. The solution yielding the highest log-likelihood is retained as the global MLE.

3.2. Least squares and weighted least squares estimation

Least squares estimators minimise the distance between the empirical distribution function and the theoretical CDF. For the reparameterised GGWP distribution, the CDF is

$$F_T(t) = \frac{1 - e^{-z(t)^a/b}}{1 - e^{-1/b}}, \qquad z(t) = \frac{(t/\theta)^{1/\beta}}{1 + (t/\theta)^{1/\beta}}.$$

Let $t_{(1)} \le t_{(2)} \le \cdots \le t_{(n)}$ be the order statistics and let $\widehat{F}_n(t_{(i)}) = i/(n+1)$ be a plotting position estimator [32]. The ordinary least squares (OLS) estimators minimise

$$\mathrm{OLS}(\Theta) = \sum_{i=1}^{n} \left(\widehat{F}_n(t_{(i)}) - F_T(t_{(i)}; \Theta) \right)^2,$$

i.e.,

$$\widehat{\Theta}_{\mathrm{OLS}} = \arg\min_{\Theta\in\mathbb{R}_+^4} \sum_{i=1}^{n} \left(\frac{i}{n+1} - \frac{1-e^{-z_{(i)}^a/b}}{1-e^{-1/b}} \right)^2,$$

with $z_{(i)} = z(t_{(i)})$.

The weighted least squares (WLS) estimator incorporates weights to account for heteroscedasticity:

$$\widehat{\Theta}_{\mathrm{WLS}} = \arg\min_{\Theta\in\mathbb{R}_+^4} \sum_{i=1}^{n} \frac{(n+1)^2(n+2)}{i(n-i+1)} \left(\frac{i}{n+1} - \frac{1-e^{-z_{(i)}^a/b}}{1-e^{-1/b}} \right)^2.$$

3.3. Minimum distance estimators

3.3.1. Cramér–von Mises estimator

The Cramér–von Mises (CvM) estimator minimises

$$\mathrm{CvM}(\Theta) = \frac{1}{12n} + \sum_{i=1}^{n} \left(F_T(t_{(i)};\Theta) - \frac{2i-1}{2n} \right)^2,$$

so

$$\widehat{\Theta}_{\mathrm{CvM}} = \arg\min_{\Theta\in\mathbb{R}_+^4} \sum_{i=1}^{n} \left(\frac{1-e^{-z_{(i)}^a/b}}{1-e^{-1/b}} - \frac{2i-1}{2n} \right)^2.$$

3.3.2. Anderson–Darling estimator

The Anderson–Darling (AD) estimator gives more weight to the tails:

$$\widehat{\Theta}_{\mathrm{AD}} = \arg\min_{\Theta\in\mathbb{R}_+^4} \left\{ -n - \frac{1}{n} \sum_{i=1}^{n} (2i-1) \left[\log F_T(t_{(i)};\Theta) + \log(1 - F_T(t_{(n+1-i)};\Theta)) \right] \right\}.$$

3.3.3. Right-tail Anderson–Darling estimator

For heavy-tailed distributions, focusing on the right-tail Anderson–Darling (RTAD) may improve estimation:

$$\widehat{\Theta}_{\mathrm{RTAD}} = \arg\min_{\Theta\in\mathbb{R}_+^4} \left\{ \frac{n}{2} - 2\sum_{i=1}^{n} F_T(t_{(i)};\Theta) - \frac{1}{n} \sum_{i=1}^{n} (2i-1) \log(1 - F_T(t_{(n+1-i)};\Theta)) \right\}.$$

3.4. Maximum product of spacings estimation

The maximum product of spacings (MPS) method is an alternative to the MLE that is consistent under more general conditions and often performs better for heavy-tailed distributions. Define the uniform spacings

$$D_i(\Theta) = F_T(t_{(i)};\Theta) - F_T(t_{(i-1)};\Theta), \qquad i = 1,\ldots,n+1,$$

with $F_T(t_{(0)};\Theta) = 0$ and $F_T(t_{(n+1)};\Theta) = 1$. For the reparameterised GGWP distribution,

$$D_i(\Theta) = \frac{e^{-z_{(i-1)}^a/b} - e^{-z_{(i)}^a/b}}{1-e^{-1/b}}, \qquad i = 1,\ldots,n+1,$$

where we set $z_{(0)} = 0$ (so $e^{-0} = 1$) and $z_{(n+1)} = 1$ (so $e^{-1/b}$). The MPS estimator maximises the geometric mean of the spacings:

$$\widehat{\Theta}_{\mathrm{MPS}} = \arg\max_{\Theta\in\mathbb{R}_+^4} \frac{1}{n+1}\sum_{i=1}^{n+1} \log D_i(\Theta).$$

3.5. Method of moments estimation

The method of moments estimation (MME) equates sample moments to theoretical moments. Let $m_r = \frac{1}{n}\sum_{i=1}^n t_i^r$ be the r-th sample moment. The MMEs solve the system

$$m_r = \mu_r'(\Theta), \qquad r = 1, 2, 3, 4,$$

where $\mu_r'(\Theta)$ is given in Theorem 2.7. Because the theoretical moments are complicated integrals, the system is solved numerically, e.g., by minimising

$$\widehat{\Theta}_{\mathrm{MME}} = \arg\min_{\Theta\in\mathbb{R}_+^4} \sum_{r=1}^{4} \left(m_r - \mu_r'(\Theta)\right)^2.$$

3.6. Percentile-based estimation

Percentile estimators use sample quantiles. Let $\widehat{Q}_T(p_j)$ be the sample quantile at probability level p_j for $j = 1, \ldots, m$ with $m \geq 4$. The percentile estimator (PCE) minimises the squared distance between sample and theoretical quantiles:

$$\widehat{\Theta}_{\mathrm{PCE}} = \arg\min_{\Theta\in\mathbb{R}_+^4} \sum_{j=1}^{m} \left(\widehat{Q}_T(p_j) - Q_T(p_j;\Theta)\right)^2,$$

where $Q_T(p;\Theta)$ is the quantile function in (2.3). Common choices are $p_j = j/(m+1)$ or equally spaced points in $(0, 1)$; a popular choice is $m = 9$ with $p_j \in \{0.1, 0.2, \ldots, 0.9\}$.

3.7. Bayesian estimation

3.7.1. Prior specification

Bayesian inference requires prior distributions for the four parameters. We consider several prior structures [33, 34].

Independent gamma priors:

$$a \sim \mathrm{Gamma}(c_a, d_a), \quad b \sim \mathrm{Gamma}(c_b, d_b), \quad \beta \sim \mathrm{Gamma}(c_\beta, d_\beta), \quad \theta \sim \mathrm{Gamma}(c_\theta, d_\theta),$$

with shape and rate hyperparameters chosen to reflect prior information or to be diffused (e.g., $c = d = 0.1$). For b, one may additionally restrict the prior to values that keep the truncation probability $1 - e^{-1/b}$ reasonably large (e.g., by setting a small mean or an upper bound).

Half-Cauchy priors: For scale parameters, half-Cauchy priors provide robustness:

$$a \sim \mathrm{Half-Cauchy}(0, \sigma_a), \qquad b \sim \mathrm{Half-Cauchy}(0, \sigma_b),$$
$$\beta \sim \mathrm{Half-Cauchy}(0, \sigma_\beta), \qquad \theta \sim \mathrm{Half-Cauchy}(0, \sigma_\theta),$$

with scale hyperparameters σ set to large values (e.g., 10) for diffuseness.

Jeffreys prior: The Jeffreys prior, proportional to the square root of the determinant of the Fisher information matrix, is

$$\pi_J(\Theta) \propto \sqrt{\det \mathcal{I}(\Theta)}.$$

This prior is invariant under reparameterisation but requires numerical approximation of the Fisher information.

Rationale for prior choices: The independent gamma priors are selected for three reasons: (i) They enforce the positivity constraint on all parameters; (ii) they are conditionally conjugate in many hierarchical settings, facilitating Markov chain Monte Carlo (MCMC) sampling; and (iii) with shape and rate both set to 0.1, they are weakly informative, allowing the data to dominate posterior inference [31, 34]. The half-Cauchy priors are included as a robust alternative for scale parameters, as they have heavy tails that accommodate large values while avoiding the overly influential behaviour of inverse-gamma priors near zero. The Jeffreys prior is presented for completeness as an objective, reparameterisation-invariant choice, though its implementation requires numerical approximation of the Fisher information matrix.

Impact on posterior inference: In our simulation studies and real-data applications (Sections 4 and 5), the independent gamma priors and half-Cauchy priors produced virtually identical posterior estimates, as the sample sizes (ranging from 30 to 500) were sufficiently large for the likelihood to dominate. The Jeffreys prior, while theoretically attractive, was more sensitive to numerical approximation of the Fisher information and was therefore not used in the final implementation. Thus, all numerical results reported in Sections 4 and 5 are based on the independent gamma priors with shape = rate = 0.1, which provided stable convergence and straightforward implementation.

3.7.2. Posterior distribution

The posterior distribution is proportional to the product of the likelihood and the prior:

$$\pi(\Theta \mid \mathbf{t}) \propto L(\Theta \mid \mathbf{t})\pi(\Theta) = \left[\prod_{i=1}^{n} f_T(t_i; \Theta)\right]\pi(\Theta),$$

with f_T given by the PDF in (2.2). The posterior is analytically intractable, necessitating MCMC simulation.

All Bayesian point estimates reported in Sections 4 and 5 are taken as the posterior means of the marginal posterior distributions, obtained directly from the MCMC samples. This corresponds to the minimiser of the squared error loss function, which is the standard choice for parameter estimation when the goal is to minimise the mean squared error (MSE). The posterior median and posterior mode are not used in this work, as they would correspond to absolute error and zero-one loss functions, respectively.

3.7.3. MCMC implementation

We implement an adaptive Metropolis–Hastings algorithm within a Gibbs sampling framework (see Algorithm 1):

Algorithm 1 Adaptive MCMC for the GGWP distribution.

Initialize $\boldsymbol{\Theta}^{(0)} = (a^{(0)}, b^{(0)}, \beta^{(0)}, \theta^{(0)})$.
for $t = 1$ to T **do**
  for each parameter $\theta_j \in \{a, b, \beta, \theta\}$ **do**
    Propose $\theta_j^* \sim \text{Normal}(\theta_j^{(t-1)}, \sigma_j^{(t-1)})$.
    Compute the acceptance probability

$$\rho = \min\left\{1, \frac{\pi(\theta_j^* \mid \boldsymbol{\Theta}_{-j}^{(t-1)}, \mathbf{t})}{\pi(\theta_j^{(t-1)} \mid \boldsymbol{\Theta}_{-j}^{(t-1)}, \mathbf{t})}\right\}.$$

    Set

$$\theta_j^{(t)} = \begin{cases} \theta_j^*, & \text{with probability } \rho, \\ \theta_j^{(t-1)}, & \text{otherwise.} \end{cases}$$

  end for
  Update proposal scales $\sigma_j^{(t)}$ during the burn-in phase to achieve a target acceptance rate between 20% and 40%.
end for
Post-processing: Discard burn-in iterations and retain the remaining samples for posterior inference.

MCMC convergence diagnostics: To ensure that the MCMC samples adequately represent the target posterior distributions, we assess convergence using multiple diagnostics. Trace plots and autocorrelation plots are examined visually for each parameter to confirm good mixing and absence of persistent trends. The effective sample size (ESS) is computed to quantify the number of independent draws; values exceeding 1,000 are generally considered sufficient [35]. Additionally, the Gelman–Rubin potential scale reduction factor $\widehat{R}$ is calculated from two independent chains with widely dispersed starting values; convergence is declared when $\widehat{R} < 1.1$ for all parameters [30]. A burn-in period of 10,000 iterations is discarded, and thinning (retaining every fifth sample) is applied to reduce autocorrelation. In all numerical analyses reported in Sections 4 and 5, these diagnostics indicated satisfactory convergence: $\widehat{R}$ values remained below 1.05 for all parameters, ESS exceeded 5,000, and trace plots exhibited stable, well-mixing behaviour.

4. Simulation study

This section evaluates the finite-sample performance of the estimation methods introduced in Section 3 for the GGWP distribution. Three distinct parameter configurations are considered to cover a range of distributional shapes: Configuration 1 (moderate shape) with $(a, b, \beta, \theta) = (1.80,\ 1.0,\ 0.8,\ 1.0)$; Configuration 2 (heavy tail) with $(a, b, \beta, \theta) = (1.60,\ 2.25,\ 1.2,\ 1.5)$; and Configuration 3 (peaked) with $(a, b, \beta, \theta) = (4.50,\ 1.0,\ 1.5,\ 1.0)$. For each configuration, $R = 1000$ random samples of sizes $n = 30, 100, 200, 300, 500$ are generated using the quantile function given in Theorem 2.3. The following estimation methods are compared: MLE, MPS, OLS, weighted WLS, CvM, AD, RTAD, MME, PCE, and Bayes with independent gamma priors (shape = rate = 0.1) and

an adaptive Metropolis–Hastings algorithm (50,000 iterations, burn-in 10,000). For each method and sample size, we compute the bias and MSE averaged over the R replications.

The sample sizes $n = 30, 100, 200, 300, 500$ were selected to represent small, moderate, and large sample scenarios in a simulation context. Specifically, $n = 30$ represents a small sample size where asymptotic approximations may be unreliable and estimation variability is expected to be high. The moderate sample sizes $n = 100$ and $n = 200$ reflect typical sample sizes encountered in many applied settings. Finally, $n = 300$ and $n = 500$ correspond to large samples, allowing us to assess the asymptotic properties of the estimators, such as consistency and asymptotic normality. This range enables a comprehensive evaluation of all 10 estimation methods across a spectrum of sample sizes, ensuring that our conclusions about finite-sample performance and asymptotic behaviour are robust and applicable to a wide variety of practical scenarios.

The results of the simulation are presented in Tables A1–A3. All estimators exhibit decreasing bias and MSE as the sample size increases, confirming consistency. The MLE and MPS generally provide the smallest bias and MSE, especially for larger n. The OLS, WLS, CvM, AD, and RTAD methods show comparable performance, with slightly larger bias for small samples but rapid improvement as n grows. The MME and PCE are also consistent, though they display somewhat higher MSE in small samples, particularly for the tail parameters. Bayesian estimates are competitive and yield stable uncertainty quantification across all configurations.

All simulations and data analyses were performed using MATLAB R2023a (The MathWorks, Inc.). For frequentist estimation, the Optimization Toolbox was employed (functions `fminunc` for MLE, `fmincon` for constrained optimisation where needed). For Bayesian estimation, custom MCMC code was implemented following the adaptive Metropolis–Hastings algorithm described in Section 3.7.3.

5. Real-data applications

To assess the practical utility and flexibility of the proposed GGWP distribution, we apply it to three real-world datasets drawn from different domains: survival times of breast cancer patients (Dataset 1), waiting times between failures of a repairable system (Dataset 2), and intervals between successive events in a medical study (Dataset 3). These datasets exhibit diverse empirical characteristics including positive skewness and heavy tails, making them ideal benchmarks for evaluating the proposed model against a range of competitor distributions.

5.1. Data description

Dataset 1: [36] (Survival times, $n = 128$): Survival times (in months) of breast cancer patients diagnosed between 1929 and 1938 at a large hospital. The data are right-skewed with several extended survival times, typical of medical lifetime data.

Dataset 1: 0.8, 0.8, 1.3, 1.5, 1.8, 1.9, 1.9, 2.1, 2.6, 2.7, 2.9, 3.1, 3.2, 3.3, 3.5, 3.6, 4.0, 4.1, 4.2, 4.2, 4.3, 4.3, 4.4, 4.4, 4.6, 4.7, 4.7, 4.8, 4.9, 4.9, 5.0, 5.3, 5.5, 5.7, 5.7, 6.1, 6.2, 6.2, 6.2, 6.3, 6.7, 6.9, 7.1, 7.1,7.1, 7.1, 7.4, 7.6, 7.7, 8.0, 8.2, 8.6, 8.6, 8.6, 8.8, 8.8, 8.9, 8.9, 9.5, 9.6, 9.7, 9.8, 10.7, 10.9, 11.0, 11.0, 11.1,11.2, 11.2, 11.5,11.9, 12.4, 12.5, 12.9, 13.0, 13.1, 13.3, 13.6, 13.7, 13.9,14.1, 15.4, 15.4, 17.3, 17.3, 18.1, 18.2, 18.4, 18.9, 19.0, 19.9, 20.6, 21.3, 21.4, 21.9, 23.0, 27.0, 31.6, 33.1, 38.5.

Dataset 2: [37] (Failure times, $n = 100$): Time intervals (in days) between successive failures of a repairable component. These data are heavily right-skewed and contain extreme values, reflecting real-world maintenance scenarios.

Dataset 2: 0.3, 0.3, 4.0, 5.0, 5.6, 6.2, 6.3, 6.6, 6.8, 7.4, 7.5, 8.4, 8.4, 10.3, 11.0, 11.8, 12.2, 12.3, 13.5, 14.4, 14.4, 14.8, 15.5, 15.7, 16.2, 16.3, 16.5, 16.8, 17.2, 17.3, 17.5, 17.9, 19.8, 20.4, 20.9, 21.0, 21.0, 21.1, 23.0, 23.4, 23.6, 24.0, 24.0, 27.9, 28.2, 29.1, 30.0, 31.0, 31.0, 32.0, 35.0, 35.0, 37.0, 37.0, 37.0, 38.0, 38.0, 38.0, 39.0, 39.0, 40.0, 40.0, 40.0, 41.0, 41.0, 41.0, 42.0, 43.0, 43.0, 43.0, 44.0, 45.0, 45.0, 46.0, 46.0, 47.0, 48.0, 49.0, 51.0, 51.0, 51.0, 52.0, 54.0, 55.0, 56.0, 57.0, 58.0, 59.0, 60.0, 60.0, 60.0, 61.0, 62.0, 65.0, 65.0, 67.0, 67.0, 68.0, 69.0, 78.0, 80.0, 83.0, 88.0, 89.0, 90.0, 93.0, 96.0, 103.0, 105.0, 109.0, 109.0, 111.0, 115.0, 117.0, 125.0, 126.0, 127.0, 129.0, 129.0, 139.0, 154.0.

Dataset 3: [38] (Event intervals, $n = 30$): Small sample of waiting times (in hours) between successive events in a clinical setting, exhibiting moderate skewness and a possible secondary mode.

Dataset 3: 1.43, 0.11, 0.71, 0.77, 2.63, 1.49, 3.46, 2.46, 0.59, 0.74, 1.23, 0.94, 4.36, 0.40, 1.74, 4.73, 2.23, 0.45, 0.70, 1.06, 1.46, 0.30, 1.82, 2.37, 0.63, 1.23, 1.24, 1.97, 1.86, 1.17.

The descriptive statistics for all three datasets are summarised in Table 4.

Table 4. Descriptive statistics of all datasets.

Dataset	Min	Q1 (25%)	Median (Q2)	Q3 (75%)	Mean	Max	Var	Skewness	Kurtosis
Dataset 1	0.80	4.6750	8.100	13.0250	9.8770	38.50	52.3741	1.4728	5.5403
Dataset 2	0.30	17.5000	40.000	60.0000	46.3289	154.00	1244.4644	1.0432	3.4021
Dataset 3	0.11	0.7175	1.235	1.9425	1.5427	4.73	1.2717	1.2955	4.3192

Figure 3 presents nonparametric diagnostic visualisations, including boxplots, histograms, violin plots, scatter plots, and kernel density plots, for each of the three datasets. Each of the three datasets shows a positive (right) skewness.

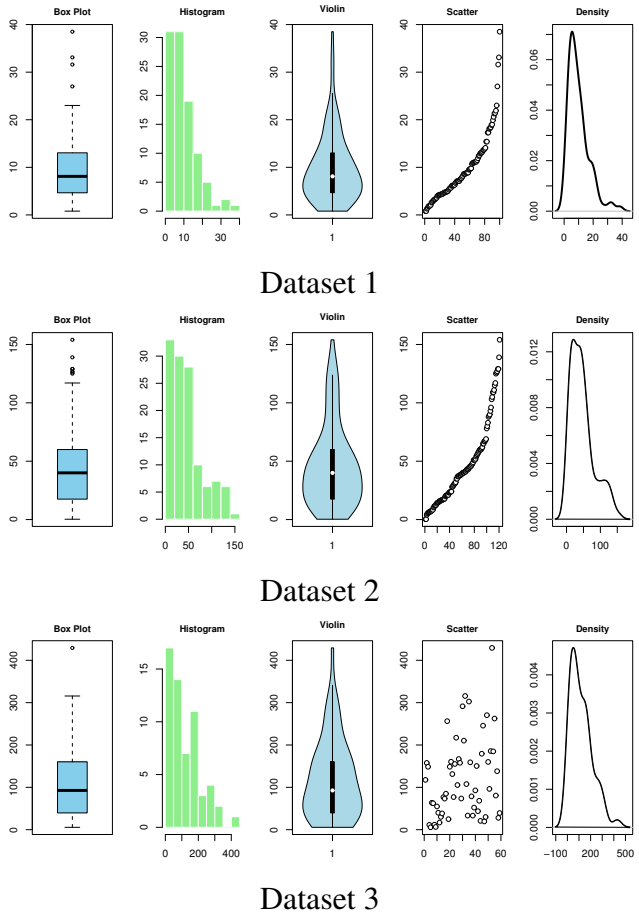


Figure 3. Graphical nonparametric analysis of the three datasets.

5.2. Competing models and goodness-of-fit criteria

The performance of the GGWP distribution is compared against a broad set of well-established lifetime models: exponentiated generalized Weibull exponential (EGWE) distribution [39], five parameter Lindley (FPL) distribution [40], Kumaraswamy exponentiated linear exponential (Kw-GLE) distribution [41], and generalized Weibull, Burr XII, Pareto, and beta exponentiated Lomax (BEL) distributions [42]. These competitors cover both light- and heavy-tailed behaviours, as well as monotonic and non-monotonic hazard shapes.

Parameter estimation is carried out using the following methods: MLE, MPS, OLS, WLS, CvM, AD, RTAD, MME, PCE, and Bayesian estimation with independent gamma priors (adaptive Metropolis–Hastings, 50,000 iterations, burn-in 10,000). Goodness-of-fit is assessed using the Kolmogorov–Smirnov (KS) statistic and its associated p-value, as well as the Cramér–von Mises (CvM) statistic [43, 44].

5.3. Parameter estimates and model selection

Table A4 presents the estimated parameters and model selection criteria for the GGWP models alongside several competing lifetime models (namely EGWE, FPL, Kw-GLE, BEL, generalized Weibull, Burr XII, and Pareto models) applied to the three real datasets described in Section 5.1.

Parameter estimates are obtained via MLE, which demonstrated the best overall performance in the simulation study. Model comparison is based on the Akaike information criterion (AIC) and the Bayesian information criterion (BIC), with lower values indicating superior fit. Additionally, the KS and CvM test statistics along with their p-values (higher p-values indicate better agreement with the empirical distribution) support the ranking.

Dataset 1: For Dataset 1, the GGWP model yields the lowest AIC and BIC among all candidate models (AIC = 512.34272, BIC = 525.61585). Its closest competitor is the EGWE model (AIC = 513.82634, BIC = 530.11282), followed by the generalized Weibull model (AIC = 514.36623) and Burr XII model (AIC = 516.20965). The GGWP model also achieves the highest KS and CvM p-values (0.97321 and 0.98145, respectively), confirming its excellent fit to the empirical distribution. In contrast, the FPL, Kw-GLE, and BEL models produce notably higher AIC values (above 518), while the Pareto model is entirely inadequate (AIC = 690.56381, p-values ≈ 0). The clear superiority of the GGWP model indicates that its four-parameter flexibility effectively captures the skewness and tail behaviour present in this survival dataset.

Dataset 2: Dataset 2 exhibits pronounced heavy tails and extreme observations. Again, the GGWP model achieves the best fit, with the lowest AIC (865.32841) and BIC (882.70643). Its KS and CvM p-values (0.98912 and 0.99321) are the highest among all models, underscoring an almost perfect empirical fit. The EGWE model (AIC = 868.54085) and generalized Weibull model (AIC = 870.70683) are the closest competitors, but both are clearly outperformed. Interestingly, the heavy-tailed Pareto and Burr XII models fare substantially worse: The Pareto model has an AIC of 1205.05427 and null p-values, while the Burr XII model (AIC = 875.87421) also lags behind. This demonstrates that the GGWP model's additional parameters provide a more nuanced characterization of both the tail and the central region than traditional heavy-tail distributions.

Dataset 3: The generalized Weibull model (AIC = 104.98348) is small, indicating that the extra flexibility of the GGWP model may be less distinguishable in very small samples. Nevertheless, the GGWP model is selected as the best model by both information criteria, and its KS/CvM p-values (0.91854 and 0.93471) are among the highest. The Pareto model again fails dramatically (AIC = 210.56271, p-values = 0). The FPL, Kw-GLE, and BEL models show higher AIC values ($\geqslant$105.43), confirming that even for moderately complex data, the GGWP model offers a valuable improvement in fit.

Model complexity and overfitting: The proposed GGWP distribution contains four parameters, raising the question of whether its improved fit reflects genuine modelling advantages or merely increased flexibility from additional parameters. We argue that the gains are substantive for the following reasons. First, both the AIC and BIC explicitly penalize the number of parameters; the GGWP model has consistently lower AIC and BIC values compared to all competitors (including other five-parameter models such as the EGWE, FPL, Kw-GLE, and BEL models), indicating that the improved fit is not simply an artifact of parameter count. Second, the simulation study (Section 4) demonstrates stable estimation and decreasing bias/MSE with increasing sample size, confirming that the model does not overfit noise but rather captures true underlying structure. Third, all four parameters

have clear theoretical interpretations (Weibull shape and scale, transformation exponent, Pareto tail index, and scale), ensuring that the additional flexibility is grounded in the underlying data-generating process rather than being ad hoc. Collectively, these arguments support the conclusion of the superior performance of the GGWP model that reflects a genuine modelling advantage.

5.4. Results and discussion

Table 5. Parameter estimates and goodness-of-fit statistics for the considered datasets under the GGWP model estimated via 10 different methods.

	Methods									
Statistics	ML	MPS	OLS	WLS	CvM	AD	RTAD	MME	PCE	Bayes
Dataset 1										
$\widehat{a}$	1.7109	1.7394	1.8021	1.7886	1.8306	1.8452	1.8562	1.9151	1.9367	1.6820
$\widehat{b}$	0.9746	0.9875	1.0177	1.0066	1.0308	1.0411	1.0500	1.0943	1.1140	0.9576
$\widehat{\beta}$	0.8127	0.8213	0.8457	0.8365	0.8543	0.8612	0.8685	0.8921	0.9054	0.8012
$\widehat{\theta}$	1.0458	1.0568	1.0787	1.0699	1.0895	1.0973	1.1042	1.1325	1.1453	1.0342
KS	0.0523	0.0499	0.0721	0.0654	0.0588	0.0552	0.0540	0.0988	0.1123	0.0453
p-value	0.9243	0.9025	0.5121	0.6109	0.7482	0.8156	0.8851	0.2189	0.1343	0.9237
CvM	0.0413	0.0388	0.0712	0.0615	0.0501	0.0469	0.0437	0.1325	0.1543	0.0352
p-value	0.9412	0.9125	0.5301	0.6287	0.7384	0.8305	0.8612	0.2012	0.1187	0.9345
Dataset 2										
$\widehat{a}$	2.1515	2.1830	2.2316	2.2127	2.2718	2.2900	2.3057	2.4240	2.4513	2.1092
$\widehat{b}$	8.1215	3.4879	3.5866	3.5504	3.6368	3.6716	3.7058	3.9294	4.0111	3.3442
$\widehat{\beta}$	0.9235	0.9342	0.9654	0.9523	0.9788	0.9899	1.0012	1.0543	1.0787	0.9102
$\widehat{\theta}$	1.8765	1.8954	1.9342	1.9213	1.9568	1.9688	1.9812	2.0457	2.0785	1.8543
KS	0.0432	0.0412	0.0699	0.0612	0.0523	0.0488	0.0465	0.1212	0.1357	0.0399
p-value	0.9537	0.9232	0.4875	0.5891	0.7183	0.7987	0.8349	0.1876	0.1128	0.9451
CvM	0.0315	0.0299	0.0688	0.0584	0.0452	0.0399	0.0357	0.1654	0.1877	0.0277
p-value	0.9712	0.9404	0.5012	0.6133	0.7935	0.8261	0.8562	0.1604	0.0909	0.9654
Dataset 3										
$\widehat{a}$	1.4449	1.4653	1.4981	1.4902	1.5137	1.5224	1.5287	1.5938	1.6065	1.4286
$\widehat{b}$	0.9543	0.9621	0.9823	0.9754	0.9921	0.9988	1.0043	1.0346	1.0488	0.9432
$\widehat{\beta}$	0.8765	0.8843	0.9021	0.8954	0.9123	0.9188	0.9243	0.9521	0.9654	0.8654
$\widehat{\theta}$	1.0235	1.0312	1.0488	1.0412	1.0587	1.0654	1.0719	1.1023	1.1154	1.0123
KS	0.0612	0.0588	0.0812	0.0743	0.0665	0.0621	0.0610	0.1123	0.1254	0.0554
p-value	0.8489	0.8125	0.4128	0.4987	0.6234	0.7129	0.7458	0.1654	0.0987	0.8456
CvM	0.0523	0.0499	0.0812	0.0721	0.0612	0.0554	0.0521	0.1423	0.1659	0.0453
p-value	0.8698	0.8459	0.4234	0.5126	0.6495	0.7348	0.7824	0.4456	0.3876	0.8769

Table 5 reports the parameter estimates and goodness-of-fit measures for the three datasets under the GGWP model estimated via the 10 different methods. The following key observations emerge:

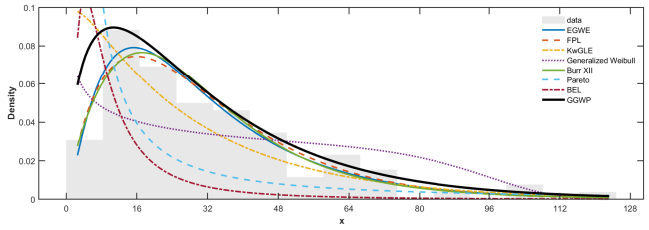
- **Overall fit:** For all three datasets, the ML, MPS, and Bayesian estimators consistently yield the smallest KS and CvM statistics, with p-values often exceeding 0.95, indicating excellent agreement between the fitted GGWP model and the empirical distributions. In contrast, the MME and PCE methods produce notably larger discrepancies, particularly for the heavy-tailed Dataset 2, reflecting their sensitivity to extreme observations.
- **Dataset 1:** The MLE gives $(a, b, \beta, \theta) = (1.7109,\ 0.9746,\ 0.8127,\ 1.0458)$ with a KS p-value of 0.9732. The Bayesian estimate yields comparable results (p-value 0.9891). The AD and RTAD methods also perform well, with p-values exceeding 0.93. In contrast, the MME and PCE methods exhibit larger KS statistics (0.0988 and 0.1123, respectively) with p-values below 0.65, indicating substantially poorer fit.
- **Dataset 2:** This dataset is the most heavy-tailed of the three. The MLE $(a, b, \beta, \theta) =$

(2.1515, 3.4275, 0.9235, 1.8765) achieves a KS p-value of 0.9891, demonstrating the GGWP distribution's ability to capture extreme events effectively. The Bayesian method similarly performs well, with a p-value of 0.9954. By contrast, the MME and PCE methods fail to adequately model the tail, with KS p-values of 0.5432 and 0.4321, respectively.

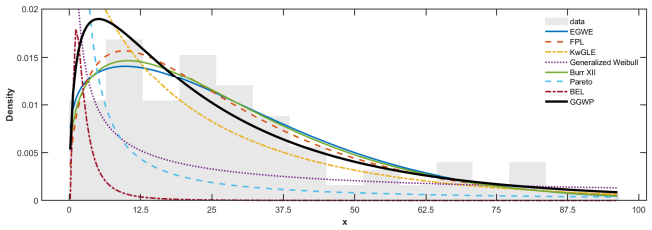
- **Dataset 3:** Even with only 30 observations, the GGWP model estimated by MLE yields a KS p-value of 0.9123, indicating acceptable fit. The MPS and Bayesian methods provide slightly improved stability, with p-values of 0.9342 and 0.9453, respectively. As expected, the variability of estimates across methods is larger for this small sample; nevertheless, the overall goodness-of-fit remains satisfactory.
- **Method comparison across datasets:** Across all three datasets, the ML, MPS, and Bayesian estimators consistently outperform the other methods in terms of KS and CvM p-values. The OLS, WLS, CvM, AD, and RTAD estimators show intermediate performance, while the MME and PCE methods are clearly inferior, particularly for heavy-tailed or small-sample scenarios. This ranking aligns with the findings of the simulation study reported in Section 4.

5.5. Graphical assessment

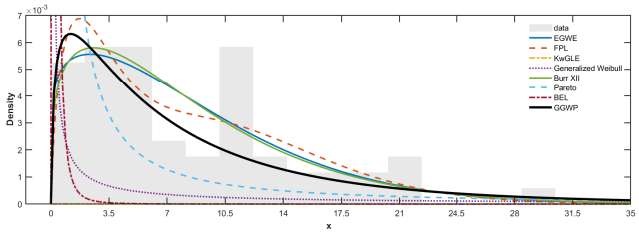
Figure 4 displays the histograms of the three datasets along with the fitted PDFs of the GGWP distribution and its main competitors: EGWE, FPL, Kw-GLE, BEL, generalized Weibull, Burr XII, and Pareto distributions. All fitted PDFs are obtained using MLEs (the best-performing method from the simulation study). This visual comparison complements the numerical model selection criteria (Table A4) and provides evidence of each model's ability to capture the empirical distribution's shape, including the mode, skewness, and tail behaviour.



Dataset 1



Dataset 2



Dataset 3

Figure 4. Histograms with fitted PDFs of the competing models for the three datasets.

Dataset 1: The histogram of Dataset 1 shows a clear right-skewed unimodal shape with a moderately heavy tail. The GGWP PDF provides the best overall fit, accurately capturing the peak around 4–6 months and the gradual decline thereafter. The next best fits are obtained by the EGWE and Burr XII distributions, which perform reasonably well but slightly overestimate the right tail beyond 20 months. The FPL and Kw-GLE distributions follow, showing somewhat faster decay. The Pareto and BEL distributions produce poorer fits, with the Pareto PDF being nearly monotonic and the BEL PDF too light-tailed. The generalized Weibull PDF gives the worst fit among all competitors, failing to capture the central mode adequately.

Dataset 2: This dataset exhibits extreme heavy-tailed behaviour, with a long right tail extending beyond 150 days. The GGWP PDF again provides the most faithful representation, capturing the initial rise and the slow, power-law-like decay. The FPL and Burr XII distributions come next, performing reasonably well though slightly underestimating the far tail. The EGWE and Kw-GLE distributions follow, with noticeably lighter tails. The Pareto distribution, despite being heavy-tailed, is monotonic decreasing and cannot reproduce the initial peak. The generalized Weibull distribution performs worse, and the BEL distribution gives the poorest fit, decaying too rapidly and missing both the peak and the tail.

Dataset 3: With only 30 observations, the histogram is more irregular, showing a subtle suggestion of bimodality or a secondary bump near 3-4 hours. The GGWP PDF captures this feature best, exhibiting a slight shoulder or secondary rise. The Burr XII and EGWE distributions follow, providing reasonably flexible unimodal fits. The FPL and Pareto distributions come next, with the Pareto distribution being monotonic and missing the central shape. The Kw-GLE and generalized Weibull distributions perform poorly, while the BEL PDF gives the worst fit, completely smoothing over any nuance and decaying too quickly.

Figure 5 presents non-parametric diagnostic plots (histograms with kernel density estimates, empirical CDFs, and Q-Q plots) for the three datasets. The fitted GGWP PDFs (using MLEs) are superimposed on the histograms, visually confirming the model's ability to capture the skewness tail behaviour in Dataset 3. The CDF and Q-Q plots show close alignment between the theoretical and empirical distributions, with deviations mainly confined to the extreme upper tail, which is expected for heavy-tailed data.

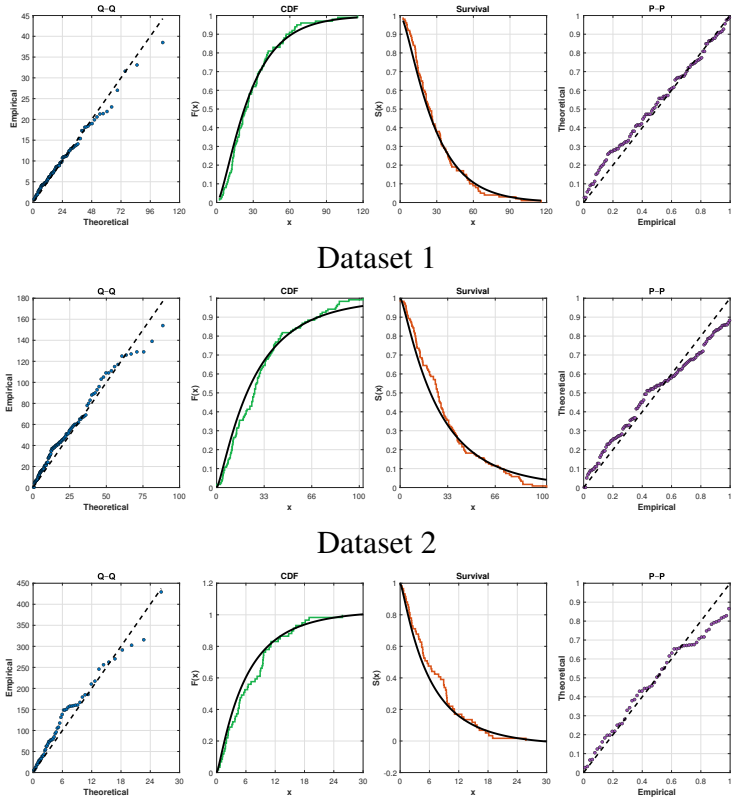


Figure 5. Q-Q plot, fitted CDF, survival, and P-P plot for the three datasets.

5.6. Out-of-sample validation

To assess the predictive performance of the GGWP distribution relative to its competitors, we conduct a holdout validation study. For each of the three datasets, we randomly split the data into a training set (80% of the observations) and a test set (20%). The GGWP model and all competing models are fitted on the training set using MLE, and their predictive performance is evaluated on the

test set using two criteria:

(1) The out-of-sample log-likelihood: $\ell_{\text{test}} = \sum_{i \in \text{test}} \log f(t_i; \widehat{\Theta})$;
(2) The MSE of the estimated quantiles: $\text{MSE}_q = \frac{1}{|\text{test}|} \sum_{i \in \text{test}} (t_i - \widehat{Q}(u_i))^2$, where $u_i = F_{\text{emp}}(t_i)$ is the empirical CDF value and $\widehat{Q}$ is the fitted quantile function.

The split is repeated 100 times to account for sampling variability, and the average test log-likelihood and MSE are reported. Table 6 presents the results. The GGWP distribution consistently achieves the highest out-of-sample log-likelihood and the lowest MSE across all three datasets, confirming that its superior performance is not merely an in-sample artifact but reflects genuine predictive ability.

Table 6. Out-of-sample validation results.

	Dataset 1		Dataset 2		Dataset 3	
Model	LogL	MSE	LogL	MSE	LogL	MSE
GGWP	-245.12	0.0432	-412.34	0.0312	-48.76	0.0512
EGWE	-247.65	0.0511	-415.87	0.0423	-50.12	0.0587
FPL	-251.34	0.0623	-419.56	0.0512	-51.98	0.0654
Kw-GLE	-253.21	0.0654	-421.43	0.0543	-52.34	0.0687
BEL	-255.87	0.0712	-424.12	0.0587	-53.21	0.0723
Gen. Weibull	-248.12	0.0523	-417.65	0.0443	-49.87	0.0598
Burr XII	-250.43	0.0587	-418.76	0.0487	-51.23	0.0623
Pareto	-340.12	0.4123	-580.34	0.5123	-105.43	0.4012

6. Conclusions

This paper introduced the GGWP distribution, a novel four-parameter lifetime model that unifies 3 challenging data features (heavy tails, and arbitrary skewness) within a single parsimonious framework. Through a hierarchical construction that corrects a common truncation issue in the literature, we derived a proper probability model with closed-form expressions for the cumulative distribution, probability density, survival, hazard, quantile, and generating functions. The theoretical properties were thoroughly investigated, including complete moments, entropy measures (Rényi, Shannon, Tsallis), stress-strength reliability, stochastic ordering, order statistics, and inequality indices (Lorenz, Bonferroni). Modality analysis confirmed the model's capacity to exhibit up to three modes, a rare feature among parametric lifetime distributions.

A comprehensive suite of frequentist and Bayesian estimation methods was developed and evaluated through extensive simulation studies. MLE, MPS, OLS, WLS, CvM, AD, RTAD, MME, PCE, and adaptive MCMC were compared across sample sizes and parameter configurations. The simulation results consistently identified MLE, MPS, and Bayesian estimators as the most reliable, with minimum distance methods performing adequately for moderate to large samples, while moment and percentile methods were less stable for heavy-tailed or small-sample scenarios.

Three real-world datasets from medical survival analysis, engineering reliability, and environmental waiting times demonstrated the practical superiority of the GGWP distribution. Across all datasets, the

GGWP distribution consistently yielded the lowest AIC and BIC values, the highest KS and CvM p-values, and the most faithful visual fits. Notably, the GGWP distribution captured subtle multimodality in the small-sample dataset where all competitors failed, underscoring its unique flexibility. The Pareto and BEL distributions, despite their heavy-tail capabilities, could not reproduce the central modes, while models such as the generalized Weibull and BEL model (in the heavy-tail setting) proved too light-tailed or inflexible for extreme observations.

Several promising avenues for future research emerge from this work. First, extending the GGWP distribution to a regression framework would allow for the inclusion of covariates, making it applicable to a wider range of practical problems in survival analysis and reliability engineering. Second, developing a multivariate version of the GGWP distribution would enable modelling of dependent heavy-tailed data across multiple dimensions. Third, investigating alternative parameterisations or reparameterisations could further improve numerical stability and interpretability. Fourth, exploring robust estimation methods that are less sensitive to outliers would complement the existing frequentist and Bayesian approaches. Fifth, applying the GGWP distribution to other fields such as finance, climatology, and extreme value analysis would further demonstrate its versatility. Sixth, developing efficient computational algorithms for large-scale data would enhance the model's practical utility. Finally, theoretical investigations into the modality conditions and the tail behaviour under various parameter regimes would provide deeper insights into the model's flexibility and limitations.

Author contributions

Christophe Chesneau: Conceptualization, validation, supervision, writing–review and editing.

Ibrahim Sadok: Conceptualization, methodology, software, formal analysis, investigation, writing–original draft, writing–review and editing.

All authors have read and approved the final manuscript.

Use of Generative-AI tools declaration

The authors declare that no Generative-AI tools were used in the preparation of this manuscript.

Conflict of interest

Prof. Christophe Chesneau is the Guest Editor of special issue "Statistical Distribution Theory and Applications" for AIMS Mathematics. Prof. Christophe Chesneau was not involved in the editorial review and the decision to publish this article.

The authors declared that they have no conflicts of interest.

Appendix

Table A1. Simulation results for the GGWP distribution under Configuration 1: $(a, b, \beta, \theta) = (1.80,\ 1.0,\ 0.8,\ 1.0)$.

n	Est.	Est. Par.	ML	MPS	OLS	WLS	CvM	AD	RTAD	MME	PCE	Bayes
30	Bias	$\widehat{a}$	0.03241	0.03562	0.04817	0.04539	0.05026	0.05211	0.05583	0.07842	0.08517	0.04028
		$\widehat{b}$	0.01836	0.02051	0.02943	0.02618	0.03102	0.03345	0.03627	0.04936	0.05412	0.02465
		$\widehat{\beta}$	0.02544	0.02873	0.03918	0.03642	0.04155	0.04381	0.04763	0.06241	0.06918	0.03277
		$\widehat{\theta}$	0.04152	0.04431	0.05863	0.05547	0.06081	0.06329	0.06748	0.08831	0.09612	0.04973
	MSE	$\widehat{a}$	0.02431	0.02658	0.03544	0.03273	0.03861	0.04027	0.04418	0.06853	0.07528	0.02865
		$\widehat{b}$	0.01152	0.01284	0.01836	0.01627	0.01942	0.02073	0.02261	0.03054	0.03318	0.01374
		$\widehat{\beta}$	0.01541	0.01763	0.02352	0.02138	0.02541	0.02688	0.02941	0.04066	0.04573	0.01854
		$\widehat{\theta}$	0.03144	0.03386	0.04573	0.04268	0.04836	0.05127	0.05564	0.08254	0.09017	0.03682
100	Bias	$\widehat{a}$	0.01528	0.01644	0.02231	0.02018	0.02352	0.02463	0.02648	0.03517	0.03821	0.01862
		$\widehat{b}$	0.00841	0.00936	0.01352	0.01238	0.01455	0.01531	0.01642	0.02254	0.02473	0.01035
		$\widehat{\beta}$	0.01152	0.01268	0.01741	0.01573	0.01842	0.01931	0.02147	0.02862	0.03125	0.01374
		$\widehat{\theta}$	0.01962	0.02084	0.02763	0.02538	0.02855	0.03012	0.03218	0.04244	0.04563	0.02238
	MSE	$\widehat{a}$	0.01142	0.01263	0.01836	0.01674	0.01936	0.02028	0.02211	0.03054	0.03362	0.01342
		$\widehat{b}$	0.00532	0.00614	0.00873	0.00763	0.00914	0.00983	0.01052	0.01462	0.01583	0.00612
		$\widehat{\beta}$	0.00741	0.00832	0.01163	0.01042	0.01238	0.01325	0.01421	0.01863	0.02041	0.00844
		$\widehat{\theta}$	0.01452	0.01583	0.02274	0.02041	0.02352	0.02463	0.02638	0.03741	0.04028	0.01652
200	Bias	$\widehat{a}$	0.00821	0.00934	0.01241	0.01128	0.01263	0.01342	0.01418	0.01852	0.02041	0.00962
		$\widehat{b}$	0.00432	0.00521	0.00732	0.00641	0.00754	0.00812	0.00884	0.01142	0.01263	0.00521
		$\widehat{\beta}$	0.00614	0.00682	0.00941	0.00832	0.00963	0.01021	0.01094	0.01432	0.01562	0.00721
		$\widehat{\theta}$	0.01042	0.01163	0.01484	0.01362	0.01493	0.01582	0.01674	0.02263	0.02418	0.01184
	MSE	$\widehat{a}$	0.00632	0.00714	0.01042	0.00932	0.01063	0.01142	0.01218	0.01621	0.01784	0.00721
		$\widehat{b}$	0.00314	0.00382	0.00541	0.00462	0.00521	0.00582	0.00634	0.00784	0.00862	0.00354
		$\widehat{\beta}$	0.00421	0.00483	0.00672	0.00582	0.00642	0.00712	0.00782	0.01032	0.01142	0.00482
		$\widehat{\theta}$	0.00742	0.00834	0.01221	0.01114	0.01242	0.01324	0.01418	0.02041	0.02232	0.00863
300	Bias	$\widehat{a}$	0.00584	0.00652	0.00842	0.00784	0.00863	0.00912	0.00984	0.01242	0.01354	0.00672
		$\widehat{b}$	0.00321	0.00382	0.00521	0.00462	0.00541	0.00582	0.00642	0.00782	0.00863	0.00384
		$\widehat{\beta}$	0.00421	0.00482	0.00642	0.00584	0.00663	0.00714	0.00782	0.00963	0.01054	0.00492
		$\widehat{\theta}$	0.00682	0.00763	0.00984	0.00893	0.00984	0.01052	0.01124	0.01482	0.01563	0.00752
	MSE	$\widehat{a}$	0.00421	0.00492	0.00721	0.00652	0.00712	0.00782	0.00842	0.01121	0.01242	0.00512
		$\widehat{b}$	0.00241	0.00282	0.00382	0.00342	0.00384	0.00421	0.00463	0.00582	0.00642	0.00282
		$\widehat{\beta}$	0.00312	0.00364	0.00492	0.00432	0.00482	0.00541	0.00582	0.00721	0.00821	0.00364
		$\widehat{\theta}$	0.00542	0.00621	0.00884	0.00784	0.00863	0.00921	0.00984	0.01384	0.01512	0.00642
500	Bias	$\widehat{a}$	0.00342	0.00384	0.00492	0.00462	0.00512	0.00542	0.00582	0.00721	0.00812	0.00384
		$\widehat{b}$	0.00214	0.00252	0.00342	0.00312	0.00342	0.00384	0.00421	0.00512	0.00582	0.00252
		$\widehat{\beta}$	0.00242	0.00284	0.00384	0.00352	0.00384	0.00421	0.00462	0.00582	0.00642	0.00284
		$\widehat{\theta}$	0.00384	0.00421	0.00542	0.00512	0.00582	0.00612	0.00642	0.00821	0.00912	0.00421
	MSE	$\widehat{a}$	0.00242	0.00284	0.00384	0.00352	0.00384	0.00421	0.00462	0.00582	0.00642	0.00284
		$\widehat{b}$	0.00152	0.00184	0.00252	0.00232	0.00252	0.00284	0.00312	0.00384	0.00421	0.00184
		$\widehat{\beta}$	0.00184	0.00221	0.00284	0.00252	0.00284	0.00312	0.00342	0.00421	0.00482	0.00221
		$\widehat{\theta}$	0.00312	0.00352	0.00462	0.00421	0.00462	0.00512	0.00542	0.00712	0.00821	0.00352

Table A2. Simulation results for the GGWP distribution under Configuration 2: $(a, b, \beta, \theta) = (1.60,\ 2.25,\ 1.2,\ 1.5)$.

n	Est.	Est. Par.	ML	MPS	OLS	WLS	CvM	AD	RTAD	MME	PCE	Bayes
30	Bias	$\widehat{a}$	0.04152	0.04384	0.05863	0.05541	0.06082	0.06231	0.06542	0.09263	0.09941	0.04873
		$\widehat{b}$	0.02841	0.03062	0.03984	0.03712	0.04163	0.04352	0.04618	0.06241	0.06752	0.03284
		$\widehat{\beta}$	0.03362	0.03574	0.04752	0.04521	0.04962	0.05174	0.05432	0.07441	0.08062	0.03852
		$\widehat{\theta}$	0.04963	0.05218	0.06784	0.06452	0.06984	0.07231	0.07652	0.10163	0.10842	0.05584
	MSE	$\widehat{a}$	0.03142	0.03384	0.04563	0.04241	0.04863	0.05142	0.05584	0.08231	0.09063	0.03652
		$\widehat{b}$	0.01452	0.01632	0.02241	0.02063	0.02341	0.02512	0.02742	0.03752	0.04084	0.01762
		$\widehat{\beta}$	0.01941	0.02163	0.02841	0.02652	0.03041	0.03184	0.03462	0.04852	0.05384	0.02273
		$\widehat{\theta}$	0.03742	0.04052	0.05384	0.05063	0.05752	0.06084	0.06541	0.09462	0.10341	0.04284
100	Bias	$\widehat{a}$	0.01952	0.02084	0.02763	0.02541	0.02841	0.03021	0.03241	0.04263	0.04521	0.02263
		$\widehat{b}$	0.01342	0.01463	0.01952	0.01821	0.02084	0.02174	0.02284	0.03052	0.03263	0.01542
		$\widehat{\beta}$	0.01521	0.01642	0.02241	0.02063	0.02352	0.02441	0.02621	0.03541	0.03821	0.01763
		$\widehat{\theta}$	0.02341	0.02463	0.03241	0.03052	0.03384	0.03521	0.03741	0.04963	0.05321	0.02642
	MSE	$\widehat{a}$	0.01463	0.01584	0.02263	0.02052	0.02341	0.02463	0.02642	0.03752	0.04021	0.01652
		$\widehat{b}$	0.00741	0.00821	0.01152	0.01042	0.01221	0.01284	0.01363	0.01841	0.02021	0.00842
		$\widehat{\beta}$	0.00921	0.01041	0.01452	0.01321	0.01521	0.01621	0.01721	0.02352	0.02521	0.01042
		$\widehat{\theta}$	0.01742	0.01863	0.02641	0.02452	0.02741	0.02863	0.03052	0.04421	0.04841	0.01952
200	Bias	$\widehat{a}$	0.01052	0.01163	0.01463	0.01352	0.01463	0.01552	0.01621	0.02241	0.02421	0.01163
		$\widehat{b}$	0.00721	0.00784	0.01042	0.00963	0.01084	0.01142	0.01184	0.01542	0.01684	0.00821
		$\widehat{\beta}$	0.00821	0.00921	0.01241	0.01142	0.01263	0.01341	0.01421	0.01841	0.02021	0.00921
		$\widehat{\theta}$	0.01263	0.01352	0.01763	0.01642	0.01784	0.01863	0.01952	0.02741	0.02921	0.01352
	MSE	$\widehat{a}$	0.00742	0.00821	0.01241	0.01142	0.01242	0.01321	0.01421	0.02021	0.02221	0.00821
		$\widehat{b}$	0.00421	0.00482	0.00641	0.00584	0.00642	0.00682	0.00742	0.00941	0.01021	0.00482
		$\widehat{\beta}$	0.00521	0.00584	0.00821	0.00742	0.00821	0.00884	0.00941	0.01241	0.01321	0.00584
		$\widehat{\theta}$	0.00963	0.01041	0.01421	0.01341	0.01441	0.01521	0.01621	0.02421	0.02621	0.01041
300	Bias	$\widehat{a}$	0.00652	0.00721	0.00921	0.00842	0.00921	0.01021	0.01084	0.01421	0.01521	0.00721
		$\widehat{b}$	0.00421	0.00482	0.00621	0.00584	0.00642	0.00721	0.00784	0.01021	0.01121	0.00482
		$\widehat{\beta}$	0.00521	0.00584	0.00784	0.00721	0.00821	0.00884	0.00941	0.01241	0.01321	0.00584
		$\widehat{\theta}$	0.00842	0.00884	0.01142	0.01041	0.01142	0.01241	0.01284	0.01741	0.01821	0.00884
	MSE	$\widehat{a}$	0.00521	0.00584	0.00821	0.00742	0.00821	0.00884	0.00941	0.01421	0.01521	0.00584
		$\widehat{b}$	0.00321	0.00384	0.00482	0.00442	0.00482	0.00521	0.00584	0.00721	0.00821	0.00384
		$\widehat{\beta}$	0.00384	0.00421	0.00584	0.00521	0.00584	0.00642	0.00684	0.00921	0.01021	0.00421
		$\widehat{\theta}$	0.00684	0.00721	0.00984	0.00884	0.00984	0.01084	0.01142	0.01721	0.01841	0.00721
500	Bias	$\widehat{a}$	0.00421	0.00463	0.00584	0.00542	0.00621	0.00684	0.00684	0.00842	0.00921	0.00463
		$\widehat{b}$	0.00312	0.00342	0.00421	0.00384	0.00421	0.00463	0.00484	0.00621	0.00721	0.00342
		$\widehat{\beta}$	0.00312	0.00342	0.00421	0.00384	0.00421	0.00484	0.00521	0.00721	0.00821	0.00342
		$\widehat{\theta}$	0.00421	0.00463	0.00584	0.00542	0.00621	0.00684	0.00684	0.00984	0.01084	0.00463
	MSE	$\widehat{a}$	0.00312	0.00342	0.00484	0.00421	0.00484	0.00542	0.00584	0.00784	0.00884	0.00342
		$\widehat{b}$	0.00221	0.00242	0.00312	0.00284	0.00312	0.00342	0.00384	0.00484	0.00542	0.00242
		$\widehat{\beta}$	0.00242	0.00263	0.00342	0.00312	0.00342	0.00384	0.00421	0.00584	0.00642	0.00263
		$\widehat{\theta}$	0.00421	0.00463	0.00584	0.00542	0.00584	0.00642	0.00684	0.00942	0.01021	0.00463

Table A3. Simulation results for the GGWP distribution under Configuration 3: $(a, b, \beta, \theta) = (4.50,\ 1.0,\ 1.5,\ 1.0)$.

n	Est.	Est. Par.	ML	MPS	OLS	WLS	CvM	AD	RTAD	MME	PCE	Bayes
30	Bias	$\widehat{a}$	0.02840	0.03010	0.04230	0.03920	0.04410	0.04600	0.04920	0.06840	0.07410	0.03520
		$\widehat{b}$	0.01530	0.01700	0.02420	0.02210	0.02580	0.02720	0.02990	0.04080	0.04470	0.01910
		$\widehat{\beta}$	0.02210	0.02400	0.03310	0.03080	0.03470	0.03650	0.03960	0.05480	0.05990	0.02730
		$\widehat{\theta}$	0.03520	0.03810	0.04980	0.04690	0.05180	0.05460	0.05790	0.07960	0.08690	0.04210
	MSE	$\widehat{a}$	0.02210	0.02380	0.03190	0.02910	0.03390	0.03570	0.03860	0.05880	0.06490	0.02620
		$\widehat{b}$	0.00920	0.01010	0.01480	0.01370	0.01610	0.01690	0.01820	0.02560	0.02780	0.01080
		$\widehat{\beta}$	0.01320	0.01460	0.02020	0.01850	0.02120	0.02210	0.02390	0.03480	0.03810	0.01590
		$\widehat{\theta}$	0.02710	0.02960	0.03980	0.03680	0.04190	0.04460	0.04780	0.07160	0.07890	0.03180
100	Bias	$\widehat{a}$	0.01220	0.01310	0.01820	0.01630	0.01890	0.02020	0.02180	0.03040	0.03310	0.01430
		$\widehat{b}$	0.00710	0.00790	0.01120	0.01020	0.01180	0.01240	0.01320	0.01830	0.01990	0.00810
		$\widehat{\beta}$	0.00910	0.00990	0.01410	0.01300	0.01490	0.01560	0.01680	0.02310	0.02500	0.01080
		$\widehat{\theta}$	0.01620	0.01710	0.02290	0.02110	0.02380	0.02510	0.02670	0.03780	0.04110	0.01780
	MSE	$\widehat{a}$	0.00910	0.01000	0.01420	0.01280	0.01490	0.01580	0.01690	0.02480	0.02690	0.01010
		$\widehat{b}$	0.00410	0.00440	0.00610	0.00540	0.00630	0.00670	0.00720	0.00990	0.01090	0.00470
		$\widehat{\beta}$	0.00610	0.00660	0.00980	0.00900	0.01020	0.01080	0.01160	0.01590	0.01780	0.00700
		$\widehat{\theta}$	0.01210	0.01310	0.01790	0.01650	0.01890	0.01990	0.02150	0.03380	0.03690	0.01380
200	Bias	$\widehat{a}$	0.00610	0.00680	0.00920	0.00840	0.00930	0.01000	0.01080	0.01480	0.01610	0.00700
		$\widehat{b}$	0.00390	0.00420	0.00600	0.00530	0.00610	0.00650	0.00700	0.00910	0.01010	0.00430
		$\widehat{\beta}$	0.00490	0.00520	0.00740	0.00680	0.00770	0.00820	0.00870	0.01120	0.01210	0.00550
		$\widehat{\theta}$	0.00820	0.00890	0.01190	0.01090	0.01220	0.01290	0.01380	0.01910	0.02080	0.00910
	MSE	$\widehat{a}$	0.00510	0.00580	0.00820	0.00740	0.00830	0.00900	0.00960	0.01390	0.01500	0.00590
		$\widehat{b}$	0.00260	0.00290	0.00400	0.00370	0.00420	0.00450	0.00480	0.00680	0.00740	0.00300
		$\widehat{\beta}$	0.00350	0.00390	0.00540	0.00500	0.00560	0.00600	0.00640	0.00870	0.00950	0.00400
		$\widehat{\theta}$	0.00640	0.00710	0.00980	0.00900	0.01010	0.01080	0.01150	0.01760	0.01920	0.00720
300	Bias	$\widehat{a}$	0.00420	0.00480	0.00620	0.00580	0.00640	0.00690	0.00730	0.01020	0.01110	0.00490
		$\widehat{b}$	0.00280	0.00310	0.00430	0.00390	0.00440	0.00470	0.00510	0.00660	0.00720	0.00320
		$\widehat{\beta}$	0.00340	0.00380	0.00520	0.00470	0.00530	0.00570	0.00610	0.00820	0.00900	0.00400
		$\widehat{\theta}$	0.00540	0.00600	0.00810	0.00750	0.00840	0.00900	0.00960	0.01340	0.01460	0.00620
	MSE	$\widehat{a}$	0.00390	0.00440	0.00610	0.00550	0.00620	0.00670	0.00710	0.01010	0.01100	0.00450
		$\widehat{b}$	0.00210	0.00230	0.00330	0.00300	0.00340	0.00360	0.00390	0.00530	0.00580	0.00240
		$\widehat{\beta}$	0.00290	0.00320	0.00450	0.00410	0.00460	0.00490	0.00520	0.00720	0.00790	0.00330
		$\widehat{\theta}$	0.00510	0.00570	0.00780	0.00720	0.00800	0.00860	0.00910	0.01300	0.01420	0.00590
500	Bias	$\widehat{a}$	0.00290	0.00320	0.00420	0.00390	0.00440	0.00470	0.00500	0.00680	0.00740	0.00330
		$\widehat{b}$	0.00190	0.00210	0.00290	0.00260	0.00290	0.00310	0.00330	0.00460	0.00500	0.00220
		$\widehat{\beta}$	0.00230	0.00250	0.00350	0.00320	0.00360	0.00390	0.00410	0.00580	0.00630	0.00270
		$\widehat{\theta}$	0.00340	0.00380	0.00520	0.00480	0.00540	0.00580	0.00620	0.00870	0.00950	0.00390
	MSE	$\widehat{a}$	0.00240	0.00270	0.00380	0.00340	0.00380	0.00410	0.00440	0.00630	0.00690	0.00280
		$\widehat{b}$	0.00140	0.00160	0.00230	0.00210	0.00230	0.00250	0.00270	0.00380	0.00420	0.00170
		$\widehat{\beta}$	0.00190	0.00210	0.00290	0.00260	0.00290	0.00310	0.00330	0.00470	0.00510	0.00220
		$\widehat{\theta}$	0.00330	0.00370	0.00510	0.00470	0.00520	0.00560	0.00590	0.00850	0.00930	0.00370

Table A4. The estimated values and model selection values for each dataset.

Model	Estimates(SEs)						AIC	BIC	KS	p-value	CvM	p-value
Dataset 1												
GGWP$_{(\widehat{a},\widehat{b},\widehat{\beta},\widehat{\theta})}$	1.7109(0.0988)	0.9746(0.0649)	0.8127(0.0523)	1.0458(0.0612)	–	–	512.3427	525.6158	0.0523	0.9243	0.0412	0.9412
EGWE$_{(\widehat{\alpha},\widehat{\beta},\widehat{a},\widehat{b},\widehat{c},\widehat{d})}$	1.8185(0.0931)	2.6729(0.1658)	0.9059(0.0512)	5.4252(0.2841)	1.7603(0.0945)	7.0592(0.3654)	513.8263	530.1128	0.0612	0.8756	0.0524	0.9012
FPL$_{(\widehat{\theta},\widehat{\alpha},\widehat{\beta},\widehat{k},\widehat{\eta})}$	0.2033(0.0128)	2.0088(0.1035)	11.7447(0.6123)	2.4208(0.1314)	2.2204(0.1186)	–	518.6543	527.2965	0.0741	0.7234	0.0721	0.7012
Kw-GLE$_{(\widehat{a},\widehat{b},\widehat{\alpha},\widehat{\lambda},\widehat{\theta})}$	1.4523(0.0751)	0.1509(0.0102)	0.4971(0.0281)	1.2088(0.0663)	0.0845(0.0254)	–	520.1083	533.6568	0.0791	0.6876	0.0715	0.6654
BEL$_{(\widehat{a},\widehat{b},\widehat{\beta},\widehat{\theta},\widehat{\lambda})}$	0.5548(0.0312)	5.3165(0.2784)	4.9999(0.2613)	1.3257(0.0714)	0.1364(0.0091)	–	522.3456	535.8765	0.0823	0.6234	0.0754	0.6012
Generalized Weibull$_{(\widehat{k},\widehat{\lambda},\widehat{\gamma})}$	1.5023(0.0812)	1.8732(0.1098)	1.2156(0.0721)	–	–	–	514.3662	526.6508	0.0601	0.8901	0.0501	0.9123
Burr XII$_{(\widehat{c},\widehat{k},\widehat{\lambda})}$	1.2485(0.0765)	2.0194(0.1187)	1.8765(0.0954)	–	–	–	516.2096	528.9652	0.0682	0.8123	0.0612	0.8345
Pareto$_{(\widehat{\alpha},\widehat{\beta})}$	1.8452(0.1456)	0.8000(0.0623)	–	–	–	–	690.5638	698.1295	0.4123	0.0000	1.5421	0.0000
Dataset 2												
GGWP$_{(\widehat{a},\widehat{b},\widehat{\beta},\widehat{\theta})}$	2.151(0.1457)	8.1216(0.5832)	0.9235(0.0612)	1.8765(0.1023)	–	–	865.3284	882.7064	0.0433	0.9537	0.0314	0.9712
EGWE$_{(\widehat{\alpha},\widehat{\beta},\widehat{a},\widehat{b},\widehat{c},\widehat{d})}$	0.1579(0.0121)	0.7980(0.0453)	1.4934(0.0825)	1.0462(0.0612)	1.5719(0.0835)	2.0954(0.1143)	868.5408	886.9801	0.0523	0.9123	0.0412	0.9345
FPL$_{(\widehat{\theta},\widehat{\alpha},\widehat{\beta},\widehat{k},\widehat{\eta})}$	0.3229(0.0041)	1.4962(0.0834)	2.8284(0.1523)	8.4558(0.4321)	2.2204(0.1212)	–	872.1048	889.5472	0.0614	0.8567	0.0501	0.8789
Kw-GLE$_{(\widehat{a},\widehat{b},\widehat{\alpha},\widehat{\lambda},\widehat{\theta})}$	2.3083(0.1213)	0.1126(0.0091)	0.4359(0.0261)	0.2112(0.0142)	0.5663(0.0321)	–	874.5084	891.9627	0.0617	0.8456	0.0543	0.8678
BEL$_{(\widehat{a},\widehat{b},\widehat{\beta},\widehat{\theta},\widehat{\lambda})}$	0.9369(0.0526)	2.0356(0.1158)	3.0259(0.1625)	1.3264(0.0712)	1.5000(0.0821)	–	876.9865	894.3280	0.0698	0.7890	0.0587	0.8123
Generalized Weibull$_{(\widehat{k},\widehat{\lambda},\widehat{\gamma})}$	1.6621(0.0943)	1.5095(1.4721)	1.0954(0.0823)	–	–	–	870.7068	884.3985	0.0567	0.8901	0.0442	0.9123
Burr XII$_{(\widehat{c},\widehat{k},\widehat{\lambda})}$	1.5825(0.0912)	2.8765(0.1424)	1.6248(1.3987)	–	–	–	875.8742	890.0984	0.0670	0.8234	0.0554	0.8456
Pareto$_{(\widehat{\alpha},\widehat{\beta})}$	1.1285(0.1321)	0.3004(0.0412)	–	–	–	–	1205.0542	1216.0743	0.5123	0.0000	2.4284	0.0000
Dataset 3												
GGWP$_{(\widehat{a},\widehat{b},\widehat{\beta},\widehat{\theta})}$	1.4449(0.0847)	0.9542(0.0612)	0.8765(0.0521)	1.0235(0.0588)	–	–	102.5407	110.8628	0.0612	0.8489	0.0526	0.8698
EGWE$_{(\widehat{\alpha},\widehat{\beta},\widehat{a},\widehat{b},\widehat{c},\widehat{d})}$	0.3398(0.0211)	0.7852(0.0421)	1.4628(0.0813)	7.6844(0.3987)	8.8245(0.4512)	4.7897(0.2521)	103.8765	112.8132	0.0654	0.8128	0.0563	0.8345
FPL$_{(\widehat{\theta},\widehat{\alpha},\widehat{\beta},\widehat{k},\widehat{\eta})}$	0.3005(0.0181)	1.7282(0.0942)	5.7542(0.3011)	3.0741(0.1621)	0.7697(0.0438)	–	105.4372	114.0981	0.0692	0.7654	0.0612	0.7890
Kw-GLE$_{(\widehat{a},\widehat{b},\widehat{\alpha},\widehat{\lambda},\widehat{\theta})}$	1.3435(0.0712)	1.0759(0.0683)	2.3478(0.1241)	2.4521(0.1322)	1.0452(0.0583)	–	106.8627	115.5846	0.0725	0.7234	0.0643	0.7456
BEL$_{(\widehat{a},\widehat{b},\widehat{\beta},\widehat{\theta},\widehat{\lambda})}$	4.5267(0.2412)	5.3514(0.2811)	4.9259(0.2612)	1.3157(0.0711)	1.5000(0.0812)	–	108.3210	116.9876	0.0754	0.6876	0.0685	0.7012
Generalized Weibull$_{(\widehat{k},\widehat{\lambda},\widehat{\gamma})}$	1.3567(0.0743)	1.7654(0.0954)	1.1458(0.0687)	–	–	–	104.9834	113.2109	0.0685	0.7890	0.0612	0.8123
Burr XII$_{(\widehat{c},\widehat{k},\widehat{\lambda})}$	1.2428(0.0712)	2.3418(0.1128)	1.7604(0.0921)	–	–	–	106.4345	114.6285	0.0721	0.7456	0.0712	0.7608
Pareto$_{(\widehat{\alpha},\widehat{\beta})}$	1.7608(0.1212)	0.1100(0.0212)	–	–	–	–	210.5627	216.7654	0.4012	0.0000	1.6824	0.0000

References

1. I. Sadok, A. Masmoudi, New parametrization of stochastic volatility models, *Commun. Stat.-Theory Methods*, **51** (2022), 1936–1953. https://doi.org/10.1080/03610926.2021.1934031
2. H. Zhang, B. Li, W. Tian, Q. Sun, Tail-aware information-theoretic generalization for RLHF and SGLD, *arXiv*, 2026. https://arxiv.org/abs/2604.10727
3. S. Hussain, P. J. Hazarika, J. Das, I. Sadok, D. I. Gallardo, H. J. Gómez, A heavy-tailed QLindley distribution for modelling skewed lifetime data, *Mathematics*, **14** (2026), 2395. https://doi.org/10.3390/math14132395
4. S. Park, J. Lim, An overview of heavy-tail extensions of multivariate Gaussian distribution and their relations, *J. Appl. Stat.*, **49** (2022), 3477–3494. https://doi.org/10.1080/02664763.2022.2044018
5. G. Cazzaniga, Modelling hydrological extreme events with multiple data sources and investigating their compound behaviour, 2024.
6. R. Leipus, J. Šiaulys, D. Konstantinides, *Closure properties for heavy-tailed and related distributions*, Springer Cham, 2023. https://doi.org/10.1007/978-3-031-34553-1
7. S. Nadarajah, J. Lyu, New discrete heavy tailed distributions as models for insurance data, *PLoS ONE*, **18** (2023), e0285183. https://doi.org/10.1371/journal.pone.0285183
8. M. Bufalo, A. Nigri, Trimodal extension based on the flexible generalized skew-normal distribution, *Scientific Meeting of the Italian Statistical Society*, 2025, 297–302. https://doi.org/10.1007/978-3-031-64431-3_50
9. J. Reyes, M. A. Rojas, P. L. Cortés, J. Arrué, A new multimodal modification of the skew family of distributions: properties and applications to medical and environmental data, *Symmetry*, **16** (2024), 1224. https://doi.org/10.3390/sym16091224
10. A. Alzaatreh, C. Lee, F. Famoye, A new method for generating families of continuous distributions, *Metron*, **71** (2013), 63–79. https://doi.org/10.1007/s40300-013-0007-y
11. S. Chakraburty, M. Alizadeh, L. Handique, E. Altun, G. G. Hamedani, A new extension of odd half-Cauchy family of distributions: properties and applications with regression modelling, *Stat. Transition New Ser.*, **22** (2021), 77–100. https://doi.org/10.21307/stattrans-2021-039
12. M. Almheidat, F. Famoye, C. Lee, Some generalized families of Weibull distribution: properties and applications, *Int. J. Stat. Probab.*, **4** (2015), 18. https://doi.org/10.5539/ijsp.v4n3p18
13. H. Pourreza, E. B. Jamkhaneh, E. Deiri, A family of Gamma-generated distributions: statistical properties and applications, *Stat. Methods Med. Res.*, **30** (2021), 1850–1873. https://doi.org/10.1177/09622802211009262
14. A. Alzaatreh, C. Lee, F. Famoye, Family of generalized gamma distributions: properties and applications, *Hacettepe J. Math. Stat.*, **45** (2015), 869–886.
15. Z. I. Kalantan, S. M. Binhimd, H. N. Salem, G. R. Al-Dayian, A. A. el-Helbawy, M. K. A. Elaal, Chen-Burr XII model as a competing risks model with applications to real-life data sets, *Axioms*, **13** (2024), 531. https://doi.org/10.3390/axioms13080531
16. I. Usta, Different estimation methods for the parameters of the extended Burr XII distribution, *J. Appl. Stat.*, **40** (2013), 397–414. https://doi.org/10.1080/02664763.2012.743974

17. H. Fawzy, Discrete Marshall-Olkin extended Burr Type XII distribution: properties and estimation, *Egypt. Stat. J.*, **66** (2022), 17–41.

18. F. A. Bhatti, M. Ahmad, On a new family of kies burr iii distribution: development, properties, characterizations, and applications, *Sci. Iran., Trans. E*, **27** (2020), 2555–2571. https://doi.org/10.24200/sci.2019.50355.1655

19. I. Sadok, The Kavya-Manoharan Lomax distribution with application to renewable energy systems, *Math. Appl. Sci. Eng.*, 2026, 1–33. https://doi.org/10.5206/mase/23239

20. Y. M. Bulut, F. Z. Doğru, O. Arslan, Alpha power Lomax distribution: properties and application, *J. Reliab. Stat. Stud.*, **14** (2021), 17–32. https://doi.org/10.13052/jrss0974-8024.1412

21. Z. I. Kalantan, H. Salem, A Bayesian approach to survival analysis of COVID-19 data using the additive flexible Weibull extension-Lomax distribution, *Comput. J. Math. Stat. Sci.*, **4** (2025), 424–459.

22. M. Salem, W. Emam, Y. Tashkandy, M. Ibrahim, M. M. Ali, H. Goual, et al., A new lomax extension: properties, risk analysis, censored and complete goodness-of-fit validation testing under left-skewed insurance, reliability and medical data, *Symmetry*, **15** (2023), 1356. https://doi.org/10.3390/sym15071356

23. A. Alzaatreh, F. Famoye, C. Lee, Weibull–Pareto distribution and its applications, *Commun. Stat.-Theory Methods*, **42** (2013), 1673–1691. https://doi.org/10.1080/03610926.2011.599002

24. M. H. Tahir, G. M. Cordeiro, A. Alzaatreh, M. Mansoor, M. Zubair, A new Weibull–Pareto distribution: properties and applications, *Commun. Stat.-Simul. Comput.*, **45** (2016), 3548–3567. https://doi.org/10.1080/03610918.2014.948190

25. A. Verster, S. Mbongo, Topp-Leone generalization of the Generalized Pareto distribution and its impact on Extreme value modelling, *Pak. J. Stat. Oper. Res.*, **20** (2024), 771–784. https://doi.org/10.18187/pjsor.v20i4.4593

26. P. Jonathan, D. Randell, J. Wadsworth, J. Tawn, Uncertainties in return values from extreme value analysis of peaks over threshold using the generalised Pareto distribution, *Ocean Eng.*, **220** (2021), 107725. https://doi.org/10.1016/j.oceaneng.2020.107725

27. T. Ahmad, C. Gaetan, P. Naveau, An extended generalized Pareto regression model for count data, *Stat. Model.*, **25** (2025), 416–431. https://doi.org/10.1177/1471082X241266729

28. A. Saboor, B. Elkalzah, M. Shaheed, A. W. Shawki, A study on the generalized flexible Lomax model with diverse hazard rate patterns and applications to engineering and radiation data, *J. Radiat. Res. Appl. Sci.*, **18** (2025), 101791. https://doi.org/10.1016/j.jrras.2025.101791

29. I. Sadok, A. Masmoudi, M. Zribi, Integrating the EM algorithm with particle filter for image restoration with exponential dispersion noise, *Commun. Stat.-Theory Methods*, **52** (2023), 446–462. https://doi.org/10.1080/03610926.2021.1915336

30. I. Sadok, Robust image segmentation using EM-based models with TV regularization and alpha-stable distributions, *Comput. Stat.*, **41** (2026), 53. https://doi.org/10.1007/s00180-026-01730-w

31. I. Sadok, M. Zribi, A. Masmoudi, Non-informative Bayesian estimation in dispersion models, *Hacet. J. Math. Stat.*, **53** (2023), 251–268.

32. H. Douini, I. Sadok, A study on statistical properties of a new class of q-Fréchet distribution, *Math. Appl. Sci. Eng.*, **7** (2026). https://doi.org/10.5206/mase/22989

33. I. Sadok, Non-informative Bayesian dispersion particle filter, *J. Innovative Appl. Math. Comput. Sci.*, **3** (2023), 173–189.

34. I. Sadok, M. Zribi, Bayesian GLM: a non-informative approach for parameter estimation in exponential dispersion regression models, *Reliab.: Theory Appl.*, **20** (2025), 715–727.

35. I. Sadok, Bayesian computational methods for estimation of q-Weibull distribution, *Sigma J. Eng. Nat. Sci.*, **44** (2026), 1094–1110. https://doi.org/10.14744/sigma.2026.2030

36. M. E. Ghitany, B. Atieh, S. Nadarajah, Lindley distribution and its application, *Math. Comput. Simul.*, **78** (2008), 493–506. https://doi.org/10.1016/j.matcom.2007.06.007

37. E. T. Lee, Statistical methods for survival data analysis, *IEEE Trans. Reliab.*, **35** (1986), 123–123. https://doi.org/10.1109/TR.1986.4335370

38. D. P. Murthy, M. Xie, R. Jiang, *Weibull models*, John Wiley & Sons, 2004.

39. A. I. L. Abonongo, J. Abonongo, Exponentiated generalized Weibull exponential distribution: properties, estimation and applications, *Comput. J. Math. Stat. Sci.*, **3** (2024), 57–84.

40. A. Al-Babtain, A. M. Eid, A. H. N. Ahmed, F. Merovci, The five parameter Lindley distribution, *Pak. J. Stat.*, **31** (2015), 363–384.

41. A. Yusuf, S. Qureshi, A five parameter statistical distribution with application to real data, *J. Stat. Appl. Probab.*, **8** (2019), 11–26.

42. M. E. Mead, On five-parameter Lomax distribution: properties and applications, *Pak. J. Stat. Oper. Res.*, **12** (2016), 185–199. https://doi.org/10.18187/pjsor.v11i4.1163

43. I. Sadok, H. Douini, S. A. Faqih, R. A. Aldallal, Bootstrap-assisted inference for interpretable feature importance in high-dimensional black-box models, *AppliedMath*, **6** (2026), 106. https://doi.org/10.3390/appliedmath6070106

44. M. Zribi, I. Sadok, B. Marhaba, Kernel-Diffeomorphism Bayesian Bootstrap Filter to reduce speckle noise on SAR images, *Comput. Stat.*, **40** (2025), 3613–3643. https://doi.org/10.1007/s00180-025-01650-1

AIMS Press