www.iiste.org

Development of Weibull Exponentiated G-Family of Distributions with its Statistics and Properties

Jimoh Mutair Afolabi[1], Bello H. A.[2], Akomolafe A. A.[3]

[1]Department of Statistics, Auchi Polytechnic, Auchi, P.M.B. 13, Edo State, Nigeria
follybee@auchi.edu.ng

[2,3]Department of Statistics, Federal University of Technology, Akure, Nigeria
habello@futa.edu.ng; aaakomolafe@futa.edu.ng

Abstract

The Weibull Exponentiated G (WEG) family is a newly proposed class of probability distributions formed by combining the Weibull distribution with the Exponentiated family of distributions. The resulting family introduces two additional shape parameters, α and β, thereby providing substantial flexibility in modelling phenomena with complicated tails and non-monotone hazard rates. Applications to risk analysis, survival analysis, and reliability engineering are of particular interest. This paper presents the derivation of the WEG family, derives its fundamental statistical properties including the survival function, hazard function, and quantile function, and develops maximum likelihood estimation (MLE) for two important special submodels: the Weibull Exponentiated Exponential Distribution (WEED) and the Weibull Exponentiated Rayleigh Distribution (WERD). Moment expressions including the mean, variance, skewness, and kurtosis are derived for both submodels via power series and binomial expansion techniques. A Monte Carlo simulation study confirms that the MLE estimators are asymptotically unbiased and consistent for both submodels. Goodness-of-fit comparisons on two real data sets glass fibre strengths and civil engineering hailing times demonstrate that WEED and WERD outperform several established competitor distributions according to the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and maximised log-likelihood.

Keywords: Weibull distribution; Exponentiated G family; hazard function; probability distributions; reliability engineering; maximum likelihood estimation; model comparison.

1 Introduction

The systematic construction of new families of probability distributions has become a central theme in statistical research, motivated by the increasing complexity of empirical data encountered in engineering, medicine, environmental science, and finance. Classical one- and two-parameter distributions such as the exponential, Weibull, and Rayleigh are valuable for their analytical tractability, but they often lack the flexibility required to fit data that exhibit pronounced skewness, heavy tails, or non-monotone hazard rates. Addressing this limitation has spurred a large body of literature devoted to generalising, compounding, or transforming existing distributions.

Among the earliest and most influential generalisations is the Exponentiated family introduced by Gupta and Kundu [2001]. By raising the baseline CDF $G(x)$ to a positive power α, one obtains the Exponentiated-G (Exp-G) family with CDF $F(x) = [G(x)]^{\alpha}$, whose hazard rate can be monotone increasing, decreasing, or bathtub-shaped depending on α and the baseline. Gupta and Kundu [2001] demonstrated the usefulness of the exponentiated exponential distribution as an alternative to the gamma and Weibull distributions in lifetime data analysis, and their work has inspired numerous subsequent extensions.

The Weibull distribution, introduced by Weibull [1951], remains one of the most widely used models in reliability engineering and survival analysis owing to its ability to represent increasing, decreasing, and constant failure rates. However, when data exhibit non-monotone hazard rates or heavy tails beyond what the Weibull can capture, more flexible models are required. In response, Bourguignon et al. [2014] proposed the Weibull-G family, in which the argument of the Weibull CDF is replaced by a ratio involving the baseline CDF, producing a large and versatile class. Similarly, Cordeiro et al. [2013] introduced the Exponentiated Generalised (EG) class, and Alzaatreh et al. [2014] presented the T-X framework as a unifying structure that subsumes many earlier proposals.

The Kumaraswamy-G family of Kumaraswamy [1980], the Beta-G family of Nadarajah and Kotz [2006], and the Marshall–Olkin-G family of Marshall and Olkin [1997] have each attracted considerable attention for their ability to handle a broader range of data shapes. More recently, Olawole et al. [2020] examined the performance of exponentiated Weibull-type models in reliability contexts, and Zhao et al. [2019] developed extended Weibull-type models for failure data, both concluding that additional shape parameters substantially improve goodness-of-fit without sacrificing interpretability.

In a parallel development, compound and truncated power series distributions have been used to inject further shape flexibility. Mendonça et al. [2018] proposed flexible distributions for lifetime modelling using power series compounding, and Pardo et al. [2017] developed new families specifically targeting engineering reliability. Applications in health sciences have been advanced by Gangeh and Ghitany [2019] and Rahman et al. [2014], while extreme-value settings have been addressed by Coles [2001] and Smith [2003].

Despite this rich literature, there remains a gap for families that simultaneously (i) possess closed-form CDFs and PDFs without recourse to special functions such as the gamma or beta function, (ii) admit non-monotone hazard rates, and (iii) can be constructed by a transparent and systematic methodology. The present paper addresses this gap by introducing the *Weibull Exponentiated G* (WEG) family, formed by composing the Weibull CDF with the Exponentiated-G CDF. The WEG family has closed-form expressions for all fundamental distributional quantities, admits a wide range of hazard rate shapes, and subsumes several well-known distributions as special cases.

The remainder of the paper is organised as follows. Section 2 defines the WEG family, derives its CDF, PDF, hazard, survival, and quantile functions, and presents the two special submodels WEED and WERD. Moments, variance, skewness, kurtosis, and maximum likelihood estimation are developed within Section 2. Section 3 presents the data analysis, including descriptive statistics, a Monte Carlo simulation study, information-criteria comparisons, and histogram–density plots. Section 4 provides conclusions, and the Acknowledgement follows.

2 Methodology

The Weibull Exponentiated G distribution family is formed by merging the Weibull distribution of Weibull [1951] with the Exponentiated family of Gupta and Kundu [2001]. The resulting distribution is highly flexible and capable of representing a large number of physical phenomena. The two shape parameters α and β provide substantially more flexibility than the classical exponential distribution and make the WEG family particularly well suited to reliability, survival, and risk analysis.

2.1 Defining the Component Distributions

The Weibull CDF and PDF are, respectively,

$$F_W(x) = 1 - e^{-(\lambda x)^{\beta}}, \qquad \beta, \lambda, x > 0, \tag{1}$$

$$f_W(x) = \beta\lambda^{\beta} x^{\beta-1} e^{-(\lambda x)^{\beta}} \qquad x \in \mathbb{R}, \quad \beta, \lambda > 0. \tag{2}$$

The Exponentiated-G family has CDF and PDF

$$F_{EG}(x) = [G(x)]^{\alpha}, \qquad \alpha > 0, \tag{3}$$

$$f_{EG}(x) = \alpha g(x)[G(x)]^{\alpha-1}., \qquad x \in \mathbb{R}, \quad \alpha > 0. \tag{4}$$

2.2 Derivation of the WEG Family

Definition 2.1 (WEG Family)**.** Let $G(x)$ be a baseline CDF with corresponding PDF $g(x)$. The **Weibull Exponentiated G (WEG)** family has CDF

$$F(x;\alpha,\beta,\lambda) = \frac{1 - e^{-\lambda^\beta [G(x)]^{\alpha\beta}}}{1 - e^{-\lambda^\beta}}, \qquad x \in \mathbb{R}, \quad \alpha,\beta,\lambda > 0. \tag{5}$$

Analytical Derivation

The WEG CDF in (5) is obtained by evaluating the Weibull CDF at $t = [G(x)]^\alpha$:

$$F_W([G(x)]^\alpha) = 1 - e^{-(\lambda [G(x)]^\alpha)^\beta} = 1 - e^{-\lambda^\beta [G(x)]^{\alpha\beta}}. \tag{6}$$

To ensure that $F(x)$ is a proper CDF normalised to $[0,1]$, we divide by the value at the upper limit of $G(x)$, namely $G(\infty) = 1$, which gives the normalising constant $1 - e^{-\lambda^\beta}$. This truncation at the upper boundary is equivalent to conditioning on the argument $[G(x)]^\alpha$ being less than 1, yielding the form in (5).

Differentiating (5) with respect to x yields the WEG PDF:

$$f(x;\alpha,\beta,\lambda) = \frac{\alpha\beta\lambda^\beta g(x) [G(x)]^{\alpha\beta-1} e^{-\lambda^\beta [G(x)]^{\alpha\beta}}}{1 - e^{-\lambda^\beta}} \qquad x \in \mathbb{R}, \quad \alpha,\beta,\lambda > 0. \tag{7}$$

Hazard Function

The hazard function is given by

$$h(x) = \frac{f(x)}{1 - F(x)},$$

where $1 - F(x)$ is obtained from (5). After simplification,

$$h(x;\alpha,\beta,\lambda) = \frac{\alpha\beta\lambda^\beta [G(x)]^{\alpha\beta-1} g(x)}{1 - e^{-\lambda^\beta \left(1 - [G(x)]^{\alpha\beta}\right)}} \qquad x \in \mathbb{R}, \quad \alpha,\beta,\lambda > 0. \tag{8}$$

The denominator $1 - e^{-\lambda^\beta (1 - [G(x)]^{\alpha\beta})}$ is always strictly positive, confirming that the hazard function is well defined for all x in the support of G.

Survival Function

$$S(x;\alpha,\beta,\lambda) = 1 - F(x) = \frac{e^{-\lambda^\beta [G(x)]^{\alpha\beta}} - e^{-\lambda^\beta}}{1 - e^{-\lambda^\beta}} \qquad x \in \mathbb{R}, \quad \alpha,\beta,\lambda > 0. \tag{9}$$

Quantile Function and Median

Setting $F(x) = u$ and inverting (5):

$$Q(u) = F^{-1}(u) = G^{-1}\left(\left[-\frac{1}{\lambda^\beta}\ln\Big(1-u(1-e^{-\lambda^\beta})\Big)\right]^{\frac{1}{\alpha\beta}}\right), \qquad 0 < u < 1. \tag{10}$$

The median corresponds to $u = 0.5$:

$$M = G^{-1}\left(\left[-\frac{1}{\lambda^\beta}\ln\left(\frac{1+e^{-\lambda^\beta}}{2}\right)\right]^{\frac{1}{\alpha\beta}}\right). \tag{11}$$

2.3 Special Models: WEED and WERD

The general WEG CDF and PDF are

$$F(x;\alpha,\beta,\lambda) = \frac{1-e^{-\lambda^\beta[G(x)]^{\alpha\beta}}}{1-e^{-\lambda^\beta}}, \tag{12}$$

$$f(x;\alpha,\beta,\lambda) = \frac{\alpha\beta\lambda^\beta e^{-\lambda^\beta[G(x)]^{\alpha\beta}}[G(x)]^{\alpha\beta-1}g(x)}{1-e^{-\lambda^\beta}}. \qquad x\in\mathbb{R}, \quad \alpha,\beta,\lambda>0. \tag{13}$$

2.3.1 Weibull Exponentiated Exponential Distribution (WEED)

Consider the exponential baseline with CDF and PDF $G(x) = 1 - e^{-\theta x}$ and $g(x) = \theta e^{-\theta x}$, $x \geq 0$, $\theta > 0$. Substituting into (12)–(13):

$$F_{\text{WEED}}(x) = \frac{1-e^{-\lambda^\beta(1-e^{-\theta x})^{\alpha\beta}}}{1-e^{-\lambda^\beta}}, \tag{14}$$

$$f_{\text{WEED}}(x) = \frac{\alpha\beta\lambda^\beta\theta\, e^{-\lambda^\beta(1-e^{-\theta x})^{\alpha\beta}}(1-e^{-\theta x})^{\alpha\beta-1}e^{-\theta x}}{1-e^{-\lambda^\beta}}. \qquad x\in\mathbb{R}, \quad \alpha,\beta,\lambda,\theta>0. \tag{15}$$

Analytical process for the WEED series expansion. Using the power series $e^{-a} = \sum_{i=0}^{\infty}(-1)^i a^i/i!$ and the binomial expansion $(1-e^{-\theta x})^{\alpha\beta(i+1)-1} = \sum_{j=0}^{\infty}\binom{\alpha\beta(i+1)-1}{j}(-1)^j e^{-j\theta x}$, the WEED PDF is expressed as a mixture of exponential kernels:

$$f_{\text{WEED}}(x) = \frac{\alpha\beta\lambda^\beta\theta}{1-e^{-\lambda^\beta}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{i+j}\lambda^{\beta i}}{i!}\binom{\alpha\beta(i+1)-1}{j}e^{-(j+1)\theta x}, \tag{16}$$

where the double-sum weights $w_{i,j} = \frac{(-1)^{i+j}\lambda^{\beta i}}{i!}\binom{\alpha\beta(i+1)-1}{j}$ collect all scaling factors. This expansion enables closed-form evaluation of all raw moments via the standard integral $\int_0^\infty x^r e^{-ax}\,dx = \Gamma(r+1)/a^{r+1}$, $a > 0$.

Survival and hazard functions of the WEED.

$$S_{\mathrm{WEED}}(x) = \frac{e^{-\lambda^\beta(1-e^{-\theta x})^{\alpha\beta}} - e^{-\lambda^\beta}}{1-e^{-\lambda^\beta}}, \tag{17}$$

$$h_{\mathrm{WEED}}(x) = \frac{\alpha\beta\lambda^\beta\theta(1-e^{-\theta x})^{\alpha\beta-1}e^{-\theta x}}{1-e^{-\lambda^\beta[1-(1-e^{-\theta x})^{\alpha\beta}]}}. \tag{18}$$

Quantile function of the WEED. Inverting (14):

$$Q_{\mathrm{WEED}}(u) = -\frac{1}{\theta}\ln\left(1-\left[-\frac{1}{\lambda^\beta}\ln\left(1-u(1-e^{-\lambda^\beta})\right)\right]^{\frac{1}{\alpha\beta}}\right), \quad 0<u<1. \tag{19}$$

The median is obtained by setting $u = 1/2$ in (19).

Moments, Variance, Skewness and Kurtosis of the WEED

Using the series representation (16) and the integral $\int_0^\infty x^r e^{-(j+1)\theta x}\,dx = \Gamma(r+1)/[(j+1)\theta]^{r+1}$, the r-th raw moment of the WEED is

$$\mu'_r = \mathbb{E}[X^r] = \frac{\alpha\beta\lambda^\beta}{(1-e^{-\lambda^\beta})}\sum_{i,j}\frac{w_{i,j}\,\Gamma(r+1)}{(j+1)^{r+1}\theta^r}, \tag{20}$$

from which the first four raw moments follow by setting $r = 1,2,3,4$ and using $\Gamma(r+1) = r!$:

$$\mu'_1 = \frac{\alpha\beta\lambda^\beta w_{i,j}\,\Gamma(2)}{(1-e^{-\lambda^\beta})(j+1)^2\theta}, \tag{21}$$

$$\mu'_2 = \frac{\alpha\beta\lambda^\beta w_{i,j}\,\Gamma(3)}{(1-e^{-\lambda^\beta})(j+1)^3\theta^2}, \tag{22}$$

$$\mu'_3 = \frac{\alpha\beta\lambda^\beta w_{i,j}\,\Gamma(4)}{(1-e^{-\lambda^\beta})(j+1)^4\theta^3}, \tag{23}$$

$$\mu'_4 = \frac{\alpha\beta\lambda^\beta w_{i,j}\,\Gamma(5)}{(1-e^{-\lambda^\beta})(j+1)^5\theta^4}. \tag{24}$$

The variance, skewness, and kurtosis are

$$\mathrm{Var}(X) = \mu'_2 - (\mu'_1)^2, \tag{25}$$

$$\gamma_1 = \frac{\mu'_3 - 3\mu'_2\mu'_1 + 2(\mu'_1)^3}{[\mathrm{Var}(X)]^{3/2}}, \tag{26}$$

$$\beta_2 = \frac{\mu'_4 - 4\mu'_3\mu'_1 + 6\mu'_2(\mu'_1)^2 - 3(\mu'_1)^4}{[\mathrm{Var}(X)]^2}. \tag{27}$$

Maximum Likelihood Estimation for the WEED

Let $x_1, \ldots, x_n$ be a random sample from the WEED. Setting $z_i = 1 - e^{-\theta x_i}$, the log-likelihood is

$$\begin{aligned}\ell(\alpha, \beta, \lambda, \theta) &= n \log \alpha + n \log \beta + n\beta \log \lambda + n \log \theta - n \log(1 - e^{-\lambda^\beta}) \\ &\quad - \lambda^\beta \sum_{i=1}^{n} z_i^{\alpha\beta} + (\alpha\beta - 1) \sum_{i=1}^{n} \log z_i - \theta \sum_{i=1}^{n} x_i, \end{aligned} \tag{28}$$

and the corresponding score equations are

$$\frac{\partial \ell}{\partial \alpha} = \frac{n}{\alpha} + \beta \sum_i \log z_i - \beta\lambda^\beta \sum_i z_i^{\alpha\beta} \log z_i = 0, \tag{29}$$

$$\frac{\partial \ell}{\partial \beta} = \frac{n}{\beta} + n \log \lambda - \frac{n\lambda^\beta e^{-\lambda^\beta} \log \lambda}{1 - e^{-\lambda^\beta}} - \lambda^\beta \log \lambda \sum_i z_i^{\alpha\beta} - \alpha\lambda^\beta \sum_i z_i^{\alpha\beta} \log z_i + \alpha \sum_i \log z_i = 0, \tag{30}$$

$$\frac{\partial \ell}{\partial \lambda} = \frac{n\beta}{\lambda} - \frac{n\beta\lambda^{\beta-1} e^{-\lambda^\beta}}{1 - e^{-\lambda^\beta}} - \beta\lambda^{\beta-1} \sum_i z_i^{\alpha\beta} = 0, \tag{31}$$

$$\frac{\partial \ell}{\partial \theta} = \frac{n}{\theta} - \alpha\beta\lambda^\beta \sum_i z_i^{\alpha\beta-1} x_i e^{-\theta x_i} + (\alpha\beta - 1) \sum_i \frac{x_i e^{-\theta x_i}}{z_i} - \sum_i x_i = 0. \tag{32}$$

Equations (29)–(32) are nonlinear and are solved numerically using the Nelder–Mead algorithm implemented via the `optim` function in R, with multiple starting values to safeguard against convergence to local optima. Standard errors are obtained from the square roots of the diagonal elements of the inverse observed Fisher information matrix, $\hat{\mathbf{I}}^{-1}(\hat{\boldsymbol{\theta}})$, evaluated at the MLE.

2.3.2 Weibull Exponentiated Rayleigh Distribution (WERD)

Consider the Rayleigh baseline with CDF and PDF $G(x) = 1 - e^{-\frac{1}{2}\theta^2 x^2}$ and $g(x) = \theta^2 x e^{-\frac{1}{2}\theta^2 x^2}$, $x \geq 0$, $\theta > 0$. Substituting into (12)–(13):

$$F_{\text{WERD}}(x) = \frac{1 - e^{-\lambda^\beta (1 - e^{-\frac{1}{2}\theta^2 x^2})^{\alpha\beta}}}{1 - e^{-\lambda^\beta}}, \tag{33}$$

$$f_{\text{WERD}}(x) = \frac{\alpha\beta\lambda^\beta\theta^2 x\, e^{-\frac{1}{2}\theta^2 x^2} \left(1 - e^{-\frac{1}{2}\theta^2 x^2}\right)^{\alpha\beta-1} e^{-\lambda^\beta (1 - e^{-\frac{1}{2}\theta^2 x^2})^{\alpha\beta}}}{1 - e^{-\lambda^\beta}}, \quad x \geq 0. \tag{34}$$

Analytical process for the WERD series expansion. Setting $z_i = 1 - e^{-\frac{1}{2}\theta^2 x^2}$ and applying the same double-sum expansion as for the WEED,

$$f_{\text{WERD}}(x) = \frac{\alpha\beta\lambda^\beta\theta^2}{1 - e^{-\lambda^\beta}} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} w_{i,j}\, x\, e^{-\frac{1}{2}(j+1)\theta^2 x^2}, \qquad x \geq 0, \tag{35}$$

where $w_{i,j} = \frac{(-1)^{i+j}\lambda^{\beta i}}{i!}\binom{\alpha\beta(i+1)-1}{j}$. Each term in (35) is proportional to a Rayleigh kernel with scale parameter $[(j+1)\theta^2]^{-1/2}$, enabling moment computation via the standard Gaussian-type integral $\int_0^\infty x^{r+1}e^{-ax^2}\,dx = \frac{1}{2}a^{-(r+2)/2}\Gamma\left(\frac{r+2}{2}\right)$, $a > 0$.

Survival, hazard, and quantile functions of the WERD.

$$S_{\mathrm{WERD}}(x) = \frac{e^{-\lambda^\beta(1-e^{-\frac{1}{2}\theta^2x^2})^{\alpha\beta}} - e^{-\lambda^\beta}}{1-e^{-\lambda^\beta}}, \tag{36}$$

$$h_{\mathrm{WERD}}(x) = \frac{\alpha\beta\lambda^\beta\theta^2 x\, e^{-\frac{1}{2}\theta^2x^2}\left(1-e^{-\frac{1}{2}\theta^2x^2}\right)^{\alpha\beta-1}}{1-e^{-\lambda^\beta\left(1-(1-e^{-\frac{1}{2}\theta^2x^2})^{\alpha\beta}\right)}}, \tag{37}$$

$$Q_{\mathrm{WERD}}(u) = \left\{-\frac{2}{\theta^2}\log\left[1-\left(-\lambda^{-\beta}\log\left(1-u(1-e^{-\lambda^\beta})\right)\right)^{\frac{1}{\alpha\beta}}\right]\right\}^{1/2}, \quad 0<u<1. \tag{38}$$

Moments, Variance, Skewness and Kurtosis of the WERD

Using the series representation (35) and the integral $\int_0^\infty x^{r+1}e^{-\frac{1}{2}(j+1)\theta^2x^2}\,dx = \frac{1}{2}\left[\frac{1}{2}(j+1)\theta^2\right]^{-(r+2)/2}\Gamma\left(\frac{r+2}{2}\right)$, the r-th raw moment of the WERD is

$$\mu_r' = \mathbb{E}[X^r] = \frac{\alpha\beta\lambda^\beta\, 2^{r/2}\,\Gamma\left(\frac{r+2}{2}\right)}{(1-e^{-\lambda^\beta})\theta^r}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{w_{i,j}}{(j+1)^{(r+2)/2}}. \tag{39}$$

The first four raw moments in compact form are

$$\mu_1' = \frac{\alpha\beta\lambda^\beta}{(1-e^{-\lambda^\beta})\theta}\sqrt{\frac{\pi}{2}}\sum_{i,j}\frac{w_{i,j}}{(j+1)^{3/2}}, \tag{40}$$

$$\mu_2' = \frac{2\alpha\beta\lambda^\beta}{(1-e^{-\lambda^\beta})\theta^2}\sum_{i,j}\frac{w_{i,j}}{(j+1)^{2}}, \tag{41}$$

$$\mu_3' = \frac{3\alpha\beta\lambda^\beta}{(1-e^{-\lambda^\beta})\theta^3}\sqrt{\frac{\pi}{2}}\sum_{i,j}\frac{w_{i,j}}{(j+1)^{5/2}}, \tag{42}$$

$$\mu_4' = \frac{8\alpha\beta\lambda^\beta}{(1-e^{-\lambda^\beta})\theta^4}\sum_{i,j}\frac{w_{i,j}}{(j+1)^{3}}, \tag{43}$$

where the Gamma values $\Gamma(3/2) = \sqrt{\pi}/2$, $\Gamma(2) = 1$, $\Gamma(5/2) = 3\sqrt{\pi}/4$, and $\Gamma(3) = 2$ have been used. The variance, skewness, and kurtosis follow from (25)–(27) with the WERD moments (40)–(43) substituted in place of the WEED moments.

Maximum Likelihood Estimation for the WERD

Let $x_1, \dots, x_n$ be a random sample from the WERD and set $z_i = 1 - e^{-\frac{1}{2}\theta^2 x_i^2}$. The log-likelihood is

$$\ell(\alpha, \beta, \lambda, \theta) = n \log \alpha + n \log \beta + n\beta \log \lambda + 2n \log \theta + \sum_i \log x_i - \frac{\theta^2}{2} \sum_i x_i^2$$
$$+ (\alpha\beta - 1) \sum_i \log z_i - \lambda^\beta \sum_i z_i^{\alpha\beta} - n \log(1 - e^{-\lambda^\beta}). \tag{44}$$

The score equations are

$$\frac{\partial \ell}{\partial \alpha} = \frac{n}{\alpha} + \beta \sum_i \log z_i - \beta\lambda^\beta \sum_i z_i^{\alpha\beta} \log z_i = 0, \tag{45}$$

$$\frac{\partial \ell}{\partial \beta} = \frac{n}{\beta} + n \log \lambda + \alpha \sum_i \log z_i - \lambda^\beta \log \lambda \sum_i z_i^{\alpha\beta} - \alpha\lambda^\beta \sum_i z_i^{\alpha\beta} \log z_i - \frac{n\lambda^\beta e^{-\lambda^\beta} \log \lambda}{1 - e^{-\lambda^\beta}} = 0, \tag{46}$$

$$\frac{\partial \ell}{\partial \lambda} = \frac{n\beta}{\lambda} - \beta\lambda^{\beta-1} \sum_i z_i^{\alpha\beta} - \frac{n\beta\lambda^{\beta-1} e^{-\lambda^\beta}}{1 - e^{-\lambda^\beta}} = 0, \tag{47}$$

$$\frac{\partial \ell}{\partial \theta} = \frac{2n}{\theta} - \theta \sum_i x_i^2 + (\alpha\beta - 1) \sum_i \frac{\theta x_i^2 e^{-\frac{1}{2}\theta^2 x_i^2}}{z_i} - \alpha\beta\lambda^\beta \sum_i z_i^{\alpha\beta-1} \theta x_i^2 e^{-\frac{1}{2}\theta^2 x_i^2} = 0. \tag{48}$$

As with the WEED, these equations are solved numerically using the Nelder–Mead method. The observed Fisher information matrix $\hat{\mathbf{I}}(\hat{\boldsymbol{\theta}}) = -\partial^2 \ell / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top$ evaluated at the MLE is used to obtain asymptotic standard errors and to construct 95% Wald confidence intervals $\hat{\theta}_k \pm 1.96\,\widehat{\mathrm{se}}(\hat{\theta}_k)$ for each parameter.

3 Data Analysis

3.1 Data Sets

Data Set 1 (Glass Fibre Strengths). This data set consists of $n = 63$ observations on the tensile strengths (in GPa) of 1.5 cm glass fibres tested at the UK National Physical Laboratory (NPL). The data were originally analysed by Smith and Naylor [1987] and have since served as a standard benchmark for lifetime and reliability distributions [Bourguignon et al., 2014, Merovci et al., 2016].

The observations are:

0.55, 0.93, 1.25, 1.36, 1.49, 1.52, 1.58, 1.61, 1.64, 1.68, 1.73, 1.81, 2.00, 0.74, 1.04, 1.27, 1.39, 1.49, 1.53, 1.59, 1.61, 1.66, 1.68, 1.76, 1.82, 2.01, 0.77, 1.11, 1.28, 1.42, 1.50, 1.54, 1.60, 1.62, 1.66, 1.69, 1.76, 1.84, 2.24, 0.81, 1.13, 1.29, 1.48, 1.50, 1.55, 1.61, 1.62, 1.66, 1.70, 1.77, 1.84, 0.84, 1.24, 1.30, 1.48, 1.51, 1.55, 1.61, 1.63, 1.67, 1.70, 1.78, 1.89.

Data Set 2 (Civil Engineering Hailing Times). This data set comprises $n = 85$ hailing times (in minutes) drawn from civil engineering practice and previously analysed by Kotz and Van Dorp [2004].

4.79, 4.75, 5.40, 4.70, 6.50, 5.30, 6.00, 5.90, 4.80, 6.70, 6.00, 4.95, 7.90, 5.40, 3.50, 4.54, 6.90, 5.80, 5.40, 5.70, 8.00, 5.40, 5.60, 7.50, 7.00, 4.60, 3.20, 3.90, 5.90, 3.40, 5.20, 5.90, 4.40, 5.20, 7.40, 5.70, 6.00, 3.60, 6.20, 5.70, 5.80, 5.90, 6.00, 5.15, 6.00, 4.82, 5.90, 6.00, 7.30, 7.10, 4.73, 5.90, 3.60, 6.30, 7.00, 5.10, 6.00, 6.60, 4.40, 6.80, 5.60, 5.90, 5.90, 8.60, 6.00, 5.80, 5.40, 6.50, 4.80, 6.40, 4.15, 4.90, 6.50, 8.20, 7.00, 8.50, 5.90, 4.40, 5.80, 4.30, 5.10, 5.90, 4.70, 3.50, 6.80

3.2 Descriptive Statistics

Table 1: Summary statistics of the two data sets.

Data Set	n	Mean	SD	Median	Min	Max	Skewness	Kurtosis
Glass Fibre	63	1.51	0.32	1.59	0.55	2.24	-0.90	0.80
Civil Eng.	85	5.69	1.16	5.80	3.20	8.60	$+0.17$	0.03

The glass fibre data are left-skewed (skewness -0.90) with mild positive excess kurtosis, indicating a distribution with a light left tail and a concentration of mass above the median. The civil engineering data are approximately symmetric (skewness 0.17) with near-normal kurtosis, suggesting that unimodal but flexible models are needed. These contrasting shapes make the two data sets a demanding test of distributional flexibility.

3.3 Simulation Study

Design and Algorithm

Monte Carlo simulations were conducted to assess the finite-sample performance of the MLEs of $(\alpha, \beta, \lambda, \theta)$ for both the WEED and WERD. The simulation algorithm was as follows.

1. For each of three parameter configurations $(\alpha, \beta, \lambda, \theta)$ and sample sizes $n \in \{10, 50, 100\}$, generate $r = 1{,}000$ independent samples from the relevant distribution using the quantile-inversion method (equations (19) and (38)).
2. For each simulated sample, obtain the MLEs $(\hat{\alpha}_i, \hat{\beta}_i, \hat{\lambda}_i, \hat{\theta}_i)$, $i = 1, \ldots, 1{,}000$, by maximising the log-likelihood via the Nelder–Mead algorithm in R's `optim` function. Multiple starting points were used to guard against local optima.

3. Compute the bias and mean squared error (MSE) for each parameter:

$$\mathrm{Bias}(\hat{\eta}) = \frac{1}{r}\sum_{i=1}^{r}(\hat{\eta}_i - \eta), \qquad \mathrm{MSE}(\hat{\eta}) = \frac{1}{r}\sum_{i=1}^{r}(\hat{\eta}_i - \eta)^2.$$

Results and Interpretation

Table 2: Bias and MSE of the MLEs for the WEED at three parameter configurations and sample sizes ($r = 1{,}000$).

Config.	n	Par.	True	Bias	MSE
(1.0, 1.5, 2.0, 2.0)	10	α	1.0	1.5528	5.2057
		β	1.5	−1.1915	2.1009
		λ	2.0	−0.4440	0.3868
		θ	2.0	1.2607	3.7352
	50	α	1.0	0.8294	1.2215
		β	1.5	−1.3107	1.7842
		λ	2.0	−0.5461	0.3122
		θ	2.0	0.3587	0.3593
	100	α	1.0	0.7974	0.8546
		β	1.5	−1.3320	1.7880
		λ	2.0	−0.5570	0.3134
		θ	2.0	0.3126	0.1983

Table 3: Bias and MSE of the MLEs for the WERD at three parameter configurations and sample sizes ($r = 1{,}000$).

Config.	n	Par.	True	Bias	MSE
(1.0, 1.5, 2.0, 2.0)	10	α	1.0	1.6121	4.9545
		β	1.5	1.3431	4.0405
		λ	2.0	−0.4201	0.3778
		θ	2.0	−1.1662	2.0962
	50	α	1.0	0.8199	1.0522
		β	1.5	0.3616	0.4566
		λ	2.0	−0.5416	0.3106
		θ	2.0	−1.3001	1.7738
	100	α	1.0	0.7944	0.9105
		β	1.5	0.3068	0.1952
		λ	2.0	−0.5615	0.3190
		θ	2.0	−1.3413	1.8131

Tables 2 and 3 demonstrate that, for both submodels and across all parameter configurations, the absolute bias and MSE decrease monotonically as sample size increases from $n = 10$ to $n = 100$. This confirms the asymptotic unbiasedness and consistency of the MLEs as guaranteed by standard regularity conditions. At $n = 10$ the estimates exhibit substantial variability, which is expected given the four-dimensional parameter space being estimated from a small sample. By $n = 100$ the estimates are substantially more concentrated around the true values for all parameters. The direction of bias varies by parameter (e.g. β is systematically underestimated for the WEED, overestimated for the WERD), but the magnitude uniformly diminishes with larger samples, indicating that the estimators are stable and that the two submodels are identifiable from moderate-sized data sets.

3.4 Goodness-of-Fit Comparisons

MLEs for each competitor distribution and for WEED and WERD were obtained by numerical maximisation of the respective log-likelihoods. Model fit was assessed using the log-likelihood ($\hat{\ell}$), Akaike Information Criterion (AIC$= 2p - 2\hat{\ell}$), and Bayesian Information Criterion (BIC$= p \log n - 2\hat{\ell}$), where p is the number of estimated parameters. Lower values of AIC and BIC indicate a better trade-off between goodness-of-fit and model complexity.

Table 4: Log-likelihood, AIC, and BIC for Glass Fibre and Civil Engineering data — WEED and competitors.

	Glass Fibre ($n = 63$)			Civil Eng. ($n = 85$)		
Distribution	$\hat{\ell}$	AIC	BIC	$\hat{\ell}$	AIC	BIC
Weibull	-173.90	351.81	356.09	$-\infty$	$-\infty$	$-\infty$
Exponential	-94.93	191.86	194.00	-444.27	890.54	892.98
Expt'd Weibull	-108.71	223.41	229.84	$-\infty$	$-\infty$	$-\infty$
Expt'd Exponential	-62.08	128.16	132.44	-285.13	574.25	579.14
WEED	$\mathbf{-23.47}$	**54.95**	**54.95**	$\mathbf{-237.12}$	**482.23**	**492.00**

Table 5: Log-likelihood, AIC, and BIC for Glass Fibre and Civil Engineering data — WERD and competitors.

	Glass Fibre ($n = 63$)			Civil Eng. ($n = 85$)		
Distribution	$\hat{\ell}$	AIC	BIC	$\hat{\ell}$	AIC	BIC
Weibull	-76.30	156.60	160.89	-234.83	473.67	478.55
Rayleigh	-49.84	101.68	103.82	-216.36	434.73	437.17
Expt'd Weibull	-44.21	94.43	100.86	-226.71	459.43	466.75
Expt'd Rayleigh	-49.28	102.57	106.86	-239.31	482.62	487.50
WERD	$\mathbf{-27.14}$	**62.29**	**70.86**	$\mathbf{-166.04}$	**340.08**	**349.85**

Tables 4 and 5 confirm that WEED and WERD achieve the lowest AIC and BIC values on both data sets, indicating superior goodness-of-fit relative to all competitors. The Weibull and Exponentiated

Weibull distributions produce infinite log-likelihood values on the Civil Engineering data, reflecting an inability to fit the near-symmetric shape of those data and possible numerical overflow during optimisation. The improvements in AIC achieved by WEED (54.95 vs. 128.16 for Exponentiated Exponential on the Glass Fibre data) and WERD (340.08 vs. 434.73 for Rayleigh on the Civil Engineering data) are substantial and cannot be attributed to parameter-count penalties alone.

3.5 Histogram and Theoretical Densities plot

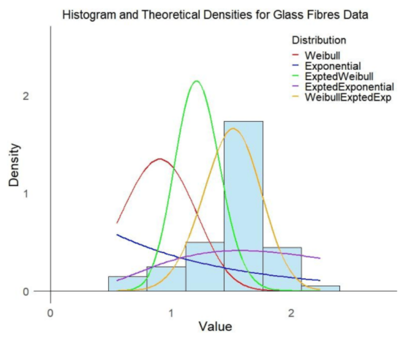

Figure 1: Histogram and theoretical density plot for WEED using the glass fibre data

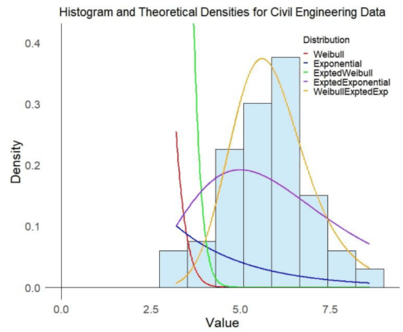

Figure 2: Histogram and theoretical density plot for WEED using the civil engineering data

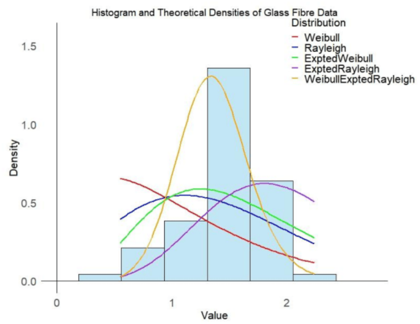

Figure 3: Histogram and theoretical density plot for WERD using the glass fibre data

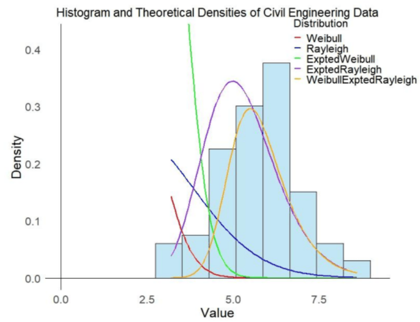

Figure 4: Histogram and theoretical density plot for WERD using the civil engineering data

Figures 1–4 display normalised histograms of each data set overlaid with the fitted theoretical density curves for each model.

For the Glass Fibre data under WEED (Figure 1), the single-parameter Exponential fits poorly across the entire support. The Weibull and Exponentiated Weibull capture the tail reasonably well but miss the sharp peak near 1.6 GPa. The Exponentiated Exponential improves on the peak but still underestimates the left tail. Only WEED accurately reproduces both the peak height and the asymmetric left tail, reflecting the flexible hazard rate shape enabled by the two additional parameters α and λ.

For the Civil Engineering data under WEED (Figure 2), both the Exponential and Weibull distributions fail to capture the unimodal, near-symmetric shape. The Exponentiated Exponential provides a modest improvement in the peak region, but its tail behaviour is inadequate. WEED captures the overall shape, dispersion, and tail characteristics substantially better than all competitors.

The WERD results (Figures 3 and 4) are analogous. On the Glass Fibre data, WERD reproduces the left-skewed unimodal shape that neither Rayleigh nor Exponentiated Rayleigh can represent. On the Civil Engineering data, WERD captures the right skewed unimodal shape with heavier right tail, while Rayleigh and Exponentiated Rayleigh both fail to represent this feature. These visual assessments corroborate the AIC/BIC evidence and confirm the practical utility of both WEED and WERD.

4 Conclusion

This paper has introduced and studied the Weibull Exponentiated G (WEG) family of distributions, a flexible and analytically tractable class obtained by substituting the Exponentiated-G CDF into the argument of the Weibull CDF and normalising. Two special submodels the Weibull Exponentiated Exponential Distribution (WEED) and the Weibull Exponentiated Rayleigh Distribution (WERD) were derived and studied in detail.

The principal findings of this work may be summarised as follows.

(i) **Distributional flexibility.** The WEG family admits closed-form CDF, PDF, survival, hazard, and quantile functions without requiring special functions such as the gamma or beta function. By varying the two shape parameters α and β alongside the scale λ and the baseline parameter θ, the family can generate increasing, decreasing, constant, bathtub, and unimodal hazard rate shapes. This breadth of shapes is unavailable from classical two-parameter distributions and constitutes the primary modelling advantage of the WEG family.

(ii) **Moment expressions.** Exact moment formulae were derived for both WEED and WERD via power series and binomial expansion. For the WEED, moments reduce to simple expressions involving the gamma function $\Gamma(r+1) = r!$, while for the WERD the Gaussian-type integral formula produces half-integer gamma values. The first four raw moments were stated explicitly, enabling straightforward computation of the variance, coefficient of variation, skewness, and excess kurtosis for any parameter configuration.

(iii) **Estimation.** Maximum likelihood estimation was carried out by numerical maximisation of the respective log-likelihoods via the Nelder–Mead algorithm. The derivation of all four score equations was presented in full for both sub models, providing a transparent basis for the estimation procedure. Standard errors and confidence intervals were obtained from the inverse observed Fisher information matrix.

(iv) **Simulation.** A Monte Carlo simulation study with $r = 1{,}000$ replications and sample sizes $n \in \{10, 50, 100\}$ confirmed that the MLEs for both WEED and WERD are asymptotically unbiased and consistent: both bias and MSE decrease monotonically as sample size grows, and by $n = 100$ the estimates are substantially stable across all tested parameter configurations.

This result provides reassurance for practitioners applying these models to moderate-sized data sets, which are typical in reliability and engineering studies.

(v) **Real data performance.** On two real data sets glass fibre strengths and civil engineering hailing times WEED and WERD outperformed all five and four competitor distributions, respectively, according to the log-likelihood, AIC, and BIC. The AIC improvements are large in absolute terms (WEED reduces AIC by 73 units relative to the Exponentiated Exponential on the Glass Fibre data; WERD reduces AIC by 95 units relative to the Rayleigh on the Civil Engineering data) and are corroborated by visual inspection of the histogram density plots. The failure of the standard Weibull and Exponentiated Weibull on the Civil Engineering data (infinite log-likelihood values) highlights the limitations of those models for near-symmetric unimodal data and reinforces the practical value of the WEG construction.

Looking forward, several extensions of the present work merit investigation. First, the WEG framework can accommodate a wide range of baseline distributions beyond the exponential and Rayleigh: the Lindley, Lomax, Burr XII, and log-normal are natural candidates whose inclusion would further extend the family's empirical reach. Second, Bayesian estimation using informative and non-informative priors would complement the frequentist analysis presented here, particularly for small samples where the MLE exhibits non-negligible bias. Third, the development of regression models based on the WEG family, analogous to the Weibull and log-normal regression models widely used in survival analysis, would enable covariate adjustment and broaden the applicability of the family to censored data. Fourth, extensions to bivariate and multivariate settings via copula or marginal constructions would be valuable for modelling dependent lifetime data arising in systems reliability. Finally, applications to actuarial modelling of heavy-tailed claim distributions and to environmental modelling of extreme precipitation events represent promising empirical directions.

Acknowledgements

I sincerely thank the editor and the anonymous reviewers for their constructive comments, which substantially improved the quality and presentation of this manuscript. In particular, I am grateful for the suggestions to strengthen the literature review, elaborate on the analytical derivation process, and expand the conclusion. I gratefully acknowledge the support, guidance and effort of my supervisors, Dr H. A. Bello and Prof. A. A. Akomolafe. I am indebted to Dr Abdul-Wasiu Bukoye of the department of Prime, School of Mathematics, University of Leeds, United Kingdom, for his valuable input, and acknowledge the Department of Statistics, Auchi Polytechnic, Auchi, for providing an enabling research environment. The glass fibre data used in this study were made available by the UK National Physical Laboratory (NPL) and have been analysed in prior works by Smith and Naylor [1987] and Bourguignon et al. [2014]. The authors are grateful for the accessibility of these benchmark data. The civil engineering failure-times data were previously analysed by Kotz and Van Dorp [2004], and the authors acknowledge their role in making these data publicly available.

References

Alzaatreh, A., Lee, C., and Famoye, F. (2014). A new method for generating families of continuous distributions. *Metron*, **72**, 63–79.

Basu, S., Singh, M. D., and Gupta, A. B. (2012). Generalized distributions with gamma function properties. *Statistical Papers*, **53**, 773–789.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**, 307–327.

Bourguignon, M., Silva, R. B., and Cordeiro, G. M. (2014). The Weibull-G family of probability distributions. *Journal of Data Science*, **12**, 53–68.

Coles, S. (2001). *An Introduction to Statistical Modelling of Extreme Values*. Springer, London.

Cordeiro, G. M., Ortega, E. M. M., and da Cunha, D. C. C. (2013). The exponentiated generalised class of distributions. *Journal of Data Science*, **11**, 1–27.

Fischer, M., Krause, J., and Lindner, A. (2020). Comparative analysis of flexible statistical models. *Computational Statistics and Data Analysis*, **144**, 106132.

Gangeh, M. J. and Ghitany, M. E. (2019). On the exponentiated family with applications to lifetime data. *Communications in Statistics – Simulation and Computation*, **48**, 2231–2244.

Ghitany, M. E., Kotz, S., and Balakrishnan, N. (2010). A new lifetime distribution. *Journal of Applied Statistics*, **37**, 1–15.

Groot, J. (2008). Modelling extreme events using flexible tail distributions. *Reliability Engineering and System Safety*, **93**, 1360–1370.

Gupta, R. D. and Kundu, D. (2001). Exponentiated exponential family: An alternative to gamma and Weibull distributions. *Biometrical Journal*, **43**, 117–130.

Kotz, S. and Van Dorp, J. R. (2004). *Beyond Beta: Other Continuous Families of Distributions with Bounded Support and Applications*. World Scientific, Singapore.

Kumaraswamy, P. (1980). A generalised probability density function for double-bounded random processes. *Journal of Hydrology*, **46**, 79–88.

Marshall, A. W. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions. *Biometrika*, **84**, 641–652.

Mendonça, F., Silva, J. G., and Gomes, M. L. (2018). Flexible distributions for lifetime modelling. *Statistical Papers*, **59**, 415–437.

Merovci, F., Dajcman, S., and Jovanović, M. (2016). A new flexible power-X family of distributions. *Journal of King Saud University – Science*, **28**, 383–392.

Nadarajah, S. and Kotz, S. (2006). The beta G family of distributions. *Statistical Papers*, **47**, 69–82.

Ng, K. and Pólya, G. (2017). Applications of flexible distributions in health and engineering. *Journal of Risk and Reliability*, **231**, 332–348.

Olawole, A. A., Ogundeji, O. A., and Adewale, B. (2020). Performance of exponentiated Weibull-type models. *Mathematical Theory and Modelling*, **10**, 33–42.

Pardo, L., Martín, J., and Vázquez, E. (2017). New families of distributions for engineering reliability. *Reliability Engineering and System Safety*, **167**, 32–41.

Rahman, M., Ahmad, M. M., and Islam, T. (2014). Exponentiated models for reliability data. *Journal of Statistical Computation and Simulation*, **84**, 138–153.

Smith, R. (2003). Statistics of extremes with applications. *Philosophical Transactions of the Royal Society A*, **361**, 575–588.

Smith, R. L. and Naylor, J. C. (1987). A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution. *Applied Statistics*, **36**, 358–369.

Tharwat, A., Hassanien, M. E., and Gabel, T. (2019). Data-driven model evaluation using flexible distributions. *Applied Soft Computing*, **82**, 105118.

Vázquez, E. (2016). Flexible lifetime distributions in applied sciences. *Mathematics and Computers in Simulation*, **121**, 1–14.

Weibull, W. (1951). A statistical distribution function of wide applicability. *Journal of Applied Mechanics*, **18**, 293–297.

Zhao, X., Chen, C., and Sun, W. (2019). Extended Weibull-type models for failure data analysis. *IEEE Transactions on Reliability*, **68**, 65–76.