



Research article**A new extension of the Rayleigh distribution: Properties, different methods of estimation, and an application to medical data****Dawlah Alsulami***

Department of Statistics, Faculty of Science, King Abdulaziz University, Jeddah 21589, Saudi Arabia

* **Correspondence:** Email: dalsulami@kau.edu.sa.

Abstract: Statistical distributions play a crucial role in modeling and analyzing real data with complex behavior. Modifying or extending traditional distributions to better capture the complex pattern of a real-world phenomenon have attracted researchers' attention. In this paper, we propose a distribution adaptable to different types of medical data: the exponentiated generalized Weibull–Rayleigh (EGWR) distribution. Its hazard function exhibits different shapes, demonstrating high flexibility in modeling different patterns. For the proposed distribution, some statistical properties, such as moments, Rényi entropy, and order statistics, are discussed. Different methods of estimation—maximum likelihood, least squares, maximum product of spacing, and Cramér–von Mises—were employed to estimate the distribution parameters. The efficiency of these methods in estimating the distribution parameters was compared in three simulation studies and three medical datasets. Furthermore, the goodness of the proposed distribution in fitting real data was examined, and the results demonstrated the efficiency and flexibility of the EGWR distribution in modeling medical data compared to other distributions.

Keywords: exponentiated generalized distribution; Weibull-G distributions; parameter estimation; medical data

Mathematics Subject Classification: 60E05, 62F10

1. Introduction

Statistical distributions play a crucial role in modeling and analyzing real data with complex behavior in many fields, such as biology, engineering, and medicine. Traditional distributions sometimes fail to adequately represent the patterns seen in real world data. Thus, researchers have introduced sophisticated alternatives to existing distributions. These new flexible distributions have high performance in modeling complex patterns, and they are in high demand as users seek to achieve accurate modeling and robust analysis.

In recent years, several approaches have been suggested to achieve highly accurate distributions

when fitting different types of data. One popular method for generating more flexible distributions involves raising the cumulative distribution function (cdf) of the base distribution using a shape parameter. This approach was proposed by [1] and is called the exponentiated method. The researchers in [2] extended this approach to more a general class of distributions, calling it the exponentiated generalized (EG) method, by adding another shape parameter to the cdf of the base distribution as follows:

$$F(x, a, b, \zeta) = [1 - [1 - G(x, \zeta)]^a]^b, \quad (1.1)$$

with the probability density function (pdf)

$$f(x, a, b, \zeta) = abg(x, \zeta) [1 - G(x, \zeta)]^{a-1} [1 - [1 - G(x, \zeta)]^a]^{b-1}, \quad (1.2)$$

where the shape parameters $a, b > 1$, and ζ is a vector of parameters for the base distribution with the cdf $G(x, \zeta)$.

Since its publication, this approach has been adopted by many researchers working to introduce new flexible distributions. Examples include the exponentiated generalized inverted Gompertz [3], exponentiated generalized exponential geometric [4], exponentiated generalized inverse flexible Weibull [5], exponentiated generalized gull alpha power Rayleigh [6], generalized exponentiated Fréchet–Weibull [7], Burr III extended exponentiated Weibull [8], exponentiated generalized Marshall Olkin exponential [9], generalized exponentiated unit Gompertz [10], and type-I heavy-tailed exponentiated generalised-log-logistic [11] distributions.

Another well-known method for attaining more versatile distributions is the transformed-transformer (T-X) method, in which any continuous distribution can be used as a generator. The researchers in [12] proposed the widely known Weibull-G class of distributions with the following cdf and pdf:

$$F(x, c, \beta, \zeta) = 1 - e^{-\left(\frac{W(G(x, \zeta))}{\beta}\right)^c}, \quad (1.3)$$

$$f(x, c, \beta, \zeta) = \frac{c}{\beta} \frac{g(x; \zeta)}{1 - G(x; \zeta)} \left(\frac{W(G(x, \zeta))}{\beta}\right)^{c-1} e^{-\left(\frac{W(G(x, \zeta))}{\beta}\right)^c}, \quad (1.4)$$

where $c, \beta > 0$ are the shape and scale parameters of the Weibull distribution, and $W(G(x, \zeta))$ is the weight function. This family with different forms of $W(G(x, \zeta))$ has been used extensively by researchers to generate more adaptable distributions capable for the fitting of lifetime data. Some of these distributions include Weibull–Gompertz Makeham [13], generalized Weibull–Lindley [14], extended Weibull Fréchet [15], Weibull Marshall–Olkin power Lomax [16], four-parameter Weibull extended Weibull [17], odd Weibull inverse Gompertz [18], Weibull-inverse Nadarajah Haghghi [19], and type-2 Gumbel Weibull exponential [20].

In this research, the EG method and Weibull-G approach are combined to introduce a new extension of the Rayleigh distribution with five parameters: the exponentiated generalized Weibull–Rayleigh (EGWR) distribution. The introduction of the EGWR model was motivated by the following objectives:

- Deliver a new extension of the Rayleigh distribution with improved efficiency and flexibility.

- Enhance the elasticity of the Rayleigh distribution by combining two attractive techniques: the EG approach and the Weibull-G class, both of which have shown great flexibility in a wide range of applications.
- Provide an adaptable distribution with great versatility as an alternative to existing distributions in modeling data with different forms as observed from the different shapes of the density function.
- Compared to competing distributions, offer a better representation of the variability of real-world data in the field of medicine.

The remaining sections of this paper are arranged as follows: In Section 2, the EGWR distribution is introduced. In Section 3, important properties of the proposed distribution are presented. In Section 4, we provide the estimators of the distribution parameters using four different methods. To test the efficiency of these methods, related simulation studies were performed, and their results are described in Section 5. In Section 6, different medical datasets are analyzed to compare various estimation methods and illustrate the flexibility of the proposed distribution. Finally, in Section 7, we conclude the paper.

2. The exponentiated generalized Weibull Rayleigh distribution

If X is a random variable from the Rayleigh distribution with scale parameter $\sigma > 0$, then the cdf and pdf are given by the following:

$$G(x; \sigma) = 1 - e^{-\frac{x^2}{2\sigma^2}}, \quad (2.1)$$

$$g(x; \sigma) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}. \quad (2.2)$$

By considering weight function $W(G(x, \zeta)) = -\log(1 - (G(x, \zeta)))$, using (2.1) in (1.3), the Weibull Rayleigh distribution is obtained with the following cdf:

$$F(x, c, \beta, \zeta) = 1 - e^{-\left(\frac{x^2}{2\sigma^2\beta}\right)^c}. \quad (2.3)$$

Substituting (2.3) in (1.1), the EGWR distribution is defined by the following cdf:

$$F(x; a, b, c, \beta, \sigma) = \left[1 - e^{-a\left(\frac{x^2}{2\sigma^2\beta}\right)^c} \right]^b. \quad (2.4)$$

The corresponding pdf is as follows:

$$f(x; a, b, c, \beta, \sigma) = \frac{abc}{(2\sigma^2\beta)^c} 2x^{2c-1} e^{-a\left(\frac{x^2}{2\sigma^2\beta}\right)^c} \left[1 - e^{-a\left(\frac{x^2}{2\sigma^2\beta}\right)^c} \right]^{b-1}. \quad (2.5)$$

Figures 1 and 2 display different shapes for the pdf and the hazard function (hf) of the EGWR distribution achieved using various parameter values. The pdf takes various forms, such as symmetric, asymmetric, increasing, J-shaped, and reverse J-shaped. In addition, the hf can model data with bathtub and upside-down bathtub shapes. This demonstrates the distribution's capacity to accommodate diverse forms of data, making it a valuable tool for modeling complex phenomena.

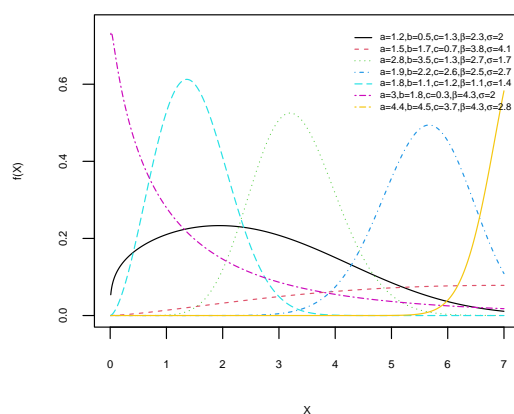


Figure 1. Plots for the EGWR pdf function with various parameter values.

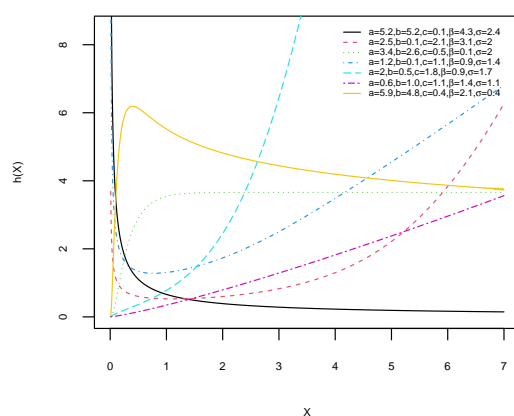


Figure 2. Plots for the EGWR hf with various parameter values.

2.1. Expansion of the density function

The binomial expansion theorem is defined as follows:

$$(1-x)^{n-1} = \sum_{i=0}^{\infty} (-1)^i \binom{n-1}{i} x^i, \quad (2.6)$$

where $|x| < 1$ and $n > 0$. Therefore,

$$\left[1 - e^{-a\left(\frac{x^2}{2\sigma^2\beta}\right)^c} \right]^{b-1} = \sum_{i=0}^{\infty} (-1)^i \binom{b-1}{i} e^{-ai\left(\frac{x^2}{2\sigma^2\beta}\right)^c}.$$

Thus, the pdf can be re-written as

$$f(x; \zeta) = \sum_{i=0}^{\infty} (-1)^i \binom{b-1}{i} \frac{2abc}{(2\sigma^2\beta)^c} x^{2c-1} e^{-a(i+1)\left(\frac{x^2}{2\sigma^2\beta}\right)^c},$$

from the following power series:

$$e^{-a(i+1)\left(\frac{x^2}{2\sigma^2\beta}\right)^c} = \sum_{j=0}^{\infty} \frac{(-1)^j}{j!} (a(i+1))^j \left(\frac{x^2}{2\sigma^2\beta}\right)^{cj}.$$

Hence, the pdf of the EGWR distribution takes the following form:

$$f(x; \zeta) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{(-1)^{i+j}}{j!} \binom{b-1}{i} a^{j+1} (i+1)^j \frac{2bc}{(2\sigma^2\beta)^{c(j+1)}} x^{2c(j+1)-1}.$$

3. Statistical properties

Some important statistical properties of the EGWR distribution—moments, incomplete moments, moment generating function, order statistics, and Rényi entropy—are presented in this section.

3.1. Quantile and median

For a random variable X from the EGWR distribution, the quantile function can be obtained by inverting (2.4) and presenting the expression in an explicit form as follows:

$$x_q = \left\{ 2\sigma^2\beta \left[-\frac{1}{a} \log \left(1 - q^{\frac{1}{b}} \right) \right]^{\frac{1}{c}} \right\}^{\frac{1}{2}}, \quad 0 < q < 1. \quad (3.1)$$

Therefore, the median is given by

$$x_{0.5} = \left\{ 2\sigma^2\beta \left[-\frac{1}{a} \log \left(1 - 0.5^{\frac{1}{b}} \right) \right]^{\frac{1}{c}} \right\}^{\frac{1}{2}}. \quad (3.2)$$

3.2. Moments

The r th moment [21] of the EGWR distribution is given by

$$\mu_r = E(X^r) = \int_0^{\infty} x^r \frac{2abc}{(2\sigma^2\beta)^c} x^{2c-1} e^{-a\left(\frac{x^2}{2\sigma^2\beta}\right)^c} \left[1 - e^{-a\left(\frac{x^2}{2\sigma^2\beta}\right)^c} \right]^{b-1} dx.$$

Using (2.6) for $\left(1 - e^{-a\left(\frac{x^2}{2\sigma^2\beta}\right)^c} \right)^{b-1}$, we have

$$\mu_r = \sum_{i=0}^{\infty} (-1)^i \binom{b-1}{i} \frac{2abc}{(2\sigma^2\beta)^c} \int_0^{\infty} x^{2c+r-1} e^{-a(i+1)\left(\frac{x^2}{2\sigma^2\beta}\right)^c} dx$$

$$= \sum_{i=0}^{\infty} (-1)^i \binom{b-1}{i} \frac{b}{a^{\frac{r}{2c}} (i+1)^{\frac{r}{2c}+1}} (2\sigma^2\beta)^{\frac{r}{2}} \Gamma\left(\frac{r}{2c} + 1\right). \quad (3.3)$$

The mean and variance of the EGWR distribution can be found from (3.3) as follows:

$$\begin{aligned} \text{Mean} &= \mu_1 = E(x), \\ \text{Variance} &= \mu_2 - \mu_1^2 = E(x^2) - E(x)^2. \end{aligned}$$

Thus, the coefficient of variation (COV), the dispersion index (DI), skewness (SK), and kurtosis (K) can be obtained as follows:

$$\begin{aligned} COV &= \frac{\sqrt{\mu_2 - \mu_1^2}}{\mu_1}, \\ DI &= \frac{\mu_2 - \mu_1^2}{\mu_1}, \\ SK &= \frac{\mu_3 - 3\mu_2\mu_1 + 2\mu_1^3}{(\mu_2 - \mu_1^2)^{\frac{3}{2}}}, \\ DI &= \frac{\mu_4 - 4\mu_3\mu_1 + 6\mu_2\mu_1^2 - 3\mu_1^4}{(\mu_2 - \mu_1^2)^2}. \end{aligned}$$

These measures can be calculated using any reputable statistical software, such as R.

Table 1 displays the measures' values for the following cases of parameters:

- I:** $a = 3.4, b = 3.0, c = 3.6, \beta = 1.3, \sigma = 3.9.$
- II:** $a = 2.3, b = 1.1, c = 3.7, \beta = 0.6, \sigma = 2.9.$
- III:** $a = 2.5, b = 1.0, c = 1.5, \beta = 2.3, \sigma = 0.3.$
- IV:** $a = 1.7, b = 0.1, c = 1.0, \beta = 3.8, \sigma = 1.8.$
- V:** $a = 3.9, b = 1.0, c = 2.4, \beta = 5.1, \sigma = 1.9.$
- VI:** $a = 1.3, b = 2.3, c = 0.6, \beta = 2.0, \sigma = 1.5.$

Table 1. Results (mean, variance, coefficient of variation [COV], dispersion index [CI], skewness [SK], and kurtosis [K]) generated by changes to the parameters.

	Mean	Variance	COV	DI	SK	K
I	5.635465	0.263489	0.091086	0.046755	-0.208116	3.102182
II	2.700400	0.165382	0.150597	0.061243	-0.477589	3.253748
III	0.423345	0.023673	0.363446	0.055921	0.168102	2.729464
IV	0.748891	1.661996	1.721458	2.219276	2.481749	9.958432
V	4.186111	0.990484	0.237745	0.236612	-0.225393	2.850532
VI	3.477175	4.145255	0.585529	1.192133	1.18643	5.137207

As shown in Table 1, the EGWR distribution can be negatively (left) skewed or positively (right) skewed; this illustrates the flexibility of the distribution in fitting various types of asymmetrical data.

Moreover, the kurtosis values indicate that the distribution can be leptokurtic or platykurtic. Hence, the distribution is capable for modeling data with sharp peaks and lighter tails or flat peaks with heavier tails. Additionally, the values of the dispersion index shows the flexibility in modeling data with various spreads.

3.3. Incomplete moments

The s th incomplete moment [22] of the EGWR distribution is defined as

$$I_s = E(X^s) = \int_0^z x^s \frac{2abc}{(2\sigma^2\beta)^c} x^{2c-1} e^{-a\left(\frac{x^2}{2\sigma^2\beta}\right)^c} \left[1 - e^{-a\left(\frac{x^2}{2\sigma^2\beta}\right)^c}\right]^{b-1} dx,$$

where from binomial expansion (2.6), we have

$$\begin{aligned} I_s &= \sum_{i=0}^{\infty} (-1)^i \binom{b-1}{i} \frac{2abc}{(2\sigma^2\beta)^c} \int_0^z x^{2c+s-1} e^{-a(i+1)\left(\frac{x^2}{2\sigma^2\beta}\right)^c} dx \\ &= \sum_{i=0}^{\infty} (-1)^i \binom{b-1}{i} \frac{b}{a^{\frac{s}{2c}} (i+1)^{\frac{s}{2c}+1}} (2\sigma^2\beta)^{\frac{s}{2}} \gamma\left(\frac{s}{2c} + 1, u_z\right), \end{aligned} \quad (3.4)$$

where $u_z = a(i+1)\left(\frac{z^2}{2\sigma^2\beta}\right)^c$.

3.4. Moment generating function

The moment generating function [23] for the EGWR distribution is obtained as follows:

$$M_X(t) = E(e^{tx}) = \int_0^{\infty} e^{tx} \frac{2abc}{(2\sigma^2\beta)^c} x^{2c-1} e^{-a\left(\frac{x^2}{2\sigma^2\beta}\right)^c} \left[1 - e^{-a\left(\frac{x^2}{2\sigma^2\beta}\right)^c}\right]^{b-1} dx.$$

By applying the binomial expansion (2.6), we obtain

$$M_X(t) = \sum_{i=0}^{\infty} (-1)^i \binom{b-1}{i} \frac{2abc}{(2\sigma^2\beta)^c} \int_0^{\infty} x^{2c-1} e^{tx} e^{-a(i+1)\left(\frac{x^2}{2\sigma^2\beta}\right)^c} dx,$$

where by using the power series expansion for e^{tx} , we have

$$\begin{aligned} M_X(t) &= \sum_{i=0}^{\infty} \sum_{r=0}^{\infty} (-1)^i \binom{b-1}{i} \frac{t^r}{r!} \frac{2abc}{(2\sigma^2\beta)^c} \int_0^{\infty} x^{2c+r-1} e^{-a(i+1)\left(\frac{x^2}{2\sigma^2\beta}\right)^c} dx \\ &= \sum_{i=0}^{\infty} \sum_{r=0}^{\infty} (-1)^i \binom{b-1}{i} \frac{t^r}{r!} \frac{b}{a^{\frac{r}{2c}} (i+1)^{\frac{r}{2c}+1}} (2\sigma^2\beta)^{\frac{r}{2}} \Gamma\left(\frac{r}{2c} + 1\right). \end{aligned} \quad (3.5)$$

3.5. Rényi entropy

The Rényi entropy [24] for the EGW-R distribution is defined as

$$H_R(x) = \frac{1}{1-R} \log \left[\int_0^{\infty} (f(x; \zeta))^R dx \right],$$

where

$$(f(x; \zeta))^R = \left(\frac{2abc}{(2\sigma^2\beta)^c} \right)^R x^{R(2c-1)} e^{-aR\left(\frac{x^2}{2\sigma^2\beta}\right)^c} \left[1 - e^{-a\left(\frac{x^2}{2\sigma^2\beta}\right)^c} \right]^{R(b-1)}.$$

Using the binomial expansion (2.6),

$$\left[1 - e^{-a\left(\frac{x^2}{2\sigma^2\beta}\right)^c} \right]^{R(b-1)} = \sum_{i=0}^{\infty} (-1)^i \binom{R(b-1)}{i} e^{-ai\left(\frac{x^2}{2\sigma^2\beta}\right)^c}.$$

Thus,

$$(f(x; \zeta))^R = \sum_{i=0}^{\infty} (-1)^i \binom{R(b-1)}{i} \left(\frac{2abc}{(2\sigma^2\beta)^c} \right)^R x^{R(2c-1)} e^{-a(R+i)\left(\frac{x^2}{2\sigma^2\beta}\right)^c},$$

and therefore,

$$\begin{aligned} \int_0^{\infty} (f(x; \zeta))^R dx &= \sum_{i=0}^{\infty} (-1)^i \binom{R(b-1)}{i} \left(\frac{2abc}{(2\sigma^2\beta)^c} \right)^R \\ &\quad \times \int_0^{\infty} x^{R(2c-1)} e^{-a(R+i)\left(\frac{x^2}{2\sigma^2\beta}\right)^c} dx \\ &= \sum_{i=0}^{\infty} (-1)^i \binom{R(b-1)}{i} (2c)^{R-1} b^R \left(\frac{1}{a} \right)^{\frac{1-R}{2c}} (2\sigma^2\beta)^{\frac{1-R}{2}} \\ &\quad \times \left[\frac{1}{(R+i)} \right]^{\frac{R(2c-1)+1}{2c}} \Gamma\left(\frac{R(2c-1)+1}{2c} \right). \end{aligned} \quad (3.6)$$

The Rényi entropy is then given by the following:

$$\begin{aligned} H_R(x) &= \frac{1}{1-R} \log \left[\sum_{i=0}^{\infty} (-1)^i \binom{R(b-1)}{i} (2c)^{R-1} b^R \left(\frac{1}{a} \right)^{\frac{1-R}{2c}} \left[\frac{1}{(R+i)} \right]^{\frac{R(2c-1)+1}{2c}} \right. \\ &\quad \left. \times (2\sigma^2\beta)^{\frac{1-R}{2}} \Gamma\left(\frac{R(2c-1)+1}{2c} \right) \right] \\ &= \frac{1}{1-R} \left\{ (R-1) \log(2c) + R \log(b) + \left(\frac{1-R}{2} \right) \log(2\sigma^2\beta) - \left(\frac{1-R}{2c} \right) \log(a) \right. \\ &\quad \left. + \log \left[\sum_{i=0}^{\infty} (-1)^i \binom{R(b-1)}{i} \left[\frac{1}{(R+i)} \right]^{\frac{R(2c-1)+1}{2c}} \Gamma\left(\frac{R(2c-1)+1}{2c} \right) \right] \right\}. \end{aligned} \quad (3.7)$$

3.6. Distribution of order statistics

Let x_1, x_2, \dots, x_n be a sample selected from the EGW-R distribution with order statistics $x_{1:n}, x_{2:n}, \dots, x_{n:n}$. The i th order statistics then have the following distribution [25]:

$$f_{i:n}(x; \zeta) = \frac{n!}{(i-1)!(n-i)!} [F(x; \zeta)]^{i-1} [1 - F(x; \zeta)]^{n-i} f(x; \zeta).$$

By using the binomial expansion (2.6),

$$[1 - F(x; \zeta)]^{n-i} = \sum_{k=0}^{n-i} (-1)^k \binom{n-i}{k} [F(x; \zeta)]^k,$$

$$f_{i:n}(x; \zeta) = \sum_{k=0}^{n-i} (-1)^k \binom{n-i}{k} \frac{n!}{(i-1)!(n-i)!} [F(x; \zeta)]^{k+i-1} f(x; \zeta),$$

where

$$[F(x; \zeta)]^{k+i-1} f(x; \zeta) = \frac{2abc}{(2\sigma^2\beta)^c} x^{2c-1} e^{-a(x^2 2\sigma^2\beta)^c} \left[1 - e^{-a\left(\frac{x^2}{2\sigma^2\beta}\right)^c} \right]^{b(k+i)-1}.$$

By the applying binomial expansion (2.6),

$$\left[1 - e^{-a\left(\frac{x^2}{2\sigma^2\beta}\right)^c} \right]^{b(k+i)-1} = \sum_{j=0}^{\infty} (-1)^j \binom{b(k+i)-1}{j} e^{-aj\left(\frac{x^2}{2\sigma^2\beta}\right)^c}.$$

Thus,

$$[F(x; \zeta)]^{k+i-1} f(x; \zeta) = \sum_{j=0}^{\infty} (-1)^j \binom{b(k+i)-1}{j} \frac{2abc}{(2\sigma^2\beta)^c} x^{2c-1} e^{-a(j+1)\left(\frac{x^2}{2\sigma^2\beta}\right)^c}.$$

By applying the power series expansion for $e^{-a(j+1)\left(\frac{x^2}{2\sigma^2\beta}\right)^c}$, we obtain

$$\begin{aligned} [F(x; \zeta)]^{k+i-1} f(x; \zeta) &= \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{j+m}}{m!} \binom{b(k+i)-1}{j} a^{m+1} (j+1)^m \\ &\quad \times \frac{2bc}{(2\sigma^2\beta)^{c(m+1)}} x^{2c(m+1)-1}. \end{aligned}$$

Therefore, the distribution takes the following form:

$$\begin{aligned} f_{i:n}(x; \zeta) &= \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \sum_{k=0}^{n-i} \frac{(-1)^{j+m+k}}{m!} \frac{n!}{(i-1)!(n-i)!} \binom{b(k+i)-1}{j} \binom{n-i}{k} \\ &\quad \times a^{m+1} (j+1)^m 2bc \left(\frac{1}{2\sigma^2\beta} \right)^{c(m+1)} x^{2c(m+1)-1}. \end{aligned} \quad (3.8)$$

4. Methods of estimation

4.1. Maximum likelihood method

The maximum likelihood (ML) method is one of the most widely applied methods of estimation in the literature [26, 27]. In this method, the values of the parameters are obtained by maximizing the log-likelihood function. Thus, if x_1, x_2, \dots, x_n follows the EGWR distribution with the vector of parameters $\theta = (a, b, c, \beta, \sigma)$, then the log-likelihood function (ℓ) is defined by

$$\begin{aligned} \ell &= n \log 2 + n \log a + n \log b + n \log c - nc \log 2 - nc \log \sigma^2 - nc \log \beta + (2c-1) \sum_{i=1}^n \log x_i \\ &\quad - a \sum_{i=1}^n \left(\frac{x_i^2}{2\sigma^2\beta} \right)^c + (b-1) \sum_{i=1}^n \log \left[1 - e^{-a\left(\frac{x_i^2}{2\sigma^2\beta}\right)^c} \right]. \end{aligned} \quad (4.1)$$

The aforementioned log-likelihood function is too complicated and difficult to solve analytically. Thus, the derivative of (4.1), with respect to each parameter, is calculated as follows:

$$\frac{\partial \ell}{\partial a} = \frac{n}{a} - \sum_{i=1}^n \left(\frac{x_i^2}{2\sigma^2\beta} \right)^c + (b-1) \sum_{i=1}^n \left[\frac{\left(\frac{x_i^2}{2\sigma^2\beta} \right)^c e^{-a \left(\frac{x_i^2}{2\sigma^2\beta} \right)^c}}{1 - e^{-a \left(\frac{x_i^2}{2\sigma^2\beta} \right)^c}} \right], \quad (4.2)$$

$$\frac{\partial \ell}{\partial b} = \frac{n}{b} + \sum_{i=1}^n \log \left[1 - e^{-a \left(\frac{x_i^2}{2\sigma^2\beta} \right)^c} \right], \quad (4.3)$$

$$\begin{aligned} \frac{\partial \ell}{\partial c} = & \frac{n}{c} - n \log 2 - n \log \sigma^2 - n \log \beta + 2 \sum_{i=1}^n \log x_i - a \sum_{i=1}^n \left(\frac{x_i^2}{2\sigma^2\beta} \right)^c \log \left(\frac{x_i^2}{2\sigma^2\beta} \right) \\ & + a(b-1) \sum_{i=1}^n \left[\frac{\left(\frac{x_i^2}{2\sigma^2\beta} \right)^c \log \left(\frac{x_i^2}{2\sigma^2\beta} \right) e^{-a \left(\frac{x_i^2}{2\sigma^2\beta} \right)^c}}{1 - e^{-a \left(\frac{x_i^2}{2\sigma^2\beta} \right)^c}} \right], \end{aligned} \quad (4.4)$$

$$\frac{\partial \ell}{\partial \beta} = -\frac{nc}{\beta} + \frac{ac}{\beta} \sum_{i=1}^n \left(\frac{x_i^2}{2\sigma^2\beta} \right)^c - \frac{ac(b-1)}{\beta} \sum_{i=1}^n \left[\frac{\left(\frac{x_i^2}{2\sigma^2\beta} \right)^c e^{-a \left(\frac{x_i^2}{2\sigma^2\beta} \right)^c}}{1 - e^{-a \left(\frac{x_i^2}{2\sigma^2\beta} \right)^c}} \right], \quad (4.5)$$

and

$$\frac{\partial \ell}{\partial \sigma} = -\frac{2nc}{\sigma} + \frac{2ac}{\sigma} \sum_{i=1}^n \left(\frac{x_i^2}{2\sigma^2\beta} \right)^c - \frac{2ac(b-1)}{\sigma} \sum_{i=1}^n \left[\frac{\left(\frac{x_i^2}{2\sigma^2\beta} \right)^c e^{-a \left(\frac{x_i^2}{2\sigma^2\beta} \right)^c}}{1 - e^{-a \left(\frac{x_i^2}{2\sigma^2\beta} \right)^c}} \right]. \quad (4.6)$$

Therefore, the MLEs of the parameter vector θ can be achieved by equating the Eqs (4.2)–(4.6) to zero and solving them in an iterative manner using any numerical method, such as the Newton–Raphson method. Alternatively, the optimization function *optim* in the statistical software R can be employed to maximize (4.1) directly and obtain the MLEs.

4.2. Ordinary least squares method

The ordinary least squares (LS) method [28] estimates the distribution parameters by minimizing the square difference between the theoretical cdf and the empirical cdf. Let $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ follow the EGWR distribution. The LS method minimizes the following sum of squares:

$$Q(\theta) = \sum_{i=1}^n \left[\left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^b - \frac{i}{n+1} \right]^2. \quad (4.7)$$

Differentiating (4.7) with respect to the vector of parameters θ results in the following equations:

$$\frac{\partial Q(\theta)}{\partial a} = \frac{2b}{(2\beta\sigma^2)^c} \sum_{i=1}^n x_{(i)}^{2c} e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^{b-1} \left[\left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^b - \frac{i}{n+1} \right], \quad (4.8)$$

$$\frac{\partial Q(\theta)}{\partial b} = 2 \sum_{i=1}^n \ln \left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right) \left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^{b-1} \left[\left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^b - \frac{i}{n+1} \right], \quad (4.9)$$

$$\frac{\partial Q(\theta)}{\partial c} = \frac{2ab}{(2\beta\sigma^2)^c} \sum_{i=1}^n x_{(i)}^{2c} \ln \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right) e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^{b-1} \left[\left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^b - \frac{i}{n+1} \right], \quad (4.10)$$

$$\frac{\partial Q(\theta)}{\partial \beta} = \frac{-2abc}{\beta(2\beta\sigma^2)^c} \sum_{i=1}^n x_{(i)}^{2c} e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^{b-1} \left[\left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^b - \frac{i}{n+1} \right], \quad (4.11)$$

$$\frac{\partial Q(\theta)}{\partial \sigma} = \frac{-4abc}{\sigma(2\beta\sigma^2)^c} \sum_{i=1}^n x_{(i)}^{2c} e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^{b-1} \left[\left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^b - \frac{i}{n+1} \right]. \quad (4.12)$$

Therefore, minimizing Eq (4.7) or equivalently solving Eqs (4.8)–(4.12) analytically or numerically using R provides the estimates of the vector of parameters $\theta = (a, b, c, \beta, \sigma)$.

4.3. Maximum product of spacing method

The maximum product of spacing (MPS) method was produced by [29] as an improvement of the ML method. Let x_1, x_2, \dots, x_n be a sample from an EGWR distribution with a vector of parameters $\theta = (a, b, c, \beta, \sigma)$ and order statistics $x_{(1)}, x_{(2)}, \dots, x_{(n)}$. The idea of the MPS is to maximize the following equation:

$$M(\theta) = \frac{1}{n+1} \sum_{i=1}^{n+1} \log D_i(\theta), \quad (4.13)$$

where $D_i(\theta) = F[x_{(i)} | \theta] - F[x_{(i-1)} | \theta]$, $i = 1, 2, \dots, n+1$ are the uniform spacings between the consecutive ordered values, $F[x_{(0)} | \theta] = 0$, and $F[x_{(n+1)} | \theta] = 1$. The estimates of the vector of parameters θ can be obtained by maximizing (4.14) or equivalently solving the system of Eqs (4.15)–(4.19) using any optimization function in R.

$$M(\theta) = \frac{1}{n+1} \sum_{i=1}^{n+1} \log \left[\left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^b - \left(1 - e^{-a \left(\frac{x_{(i-1)}^2}{2\beta\sigma^2} \right)^c} \right)^b \right], \quad (4.14)$$

$$\frac{\partial M(\theta)}{\partial a} = \frac{1}{n+1} \frac{b}{(2\beta\sigma^2)^c} \sum_{i=1}^{n+1} \frac{1}{A_i} \left[x_{(i)}^{2c} e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^{b-1} - x_{(i-1)}^{2c} e^{-a \left(\frac{x_{(i-1)}^2}{2\beta\sigma^2} \right)^c} \left(1 - e^{-a \left(\frac{x_{(i-1)}^2}{2\beta\sigma^2} \right)^c} \right)^{b-1} \right], \quad (4.15)$$

$$\frac{\partial M(\theta)}{\partial b} = \frac{1}{n+1} \sum_{i=1}^{n+1} \frac{1}{A_i} \left[\left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^b \ln \left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right) - \left(1 - e^{-a \left(\frac{x_{(i-1)}^2}{2\beta\sigma^2} \right)^c} \right)^b \ln \left(1 - e^{-a \left(\frac{x_{(i-1)}^2}{2\beta\sigma^2} \right)^c} \right) \right], \quad (4.16)$$

$$\begin{aligned} \frac{\partial M(\theta)}{\partial c} = & -\frac{1}{n+1} \frac{ab}{(2\beta\sigma^2)^c} \sum_{i=1}^{n+1} \frac{1}{A_i} \left[x_{(i)}^{2c} \ln \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right) e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^{b-1} \right. \\ & \left. - x_{(i-1)}^{2c} \ln \left(\frac{x_{(i-1)}^2}{2\beta\sigma^2} \right) e^{-a \left(\frac{x_{(i-1)}^2}{2\beta\sigma^2} \right)^c} \left(1 - e^{-a \left(\frac{x_{(i-1)}^2}{2\beta\sigma^2} \right)^c} \right)^{b-1} \right], \end{aligned} \quad (4.17)$$

$$\begin{aligned} \frac{\partial M(\theta)}{\partial \beta} = & -\frac{1}{n+1} \frac{abc}{\beta(2\beta\sigma^2)^c} \sum_{i=1}^{n+1} \frac{1}{A_i} \left[x_{(i)}^{2c} e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^{b-1} \right. \\ & \left. - x_{(i-1)}^{2c} e^{-a \left(\frac{x_{(i-1)}^2}{2\beta\sigma^2} \right)^c} \left(1 - e^{-a \left(\frac{x_{(i-1)}^2}{2\beta\sigma^2} \right)^c} \right)^{b-1} \right], \end{aligned} \quad (4.18)$$

$$\begin{aligned} \frac{\partial M(\theta)}{\partial \sigma} = & -\frac{1}{n+1} \frac{2abc}{\sigma(2\beta\sigma^2)^c} \sum_{i=1}^{n+1} \frac{1}{A_i} \left[x_{(i)}^{2c} e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^{b-1} \right. \\ & \left. - x_{(i-1)}^{2c} e^{-a \left(\frac{x_{(i-1)}^2}{2\beta\sigma^2} \right)^c} \left(1 - e^{-a \left(\frac{x_{(i-1)}^2}{2\beta\sigma^2} \right)^c} \right)^{b-1} \right], \end{aligned} \quad (4.19)$$

$$\text{where } A_i = \left[\left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^b - \left(1 - e^{-a \left(\frac{x_{(i-1)}^2}{2\beta\sigma^2} \right)^c} \right)^b \right].$$

4.4. Cramér–von Mises method

The Cramér–von Mises method, also called the minimum distance estimator, was proposed by the researchers in [30] to estimate the values of the parameters by minimizing the following objective function:

$$C(\theta) = \frac{1}{12n} + \sum_{i=1}^n \left[\left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^b - \frac{2i-1}{2n} \right]^2. \quad (4.20)$$

Minimizing (4.20) or solving the following system of nonlinear equations (4.21)–(4.25) numerically yields the estimate of the vector of parameters θ :

$$\frac{\partial C(\theta)}{\partial a} = \frac{2b}{(2\beta\sigma^2)^c} \sum_{i=1}^n x_{(i)}^{2c} e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^{b-1} \left[\left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^b - \frac{2i-1}{2n} \right], \quad (4.21)$$

$$\frac{\partial C(\theta)}{\partial b} = 2 \sum_{i=1}^n \ln \left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right) \left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^b \left[\left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^b - \frac{2i-1}{2n} \right], \quad (4.22)$$

$$\frac{\partial C(\theta)}{\partial c} = \frac{2ab}{(2\beta\sigma^2)^c} \sum_{i=1}^n x_{(i)}^{2c} \ln \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right) e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^{b-1} \left[\left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^b - \frac{2i-1}{2n} \right], \quad (4.23)$$

$$\frac{\partial C(\theta)}{\partial \beta} = \frac{-2abc}{\beta(2\beta\sigma^2)^c} \sum_{i=1}^n x_{(i)}^{2c} e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^{b-1} \left[\left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^b - \frac{2i-1}{2n} \right], \quad (4.24)$$

$$\frac{\partial C(\theta)}{\partial \sigma} = \frac{-4abc}{\sigma(2\beta\sigma^2)^c} \sum_{i=1}^n x_{(i)}^{2c} e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^{b-1} \left[\left(1 - e^{-a \left(\frac{x_{(i)}^2}{2\beta\sigma^2} \right)^c} \right)^b - \frac{2i-1}{2n} \right]. \quad (4.25)$$

5. Simulation study

To assess and evaluate the performance of the four methods of estimation in determining the parameters of the EGWR distribution, three cases of simulation studies are presented. A number of iterations equal to $nsim = 10,000$ were applied with different sample sizes ($n = 50, 100, 300$, and 500) for the following three cases of true values of the vector of parameters θ_{tr} :

Case I: $a = 3.5, b = 4.1, c = 2.7, \beta = 3.9, \sigma = 3.1,$

Case II: $a = 2.1, b = 3.7, c = 2.5, \beta = 1.1, \sigma = 2,$

Case III: $a = 1.6, b = 2.8, c = 3.1, \beta = 0.6, \sigma = 2.2.$

The root mean square error (RMSE) measure was used to assess the efficiency of estimation methods selected for each estimator $\hat{\theta}$. The RMSE can be obtained as follows:

$$RMSE(\hat{\theta}) = \sqrt{\frac{\sum_{i=1}^{nsim} (\hat{\theta}_i - \theta_{tr})^2}{nsim}}. \quad (5.1)$$

The estimates obtained by the four methods of estimation together with the RMSE are shown in Tables 2–4 and Figures 3–5.

Table 2. The estimates and RMSE across sample sizes ($N = 50, 150, 300, 500$) for all methods of estimation (case I).

Sample size	Parameter	ML		LS		MPS		CVM	
		Estimate	RMSE	Estimate	RMSE	Estimate	RMSE	Estimate	RMSE
N = 50	a	3.6995	5.1316	3.6385	3.1513	3.8445	5.9930	3.6248	3.2245
	b	7.7331	14.1917	6.3047	7.0769	7.7140	19.3329	6.4418	7.3492
	c	3.1177	1.6661	3.2465	1.8193	2.9335	1.4354	3.3209	1.8888
	β	3.8604	1.8654	3.9251	1.4071	3.9385	2.1858	3.9085	1.3842
	σ	3.0584	0.7980	3.0365	0.5864	3.0872	0.8721	3.0320	0.5879
N = 150	a	3.5902	1.9735	3.5586	2.1011	3.6003	2.2346	3.5788	2.1639
	b	5.2776	4.7973	5.5309	4.8928	5.2036	4.8694	5.5530	4.8901
	c	2.8033	0.6916	2.9100	1.0184	2.7373	0.6651	2.9319	1.0278
	β	3.9164	0.9357	3.9108	0.9627	3.9179	0.9034	3.9242	0.9478
	σ	3.0759	0.3867	3.0650	0.4082	3.0881	0.3944	3.0599	0.4122
N = 300	a	3.5891	1.1388	3.5881	1.6243	3.6112	1.0054	3.5697	1.5720
	b	4.6108	2.3514	5.0208	3.4684	4.5656	2.2800	5.0395	3.4926
	c	2.7453	0.4589	2.7992	0.7028	2.7074	0.4483	2.8098	0.7084
	β	3.9256	0.5161	3.9166	0.7280	3.9496	0.5470	3.9176	0.7181
	σ	3.0916	0.2252	3.0869	0.3065	3.0945	0.2212	3.0801	0.3017
N = 500	a	3.5909	0.7457	3.5286	1.1656	3.5945	0.7019	3.5446	1.1271
	b	4.3870	1.5444	4.7266	2.6295	4.3670	1.5292	4.7443	2.6457
	c	2.7268	0.3514	2.7568	0.5449	2.7008	0.3456	2.7619	0.5467
	β	3.9320	0.3508	3.8970	0.5732	3.9356	0.3645	3.9042	0.5650
	σ	3.0964	0.1568	3.0938	0.2427	3.1001	0.1417	3.0916	0.2408

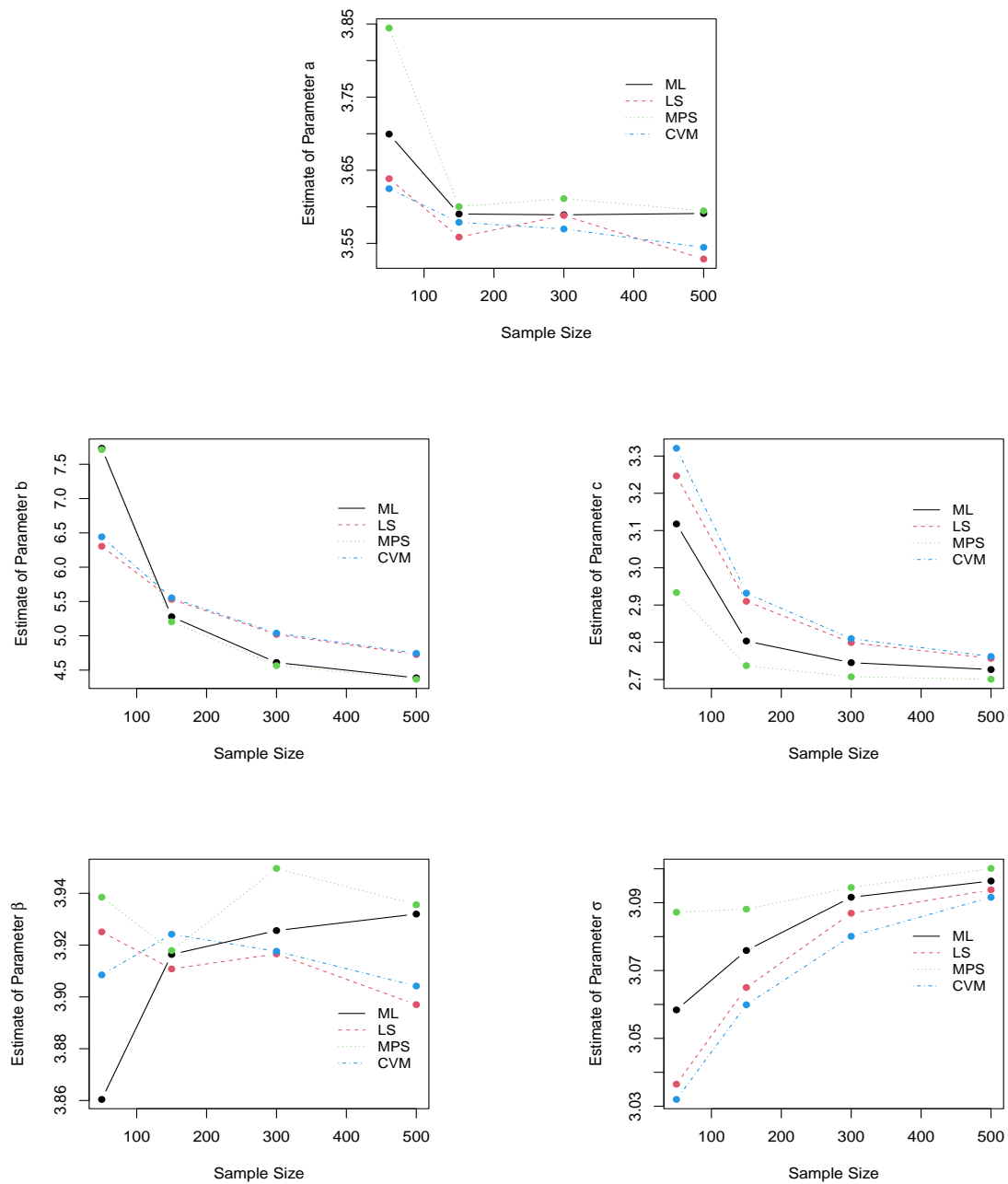


Figure 3. The values of estimates using the ML, LS, MPS, and CVM methods for parameters: a , b , c , β , and σ , respectively (case I).

Table 3. The estimates and RMSE across sample sizes ($n = 50, 150, 300, 500$) for all methods of estimation (case II).

Sample size	Parameter	ML		LS		MPS		CVM	
		Estimate	RMSE	Estimate	RMSE	Estimate	RMSE	Estimate	RMSE
n = 50	a	2.3007	3.6239	2.4348	2.8435	2.6361	5.9388	2.3934	2.8937
	b	6.7193	11.3238	5.8464	7.0627	7.1463	16.7620	5.9316	7.3497
	c	2.8859	1.4829	3.0265	1.7657	2.7074	1.2645	3.0893	1.7993
	β	1.1382	0.6712	1.1378	0.4795	1.1467	0.7854	1.1447	0.4877
	σ	1.9454	0.5375	1.9821	0.4565	1.9902	0.5900	1.9666	0.4616
n = 150	a	2.1739	1.4165	2.2269	1.6606	2.2102	1.6025	2.2370	1.6667
	b	4.6880	3.9955	4.9683	4.5144	4.6579	5.4461	4.9871	4.4469
	c	2.5963	0.6389	2.6916	0.9416	2.5340	0.6131	2.7117	0.9522
	β	1.1040	0.2837	1.1097	0.3115	1.1025	0.2612	1.1089	0.3091
	σ	1.9890	0.2540	1.9905	0.3235	1.9955	0.2541	1.9909	0.3175
n = 300	a	2.1724	0.6388	2.1979	1.1787	2.1682	0.6165	2.1942	1.1846
	b	4.1214	1.9782	4.4892	3.0452	4.0889	1.9464	4.5108	3.0684
	c	2.5425	0.4221	2.5894	0.6474	2.5064	0.4120	2.5975	0.6502
	β	1.1058	0.1646	1.1026	0.2326	1.1048	0.1545	1.1045	0.2388
	σ	1.9995	0.1377	2.0031	0.2450	2.0016	0.1321	1.9992	0.2511
n = 500	a	2.1665	0.4570	2.1914	0.9402	2.1677	0.4323	2.1809	0.9069
	b	3.9390	1.3285	4.2102	2.2266	3.9223	1.3106	4.2180	2.2194
	c	2.5252	0.3230	2.5527	0.4977	2.5004	0.3169	2.5584	0.5009
	β	1.1053	0.1081	1.1013	0.1951	1.1055	0.1051	1.1033	0.1920
	σ	2.0032	0.0934	2.0093	0.1570	2.0048	0.0970	2.0055	0.1958

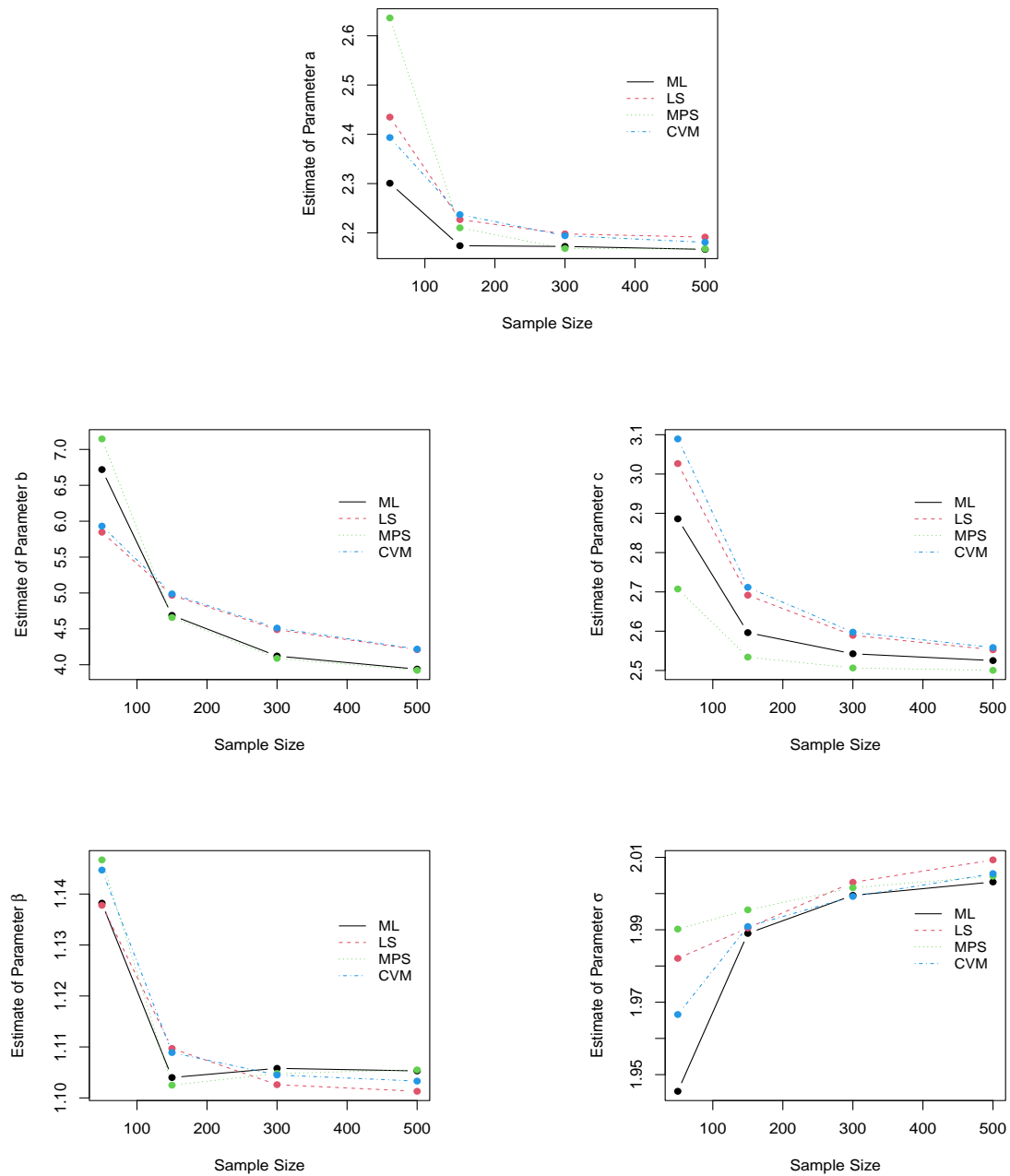


Figure 4. The values of estimates using the ML, LS, MPS, and CVM methods for parameters: a , b , c , β , and σ , respectively (case II).

Table 4. The estimates and RMSE across sample sizes ($n = 50, 150, 300, 500$) for all methods of estimation (case III).

Sample size	Parameter	ML		LS		MPS		CVM	
		Estimate	RMSE	Estimate	RMSE	Estimate	RMSE	Estimate	RMSE
n = 50	a	1.7729	2.9025	1.9031	2.3260	1.9478	4.2473	1.8835	2.3541
	b	5.0394	11.6821	4.4460	5.3990	5.1186	12.6435	4.4981	5.5023
	c	3.5806	1.8801	3.7252	2.1969	3.3497	1.5768	3.8168	2.3042
	β	0.6295	0.3288	0.6263	0.2435	0.6244	0.3501	0.6306	0.2435
	σ	2.1310	0.5073	2.1771	0.4933	2.1714	0.5138	2.1602	0.4936
n = 150	a	1.6793	0.9946	1.7432	1.2346	1.7066	0.8836	1.7749	1.4033
	b	3.3994	2.5926	3.6631	3.0490	3.3995	5.2307	3.7503	3.3148
	c	3.2226	0.7839	3.3299	1.1712	3.1400	0.7469	3.3428	1.1677
	β	0.6093	0.1209	0.6112	0.1555	0.6067	0.1216	0.6146	0.1576
	σ	2.1819	0.2386	2.1870	0.3260	2.1930	0.2262	2.1841	0.3395
n = 300	a	1.6806	0.4577	1.7269	0.9422	1.7020	0.4527	1.7252	0.9100
	b	3.0529	1.2873	3.3295	2.1153	3.0367	1.2752	3.3560	2.1439
	c	3.1548	0.5147	3.2056	0.7994	3.1066	0.5002	3.2138	0.8043
	β	0.6082	0.0653	0.6091	0.1183	0.6096	0.0687	0.6092	0.1147
	σ	2.1946	0.1273	2.1980	0.2464	2.1972	0.1259	2.1959	0.2446
n = 500	a	1.6776	0.3247	1.6874	0.6620	1.6876	0.3380	1.6951	0.6742
	b	2.9437	0.8820	3.1310	1.5109	2.9391	0.8765	3.1472	1.5431
	c	3.1326	0.3930	3.1627	0.6128	3.0988	0.3845	3.1665	0.6141
	β	0.6080	0.0487	0.6068	0.0907	0.6080	0.0492	0.6069	0.0921
	σ	2.1981	0.0896	2.1998	0.1917	2.1999	0.0870	2.2007	0.1892

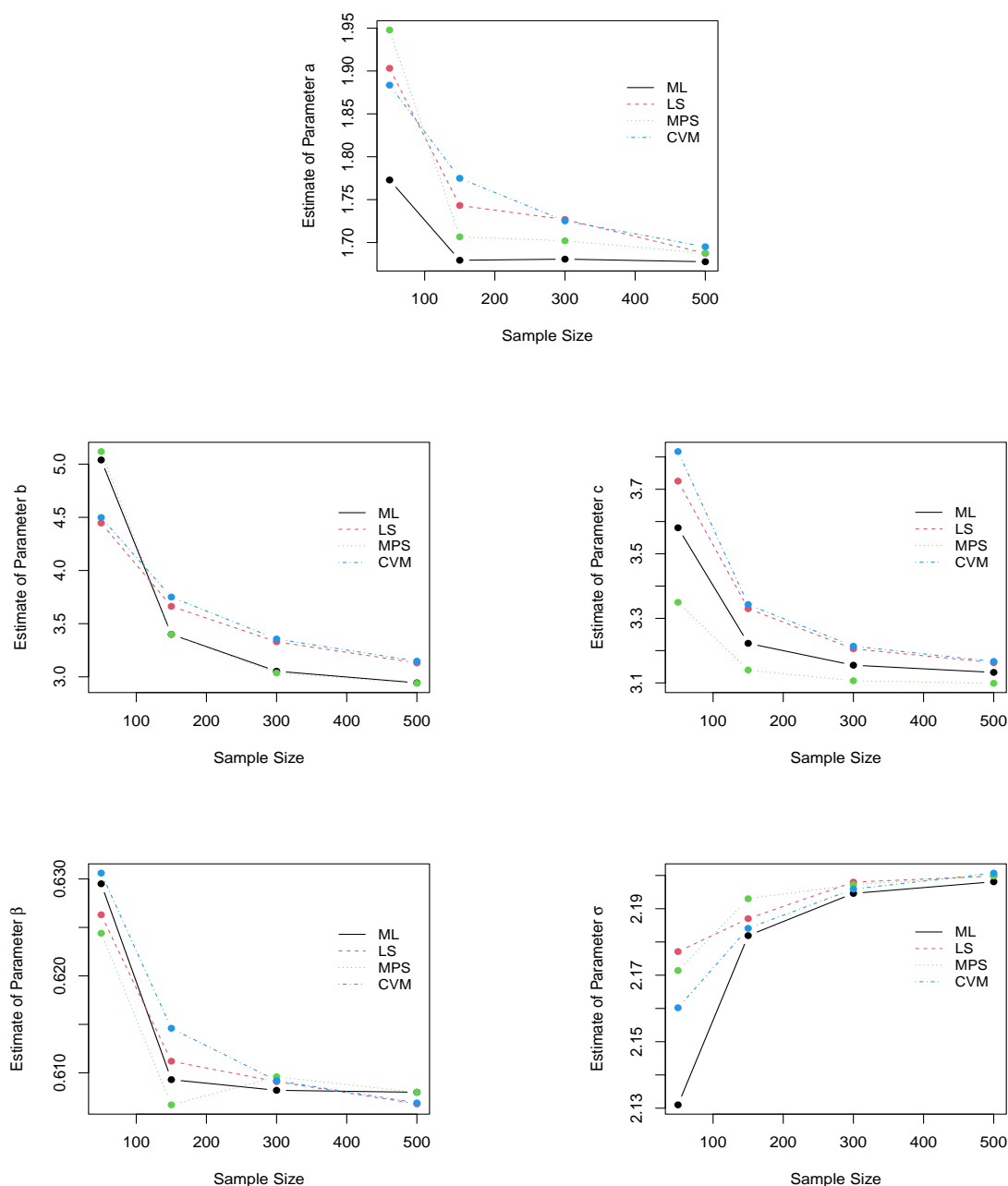


Figure 5. The values of estimates using the ML, LS, MPS, and CVM methods for parameters: a , b , c , β , and σ , respectively (case III).

As can be seen from the above tables and figures, all of the examined estimation methods achieved consistency. When the sample size n increased, the RMSE decreased, and the estimate approached its true value. Moreover, For sample size $n = 50$, except for parameter c , the LS method provided the closest estimate to the true value for all parameters. Thus, it was considered the most precise method of estimation, followed by the CVM method, while MPS was considered the worst out of all examined methods. Furthermore, when the sample size increased and reached $n = 500$, the MPS and

ML methods outperformed the others. These methods produced the lowest RMSE values, with the MPS method having a slight advantage in most cases.

6. Real application

In this section, three medical datasets were used to provide insights into which method had the best performance in estimating the parameters of the EGWR distribution.

Dataset 1

Provided by a Saudi Ministry of Health hospital, this dataset was reported in [31] and represents the lifetimes (in years) of 40 individuals who were diagnosed with leukemia. The data are listed below: 0.315, 0.496, 0.616, 1.145, 1.208, 1.263, 1.414, 2.025, 2.036, 2.162, 2.211, 2.370, 2.532, 2.693, 2.805, 2.910, 2.912, 3.192, 3.263, 3.348, 3.348, 3.427, 3.499, 3.534, 3.767, 3.751, 3.858, 3.986, 4.049, 4.244, 4.323, 4.381, 4.392, 4.397, 4.647, 4.753, 4.929, 4.973, 5.074, 5.381.

Dataset 2

The second dataset was reported by [32] and lists the periods of symptom relief for 36 bladder cancer patients. The durations (in days) were as follows: 0.08, 0.2, 0.4, 0.5, 0.51, 0.81, 0.87, 0.9, 1.05, 1.19, 1.26, 1.35, 1.4, 1.46, 1.76, 2.02, 2.02, 2.07, 2.09, 2.23, 2.26, 2.46, 2.54, 2.62, 2.64, 2.69, 2.69, 2.75, 2.83, 2.87, 3.02, 3.02, 3.25, 3.31, 3.36, 3.36.

Dataset 3

The last dataset was reported in [33] and consists of the relief times (in minutes) of 20 patients who received a pain reliever as follows: 1.1, 1.4, 1.3, 1.7, 1.9, 1.8, 1.6, 2.2, 1.7, 2.7, 4.1, 1.8, 1.5, 1.2, 1.4, 3, 1.7, 2.3, 1.6, 2.

6.1. Comparison between methods of estimation

The accuracy of the ML, LS, MPS, and CVM methods in estimating the distribution parameters and fitting the data were compared. Tables 5–7 offer the estimate and standard error for each parameter along with the Kolmogorov–Smirnov (KS) test statistic and P-value for all the methods of estimation.

Table 5. Parameter estimates, standard errors (in parentheses), and KS results for Dataset 1.

Parameter	ML	LS	MPS	CVM
a	2.5242 (0.0054)	3.4949 (1.3842)	2.7979 (0.0789)	2.9447 (2.3129)
b	0.1273 (0.0201)	0.1335 (0.0988)	0.1188 (0.0199)	0.1355 (0.1125)
c	6.3220 (0.0054)	6.0953 (1.3793)	6.1141 (0.0827)	6.2266 (2.3532)
β	1.6975 (0.0054)	3.2830 (1.3237)	1.6127 (0.3256)	4.0251 (2.2503)
σ	2.8987 (0.0054)	2.1829 (0.6336)	3.1104 (0.3215)	1.9152 (0.6911)
KS	0.0659	0.0610	0.0835	0.0557
P-value	0.9950	0.9983	0.9428	0.9996

Table 6. Parameter estimates, standard errors (in parentheses), and KS results for Dataset 2.

Parameter	ML	LS	MPS	CVM
a	4.0014 (0.0033)	4.6811 (0.4829)	3.2677 (0.0004)	5.1672 (0.1580)
b	0.1372 (0.0229)	0.1322 (0.0902)	0.0940 (0.0173)	0.1308 (0.0655)
c	4.8434 (0.0033)	4.7446 (0.4824)	6.6395 (0.0004)	4.9779 (0.1578)
β	0.4352 (0.0033)	0.5665 (0.2969)	0.9050 (0.0004)	0.7319 (0.1475)
σ	3.9916 (0.0033)	3.8530 (0.4696)	2.7204 (0.0004)	3.3434 (0.1558)
KS	0.1193	0.0874	0.1152	0.0848
P-value	0.6847	0.9460	0.7253	0.9580

Table 7. Parameter estimates, standard errors (in parentheses), and KS results for Dataset 3.

Parameter	ML	LS	MPS	CVM
a	0.0051 (0.0017)	0.0902 (1.2776)	0.0635 (1.2752)	0.0611 (0.6350)
b	6.4327 (4.6024)	7.3693 (55.6305)	6.4930 (18.9398)	6.5356 (57.7738)
c	0.7218 (0.1798)	0.7826 (2.2306)	0.5592 (0.6037)	0.9024 (3.0946)
β	2.3471 (5.7104)	1.4198 (31.5817)	0.0169 (0.0539)	0.0422 (0.2554)
σ	0.0122 (0.0056)	0.1267 (1.0148)	0.4143 (5.1380)	0.8023 (7.9693)
KS	0.1595	0.1044	0.1932	0.1101
P-value	0.6887	0.9812	0.4437	0.9686

As displayed in Tables 5–7, from a precision perspective, the ML method provided the smallest values of standard error and would be the most suitable method of estimation for the leukemia dataset. The CVM method would be considered the worst of those examined as it had the highest standard error values. A similar pattern can be seen with the pain relief dataset, except for parameter β . However, for the bladder cancer dataset, MPS slightly outperformed ML although both methods had the lowest standard error, while the worst method for estimation of those examined was the LS method. From a goodness-of-fit perspective, the CVM method produced good fit for the leukemia and the bladder cancer datasets as it provided higher P-values and lower KS statistics among all the examined methods, while LS offered a more suitable fit for the pain relief dataset.

6.2. Comparison between the EGWR and other distributions

In this section, the ML method was used to assess the performance of the EGWR distribution in fitting the aforementioned datasets. A comparison was made between the EGWR distribution and some competitors: the exponentiated Weibull Rayleigh (EWR), based on the idea of the researchers in [34]; the Marshall–Olkin Weibull Rayleigh (MOWR), a sub-model from the general class proposed by the researchers in [35]; the Weibull Rayleigh (WR) by [34]; the Kumaraswamy Rayleigh (KumR) presented by the researchers in [36]; the Marshall–Olkin alpha power Rayleigh (MOAPR) by the researchers in [37]; and the generalized Rayleigh (GR) explored by the researchers in [38]. To pick the most suitable of these models, the Akaike information criterion (AIC), Bayesian information criterion (BIC), Hannan–Quinn information criterion (HQIC), and KS statistic with P-value were

computed. Tables 8–10 display the values of the MLEs and goodness-of-fit measures for the EGWR and competing distributions, while the theoretical and empirical cdf and pdf are displayed in Figures 6–8.

Table 8. The ML estimates and goodness-of-fit measures for all distributions for Dataset 1.

Model	EGWR	EWR	MOWR	WR	KumR	MOAPR	GR
MLEs	$\hat{a} = 2.5242$	$\hat{a} = 2.2206$	$\hat{a} = 5.8178$	$\hat{c} = 1.2499$	$\hat{a} = 0.8702$	$\hat{\theta} = 0.9681$	$\hat{a} = 1.1726$
	$\hat{b} = 0.1273$	$\hat{c} = 0.8071$	$\hat{b} = 1.0348$	$\hat{\beta} = 0.0024$	$\hat{b} = 0.0714$	$\hat{\alpha} = 5.8756$	$\hat{\sigma} = 2.3002$
	$\hat{c} = 6.3220$	$\hat{\beta} = 3.4238$	$\hat{c} = 0.9117$	$\hat{\sigma} = 50.2894$	$\hat{\sigma} = 0.6457$	$\hat{\sigma} = 2.0199$	--
	$\hat{\beta} = 1.6975$	$\hat{\sigma} = 0.4262$	$\hat{\sigma} = 1.5696$	--	--	--	--
	$\hat{\sigma} = 2.8987$	--	--	--	--	--	--
Goodness-of-Fit Measures							
AIC	139.8385	141.5956	143.6428	145.1159	147.3771	143.5203	145.0313
BIC	148.2829	148.3512	150.3983	150.1826	152.4437	148.5869	148.4091
HQIC	142.8917	144.0382	146.0854	146.9479	149.2090	145.3522	146.2526
KS	0.0659	0.1029	0.0913	0.1183	0.1615	0.1029	0.1439
P-value	0.9950	0.7904	0.8922	0.6301	0.2475	0.7898	0.3786

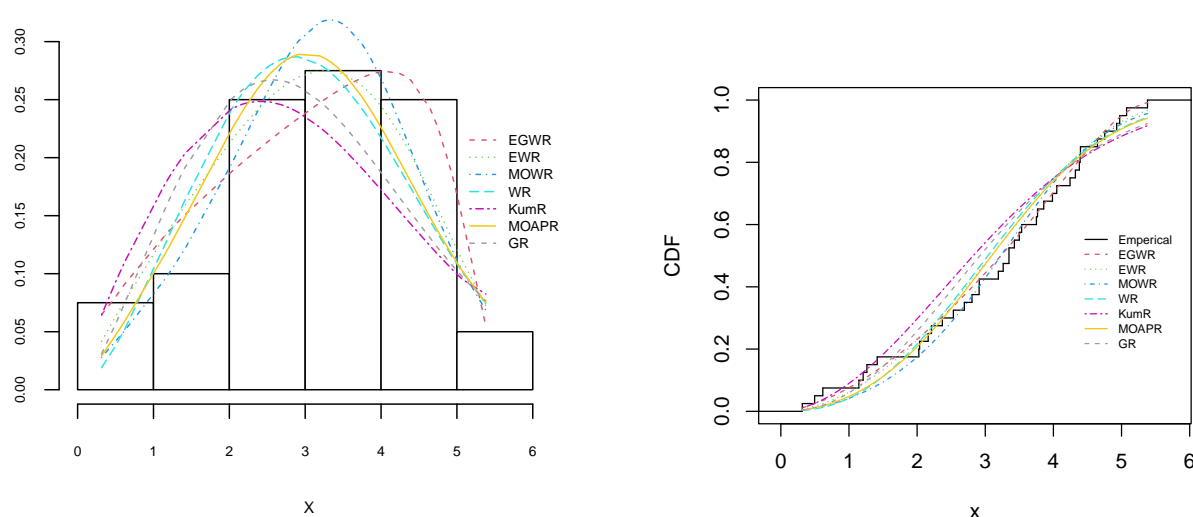


Figure 6. The theoretical and empirical pdf (right) and cdf (left) for all distributions for Dataset 1.

Table 9. The ML estimates and goodness-of-fit measures for all distributions for Dataset 2.

Model	EGWR	EWR	MOWR	WR	KumR	MOAPR	GR
MLEs	$\hat{a} = 0.4211$	$\hat{a} = 0.3133$	$\hat{a} = 0.5504$	$\hat{c} = 0.8673$	$\hat{a} = 0.4873$	$\hat{\theta} = 1.6663$	$\hat{a} = 1.6272$
	$\hat{b} = 0.1402$	$\hat{c} = 2.2488$	$\hat{b} = 0.0659$	$\hat{\beta} = 2.5817$	$\hat{b} = 0.0905$	$\hat{a} = 0.7080$	$\hat{\sigma} = 0.8335$
	$\hat{c} = 4.6449$	$\hat{\beta} = 0.4496$	$\hat{c} = 1.1090$	$\hat{\sigma} = 0.9295$	$\hat{\sigma} = 0.4770$	$\hat{\sigma} = 1.4234$	--
	$\hat{\beta} = 1.4346$	$\hat{\sigma} = 3.1404$	$\hat{\sigma} = 6.8985$	--	--	--	--
	$\hat{\sigma} = 1.7627$	--	--	--	--	--	--
Goodness-of-Fit Measures							
AIC	100.7827	103.1531	113.4518	109.4460	106.3240	107.9504	105.9897
BIC	108.7003	109.4872	119.7859	114.1966	111.0745	112.7010	109.1567
HQIC	103.5462	105.3638	115.6626	111.1041	107.9820	109.6085	107.0950
KS	0.1193	0.1466	0.1928	0.1870	0.1674	0.1369	0.1790
P-value	0.6847	0.4213	0.1375	0.1611	0.2652	0.5089	0.1986

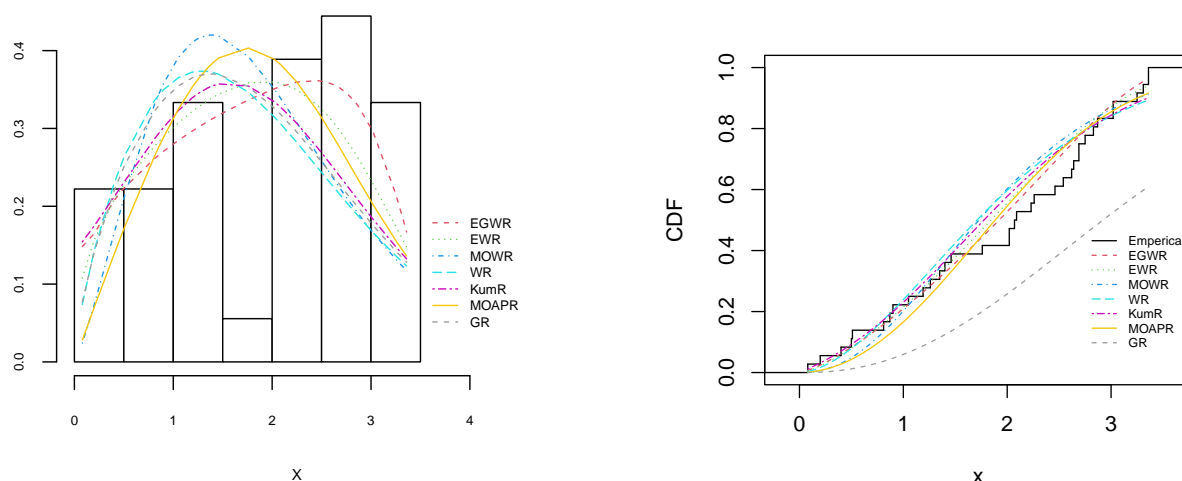
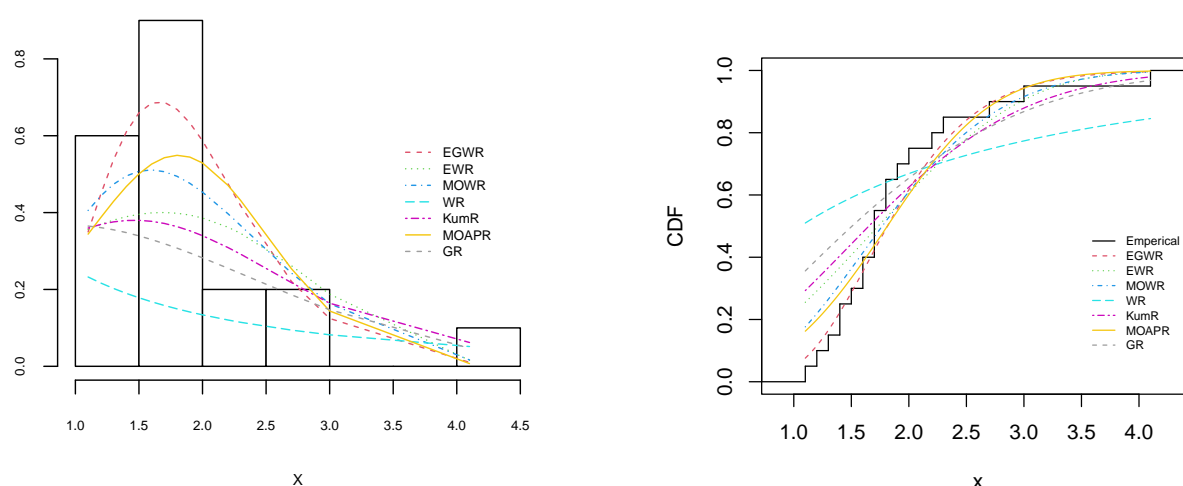
**Figure 7.** The theoretical and empirical pdf (right) and cdf (left) for all distributions for Dataset 2.

Table 10. The ML estimates and goodness-of-fit measures for all distributions for Dataset 1.

Model	EGWR	EWR	MOWR	WR	KumR	MOAPR	GR
MLEs	$\hat{a} = 0.0051$	$\hat{a} = 0.4607$	$\hat{a} = 0.4710$	$\hat{c} = 0.3657$	$\hat{a} = 0.4991$	$\hat{\theta} = 1.2453$	$\hat{a} = 1.6766$
	$\hat{b} = 6.4327$	$\hat{c} = 1.7105$	$\hat{b} = 0.8595$	$\hat{\beta} = 21.2434$	$\hat{b} = 0.1073$	$\hat{\alpha} = 7.6572$	$\hat{\sigma} = 0.6289$
	$\hat{c} = 0.7218$	$\hat{\beta} = 15.5398$	$\hat{c} = 1.4638$	$\hat{\sigma} = 0.2678$	$\hat{\sigma} = 0.4865$	$\hat{\sigma} = 1.0740$	--
	$\hat{\beta} = 2.3471$	$\hat{\sigma} = 0.4656$	$\hat{\sigma} = 1.8670$	--	--	--	--
	$\hat{\sigma} = 0.0122$	--	--	--	--	--	--
Goodness-of-Fit Measures							
AIC	44.7134	54.7442	48.2518	83.2277	52.9718	46.9752	56.3742
BIC	49.6920	58.7271	52.2347	86.2149	55.9590	49.9624	58.3657
HQIC	45.6853	55.5217	49.0293	83.8109	53.5549	47.5584	56.7629
KS	0.1595	0.2552	0.1764	0.5229	0.2929	0.1624	0.3561
P-value	0.6887	0.1477	0.5624	<0.001	0.0646	0.6668	0.0125

**Figure 8.** The theoretical and empirical pdf (right) and cdf (left) for all distributions for Dataset 3.

The above tables and figures show the potential flexibility of the EGWR distribution in modeling medical data compared to other competing distributions.

7. Conclusions

We extended the Rayleigh distribution to obtain a more flexible distribution capable of modeling complex patterns in medical data. The hazard function of the proposed distribution can take on different patterns, which shows great flexibility as required by many lifetime datasets. The statistical

characteristics of the new distribution, such as quantile, median, moment-generating function, Rényi entropy, and order statistics, were studied in this research. The distribution's parameters were estimated using ML, LS, MPS, and CVM, and the efficiency of the estimators obtained by the different methods was assessed in the reported Monte Carlo simulation studies. The results revealed that for the three simulated cases, the LS method had better estimation for smaller samples, while the ML and MPS methods performed better for larger samples. The efficiency of the estimation methods was also evaluated using three medical datasets. The results demonstrated the ability of ML and MPS in estimating the model parameters, as they provided lower standard error values compared to other examined methods in most cases. Finally, the EGWR distribution was compared with a selection of competitive distributions to investigate its ability in modeling the provided medical data, and the results clearly demonstrated that the EGWR outperformed other distributions in analyzing the data. Thus, it can provide insights for scientists to better understand and address complex patterns in healthcare data. As future work, more advanced methods of estimation, such as the Bayesian method, could be employed to enhance the estimation of the parameters. Moreover, the development of a discretised version of the EGWR distribution would offer crucial tools for describing, evaluating, and modeling different phenomena with countable and distinct outcomes, which would lead to precise decision-making and offer better understanding of data across a variety of fields.

Use of Generative-AI tools declaration

The author declares she has not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

For the editor-in-chief and anonymous reviewers, the author extends great gratitude for their insightful comments, which refined the quality of this work.

Conflict of interest

The author declares that she has no conflicts of interest.

References

1. R. Gupta, P. Gupta, R. Gupta, Modeling failure time data by Lehman alternatives, *Commun. Stat. Theory Methods*, **27** (1998), 887–904. <https://doi.org/10.1080/03610929808832134>
2. G. Cordeiro, E. Ortega, D. da Cunha, The exponentiated generalized class of distributions, *J. Data Sci.*, **11** (2013), 1–27. [https://doi.org/10.6339/JDS.2013.11\(1\).1086](https://doi.org/10.6339/JDS.2013.11(1).1086)
3. M. El-Morshedy, A. El-Faheem, A. Al-Bossly, M. El-Dawoody, Exponentiated generalized inverted Gompertz distribution: Properties and estimation methods with applications to symmetric and asymmetric data, *Symmetry*, **13** (2021), 1868. <https://doi.org/10.3390/sym13101868>
4. L. Telee, M. Karki, V. Kumar, Exponentiated generalized exponential geometric distribution: Model, properties and applications, *Interdisc. J. Manag. Soc. Sci.*, **3** (2022), 37–60. <https://doi.org/10.3126/ijmss.v3i2.50261>

5. M. El-Morshedy, M. Eliwa, A. El-Gohary, E. Almetwally, R. El-Desokey, Exponentiated generalized inverse flexible Weibull distribution: Bayesian and non-Bayesian estimation under complete and type II censored samples with applications, *Commun. Math. Stat.*, **10** (2002), 413–434. <https://doi.org/10.1007/s40304-020-00225-4>
6. M. Kilai, G. Waititu, W. Kibira, M. Abd El-Raouf, T. Abushal, A new versatile modification of the Rayleigh distribution for modeling COVID-19 mortality rates, *Results Phys.*, **35** (2002), 105260. <https://doi.org/10.1016/j.rinp.2022.105260>
7. H. Klakattawi, A. Khormi, L. Baharith, The new generalized exponentiated Fréchet–Weibull distribution: Properties, applications, and regression model, *Complexity*, **2003** (2003), 2196572. <https://doi.org/10.1155/2023/2196572>
8. S. Hussain, M. Ul Hassan, M. Rashid, R. Ahmed, Families of extended exponentiated generalized distributions and applications of medical data using Burr III extended exponentiated Weibull distribution, *Mathematics*, **11** (2003), 3090. <https://doi.org/10.3390/math11143090>
9. N. Temraz, Inference on the stress strength reliability with exponentiated generalized Marshall Olkin-G distribution, *PLOS One*, **18** (2023), e0280183. <https://doi.org/10.1371/journal.pone.0280183>
10. T. Sindhu, A. Shafiq, Z. Huassian, Generalized exponentiated unit Gompertz distribution for modeling arthritic pain relief times data: Classical approach to statistical inference, *J. Biopharm. Stat.*, **34** (2024), 323–348. <https://doi.org/10.1080/10543406.2023.2210681>
11. G. Lekono, B. Oluyede, L. Gabaitiri, A new heavy-tailed exponentiated generalised-G family of distributions: Properties and applications, *Int. J. Math. Oper. Res.*, **27** (2024), 1–34. <https://doi.org/10.1504/IJMOR.2024.136860>
12. 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>
13. P. Koleoso, A. Chukwu, A Weibull-Gompertz Makeham distribution with properties and application to cancer data, *J. Adv. Math. Comput. Sci.*, **34** (2019), 1–17. <https://doi.org/10.9734/jamcs/2019/v34i530223>
14. P. Marthin, G. Rao, Generalized Weibull–Lindley (GWL) distribution in modeling lifetime data, *J. Math.*, **2020** (2020), 2049501. <https://doi.org/10.1155/2020/2049501>
15. E. Hussein, H. Aljohani, A. Afify, The extended Weibull–Fréchet distribution: Properties, inference, and applications in medicine and engineering, *AIMS Mathematics*, **7** (2022), 225–246. <https://doi.org/10.3934/math.2022014>
16. E. Moradi, Z. S. Ghazani, A novel Weibull Marshall–Olkin power Lomax distribution: Properties and applications to medicine and engineering, *J. Appl. Math. Inform.*, **41** (2023), 1275–1301. <https://doi.org/10.14317/jami.2023.1275>
17. G. Cordeiro, E. Biazatti, L. de Santana, A new extended Weibull distribution with application to influenza and hepatitis data, *Stats*, **6** (2023), 657–673. <https://doi.org/10.3390/stats6020042>
18. L. Baharith, New generalized Weibull inverse Gompertz distribution: Properties and applications, *Symmetry*, **16** (2024), 197. <https://doi.org/10.3390/sym16020197>

19. A. Ahmad, N. Alsadat, A. Rather, M. Meraou, M. Mohie El-Din, A novel statistical approach to COVID-19 variability using the Weibull-inverse Nadarajah Haghighi distribution, *Alex. Eng. J.*, **107** (2024), 950–962. <https://doi.org/10.1016/j.aej.2024.08.008>
20. X. Chen, Z. Shi, Y. Xie, Z. Zhang, A. Cohen, S. Pu, Advancing continuous distribution generation: An exponentiated odds ratio generator approach, *Entropy*, **26** (2024), 1006. <https://doi.org/10.3390/e26121006>
21. G. Kass, A. Mood, F. Graybill, D. Boes, Introduction to the theory of statistics, 3 Eds., *J. Am. Stat. Assoc.*, **69** (1974), 1050–1051. <https://doi.org/10.2307/2286195>
22. J. Patel, C. Kapadia, D. Owen, *Handbook of statistical distributions*, New York: Marcel Dekker, 1976.
23. W. Feller, *An introduction to probability theory and its applications*, Wiley, 1991.
24. A. Rényi, On measures of entropy and information, In: *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*, 1961, 547–561.
25. H. David, H. Nagaraja, *Order statistics*, Wiley, 2004.
26. R. Fisher, On the mathematical foundations of theoretical statistics, In: *Philosophical transactions of the Royal Society of London: Series A, containing papers of a mathematical or physical character*, **222** (1922), 309–368. <https://doi.org/10.1098/rsta.1922.0009>
27. H. Wang, W. Chen, B. Li, Large sample properties of maximum likelihood estimator using moving extremes ranked set sampling, *J. Korean Stat. Soc.*, **53** (2024), 398–415. <https://doi.org/10.1007/s42952-023-00251-2>
28. J. Swain, S. Venkatraman, J. Wilson, Least-squares estimation of distribution functions in Johnson's translation system, *J. Stat. Comput. Simul.*, **29** (1988), 271–297. <https://doi.org/10.1080/00949658808811068>
29. R. Cheng, N. Amin, *Maximum product of spacings estimation with application to the lognormal distribution*, Cardiff: University of Wales Institute of Science and Technology, 1979.
30. K. Choi, W. Bulgren, An estimation procedure for mixtures of distributions, *J. R. Stat. Soc. B*, **30** (1968), 444–460. <https://doi.org/10.1111/j.2517-6161.1968.tb00743.x>
31. Z. Al-Saiary, R. Bakoban, The Topp-Leone generalized inverted exponential distribution with real data applications, *Entropy*, **22** (2020), 1144. <https://doi.org/10.3390/e22101144>
32. R. Hibatullah, Y. Widyaningsih, S. Abdullah, Marshall–Olkin extended power Lindley distribution with application, *JRAM*, **2** (2018), 84–92. <https://doi.org/10.26740/jram.v2n2.p84-92>
33. A. Gross, V. Clark, *Survival distributions: Reliability applications in the biomedical sciences*, Wiley, 1975.
34. A. Ahmad, S. Ahmad, A. Ahmed, Characterization and estimation of Weibull–Rayleigh distribution with applications to life time data, *Appl. Math. Inf. Sci. Lett.*, **5** (2017), 71–79. <https://doi.org/10.18576/amisl/050206>
35. H. Klakattawi, D. Alsulami, M. Elaal, S. Dey, L. Baharith, A new generalized family of distributions based on combining Marshal–Olkin transformation with T-X family, *PLOS One*, **17** (2022), e0263673. <https://doi.org/10.1371/journal.pone.0263673>

-
36. N. Rashwan, A note on Kumaraswamy exponentiated Rayleigh distribution, *J. Stat. Theory Appl.*, **15** (2016), 286–295. <http://doi.org/10.2991/jsta.2016.15.3.8>
37. E. Almetwally, A. Afify, G. Hamedani, Marshall–Olkin alpha power Rayleigh distribution: Properties, characterizations, estimation and engineering applications, *Pak. J. Stat. Oper. Res.*, **17** (2021), 745–760. <http://dx.doi.org/10.18187/pjsor.v17i3.3473>
38. J. Surles, W. Padgett, Inference for reliability and stress-strength for a scaled Burr type X distribution, *Lifetime Data Anal.*, **7** (2001), 187–200. <http://doi.org/10.1023/a:1011352923990>



AIMS Press

© 2025 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>)