Statistics
Innovation

ARTICLE

Open Access

https://doi.org/10.48130/stati-0026-0010
Statistics Innovation **2026**, 3: e012

Rayleigh Generated Weibull distribution: properties and performance analysis

Arjun Kumar Gaire*

Department of Science and Humanities, Khwopa Engineering College, Purbanchal University, Bhaktapur 44800, Nepal
* Correspondence: gaire.arjun@khec.edu.np or arjun.gaire@gmail.com (Gaire AK)

Abstract

This study presents a new three-parameter continuous distribution, termed the Rayleigh Generated Weibull distribution, derived within the framework of the Rayleigh Generated family of distributions. Some important statistical properties of the proposed distribution are derived, including its probability density function, cumulative distribution function, hazard rate function, and a random number generation procedure. Parameter estimation is performed using maximum likelihood. To demonstrate the model's flexibility and practical applicability, two real datasets are analyzed: customer waiting times at a bank before receiving service, and remission times for patients with bladder cancer. In addition, a simulation study is conducted to evaluate the performance of the maximum likelihood estimators under different parameter settings. The empirical analyses and simulation results indicate that the proposed RGW distribution provides greater modeling flexibility and improved goodness-of-fit compared with several tractional benchmark models for both data sets.

Keywords: Rayleigh Generated, Waiting time, Remission times, Hazard rate, Simulation

Citation: Gaire AK. 2026. Rayleigh Generated Weibull distribution: properties and performance analysis. *Statistics Innovation* 3: e012 https://doi.org/10.48130/stati-0026-0010

Introduction

The development of modern statistical distributional models often involves integrating additional structural parameters into baseline distributions to enhance their empirical robustness. The Weibull distribution is one of the most widely used models in reliability analysis, survival studies, and engineering applications due to its mathematical tractability and ability to represent increasing or decreasing failure rates. However, real-world problems often exhibit more complex patterns that the classical Weibull model cannot adequately capture. In particular, empirical data sets often exhibit non-monotonic hazard behavior across varying tail characteristics, which motivates the development of more flexible generalizations capable of accommodating diverse distributional shapes. To address these limitations, several families of generalized distributions have been proposed by introducing additional parameters through generator mechanisms. One such approach is the Rayleigh Generated (RG) family of distributions, which has been shown to enhance modeling flexibility while preserving analytical tractability. Motivated by these advantages, this study introduces a methodology for formulating a three-parameter Rayleigh Generated Weibull (RGW) distribution by applying the RG mechanism introduced by Gaire & Gurung[1] to the Weibull baseline model. The proposed RGW distribution possesses unique statistical properties and a more flexible hazard rate structure than the classical Weibull model and several existing extensions. In particular, the additional shape parameter introduced by the RG-family allows the model to capture complex failure-rate patterns and provides improved adaptability for modeling lifetimes and reliability data. Furthermore, the RGW distribution is investigated by deriving its statistical properties and estimating its parameters using the maximum likelihood method: simulation analysis and application to real datasets demonstrating its practical usefulness and modeling capability.

The remaining research work is organized as follows. In section *'Rayleigh Generated Weibull Distribution'*, a three-parameter RGW probability distribution is formulated, and some statistical properties and the rule for random number generation are derived. Section '*Method of parameter estimation using the method of maximum likelihood*' presents the methods of parameter estimation, while section '*Numerical application*' illustrates the numerical application and validation of the model using two real datasets. Section '*Simulation study*' presents the simulation study of the model, and the last section concludes the paper.

Rayleigh Generated Weibull distribution

Researchers have proposed numerous new families of probability distributions using transformation and generator techniques to enhance modeling flexibility and better capture complex data behaviors. Several generator-based distribution families have been introduced in the literature, including the Transmutation-G family[2], Exponentiated-G family[3], Marshall-Olkin-G family[4], Zografos-Balakrishnan-G family[5], Beta-G family[6], T-X-G family[7], and Exponential Arctan-G[8], among others. These generator mechanisms extend baseline distributions by introducing additional shape parameters, thereby providing greater flexibility in modeling real-world phenomena such as actuarial data[9], age-specific fertility rates[10], age at first marriage[11], reliability engineering[12], and other demographic datasets. These applications demonstrate the flexibility and usefulness of generator-based distributions in capturing diverse patterns in empirical data. Motivated by these developments, the RG-family of distributions is adopted in this study to construct a new sub-model, referred to as the RGW distributions. The cumulative distribution function (CDF) and corresponding probability density function (PDF) of the RG-family of distribution are given as follows:

$$F(x,\theta) = \frac{1-\exp\left\{-\theta(\mathrm{G}(\mathrm{x})^2\right\}}{(1-\exp(-\theta))}, \text{ for } x>0,\ \theta>0 \tag{1}$$

$$f(x,\theta) = \frac{2\theta g(x)\,\mathrm{G}(\mathrm{x})\exp\left\{-\theta(\mathrm{G}(\mathrm{x})^2\right\}}{(1-\exp(-\theta))}, \text{ for } x>0,\ \theta>0 \tag{2}$$

where, the $g(x)$ and the $G(x)$ are the PDF and the CDF of the base distribution to be chosen. Here $F(x,\theta)$ is a valid CDF since it satisfies the boundary conditions; the baseline CDF $G(x)$ is defined in [0,1], as $G(x)\to 0, F(x)\to 0$ and $G(x)\to 1, F(x)\to 1$.

Here, a flexible and widely used two-parameter Weibull distribution is chosen as the base; detailed applications of this distribution in the real world can be found in[13,14]. The two-parameter Weibull distribution has a PDF and a CDF given as:

$$g(x) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1}\exp\left\{-\left(\frac{x}{\beta}\right)^{\alpha}\right\}, \text{ for } x>0 \tag{3}$$

$$G(x) = 1-\exp\left\{-\left(\frac{x}{\beta}\right)^{\alpha}\right\} \tag{4}$$

where, $\alpha > 0$ is the shape parameter, and $\beta > 0$ is the scale parameter.

Due to the simplicity, flexibility, and practical usefulness in modeling lifetime data, the Weibull distribution has been widely adopted as a baseline distribution for developing various generalized and modified models. Consequently, numerous extensions of the Weibull distribution have been proposed in the literature[15–28]. Among these, some of the models have gained prominence by introducing two or three additional shape parameters. While these generalizations offer remarkable flexibility; they frequently encounter the problem of over-parameterization, where the increased parameter space leads to computational instability, over-parameterization, and significant challenges in maximum likelihood estimation for small to moderate sample sizes. On the other end of the spectrum, three-parameter models provide better parsimony but may lack the structural depth required to capture the complex tail behavior and specific skewness found in real-world data. The RGW distribution is strategically positioned to bridge this gap in the existing literature. The RGW maintains a parsimonious three-parameter structure, introducing only a single additional shape parameter θ through a Rayleigh Generated transformation. This configuration allows the model to emulate the high-level flexibility of four and five-parameter distributions—specifically in capturing unimodal and heavy-tailed characteristics—while retaining the analytical tractability of simpler models. By providing a closed-form cumulative distribution function and a stable numerical framework, the RGW offers a computationally efficient alternative for researchers seeking a balance between mathematical precision and model parsimony, where data may be limited. Further discussions on extensions of the Weibull distribution can be found in Mohammad & Gaire[8]. Motivated by these developments, this study introduces the RGW distribution by applying the RG-family of distributions to the Weibull distribution.

Probability density function of RGW distribution

The RGW distribution is obtained by applying the generator mechanism of the RG-family of distribution proposed in[1] to a baseline Weibull distribution. According to the RG generator defined in Eqs (1) and (2), a new class of distribution can be constructed by transforming the baseline distribution through the Rayleigh Generated mechanism. By substituting the CDF and PDF of the Weibull distribution into Eqs (1) and (2), the CDF and PDF of the RGW distributions are derived and expressed in Eqs (5) and (6), respectively.

$$F(x) = \frac{1-\exp\left[-\theta\left(1-\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)\right)^2\right]}{1-\exp(-\theta)}, \text{ for } x>0 \tag{5}$$

$$f(x) = \frac{2\alpha\theta\left(\frac{x}{\beta}\right)^{\alpha-1}\left[\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)-\exp\left(-2\left(\frac{x}{\beta}\right)^{\alpha}\right)\right]\exp\left[-\theta\left(1-\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)\right)^2\right]}{\beta(1-\exp(-\theta))}, \quad for\ \ x>0 \tag{6}$$

Although the generator structure follows[1], the resulting distribution exhibits distinct statistical properties and hazard-rate behavior when the Weibull distribution is used as the baseline. The CDF and PDF of the RGW distribution are legitimate probability distributions since they satisfy the following defining properties of a probability distribution.

i. **The boundary conditions:** As $x\to 0$, $F(x)\to 0$, and as $x\to\infty$, $F(x)\to 1$. Therefore, the CDF of the RGW distribution satisfies the boundary conditions.

ii. **The non-negativity of the PDF:** The PDF of the RGW distribution defined in Eq. (6) has three non-negative parameters. Here, $\left(\frac{x}{\beta}\right)^{\alpha-1}\ge 0$, for $x>0$; $\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)$ and $\exp\left(-2\left(\frac{x}{\beta}\right)^{\alpha}\right)$ are positive and $\left[\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)-\exp\left(-2\left(\frac{x}{\beta}\right)^{\alpha}\right)\right]\ge 0$. Since every component of $f(x)$ is non-negative, it follows that $f(x)\ge 0$.

iii. **Normalization condition**:

Since $f(x)=\frac{d}{dx}F(x)$, which implies $\int_0^{\infty} f(x)\,dx = F(\infty)-F(0)=1$.

iv. **Monotonicity:**

As $f(x)=\frac{d}{dx}F(x)\ge 0$, which proves that the CDF is a non-decreasing function.

The graphical illustrations of the PDF and CDF of the RGW distribution for different parameter values are presented in Figs 1 and 2, respectively.

Reliability analysis of RGW distribution

The probability that an event does not fail before time x is a reliability function. It is simply the complement of CDF (Rodrıguez[29]), defined as $R(x)=P(X>x)=\ 1-P(X\le x)=1-F(x)$, where $F(x)$ is the CDF of x.

Thus, the reliability function of the three-parameter RGW distribution is given by Eq. (7), and its graphical representation is shown in Fig. 3 for selected parameter values.

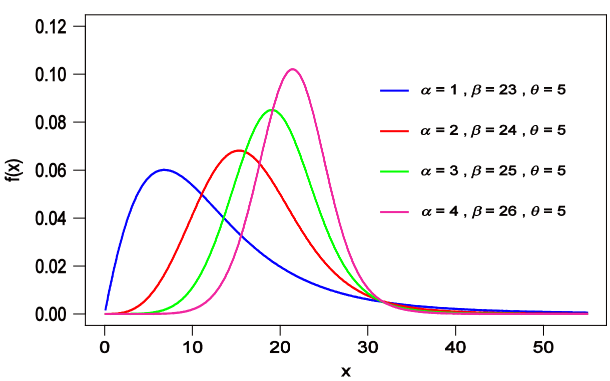

Fig. 1 Graph of PDF of RGW distribution for different values of parameters.

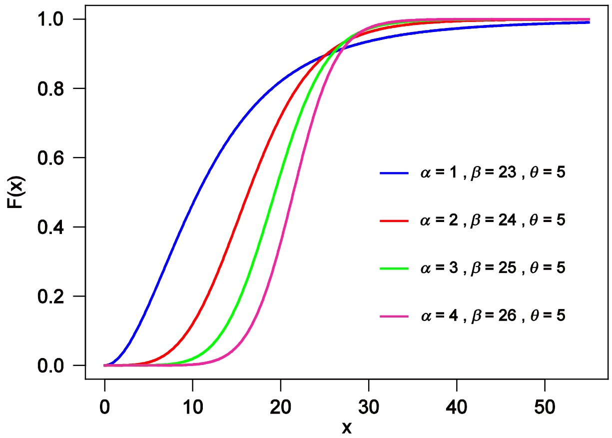

Fig. 2 Graphs CDF of RGW distribution for different values of parameters.

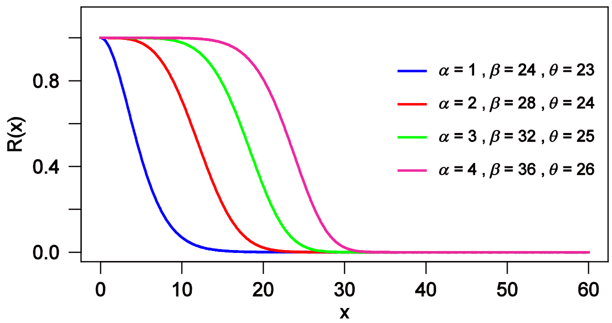

Fig. 3 The reliability function of the RGW distribution.

$$R(x)=\frac{\exp\left(-\theta\left(1-\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)\right)^{2}\right)-\exp(-\theta)}{1-\exp(-\theta)} \tag{7}$$

Hazard, inverse hazard, and cumulative hazard rate function

The hazard rate function, which is the probability of failure of the random variable X, given that it has survived up to the time x, is defined as $h(x)=\frac{f(x)}{1-F(x)}$. Mathematically, the hazard rate function for the RGW distribution is well-defined based on Eq. (8). The hazard rate function is unique, increasing, and fluctuating after a certain point, and remains undefined beyond this point.

$$h(x)=\frac{2\alpha\theta}{\beta}\frac{\left(\frac{x}{\beta}\right)^{\alpha-1}\left[\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)-\exp\left(-2\left(\frac{x}{\beta}\right)^{\alpha}\right)\right]\exp\left(-\theta\left(1-\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)\right)^{2}\right)}{\left[\exp\left(-\theta\left(1-\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)\right)^{2}\right)-\exp(-\theta)\right]} \tag{8}$$

Figure 4 represents the hazard rate functions of the RGW distribution for selected parameter values. The curves exhibit diverse shapes, including nearly constant, increasing, and non-monotonic patterns. Notably, for some parameter combinations, the hazard fluctuates and terminates after a finite time, reflecting the model's ability to represent lifetime distributions with natural upper bounds. For the blue curve ($\alpha = 1$, $\beta = 23$, $\theta = 5$), the hazard rate remains nearly flat, suggesting an approximately constant failure rate, similar to exponential behavior. For the red curve ($\alpha = 2$, $\beta = 24$, $\theta = 5$), the hazard increases gradually after an initial low level, reflecting a

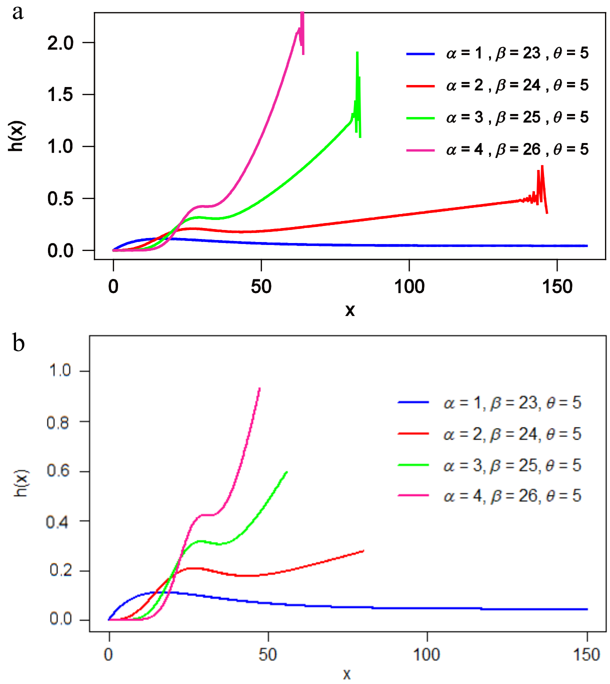

Fig. 4 (a) Plot of the hazard rate function of the RGW distribution. (b) Clean plot of the hazard rate function of the RGW distribution.

monotonically increasing failure risk. For the green curve ($\alpha = 3$, $\beta = 25$, $\theta = 5$), the hazard starts small, rises moderately, then steepens, showing a non-monotonic pattern. Finally, for the pink curve ($\alpha = 4$, $\beta = 26$, $\theta = 5$): the hazard rises sharply after a delay, exhibiting a bathtub-like or late-accelerating hazard, often seen in biological survival data or in mechanical systems with wear-out phases.

Furthermore, the hazard-rate behavior of the RGW distribution demonstrates its flexibility compared to classical extensions of the Weibull model. As shown in Fig. 4, different parameter combinations allow the RGW to represent nearly constant hazards (resembling the exponential distribution), monotonically increasing hazards, and non-monotonic, bathtub-shaped hazards. This versatility is particularly valuable in survival and reliability analysis, where real-world systems often exhibit hazard functions that change direction over time. Unlike the classical Weibull, which is restricted to strictly increasing or decreasing hazard forms, the RGW can capture more realistic risk dynamics, thereby enhancing its practical applicability.

To rigorously evaluate the behavior of the hazard function $h(x)$, as $x \to \infty$, we examine the limit of Eq. (8). Let, $z=\left(\frac{x}{\beta}\right)^{\alpha}$. As $x \to \infty$, it follows that $z \to \infty$ and $\exp(-z) \to 0$. The hazard function can be expressed as:

$$h(x)=\frac{2\alpha\theta}{\beta}\frac{\left(\frac{x}{\beta}\right)^{\alpha-1}\left[\exp(-z)-\exp(-2z)\right]\exp\left(-\theta(1-\exp(-z))^{2}\right)}{\left[\exp\left(-\theta(1-\exp(-z))^{2}\right)-\exp(-\theta)\right]}$$

As x becomes large, we apply the Taylor expansion for the term $(1-\exp(-z))^2 \approx (1-2\exp(-z))$. Substituting this into the exponential term in the denominator:

$$\exp\left(-\theta(1-\exp(-z))^{2}\right)\approx\exp(-\theta(1-2\exp(-z)))= \exp(-\theta)(\exp(2\theta\exp(-z)))$$

Using the expansion $\exp(1+\epsilon)\approx 1+\epsilon$ for small ϵ, the denominator becomes:

$$\exp(-\theta)(1+2\theta\exp(-z))-\exp(-\theta)=2\theta\exp(-z)$$

Substituting this back into the limit for $h(x)$:

$$\lim_{x\to\infty} h(x)=\frac{2\alpha\theta}{\beta}\frac{\left(\frac{x}{\beta}\right)^{\alpha-1}\exp(-z)\exp(-\theta)}{2\theta\exp(-z)\exp(-\theta)}$$

After canceling the common term, the expression simplifies to:

$$\lim_{x\to\infty} h(x)\propto\frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1}$$

This mathematical derivation confirms that:

• For $\alpha>1$, $\lim_{x\to\infty} h(x)=\infty$, indicating a strictly increasing hazard, meaning the failure is more likely as time progresses; this is characterizing systems that undergo wear-out or aging, where the risk of failure increases over time.

• For $\alpha=1$, the hazard approaches a constant. This represents the memoryless property, where failure is independent of the system's age.

• For $\alpha<1$, this represents a decreasing hazard rate, often seen in 'infant mortality' scenarios, where the probability of failure decreases as the entity survives the initial high-risk period.

These conditions prove that the fluctuations observed in Fig. 4a are computational iterations that were strictly numerical artifacts caused by the subtraction of nearly equal floating-point values in the denominator $\left(-\theta(1-\exp(-z))^2\right)-\exp(-\theta)$ as they both converged to $\exp(-\theta)$. By identifying that the power-law term governs the limit $\left(\frac{x}{\beta}\right)^{\alpha}$. It demonstrates that the model is mathematically stable and physically interpretable throughout its entire support.

In Fig. 4a, the hazard function appeared to fluctuate in the extreme right tail. A rigorous examination revealed that these fluctuations were not a theoretical property of the distribution, but were caused by catastrophic cancellation in the floating-point subtraction within the denominator. As x increases, the survival function approaches zero, and the computer is forced to calculate the difference between two values that are identical within machine precision approximately 10^{-16}.

To ensure scientific accuracy, the visual representation in Fig. 4b has been adjusted. We improved the numerical stability of the computations by using Python's expm1 function and truncating the curves where the survival function falls below 10^{-7}. Figure 4b now displays the corrected hazard rate functions. By removing the unstable terminal noise, the figure clearly demonstrates the smooth, monotonic increase in the hazard for $\alpha>1$. These results confirm that the model is numerically stable and physically interpretable across the distribution's statistically significant support, successfully capturing the intended reliability characteristics without interference from computational artifacts.

Similarly, the reverse (or inverse) hazard rate function of a continuous distribution is defined as $rh(x)=\frac{f(x)}{F(x)}, x>0$. For the RGW distribution, substituting the expression of $f(x)$ and $F(x)$ given in Eqs (6) and (5) into this definition, and after simplification, the reverse hazard rate function is expressed in Eq. (9) as:

$$rh(x)=\frac{2\alpha\theta}{\beta}\frac{\left(\frac{x}{\beta}\right)^{\alpha-1}\left[\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)-\exp\left(-2\left(\frac{x}{\beta}\right)^{\alpha}\right)\right]\exp\left(-\theta\left(1-\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)\right)^2\right)}{1-\exp\left(-\theta\left(1-\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)\right)^2\right)} \quad (9)$$

The cumulative hazard rate function of the RGW distribution is given by Eq. (10). Furthermore, Fig. 5 shows cumulative hazard rate functions for different parameter values.

$$H(x)=-\ln R(x)=\int_0^x h(x)\,dx=-\ln\left\{\frac{\exp\left(-\theta\left(1-\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)\right)^2\right)-\exp(-\theta)}{\exp(-\theta)}\right\}$$

$$H(x)=\ln(1-\exp(-\theta))-\ln\left\{\exp\left(-\theta\left(1-\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)\right)^2\right)-\exp(-\theta)\right\} \quad (10)$$

Random number generation of the RGW distribution

To generate the set of random variables that follows the RGW distribution, a method of inversion from the CDF has been used; for this, assume $F(x)=u$, where u is a function that follows the uniform distribution in [0, 1].

$$F(x)=\frac{1-\exp\left(-\theta\left(1-\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)\right)^2\right)}{1-\exp(-\theta)}=u$$

$$1-\exp\left(-\theta\left(1-\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)\right)^2\right)=u(1-\exp(-\theta))$$

$$\exp\left(-\theta\left(1-\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)\right)^2\right)=1-u(1-\exp(-\theta))$$

$$\left(1-\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)\right)^2=-\frac{1}{\theta}\ln(1-u(1-\exp(-\theta)))$$

$$\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)=1-\left\{-\frac{1}{\theta}\ln(1-u(1-\exp(-\theta)))\right\}^{\frac{1}{2}}$$

Further, simplification from this equation gives the value of the random variable X that follows the RGW distribution as derived,

$$X=\beta\left[-\ln\left\{1-\left\{-\frac{1}{\theta}\ln(1-u(1-\exp(-\theta)))\right\}^{\frac{1}{2}}\right\}\right]^{\frac{1}{\alpha}} \quad (11)$$

Equation (11) can be used to generate the set of random variates X for the RGW distribution from the given value of parameters α, β and θ. Similarly, this equation also serves as the quantile function of the RGW distribution.

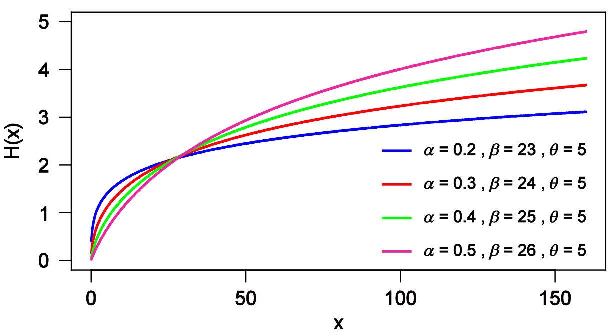

Fig. 5 Plot of the cumulative hazard rate function of the RGW distribution.

Method of parameter estimation using the method of maximum likelihood

Let, $X_1, X_2, \cdots, X_n$ to be a random sample of size n from the RGW distribution, having a PDF defined in Eq. (5). The likelihood function of the RGW has been expressed as:

$$L = \left(\frac{2\alpha\theta}{\beta(1-\mathrm{e}^{-\theta})}\right)^n \prod_{i=1}^{n}\sum_{i=1}^{n}\ln\left(\frac{x_i}{\beta}\right)^{\alpha-1}\prod_{i=1}^{n}\left[\exp\left(-\left(\frac{x_i}{\beta}\right)^{\alpha}\right)-\exp\left(-2\left(\frac{x_i}{\beta}\right)^{\alpha}\right)\right] \prod_{i=1}^{n}\exp\left(-\theta\left(\left(1-\exp\left(-\left(\frac{x_i}{\beta}\right)^{\alpha}\right)\right)^2\right)\right) \tag{12}$$

To determine the most likely parameter values, we operate on the logarithm of the likelihood function to get the Log-likelihood function, which is formulated as follows:

$$\ln L = n\ln\left(\frac{2\alpha\theta}{\beta(1-\exp(-\theta))}\right)+(\alpha-1)\sum_{i=1}^{n}\ln\left(\frac{x_i}{\beta}\right)+ \sum_{i=1}^{n}\ln\left\{\exp\left(-\left(\frac{x_i}{\beta}\right)^{\alpha}\right)-\exp\left(-2\left(\frac{x_i}{\beta}\right)^{\alpha}\right)\right\}- \theta\sum_{i=1}^{n}\left(\left(1-\exp\left(-\left(\frac{x_i}{\beta}\right)^{\alpha}\right)\right)^2\right) \tag{13}$$

The likelihood function is obtained by multiplying the density of the independent observations. For computational convenience, the natural logarithm of the likelihood function is taken, converting the product into a sum and simplifying the derivation of the maximum likelihood estimators (MLEs). To estimate the MLE of parameters involved in the RGW distribution, the components of the score vectors can be generated by taking the first derivative of the $\ln L$ with respect to the parameters as:

$$\frac{\partial \ln L}{\partial\theta} = \frac{n}{\theta}-\frac{n\exp(-\theta)}{1-\exp(-\theta)}-\sum_{i=1}^{n}\left(\left(1-\exp\left(-\left(\frac{x_i}{\beta}\right)^{\alpha}\right)\right)^2\right) \tag{14}$$

$$\frac{\partial \ln L}{\partial\alpha} = \frac{n}{\alpha}+\sum_{i=1}^{n}\ln\left(\frac{x_i}{\beta}\right)- \sum_{i=1}^{n}\frac{\left\{\exp\left(-\left(\frac{x_i}{\beta}\right)^{\alpha}\right)-2\exp\left(-2\left(\frac{x_i}{\beta}\right)^{\alpha}\right)\right\}\left(\frac{x_i}{\beta}\right)^{\alpha}\ln\left(\frac{x_i}{\beta}\right)}{\exp\left(-\left(\frac{x_i}{\beta}\right)^{\alpha}\right)-\exp\left(-2\left(\frac{x_i}{\beta}\right)^{\alpha}\right)}- 2\theta\sum_{i=1}^{n}\left(\exp\left(-\left(\frac{x_i}{\beta}\right)^{\alpha}\right)-\exp\left(-2\left(\frac{x_i}{\beta}\right)^{\alpha}\right)\right)\left(\frac{x_i}{\beta}\right)^{\alpha}\ln\left(\frac{x_i}{\beta}\right) \tag{15}$$

$$\frac{\partial \ln L}{\partial\beta} = -\frac{n}{\beta}-\left(\frac{n(\alpha-1)}{\beta}\right)- \frac{\alpha}{\beta}\sum_{i=1}^{n}\frac{\left\{\exp\left(-\left(\frac{x_i}{\beta}\right)^{\alpha}\right)-2\exp\left(-2\left(\frac{x_i}{\beta}\right)^{\alpha}\right)\right\}\left(\frac{x_i}{\beta}\right)^{\alpha}}{\exp\left(-\left(\frac{x_i}{\beta}\right)^{\alpha}\right)+\exp\left(-2\left(\frac{x_i}{\beta}\right)^{\alpha}\right)}- \frac{2\alpha\theta}{\beta}\sum_{i=1}^{n}\left(\exp\left(-\left(\frac{x_i}{\beta}\right)^{\alpha}\right)-\exp\left(-2\left(\frac{x_i}{\beta}\right)^{\alpha}\right)\right)\left(\frac{x_i}{\beta}\right)^{\alpha} \tag{16}$$

After equating these partial derivatives to zero, systems of nonlinear equations will be obtained. A simple calculation cannot produce the MLE for the unknown parameters of the GRW distribution. Suitable numerical nonlinear optimization algorithms, such as the Newton-Raphson method are needed to maximize the likelihood function.

Numerical application

Two real data sets were taken to test the proposed model's flexibility, suitability, and validity. We initially validate our model using a well-known 100-point sample representing bank customer service intervals, as documented by Ghitany et al.[30], and also utilized by Gaire & Gurung[31]. The second dataset comprises 128 data points on the remission times (in months) of bladder cancer patients. In this dataset, remission time is defined as the decrease or disappearance of cancer after treatment, as described by Lee & Wang[32]. Goodness-of-fit and parameter estimates for the fitted distributions were obtained using the R programming language.

The validity tools applied to validate the model fitting are the negative log-likelihood (NLL), Akaike information criterion (AIC), Bayesian information criterion (BIC), Anderson-Darling (AD) test, and Cramér–von Mises (CVM) test. The OPTIM function available in the R programming language[33] using the Adequacy Model[34] has been used to estimate the parameter associated with the RGW distribution. The algorithm used to estimate the parameters were the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method.

The RGW model was applied to fit two real-world data sets, and the parameters were estimated using MLE, the expression derived above in this study. The fitted results using an RGW distribution for both data sets are compared with those from a two-parameter Weibull distribution (Eq. [1]), a three-parameter transmuted Weibull (TrW) distribution (Eq. [17]), a four-parameter Kumaraswamy Weibull (KuW) distribution (Eq. [18]), a Beta Weibull (Beta-W) distribution[18], and a Marshall-Olkin Weibull (MO-W) distribution[4]. Estimated parameter values and test statistics for the waiting time data under the different models are presented in Table 1. Further, Fig. 6 presents histograms of the empirical and fitted values of customers' waiting time for different models.

$$f(x)_{TrW} = \left(\frac{\alpha\left(\frac{x}{\beta}\right)^{\alpha-1}\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)\left(1-\theta-2\theta\left(1-\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)\right)\right)}{\beta}\right), -1\le\theta\le 1 \tag{17}$$

$$f(x)_{KuW} = \left(\frac{ab\alpha(x)^{\alpha-1}\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)\left(1-\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)\right)^{a-1}\left(1-\left(1-\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)\right)^{a}\right)^{b-1}}{\beta^{\alpha}}\right) \tag{18}$$

Table 1. Estimated parameters and test statistics for waiting time.

Models	α	β	θ	a	b	NLL	AIC	BIC	KS (sig.)	AD	CVM
Weibull	1.461	10.952	–	–	–	318.618	641.236	646.45	0.0574 (0.897)	0.4031	0.0607
Beta-W	0.724	49.484	–	3.610	10.959	316.946	641.891	652.31	0.0368 (0.999)	0.1288	0.0177
MO-W	2.219	50.575	0.016	–	–	318.688	643.376	651.19	0.0478 (0.976)	0.2626	0.0295
KuW	1	0.367	47.18	5.369	81.733	317.482	644.965	657.99	0.0453 (0.987)	0.2204	0.0337
TrW	1.571	14.029	0.611	–	–	317.796	641.591	649.41	0.0488(0.972)	0.2595	0.0384
RGW	1.085	8.095	0.841	–	–	316.945	639.890	647.71	0.0376 (0.999)	0.1271	0.0176

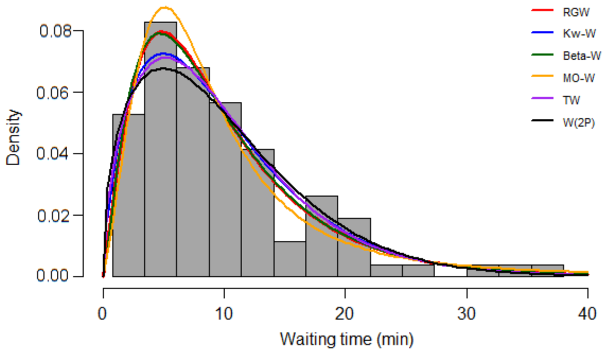

Fig. 6 Histogram of observed and fitted lines of values of the number of customers for waiting time.

Figure 7 presents P-P plots of waiting-time data to assess the visual adequacy of the RGW model. All models provided a reasonable fit; the RGW model, the greed one, exhibited the closest adherence to the diagonal reference line, particularly in the median and upper quantile regions, as indicated by the different test result matrices.

Based on the goodness-of-fit results, the RGW distribution provides the best fit among the considered models for customers' waiting time. It has the lowest NLL (316.95), AIC (639.89), and BIC (647.70). The smallest values of NLL, AIC, BIC, AD, and CVM test values for the RGW distribution indicate superior model performance in balancing overall goodness of fit and parsimony among the comparative models. Additionally, the KS test results indicate that the RGW distribution yields the smallest KS statistic (0.0376), the highest p-value (0.9989), the lowest AD test value (0.127), and the lowest CVM test value (0.017). These KS results, which are in strong agreement with the RGW-fitted results, are in good agreement with the observed data. The AD test suggests that the RGW model captures the data's extremes better than the other models. Similarly, the CVM test result indicates that the overall distance between the fitted result and observed distribution is smallest for RGW. Thus, the RGW distribution is the most appropriate model for describing customer waiting times in this dataset among the compared models.

In the same manner, estimated parameter values and test results for different statistics on the remission time data are presented in Table 2. The histograms of the observed and fitted patient counts are shown in Fig. 8.

A similar analysis of remission time data indicates that the RGW distribution is the best-fitting distribution among those considered. It produces the lowest NLL (409.824), AIC (825.6489), and BIC (831.3846), indicating its superiority in balancing model complexity and goodness of fit, and providing the best model fit among the four distributions. Additionally, the KS, AD, and CVM test results confirm strong agreement with the observed data, as the RGW model has the lowest KS statistic (0.0344), the highest p-value (0.9981), the lowest AD (0.127), and the lowest CVM (0.017), indicating an excellent fit. The RGW distribution is the most appropriate model for accurately describing remission times among the selected distributions.

Figure 9 presents the P-P plots of remission times for patients with bladder cancer to provide a graphical qualitative assessment of the model. In this case, the RGW model exhibits the most consistent adherence to the diagonal reference line across all quantiles, demonstrating that it effectively captures the underlying distribution of remission times and further validating its superiority, as suggested by multiple validity test result matrices.

Simulation study

The performance of the MLEs discussed in parameter estimation section for the RGW distribution was evaluated via a Monte Carlo simulation. Three distinct sets of true parameter values were tested across sample sizes n = 100, 200, and 500, with N = 1,000 iterations per sample size. The random numbers were generated using Eq. (11). The numerical results of the evaluation, Tables 3–5, were summarized and included the estimated mean, bias, and mean squared error (MSE).

The simulation study results for three sets of parameter values provide evidence of the asymptotic consistency of the MLEs for the

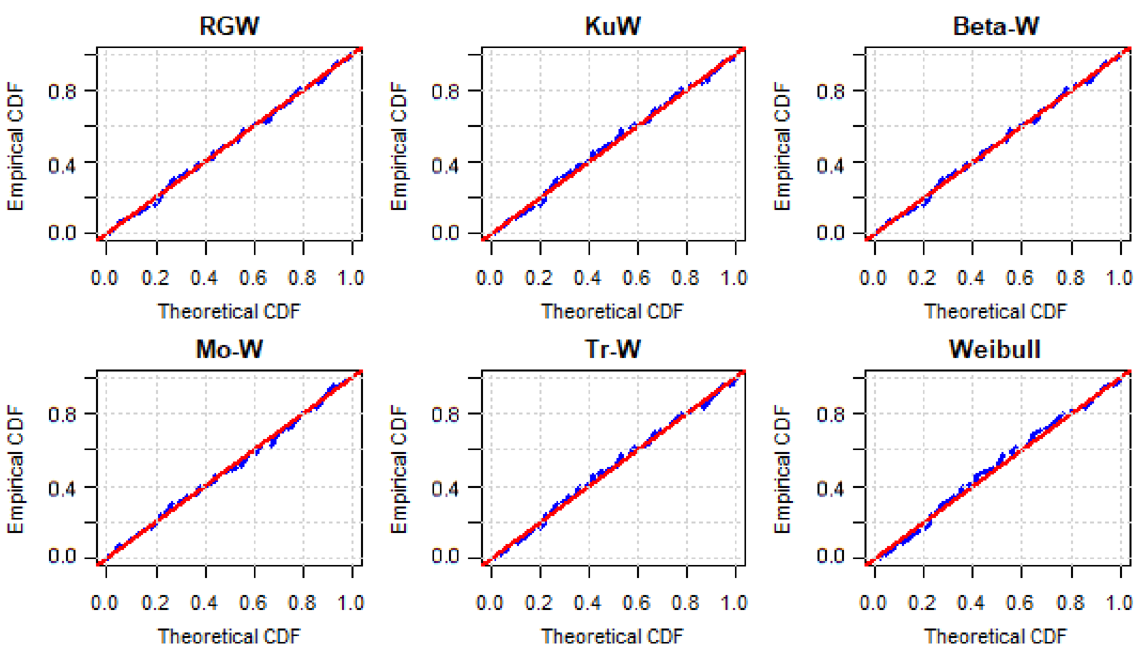

Fig. 7 P-P plot of customers for waiting time.

Table 2. Estimated parameters and test statistics for remission time of bladder cancer.

Models	α	β	θ	a	b	NLL	AIC	BIC	KS(Sig.)	AD	CVM
Weibull	1.048	9.5607	–	–	–	414.087	832.174	837.878	0.0700 (0.557)	0.958	0.154
Beta-W	0.528	49.643	–	3.672	8.679	410.839	829.678	841.086	0.0466 (0.944)	0.297	0.045
MO-W	1.660	49.984	–	–	–	410.321	826.642	835.198	0.0357 (0.997)	0.186	0.019
KuW	1.000	0.335	–	4.243	48.68	411.537	833.075	847.335	0.056(0.820)	0.488	0.079
TrW	1.133	14.619	0.745	–	–	411.958	829.916	838.473	0.0587(0.769)	0.560	0.088
RGW	0.786	14.302	4.082	–	–	409.824	825.649	834.205	0.0344 (0.998)	0.127	0.017

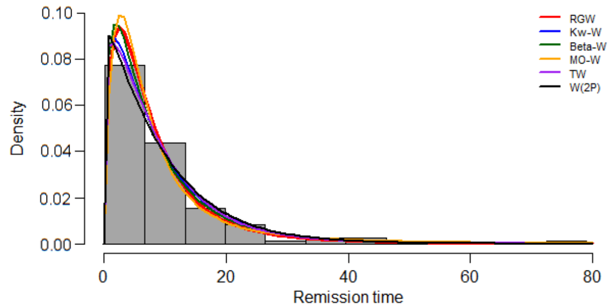

Fig. 8 Empirical and fitted number of patients against remission time.

Table 3. Simulation results for Set I.

Sample size	Parameter	True value	Estimate	Bias	MSE
100	θ	1.5	1.954	0.454	0.257
	α	2.0	2.570	0.570	0.368
	β	0.8	1.199	0.399	0.169
200	θ	1.5	1.922	0.422	0.187
	α	2.0	2.549	0.549	0.323
	β	0.8	1.194	0.394	0.158
500	θ	1.5	1.905	0.405	0.166
	α	2.0	2.551	0.551	0.311
	β	0.8	1.186	0.385	0.150

RGW distribution. The initial observations indicate that the MSE consistently decreases as the sample size increases from 100 to 500 across almost all parameters. Such downward trends in MSE indicate that the estimators become more precise and converge toward the true parameter values as more data become available. In the presence of bias, the estimators exhibit a general trend of convergence toward zero as the sample size increases, though some parameters, such as theta in the third set, show persistent negative bias in smaller samples. The shape parameter alpha tended to show a slight overestimation across all sets, while the scale parameter beta remained relatively stable with low MSE. In general, as the sample size increases, both bias and MSE decrease, which supports the use of MLE to estimate the parameters of the RGW distribution.

Figures 10 and 11 present the visual comparisons of bias and MSE for the simulation results across the three parameter sets. The bias plot shows that the bias for each parameter changes with increasing sample size across the three sets. It is observed that most parameters converge toward the zero line as the sample size increases, suggesting that the estimators are asymptotically unbiased. Similarly, the MSE plots clearly show a downward slope for all parameter selections. As the sample size increases, the variation in parameter estimates decreases, indicating that the model is statistically consistent. For the third parameter, theta shows higher sensitivity, suggesting that for very large theta values, even a larger sample size might be required to achieve significant convergence.

Conclusions

The RGW distribution, a sub-model of the RG-family of distributions, has been presented. The relationships for estimating the

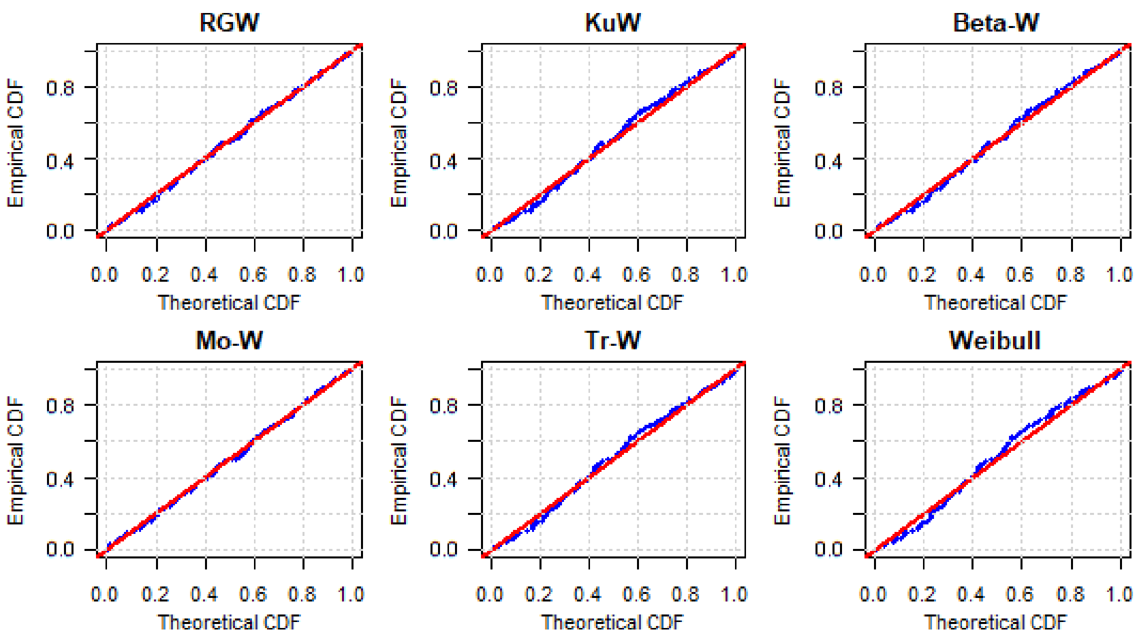

Fig. 9 P-P plot of remission time.

Table 4. Simulation results for Set II.

Sample size	Parameter	True value	Estimate	Bias	MSE
100	θ	2.0	1.925	−0.075	0.039
	α	3.0	3.852	0.852	0.824
	β	1.8	2.261	0.461	0.224
200	θ	2.0	1.900	−0.100	0.017
	α	3.0	3.819	0.819	0.717
	β	1.8	2.257	0.457	0.213
500	θ	2.0	1.887	−0.113	0.015
	α	3.0	3.819	0.819	0.687
	β	1.8	2.490	0.449	0.203

Table 5. Simulation results for Set III.

Sample size	Parameter	True value	Estimate	Bias	MSE
100	θ	3.0	2.919	−0.0807	11.577
	α	3.0	2.999	−0.0012	0.0705
	β	3.0	2.940	−0.0596	0.1979
200	θ	3.0	2.912	−0.0882	4.2503
	α	3.0	2.967	−0.0313	0.0396
	β	3.0	2.966	−0.0338	0.1346
500	θ	3.0	3.013	0.0127	2.2105
	α	3.0	2.975	−0.0246	0.0184
	β	3.0	2.993	−0.0068	0.0755

parameter using the MLE and random number generation are derived, and certain statistical features of the distribution are examined. The hazard curves of this distribution exhibit a variety of forms, including non-monotonic, rising, and almost constant patterns. The simulation results verify that as the sample size increases, the MLE consistently yields estimates of the RGW distribution parameters that are progressively more accurate. To assess the performance of the RGW distribution, two datasets were employed: bank queue length, and patient remission intervals. As validity tools, a variety of test statistics have been employed, including AIC, BIC, KS, AD, and CVM. From the diverse validity test statistics and the P-P plot, it was found that the RGW model is sufficiently flexible to better fit the data than the other compared models.

Author contributions

The author confirms sole responsibility for all aspects of this study and approved the final version of the manuscript.

Data availability

The two datasets used to analyze during this study are available from Rodriguez and Gaire & Gurung[29,31].

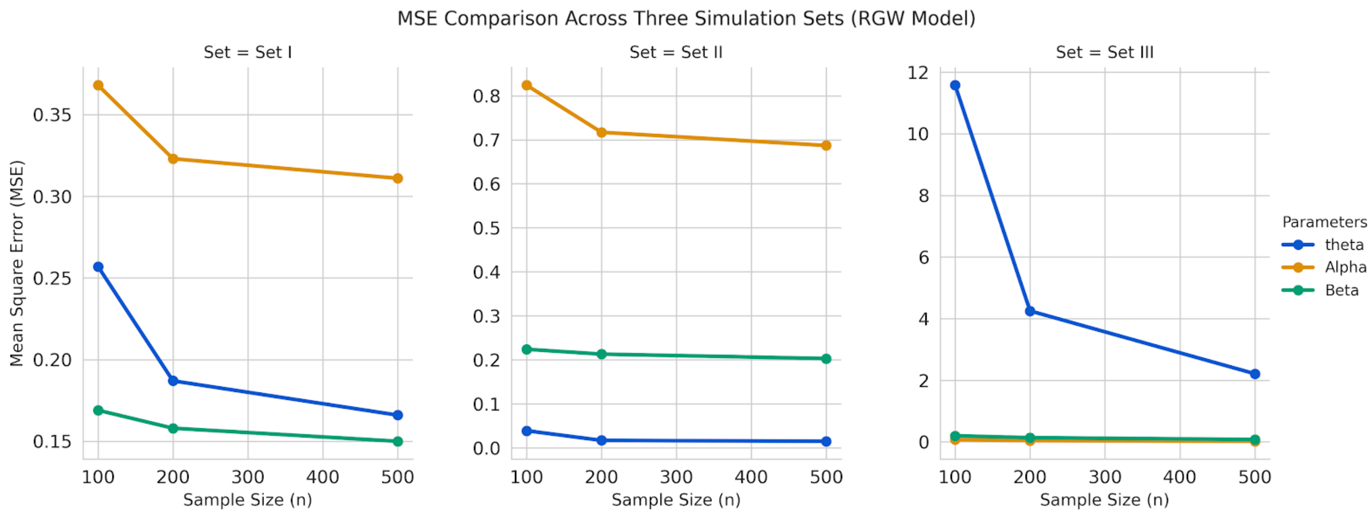

Fig. 10 MSE comparison across parameter sets for RGW simulation.

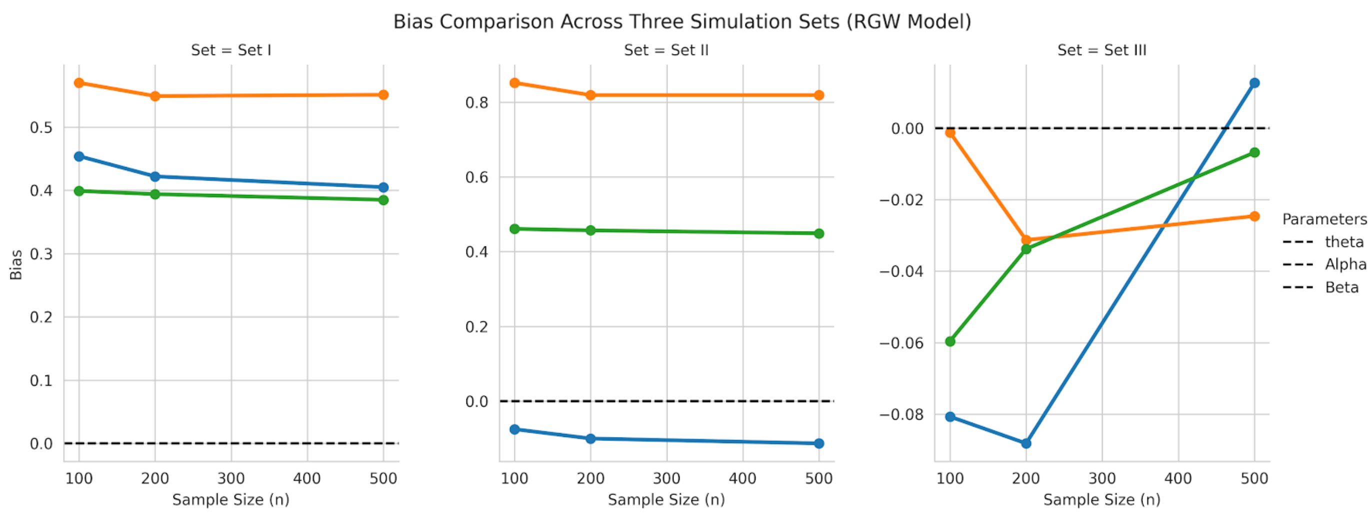

Fig. 11 Bias comparison across parameter sets for RGW simulation.

Acknowledgments

The author received no financial support for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that there is no conflict of interest regarding this research and that no ethical violations occurred.

Dates

Received 9 October 2025; Revised 10 April 2026; Accepted 21 April 2026; Published online 26 August 2026

References

[1] Gaire AK, Gurung YB. 2024. Rayleigh generated log–logistic distribution properties and performance analysis. *Istatistik Journal of the Turkish Statistical Association* 15(1):13–28
[2] Shaw WT, Buckley IRC. 2007. The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. *Papers 0901.0434*
[3] Gupta RC, Gupta PL, Gupta RD. 1998. Modeling failure time data by Lehman alternatives. *Communications in Statistics - Theory and Methods* 27:887–904
[4] Marshall AW, Olkin I. 1997. A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika* 84(3):641–652
[5] Zografos K, Balakrishnan N. 2009. On families of beta- and generalized gamma-generated distributions and associated inference. *Statistical Methodology* 6(4):344–362
[6] Eugene N, Lee C, Famoye F. 2002. Beta-normal distribution and its applications. *Communications in Statistics - Theory and Methods* 31(4):497–512
[7] Alzaatreh A, Lee C, Famoye F. 2013. A new method for generating families of continuous distributions. *METRON* 71(1):63–79
[8] Mohammad S, Gaire AK. 2025. Exponential arctan-G family of distribution with properties and applications. *Journal of Probability and Statistics* 25(1):3643496
[9] Mohammad S, Cooray K. 2025. Modeling actuarial data using iterated trigonometric distributions. *Communications in Statistics - Theory and Methods* 5112–5128
[10] Gaire AK, Gurung YB, Bhusal TP. 2024. Fertility model evolution: a survey on mathematical models of age-specific fertility with application to Nepalese and Malaysian data. *Global Health Economics and Sustainability* 3(1):222–234
[11] Gaire AK, Gurung YB, Bhusal TP. 2023. Age at first marriage of Nepalese women: a statistical analysis (status, differential, determinants, and distributional pattern). *Journal of Population and Social Studies* 32:308–328
[12] Xu M, Mao H. 2024. q-Weibull distributions: perspectives and applications in reliability engineering. *IEEE Transactions on Reliability* 74(3):3112–3125
[13] Johnson NL, Kotz S, Balakrishnan N. 1994. *Continuous Univariate Distributions*, Volume 1, 2nd Edition. New York: John Wiley. pp. 25–26
[14] Lai CD, Murthy DN, Xie M. 2006. Weibull distributions and their applications. In *Springer Handbook of Engineering Statistics*, ed. Pham H. London: Springer. pp. 63–78 doi: 10.1007/978-1-84628-288-1_3
[15] Drapella A. 1993. The complementary Weibull distribution: unknown or just forgotten? *Quality and Reliability Engineering International* 9(4):383–385
[16] Mudholkar GS, Kollia GD. 1994. Generalized Weibull family: a structural analysis. *Communications in Statistics - Theory and Methods* 23(4):1149–1171
[17] Cordeiro GM, Ortega EM, Nadarajah S. 2010. The Kumaraswamy Weibull distribution with application to failure data. *Journal of the Franklin Institute* 347(8):1399–1429
[18] Lai CD, Xie M, Murthy DNP. 2003. A modified Weibull distribution. *IEEE Transactions on Reliability* 52(1):33–37
[19] Lee C Famoye F, Olumolade O. 2007. Beta-Weibull distribution: some properties and applications to censored data. *Journal of Modern Applied Statistical Methods* 6(1):173–186
[20] Aryal, G. R. and Tsokos, C. P. 2011. Transmuted Weibull distribution: a generalization of the Weibull probability distribution. *European Journal of Pure and Applied Mathematics* 4(2):89–102
[21] Ghitany ME, Al-Hussaini EK, Al-Jarallah RA. 2005. Marshall–Olkin Extended Weibull distribution and its application to censored data. *Journal of Applied Statistics* 32(10):1025–1034
[22] Nadarajah S, Kotz S. 2005. The Beta exponential distribution. *Reliability Engineering and System Safety* 91:689–697
[23] Carrasco JMF, Ortega EMM, Cordeiro GM. 2008. A generalized modified Weibull distribution for lifetime modeling. *Computational Statistics & Data Analysis* 53(2):450–462
[24] Silva GO, Ortega EM, Cordeiro GM. 2010. The Beta modified Weibull distribution. *Lifetime Data Analysis* 16(3):409–430
[25] Nadarajah S, Cordeiro GM, Ortega EMM. 2011. General results for the beta-modified Weibull distribution. *Journal of Statistical Computation and Simulation* 81(10):1211–1232
[26] Pinho LGB, Cordeiro GM, Nobre JS. 2012. The Gamma-Exponentiated Weibull distribution. *Journal of Statistical Theory and Applications* 11(4):379–395
[27] Khan MS, King R. 2013. Transmuted modified Weibull distribution: a generalization of the modified Weibull probability distribution. *European Journal of Pure and Applied Mathematics* 6(1):66–88
[28] Oluyede BO, Huang S, Yang T. 2015. A new class of generalized modified Weibull distribution with applications. *Austrian Journal of Statistics* 44(3):45–68
[29] Rodrıguez G. 2010. *Parametric survival models*. Rapport technique. Princeton: Princeton University
[30] Ghitany ME, Atieh B, Nadarajah S. 2008. Lindley distribution and its application. *Mathematics and Computers in Simulation* 78(4):493–506
[31] Gaire AK, Gurung YB. 2024. Skew log-logistic distribution: properties and application. *Statistics in Transition New Series* 25(1):43–62
[32] Lee ET, Wang JW. 2003. *Statistical methods for survival data analysis*. Vol. 476. US: John Wiley & Sons. doi: 10.1002/0471458546
[33] R Core Team. 2025. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria
[34] Marinho PRD, Silva RB, Bourguignon M, Cordeiro GM, Nadarajah S. 2019. Adequacy model: an R package for probability distributions and general-purpose optimization. *PLoS One* 14(8):e0221487