Tanzania Journal of Science 51(2): 407-428, 2025
ISSN 0856-1761, e-ISSN 2507-7961

Probability Modeling of Carbon Dioxide Emissions in Tanzania

Twahil Hemed Shakiru
Department of Statistics, University of Dar es Salaam, Tanzania
Email: *hemedtwahil@gmail.com, hemed.twahil@udsm.ac.tz*
Received 14 Nov 2024; Reviewed 15 May 2025; Accepted 30 May2025; Publ. 30 June 2025
https://dx.doi.org/10.4314/tjs.v51i2.12

Abstract

Carbon dioxide emissions are a significant driver of climate change, impacting ecosystems, human health, and economies. In Tanzania, increasing carbon dioxide emissions from fossil fuel use, deforestation, and industrial growth contribute to environmental challenges such as air pollution, agricultural disruption, and changing weather patterns. This study models Tanzania's carbon dioxide emissions using the Generalized Log-Logistic distribution and compares it to other models, including Burr XII, log-logistic, Weibull, and log-normal distributions. The results show that the Generalized Log-Logistic distribution provides the best fit to the data, outperforming other models in goodness-of-fit, log-likelihood values, and information criteria. Three estimation methods such as maximum likelihood, least squares, and weighted least squares were applied, with maximum likelihood yielding the lowest mean squared error, making it the most effective for parameter estimation. The likelihood ratio test further confirmed that the Generalized Log-Logistic model offers a superior fit compared to its sub-models, demonstrating its robustness in representing carbon dioxide emissions data. These findings establish the Generalized Log-Logistic model as a valuable tool for monitoring carbon dioxide emission patterns in Tanzania, essential for managing rising emissions. This study underscores the importance of reliable probability models in addressing environmental challenges and informing strategies to mitigate carbon dioxide emissions.

Keywords: Generalized Log-Logistic; Carbon dioxide Emissions; Sub-models; Parameter Estimation.

Introduction

The choice of appropriate probability models is crucial in environmental pollution studies, as they provide a robust framework for understanding and managing variability in pollutant concentrations. These models enable accurate parameter estimation, facilitate health risk assessment, and support adaptive environmental management strategies. Selecting suitable probability distributions allows researchers to effectively model contaminant behavior, predict their spread, and inform regulatory decisions to mitigate environmental and public health impacts. However, pollution data often exhibit right-skewed distributions, making normal distribution models unsuitable. A common approach to address this is to transform the data, often using a logarithmic conversion. Nevertheless, environmental studies prioritize estimating statistics on the original measurement scale rather than relying on transformed data (Singh et al. 2001).

Modeling environmental pollution is complicated by the challenge of handling observations that fall below the detection limit (DL), often reported as non-detect (ND) or less than the detection limit (LDL). These observations lack precise numerical values and traditional methods become problematic with left-censored data (Aryal 2013). Analysts frequently exclude observations below the DL or substitute them with values like zero or DL/2, assuming uniform distribution between zero and DL. However, both methods can introduce bias, which intensifies as censoring increases (Newman et al. 1989). This study recommends exploring more robust methods for handling censored data in environmental pollution research.

The log-normal distribution has been widely used to model pollution concentrations, but alternative distributions may provide a better fit. Larsen (1974) introduced a third parameter to the log-normal distribution, which helps transform a log-probability plot into a nearly straight line. This additional parameter increases the model's flexibility and was further developed into the censored three-parameter log-normal model by Mage and Ott (1978). However, they cautioned against automatically using any specific model without validating its suitability, as this could lead to misleading results.

Studies have shown that alternative distributions, such as the gamma distribution, can outperform the log-normal in some cases. For example, Berger et al. (1982) found that the gamma distribution provided a better fit for atmospheric sulfur dioxide (SO_2) concentrations in Belgium. Jakeman and Taylor (1985) also demonstrated that gamma models were more effective than log-normal models for acid-gas concentrations in industrial airsheds. These findings suggest that exploring various probability models may better capture the unique characteristics of environmental data.

Adeyinka (2019) argued that no single probability model, including the log-normal distribution, is universally superior. The generalized log-logistic (GLL) distribution emerges as a promising option for environmental pollution data analysis (Lima and Cordeiro 2017). As an extension of the log-logistic distribution, the GLL shares similarities with the log-normal, but its enhanced mathematical simplicity makes it particularly suitable for censored data (Singh et al. 1994). This flexibility makes the GLL distribution well-suited for handling environmental data.

The selection of an appropriate model often begins with a general model that can incorporate specific cases (Aldahlan 2020). Although less common in environmental studies, the skewness and heavy tails of GLL distributions make them suitable for a variety of datasets (Malik and Ahmad 2020). The GLL family includes well-known sub-models like the log-logistic, Weibull, and Burr XII distributions, which are commonly used for pollutant concentration data. Given this flexibility, the GLL distribution is expected to be an effective model for environmental data analysis. This study proposes the use of the GLL family to analyze carbon dioxide emissions data in Tanzania, as it can accommodate deviations from log-normality, such as skewness and kurtosis.

This research stands apart by applying the generalized log-logistic (GLL) distribution as a general model for environmental pollution data, specifically focusing on CO_2 emissions in Tanzania. Unlike previous studies, it compares a wide range of distributions, including the log-normal and various GLL sub-models such as log-logistic, Weibull, exponential, and Burr XII distributions. The study emphasizes the simplicity of the GLL distribution's formulas and explores its mathematical and statistical properties. The versatility of the GLL model and its compatibility with modern computational tools provide new insights into Tanzania's environmental challenges, illustrating the applicability of advanced probability models in this setting. To the best of current knowledge, this is the first study to apply GLL distributions to model CO_2 emissions data specifically in Tanzania.

Overview of Carbon Dioxide Emission in Tanzania

According to Global Carbon Budget data, Tanzania's CO_2 emissions reached approximately 15.57 megatons in 2022, making it the second-largest emitter in East Africa, following Kenya. This rise highlights the country's growing contribution to global warming, driven by fossil fuel consumption (coal, oil, and natural gas), deforestation linked to agricultural expansion and infrastructure development, and agricultural practices such as crop burning and livestock farming (Hafner et al. 2019). Among these sources, oil has been the largest contributor, with emissions steadily increasing since the 1960s (see Figure 1). Gas emissions have risen significantly since the early 2000s, while coal and cement production have also contributed, with emissions sharply increasing since the 1990s. Although flaring contributes less, its emissions have also gradually risen.

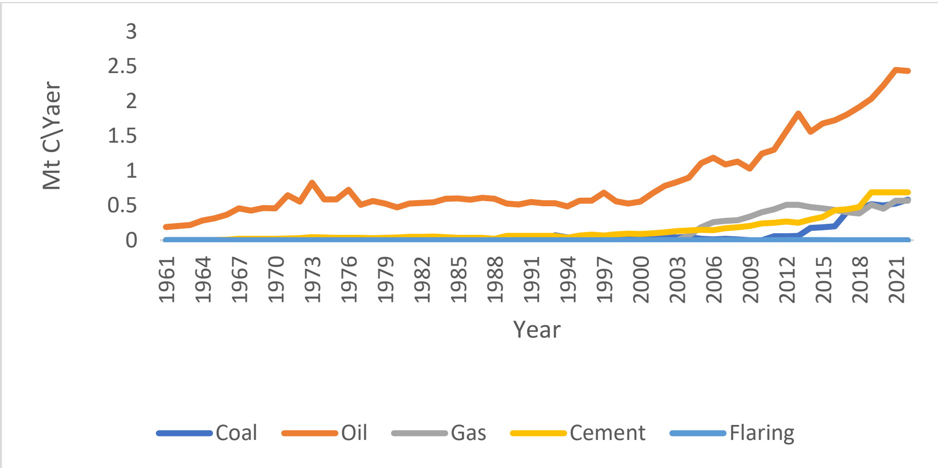

Figure 1: Carbon dioxide Emissions by Fuel/Industry in Tanzania
Source: Global Carbon Budget (2023) data

The trends in CO_2 emissions for Tanzania show a steady increase, with a notable peak in 2022, as depicted in Figure 2. A blue trend line illustrates the rise in emissions over time, while a red horizontal line indicates the average CO_2 emission level of approximately 4.35 megatons. A green vertical line marks 2022 as the year with the highest emissions. This data underscores the urgent need for Tanzania to adopt more effective strategies for managing CO_2 emissions, including transitioning to renewable energy sources, conserving forests, and promoting sustainable agricultural practices. Addressing these issues is critical for mitigating environmental impacts, such as changes in temperature, sea levels, and rainfall patterns, which could disrupt ecosystems, water supplies, and agriculture, ultimately affecting socioeconomic development and food security in the country.

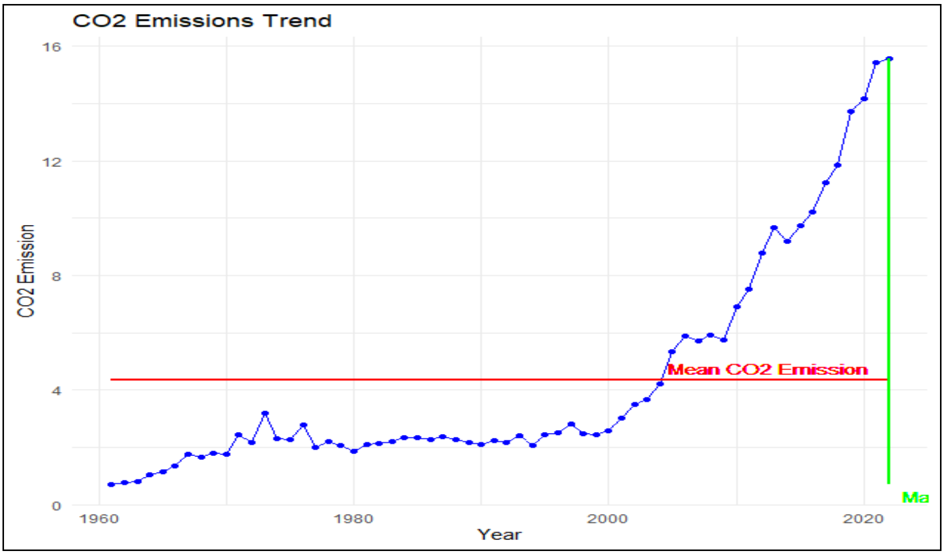

Figure 2: Carbon Dioxide Emission Trend in Tanzania
Source: Computed by Author based on Global Carbon Budget (2023) data

Despite the upward trend in Tanzania's CO_2 emissions, its global contribution remains low at approximately 0.02% (see Figure 3), underscoring its minimal role in global climate change compared to major emitters like the United States (24.08%), the European Union (16.69%), and China (14.7%), with the United Kingdom and India contributing 4.45% and 3.37%, respectively, and Africa as a whole accounting for 2.288%. However, Tanzania's emissions reached to 15.57 megatons in 2022 (Figure 2), position it as the second-largest emitter in East Africa, emphasizing its regional significance. This trend, combined with Tanzania's vulnerability to climate impacts like flooding and droughts, justifies the importance of studying and mitigating its emissions to support regional climate resilience, sustainable development, and local environmental health, even if its global impact is limited.

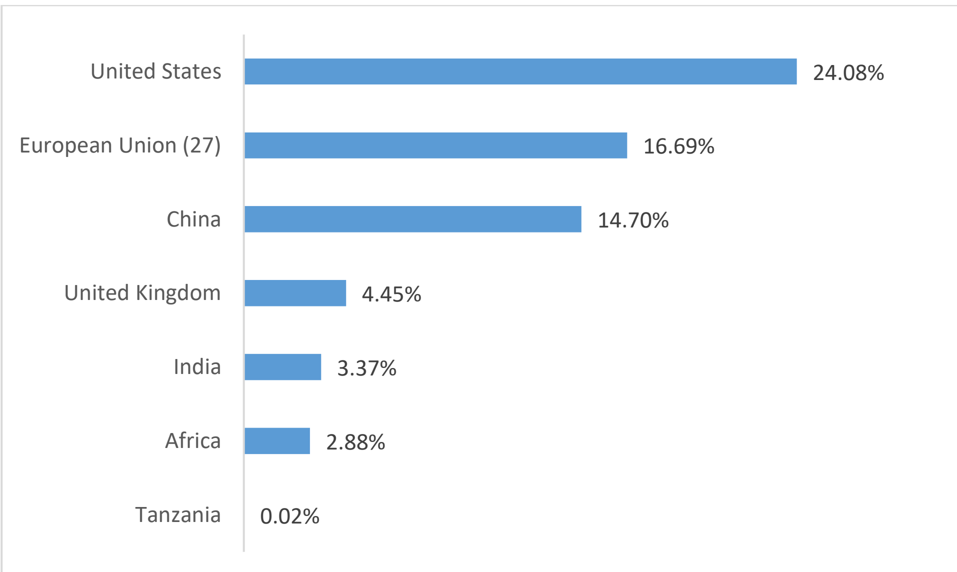

Figure 3: Comparison of Global CO_2 Emission Contributions in 2022

Source: Computed by Author based on Global Carbon Budget (2023) data

Materials and Methods

Materials

This study used a dataset on CO_2 emissions in Tanzania from the Global Carbon Budget (2023), covering the period from 1961 to 2022. The dataset includes CO_2 emission values in megaton, offering a detailed view of the trends and changes in emissions over six decades.

Methods

Log-normal Distribution

A random variable $X > 0$ follows a two-parameter log-normal (LN) distribution if its logarithm, $Y = (InX)$, is normal distribution with mean (μ) and variance (σ^2). In other words, if (X) has a log-normal distribution, then the logarithmic transformation of X (i.e., $Y = (InX)$) will be normally distributed with the parameters μ and σ^2. The probability density function (PDF) for (X) depends on the parameters μ (the mean of the log-transformed variable), σ (the standard deviation of the log-transformed variable), and x (the value of the random variable (X), and is given by:

$$f(x, \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} exp\left(\frac{-(lnx-\mu)^2}{2\sigma^2}\right) \quad , \quad x < 0.$$
(1)

The Generalized log-logistic Distribution

The Generalized Log-Logistic (GLL) distribution is a continuous probability distribution commonly used to model random variables. The Generalized Log-Logistic (GLL) distribution is a continuous probability distribution often used to model random variables. It extends the log-logistic (LL) distribution and is defined by three key parameters: shape (α), scale (β), and location (γ). The GLL distribution's flexibility in modeling a wide range of data patterns has made it popular across various disciplines, including environmental science, reliability analysis, and finance. Singh (1989) first introduced the GLL distribution as an extension of the LL distribution to model data related to lung cancer and other types of cancer. He demonstrated its adaptability by applying it to lung cancer survival data. Building on this work, Singh et al. (1994) highlighted the GLL distribution's utility in modeling breast cancer survival data, further proving its versatility in handling complex datasets in medical research.

If X is a random variable following the Generalized Log-Logistic (GLL) distribution, its probability density function (PDF) can be expressed using Equation (3), which includes three parameters: shape, scale, and location. Each of these parameters plays a role in shaping the distribution's characteristics and behavior. The PDF is initially represented as:

$$f(x, \alpha, \beta, \lambda) = k(x, \alpha, \beta, \lambda) exp\left\{-\int_0^x k(x)dx\right\}. \quad (2)$$

where, $f(x, \alpha, \beta, \lambda) = \frac{\alpha\beta^\alpha x^\alpha}{1+(\lambda x)^\alpha}$. Simplifying Equation (2) yields (3):

$$f(x, \alpha, \beta, \lambda) = \frac{\alpha\beta(\beta x)^{\alpha-1}}{[1+(\lambda x)^\alpha]^{\left(\frac{\beta^\alpha}{\lambda^\alpha}\right)+1}}, \qquad x \geq 0, \alpha, \beta, \lambda > 0. \quad (3)$$

Here, the function k is incorporated into the final probability density function through the integral $\int_0^x k(t)dt = \frac{\beta^\alpha}{\lambda^\alpha} ln(1 + (\lambda x)^\alpha)$, which produces the exponential term $[1 + (\lambda x)^\alpha]^{-\frac{\beta^\alpha}{\lambda^\alpha}}$, resulting in the standard GLL probability density function form expressed in Equation (3). The parameters α, β and λ are defined as follows:

α indicates the shape parameter that affects the GLL distribution curve's shape.
β denotes the scale parameter, which establishes the distribution's scale or spread.
λ denotes the location parameter and permits a shift or displacement along the x-axis.

The curves of the probability density function for the generalized log-logistic distribution, based on various parameter combinations, are illustrated in Figure 4. This figure illustrates how the distribution tends to be more dispersed and has wider tails and a flatter overall shape when the scale parameter is small. On the other hand, a large-scale parameter causes the distribution to become more peaked and concentrated, with smaller tails. Additionally, a location parameter that is extremely small or negative compresses the distribution around smaller values, highlighting the distribution's lighter side.

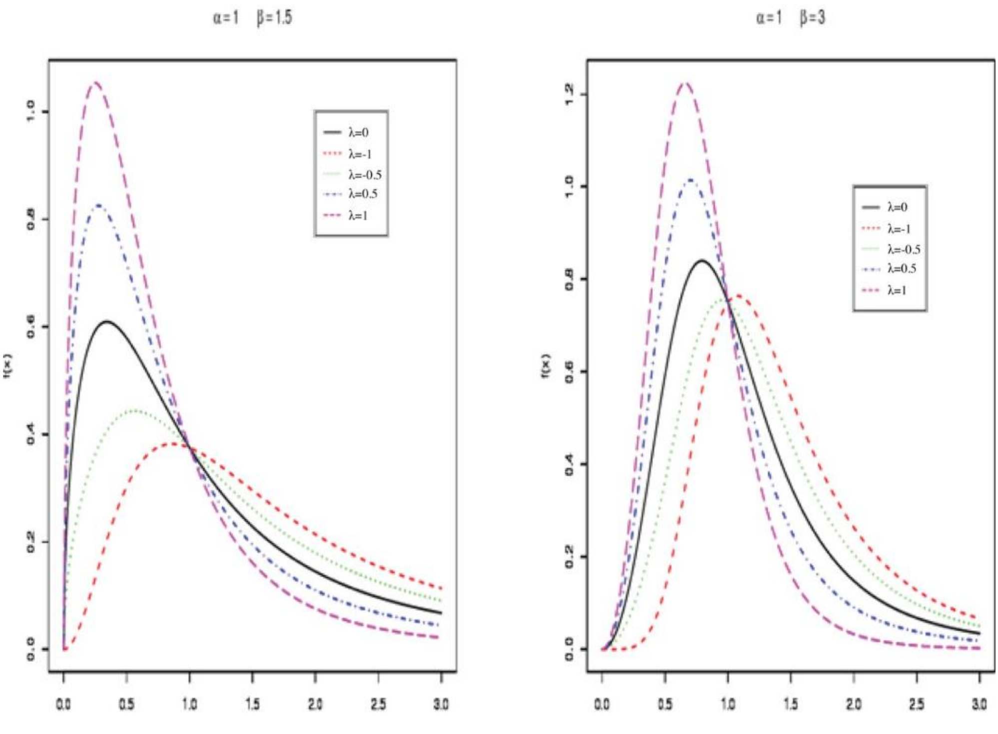

Figure 4: Probability Density Function Curves of the Generalized Log-Logistic (GLL) Distribution Under Different Parameter Configurations.

The cumulative distribution function of the Generalized Log-Logistic (GLL) distribution is described by:

$$F(x,\alpha,\beta,\lambda) = \frac{[1+(\lambda x)^{\alpha}]^{\left(\frac{\beta^{\alpha}}{\lambda^{\alpha}}\right)} - 1}{[1+(\lambda x)^{\alpha}]^{\left(\frac{\beta}{\lambda}\right)^{\alpha}}}, \quad x \geq 0, \alpha, \beta, \lambda > 0. \tag{4}$$

The reliability function of the Generalized Log-Logistic (GLL) distribution, which denotes the probability that a variable or component will continue to function without failure up to a specified time, is given by Equation (6). This function provides insight into the likelihood of survival or continued operation over time, capturing the distribution's behavior in terms of reliability and longevity.

$$R(x,\alpha,\beta,\lambda) = \frac{f(x,\alpha,\beta,\lambda)}{k(x,\alpha,\beta,\lambda)} \tag{5}$$

Upon simplification Equation (5) gives

$$R(x,\alpha,\beta,\lambda) = [1+(\lambda x)^{\alpha}]^{-\left(\frac{\beta}{\lambda}\right)^{\alpha}}, x \geq 0, \alpha, \beta, \lambda > 0. \tag{6}$$

Figure 5 illustrates how the reliability characteristics of the Generalized Log-Logistic (GLL) distribution are affected as the parameter λ changes from -1 to 1. The figure highlights that the shape and spread of the distribution, along with the rate at which reliability declines over time, are determined by the value of the parameter β. Specifically, when β is higher, the distribution becomes more peaked, and the rate of reliability decline increases. In other words, a larger scale parameter β results in a narrower and more concentrated distribution, which leads to a more pronounced drop in reliability as time t rises.

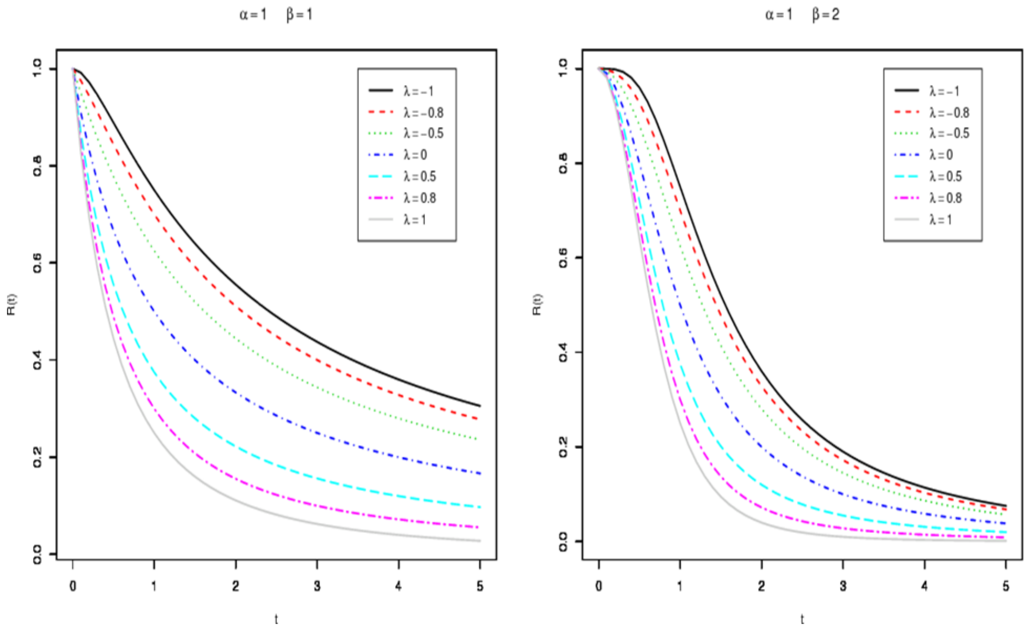

Figure 5: Reliability Function of the Generalized Log-Logistic (GLL) Distribution

Thus, the Generalized Log-Logistic (GLL) distribution is defined by three parameters: the shape parameter (α), the scale parameter (β), and the location parameter (λ), represented as GLL (α, β, λ). The hazard function, which indicates the instantaneous failure rate at any time t, is expressed in Equation (7). This function provides critical insight into how the risk of failure varies over time based on the distribution's parameters.

$$H(x,\alpha,\beta,\lambda) = \frac{\alpha\beta(\beta x)^{\alpha-1}}{[1+(\lambda x)^{\alpha}]}, x \geq 0, \beta, \alpha, \lambda > 0, \quad (7)$$

Equation (7) clearly shows that the behavior of the hazard rate function is influenced by the value of the shape parameter α. Specifically:

1. When α ≤ 0, the hazard rate function decreases steadily over time, meaning that the risk of failure consistently reduces without any rise as time progresses.
2. When α > 1, the hazard rate function exhibits a unimodal pattern. In this case, the function increases initially, reaches a peak at a specific time t, and then gradually diminishes toward zero as time t draws near infinity.

The time at which the hazard rate function reaches its maximum is expressed as $t = \left[\frac{\alpha-1}{\lambda^{\alpha}}\right]^{\frac{1}{\alpha}}$. At this point, the hazard rate function is at its highest value, after which it decreases steadily and eventually approaches zero as time goes on.

Sub-models

The Generalized Log-Logistic (GLL) distribution includes several key sub-models, such as the exponential, Burr XII, Weibull, log-logistic, and standard log-logistic distributions, each derived by setting specific parameters. For example, the exponential distribution emerges from the GLL when the shape parameter α approaches 1 and the location parameter λ is set to zero. Adjusting the GLL's parameters also yields the Burr XII distribution for heavy-tailed data and the Weibull distribution for varying hazard rates. The log-logistic and standard log-logistic distributions are used in survival analysis, with the former being more general and the latter a simplified version. Additionally, the log-shaped distribution is another variant with a different parameterization influencing the hazard function. These variations highlight the GLL distribution's flexibility in modeling diverse data types (Malik and Ahmad 2020).

Log-Logistic Distribution

Consider a random variable X that adheres to a generalized log-logistic distribution, denoted as $X \sim GLL(\alpha, \beta, \lambda)$. If the parameter λ depends on β such that $\beta = \lambda$, the hazard rate function of the GLL distribution simplifies to that of a standard log-logistic distribution.

The hazard rate function for the GLL distribution is defined as:

$$H(x, \alpha, \beta, \lambda) = \frac{\alpha\beta(\beta x)^{\alpha-1}}{[1+(\lambda x)^{\alpha}]}$$

By substituting $\lambda = \beta$ into the above formula, we derive:

$$H(x, \alpha, \beta) = \frac{\alpha\beta(\beta x)^{\alpha-1}}{[1+(\beta x)^{\alpha}]} = \frac{\alpha\beta(\beta x)^{\alpha-1}}{[1+(\beta x)^{\alpha}]}, \tag{8}$$

This expression matches the hazard rate function of a two-parameter log-logistic distribution, which is characterized by the shape parameter α and scale parameter β. From the simplified hazard rate function, $H(x, \alpha, \beta)$, it can be analyzed that for $0 < \alpha \leq 1$, the hazard rate is monotonically decreasing as x increases. This implies that the likelihood of failure or occurrence decreases over time. Conversely, for $\alpha > 1$, the hazard rate function is unimodal, meaning it initially increases, reaches a peak, and then decreases. The peak or maximum hazard rate occurs at $x = \frac{1}{\beta}(\alpha - 1)^{\frac{1}{\alpha}}$, indicating the most likely point of occurrence before the rate begins to decline. Thus, the generalized log-logistic distribution reduces to the classic log-logistic form with two parameters, providing insight into the behavior of the hazard rate under different values of α

Standard Log-Logistic Distribution

Consider a random variable X that follows a generalized log-logistic distribution, denoted as $X \sim GLL(\alpha, \beta, \lambda)$. Suppose that the parameters β and λ are such that $\beta = \lambda = 1$. Under this specific condition, the hazard rate function of the generalized log-logistic (GLL) distribution simplifies to that of the standard log-logistic distribution.

By substituting $\beta = \lambda = 1$ into Equation (7), (Hazard function of GLL), the hazard rate function reduces to:

$$H = \frac{\alpha \times 1 \times (1 \times x)^{\alpha-1}}{[1+(1 \times x)^{\alpha}]} = \frac{\alpha(x)^{\alpha-1}}{[1+(x)^{\alpha}]}, \tag{9}$$

This expression, given by Equation (9), is recognized as the hazard rate function of the standard log-logistic distribution with only one parameter, α, which serves as the shape parameter. The variable $x > 0$, defines the support of the distribution, meaning that the function is only valid for positive values of x. It is crucial to observe the behavior of the hazard rate function for different values of α. When $0 < \alpha \leq 1$, the hazard rate function is monotonically decreasing, which implies that the probability of the event decreases as x increases. In contrast, when $\alpha > 1$, the hazard rate function is unimodal, indicating a non-monotonic pattern where the function initially increases to a peak and then decreases. The maximum hazard rate occurs at the point $x = (\alpha - 1)^{\frac{1}{\alpha}}$), representing the most likely time of occurrence before the hazard rate begins to decline.

This simplification highlights that under the constraints $\beta = \lambda = 1$, the GLL distribution converges to the standard log-logistic distribution form. The behavior of the hazard rate in this setting is determined solely by the shape parameter α, offering a clear understanding of the distribution's dynamics across different values of α.

Burr XII Distribution

Consider a random variable X that follows a generalized log-logistic distribution, denoted by X~GLL (α, β, λ). Suppose that the parameter λ is related to β through the expression $\lambda = \beta\tau^{-\left(\frac{1}{\alpha}\right)}, \tau > 0$, where $\tau > 0$ is a constant. Under this relationship, the hazard rate function of the generalized log-logistic (GLL) distribution transforms into the hazard rate function of the Burr XII distribution.

By substituting expression $\lambda = \beta\tau^{-\left(\frac{1}{\alpha}\right)}$ into the hazard rate function (Equation (7)) we obtain:

$$H(x,\alpha,\beta) = \frac{\alpha\beta(\beta x)^{\alpha-1}}{\left[1+\left(\beta\tau^{-\left(\frac{1}{\alpha}\right)}x\right)^{\alpha}\right]} = \frac{\alpha\beta(\beta x)^{\alpha-1}}{\left[1+\left(\beta\tau^{-\left(\frac{\alpha}{\alpha}\right)}x^{\alpha}\right)\right]} = \frac{\alpha\beta(\beta x)^{\alpha-1}}{[1+x^{\alpha}]} \quad (10)$$

This result is represented by Equation (10), which is the hazard rate function of the two-parameter Burr XII distribution. Here, α is the shape parameter controlling the distribution's skewness and tail behavior, while β is a scale parameter, influencing the spread of the distribution. This derivation demonstrates that, under the specified substitution, the GLL hazard rate simplifies to the Burr XII hazard rate, establishing a direct connection between the two distributions.

Analyzing the behavior of the Burr XII hazard rate function reveals interesting characteristics. When $\alpha \leq 1$, the hazard rate function is monotonically decreasing as x increases. This implies that the likelihood of the event occurring diminishes over time. In contrast, when $\alpha > 1$, the hazard rate function exhibits an "upside-down bathtub" shape. This means the hazard rate initially increases, reaches a peak, and subsequently decreases to zero as x approaches infinity (Adeyinka 2019). The maximum hazard rate occurs at $x = (\alpha - 1)^{\frac{1}{\alpha}}$, indicating the most probable time of failure or event occurrence before the rate starts to decline. This transformation of the GLL hazard rate function into the Burr XII hazard rate function under the specified parameter condition demonstrates the flexibility and generality of the GLL distribution. By adjusting the relationship between λ and β, the GLL distribution can model various hazard rate behaviors, including those of the Burr XII distribution, allowing for robust modeling of diverse datasets.

Weibull Distribution

Consider a random variable X that follows a generalized log-logistic distribution, denoted as $X \sim GLL(\alpha, \beta, \lambda)$. Suppose the parameter λ is such that $\lambda^{\alpha} \to 0$. Under this condition, the hazard rate function of the generalized log-logistic (GLL) distribution converges to the hazard rate function of the Weibull distribution.

The hazard rate function of the GLL distribution is expressed as:

$$H(x,\alpha,\beta,\lambda) = \frac{\alpha\beta(\beta x)^{\alpha-1}}{[1+(\lambda x)^{\alpha}]},$$

When $\lambda^{\alpha} \to 0$, the term $(\lambda x)^{\alpha}$ becomes negligible, simplifying the hazard rate function to:

$$H(x,\alpha,\beta) = \frac{\alpha\beta(\beta x)^{\alpha-1}}{1+0} = \alpha\beta(\beta x)^{\alpha-1}, \quad (11)$$

The expression, shown in Equation (11), is precisely the hazard rate function of a Weibull distribution with shape parameter α and scale parameter β. This result illustrates a key property of the GLL distribution, highlighting its ability to approximate the behavior of the Weibull distribution when λ^{α} is very close to zero. The GLL distribution becomes particularly effective in modeling hazard rates that are monotonically increasing when $\alpha > 1$ and τ $(where, \tau = \lambda^{1/\alpha})$ is close to zero, indicating a very small value of λ.

The behavior of the Weibull hazard rate function is determined by the value of the shape parameter α. $0 < \alpha < 1$, the hazard rate function decreases as x increases, indicating a diminishing likelihood of an event or failure over time, which is typical of situations where the risk decreases with age or usage. For $\alpha > 1$, the hazard rate function is monotonically increasing, suggesting that the likelihood of the event or failure becomes higher as x increases, which is common in aging processes where the risk increases over time. When $\alpha = 1$, the hazard rate function remains constant, corresponding to the exponential distribution, which represents a memoryless process where the probability of occurrence is the same regardless of the elapsed time, implying that past events do not affect future risks. Thus, the ability of the GLL distribution to converge to the Weibull distribution under certain parameter constraints, such as $\lambda^{\alpha} \rightarrow 0$, provides it with flexibility in handling different types of hazard rate behaviors, whether decreasing, constant, or increasing. This adaptability makes it a powerful tool in analyzing different datasets (Lima and Cordeiro 2017).

Exponential Distribution

Starting with the hazard rate function of the Weibull distribution given by Equation (11),

$H(x, \alpha, \beta) = \alpha\beta(\beta x)^{\alpha-1}$,

we observe that when the shape parameter α is set to 1, this function simplifies to the hazard rate function of the exponential distribution. By substituting $\alpha = 1$ into the Weibull hazard rate equation, the expression becomes:

$H(t, \beta) = \beta \times 1 \times (\beta \times t)^{1-1} = \beta$

Thus, the hazard rate function reduces to:

$$H(t, \beta) = \beta \quad (12)$$

Equation (12) represents the hazard rate function of the exponential distribution, characterized by a constant hazard rate β. This indicates that the probability of an event occurring remains constant over time, independent of any previous occurrences. The exponential distribution is commonly used to model the time between independent events that happen at a constant rate, such as the time until a machine fails or the time between customer arrivals in a queuing system (Georgopoulos and Seinfeld 1982). However, due to its inherent assumption of a constant hazard rate, the exponential distribution is not suitable for modeling CO_2 emissions data. CO_2 emissions are influenced by various dynamic factors such as economic growth, energy consumption, regulatory policies, technological advancements, and seasonal variations which result in non-constant and often complex patterns over time. The exponential distribution's simplicity does not account for trends, cycles, or other temporal dependencies that are typical in CO_2 emissions data.

Table 1 illustrates how the Generalized Log-Logistic (GLL) distribution can be reduced to its various sub-models under specific parameter conditions.

Table 1: Sub-Models Derived from the Generalized Log-Logistic Distribution

Probability Distributions	$\boldsymbol{\alpha}$	$\boldsymbol{\gamma}$	$\boldsymbol{\beta}$
Burr XII Distribution	α	$\lambda = \beta\tau^{-\left(\frac{1}{\alpha}\right)}, \tau > 0$	$\lambda = \beta\tau^{-\left(\frac{1}{\alpha}\right)}, \tau > 0$
Weibull Distribution	$\lambda^{\alpha} \rightarrow 0$	$\lambda^{\alpha} \rightarrow 0$	β
Standard log-logistic Distribution	α	$\lambda = \beta = 1$	$\beta = \lambda = 1$
Log-logistic Distribution	α	$\lambda = \beta$	$\beta = \lambda$
Exponential Distribution	$\alpha = 1$	$\lambda \rightarrow 0$	β

Statistical Properties of GLL Distribution

This section details important statistical and mathematical attributes of the GLL distribution, including its moments, moment generating function, mode, quantile function, median, skewness, and kurtosis. It also provides insights into how these properties influence the distribution's behavior and characteristics.

Quantiles Function

The quantile function, serving as the inverse of the cumulative distribution function (CDF), is a fundamental tool in statistical and quantitative data analysis CDF (Midhu et al. (2013)). It provides a way to understand the distribution of data by specifying the value below which a given percentage of observations fall. Probability distributions can be described either through the quantile function or the. Quartiles are specific percentiles that divide the data into four equal segments, helping to summarize the distribution. Specifically, the first quartile (Q1) represents the 25th percentile, indicating that 25% of the data falls below this value. The second quartile (Q2), also known as the median, is the 50th percentile, marking the midpoint of the data where half of the values are below and half are above. The third quartile (Q3) corresponds to the 75th percentile, showing that 75% of the data lies below this value. These quartiles are crucial for understanding the spread and central tendency of the data, providing information into the distribution's variability and skewness.

Theorem 1: For a random variable X that follows a Generalized Log-Logistic (GLL) distribution with parameters $\alpha, \beta\ and\ \lambda$, the quantile function can be used to compute various quantiles of the distribution. Specifically, the quantile function for the lower quantile, median, and upper quantile are given by Equations (14), (15), and (16), respectively. These equations provide the values below which a specified proportion of observations fall, enabling precise determination of the distribution's lower and upper bounds, as well as its central tendency.

$$X_q = F^{-1}(q, \alpha, \beta, \lambda) = \frac{\left[\left(\frac{1}{1-p}\right)^{\left(\frac{\lambda}{\beta}\right)^{\alpha}} - 1\right]^{\left(\frac{1}{\alpha}\right)}}{\lambda} \tag{13}$$

$$X_{q1} = \frac{\left[\left(\frac{4}{3}\right)^{\left(\frac{\lambda}{\beta}\right)^{\alpha}} - 1\right]^{\left(\frac{1}{\alpha}\right)}}{\lambda} \tag{14}$$

$$X_{q2} = \frac{\left[(2)^{\left(\frac{\lambda}{\beta}\right)^{\alpha}} - 1\right]^{\left(\frac{1}{\alpha}\right)}}{\lambda} \tag{15}$$

$$X_{q3} = \frac{\left[(4)^{\left(\frac{\lambda}{\beta}\right)^{\alpha}} - 1\right]^{\left(\frac{1}{\alpha}\right)}}{\lambda} \tag{16}$$

To derive the quantile function of the Generalized Log-Logistic (GLL) distribution, we start with the cumulative distribution function (CDF):

$$F(x) = 1 - \{1 + ((\lambda x)^{\alpha})\}^{-\left(\frac{\beta}{\lambda}\right)^{\alpha}}$$

We set $F(x) = p$, where $p \in [0,1]$ to find the quantile $X_q = F^{-1}(p, \alpha, \beta, \lambda)$:

$$1 - \{1 + ((\lambda x)^{\alpha})\}^{-\left(\frac{\beta}{\lambda}\right)^{\alpha}} = p$$

Isolate the exponential term:

$$\{1 + ((\lambda x)^{\alpha})\}^{-\left(\frac{\beta}{\lambda}\right)^{\alpha}} = 1 - p$$

Take the reciprocal and adjust the exponent:

$$\{1 + ((\lambda x)^{\alpha})\}^{\left(\frac{\beta}{\lambda}\right)^{\alpha}} = \frac{1}{1-p}$$

Raise both sides to the power $\left(\frac{\beta}{\lambda}\right)^{\alpha}$:

$$1 + (\lambda x)^{\alpha} = \left(\frac{1}{1-p}\right)^{\left(\frac{\beta}{\lambda}\right)^{\alpha}}$$

Subtract 1 from both sides:

$$(\lambda x)^{\alpha} = \left(\frac{1}{1-p}\right)^{\left(\frac{\beta}{\lambda}\right)^{\alpha}} - 1$$

Take the $\alpha - th$ root:

$$\lambda x = \left[\left(\frac{1}{1-p}\right)^{\left(\frac{\beta}{\lambda}\right)^{\alpha}} - 1\right]^{\frac{1}{\alpha}}$$

Solve for x:

$$x = \frac{1}{\lambda}\left[\left(\frac{1}{1-p}\right)^{\left(\frac{\beta}{\lambda}\right)^{\alpha}} - 1\right]^{\frac{1}{\alpha}}$$

Thus, the quantile function is:

$$F^{-1}(p,\alpha,\beta,\lambda) = \left[\left(\frac{1}{1-p}\right)^{\left(\frac{\beta}{\lambda}\right)^{\alpha}} - 1\right]^{\frac{1}{\alpha}}, p \in [0,1], \alpha > 0, \beta > 0, \lambda > 0$$

Similarly, Equations (14)– (16) can be derived by applying the following values: the lower quartile is set to 1/4, the median to 2/4 =1/2, and the upper quartile to 3/4.

Skewness and Kurtosis

Skewness and kurtosis are important statistical metrics that characterize the shape and symmetry of probability distributions. Skewness indicates the degree and direction of asymmetry relative to a perfectly symmetrical distribution. Positive skewness points to a longer or fatter tail on the right side (right-skewed or positively skewed), while negative skewness indicates a longer or fatter tail on the left side (left-skewed or negatively skewed). Kurtosis measures the extent of tails in a distribution. Positive kurtosis (leptokurtic) signifies heavier tails and more extreme values compared to a normal distribution, whereas negative kurtosis (platykurtic) reflects lighter tails with fewer extreme values than a normal distribution (mesokurtic). These metrics are crucial for understanding the underlying data characteristics and selecting appropriate statistical models.

The following Equations (17) define the Galton Skewness and Moors Kurtosis for the Generalized Log-Logistic (GLL) model with three parameters. The following relationship defines the mathematical form of the Galton Skewness and Moors Kurtosis of the GLL model with three parameters:

$$\left.\begin{aligned} S_K &= \frac{Q(3/4)+Q(1/4)-2Q1/2}{Q(3/4)-Q(1/4)}, \\ K_M &= \frac{Q(7/8)+Q(3/8)-Q(5/8)-Q(1/8)}{Q(3/4)-Q(1/4)} \end{aligned}\right\} \qquad (17)$$

where Q represents different quantile values. These expressions can be derived as functions of the GLL quantile function. The benefit of these measures lies in their reduced sensitivity to outliers and their applicability even when the distribution lacks finite moments

The rth moments

The rth moment of a random variable is a measure that captures the distribution of the variable's values relative to a particular point. It is determined by raising each value of the random variable to the power r and then calculating their average. Moments are essential in statistical analysis as they offer valuable information about the distribution's shape and dispersion. Key moment functions, such as the rth moment, rth central moment, mean, variance, skewness, and kurtosis, are critical for understanding the proposed distribution's properties and behavior.

Theorem 2:

For a random variable X following the Generalized Log-Logistic (GLL) distribution with parameters α, β and λ, the rth power moments, negative moments, and logarithmic moments are given by specific functions of the distribution parameters:

$$E(X^r) = \frac{\beta^{\alpha}}{\lambda^{\alpha+r}} \times \frac{\Gamma\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}\right)-\left(\frac{r}{\alpha}\right)\Gamma\left(\left(\frac{r}{\alpha}\right)+1\right)}{\Gamma\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}+1\right)}, \text{ for } \frac{\alpha\beta^{\alpha}}{\lambda^{\alpha}} > r \qquad (18)$$

$$E(X^{-r}) = \frac{\tau^{\alpha+r}}{\beta^r} \times \frac{\Gamma\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}\right)+1}{\Gamma\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}-\left(\frac{r}{\alpha}\right)\right)\Gamma\left(\left(\frac{r}{\alpha}\right)+1\right)}, \tag{19}$$

Equation (19) can be demonstrated as follows:

$$E(X^r) = \int_0^\infty x^r (x;\beta,\alpha,\lambda)dx = \int_0^\infty x^r \frac{\alpha\beta(\beta x)^{\alpha-1}}{[1+(\lambda x)^{\alpha}]^{\left(\frac{\beta^\alpha}{\lambda^\alpha}\right)+1}} dx = \frac{\alpha\beta}{\Gamma\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}+1\right)} \int_0^\infty x^r \frac{(\beta x)^{\alpha-1}}{1+(\lambda x)^{\alpha}} dx$$

$$= \frac{\beta^\alpha}{\lambda^{\alpha+r}} \times \frac{\Gamma\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}\right)-\left(\frac{r}{\alpha}\right)\Gamma\left(\left(\frac{r}{\alpha}\right)+1\right)}{\Gamma\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}+1\right)}, \text{ for } \frac{\alpha\beta^\alpha}{\lambda^\alpha} > r. \tag{20}$$

Using the moment-generating function, we can determine the mean and variance of the random variable X with a Generalized Log-Logistic (GLL) distribution having parameters α, β and λ as follows:

To obtain the mean, we substitute $r = 1$ into Equation (18) to get Equation (21)

$$E(X) = \mu = \frac{\beta^\alpha}{\lambda^\alpha} \times \frac{\Gamma\left(\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}\right)-(1/\alpha)\Gamma(1/\alpha)+1\right)}{\Gamma\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}+1\right)}, \alpha\beta^\alpha/\lambda^\alpha > 1. \tag{21}$$

The variance of the random variable $X \sim GLL(\alpha,\beta,\lambda)$ is given by:

$$V(X) = \sigma^2 = E(X^2) - \left(E(X)\right)^2$$

$$= \frac{\beta^\alpha}{\lambda^{\alpha+2}} \times \frac{\Gamma\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}\right)-\left(\frac{2}{\alpha}\right)\Gamma\left(\left(\frac{2}{\alpha}\right)+1\right)}{\Gamma\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}+1\right)} - \left(\frac{\beta^\alpha}{\lambda^\alpha} \times \frac{\Gamma\left(\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}\right)-(1/\alpha)\Gamma(1/\alpha)+1\right)}{\Gamma\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}+1\right)}\right)^2, \alpha\beta^\alpha/\lambda^\alpha > 2. \tag{22}$$

The r^{th} Central Moments

Central moments are a set of statistical measures used to assess the symmetry and dispersion of the data in order to characterize the form and properties of a probability distribution. Central moments are computed using deviations from the distribution's mean as opposed to raw moments, which are based on departures from a predetermined point. The first, second and r^{th} central moments of the GLL distribution presented as follows;

First central moment is the same as the average (mean). Thus,

$$c_1 = \mu_1' = E(X) = \frac{\beta^\alpha}{\lambda^\alpha} \times \frac{\Gamma\left(\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}\right)-(1/\alpha)\Gamma(1/\alpha)+1\right)}{\Gamma\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}+1\right)},$$

The second central moment, which measures the dispersion of a distribution, is equivalent to the variance. It quantifies the extent to which the values of a random variable deviate from the mean.

$$c_2 = E(X^2) - \left(E(X)\right)^2 = \frac{\beta^\alpha}{\lambda^{\alpha+2}} \times \frac{\Gamma\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}\right)-\left(\frac{2}{\alpha}\right)\Gamma\left(\left(\frac{2}{\alpha}\right)+1\right)}{\Gamma\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}+1\right)} - \left(\frac{\beta^\alpha}{\lambda^\alpha} \times \frac{\Gamma\left(\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}\right)-(1/\alpha)\Gamma(1/\alpha)+1\right)}{\Gamma\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}+1\right)}\right)^2$$

$$c_r = \mu_r' - \sum_{n=1}^{r-1}\binom{r-1}{n-1} c_n \mu_{r-n}' = \frac{\beta^\alpha}{\lambda^{\alpha+r}} \times \frac{\Gamma\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}\right)-\left(\frac{r}{\alpha}\right)\Gamma\left(\left(\frac{r}{\alpha}\right)+1\right)}{\Gamma\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}+1\right)} -$$

$$-\sum_{n=1}^{r-1}\binom{r-1}{n-1} c_n \frac{\beta^\alpha}{\lambda^{\alpha+(r-n)}} \times \frac{\Gamma\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}\right)-\left(\frac{r-n}{\alpha}\right)\Gamma\left(\left(\frac{r-n}{\alpha}\right)+1\right)}{\Gamma\left(\left(\left(\frac{\beta}{\lambda}\right)^{\alpha}\right)+1\right)} \tag{23}$$

Consequently, we can determine the generalized log logistic distribution's skewness and kurtosis as follows;

$$Skewness = \frac{c_3}{(\sigma^2)^{3/2}},$$

$$Kurtosis = \frac{C_4}{(\sigma^2)^2} \tag{24}$$

Estimation of the GLL Distribution Parameters

The parameters of the Generalized Log-Logistic (GLL) distribution (α, β and λ) can be estimated using various techniques, including percentile-based methods, weighted least squares (WLS), maximum likelihood estimation (MLE), and ordinary least squares (OLS). The literature has extensively explored and compared these estimation techniques for different probability distributions. Key references providing in-depth analyses include works by Kundu and Raqab 2005, Alkasasbeh and Raqab 2009, Mazucheli et al. 2013, do Espirito Santo and Mazucheli 2015, and Dey et al. 2016. This study focuses on a detailed examination of the MLE method, which serves as a benchmark for other estimation techniques, as well as the weighted least squares (WLS) and ordinary least squares (OLS) methods.

Maximum likelihood estimation Method

The maximum likelihood estimation (MLE) method is widely favored in research due to its desirable properties, such as consistency, asymptotic efficiency, and invariance (Dey et al. 2016). Consistency ensures that as the sample size increases, the estimates converge to the true parameter values. Asymptotic efficiency guarantees that MLE provides the minimum variance among unbiased estimators in large samples. The invariance property means that if a parameter transformation is applied, the MLE of the transformed parameters can be derived from the MLE of the original parameters. For a random sample of size $X_1, X_2, X_3, \ldots, X_n$, drawn from a Generalized Log-Logistic (GLL) distribution, the MLE for the parameters is obtained through a specific estimation process.

$L = \prod_{i=1}^{n} f(x_i, \alpha, \beta, \lambda)$,

$$L(x;\alpha,\beta,\lambda) = \prod_{i=1}^{n} \frac{\alpha\beta(\beta x_i)^{\alpha-1}}{[1+(\lambda x_i)^{\alpha}]^{((\lambda)^{\alpha})+1}} \tag{25}$$

Let $\ell = \text{InL}$

$$\ell = \text{nIn}(\alpha\beta) + (\alpha-1)\sum_{i=1}^{n} In(\beta x_i) - \sum_{i=1}^{n} In[1+(\lambda x_i)^{\alpha}] - \left({}^{\beta}/_{\lambda}\right)\sum_{i-=1}^{n} In(1+(\lambda x_i)^{\alpha}).$$

(27)

$$\frac{\partial\ell}{\partial\alpha} = 0 \Rightarrow \frac{n}{\alpha} + \sum_{i=1}^{n} \ln(\beta x_i) - \sum_{i=1}^{n}\left\{\frac{(\gamma x_i)^{\alpha}\ln(\lambda x_i)}{1+(\lambda x_i)^{\alpha}}\right\} - \left({}^{\beta}/_{\lambda}\right)\sum_{i=1}^{n}\left\{\frac{(\gamma x_i)^{\alpha}\ln(\lambda x_i)}{1+(\lambda x_i)^{\alpha}}\right\}, \tag{26}$$

$$\frac{\partial\ell}{\partial\beta} = 0 \Rightarrow \frac{n}{\beta} + n\beta(\alpha-1) - {}^{1}/_{\lambda}\sum_{i=1}^{n} \ln(1+\lambda x_i), \tag{27}$$

$$\frac{\partial\ell}{\partial\lambda} = 0 \Rightarrow -\sum_{i=1}^{n}\left\{\frac{(\gamma x_i)^{\alpha}\ln(\gamma x_i)}{1+(\lambda x_i)^{\alpha}}\right\} - \left({}^{\beta}/_{\lambda^2}\right)\sum_{i=1}^{n}\left\{\frac{(\gamma x_i)^{\alpha}\ln(\lambda x_i)}{1+(\lambda x_i)^{\alpha}}\right\}. \tag{28}$$

The maximum likelihood estimates (MLEs) for parameters α, β, and λ can be obtained by equating the derivative of the likelihood function to zero and solving the resulting set of nonlinear equations numerically. Given the complexity involved in the expected information matrix, the observed information matrix, denoted as J(θ), is typically used instead to derive confidence intervals for the model parameters. The observed information matrix is expressed as follows:

$$\mathrm{J}(\theta) = -\begin{bmatrix} \frac{\partial^2\ell}{\partial^2\alpha} & \frac{\partial^2\ell}{\partial\alpha\,\partial\beta} & \frac{\partial^2\ell}{\partial\alpha\,\partial\lambda} \\ . & \frac{\partial^2\ell}{\partial^2\beta} & \frac{\partial^2\ell}{\partial\beta\,\partial\,\partial} \\ . & . & \frac{\partial^2\ell}{\partial^2\lambda} \end{bmatrix}, \tag{29}$$

where: $\theta = (\alpha, \beta, \lambda)'$. Under standard regularity conditions, when the parameters are within the interior of the parameter space (excluding the boundary), the scaled difference $\sqrt{n}(\cong \theta - \theta)$, where θ is the true parameter vector and ≅θ is its estimator, converges in distribution to a multivariate normal distribution, denoted as $N_3\left(0, I^{-1}(\theta)\right)$.

The expected Fisher information matrix, I(θ), provides a measure of the information that an observable random variable carries about an unknown parameter. However, due to practical difficulties in computing I(θ), it is often replaced by the observed information matrix, J(θ), which is derived from the second derivatives of the log-likelihood function with respect to the parameters. Despite this substitution, the asymptotic normality of the distribution remains valid. Consequently, the asymptotic distribution $N_3\left(0, J^{-1}(\theta)\right)$ can be applied to construct two-sided confidence intervals for the model parameters at a chosen confidence level of 100(1 - τ) %, where τ is the significance level (Mazucheli et al. 2013). This approach is commonly employed in statistical inference to provide robust interval estimates for parameter uncertainty in complex models.

Standard and Weighted Least Square Estimation

Weighted Least Squares Estimation (WLSE) is a statistical method designed to handle heteroscedasticity, or unequal error variability, when estimating parameters of a statistical model (Kundu and Raqab 2005). Unlike Standard Least Squares Estimation (LSE), which minimizes the sum of squared differences between predicted and observed values by treating all observations equally, WLSE assigns different weights to observations to reflect varying levels of accuracy or reliability associated with each data point (Swain et al. 1988). By minimizing the weighted sum of squared differences, WLSE gives greater importance to more precise data points, making it particularly useful in situations where error variability differs significantly across different parts of the dataset.

For the generalized log-logistic (GLL) distribution, the WLSE involves minimizing the weighted squared differences between the model's predicted values and the actual observations. Consider a random variable x with a sample size of n denoted as $(x_1, x_2, \dots, x_n)$, drawn from a GLL distribution. The WLSE for this distribution can be formulated by assigning appropriate weights to the squared differences and solving for the model parameters that minimize this weighted sum.

$$f(x_i; \alpha, \beta, \lambda) = {}^{\alpha}/_{\beta}\left(1 + \left(\frac{x_i - \lambda}{\beta}\right)^{\alpha+1}\right)^{-1}$$

The WLSE seeks to minimize the weighted sum of squared differences, which is represented by:

$$S = WLSE(\alpha, \beta, \lambda) = \sum_{i=1}^{n} w_i \left(x_i - f(x_i; \alpha, \beta, \lambda)\right)^2 \tag{30}$$

Where w_i represents the significance weight assigned to each observation in the dataset. To determine the Weighted Least Squares Estimates (WLSE) for the parameters (α, β and λ), the partial derivatives of Equation (32) with respect to each of these unknown parameters must be calculated. After taking these derivatives, the resulting system of equations is set to zero, and then solved to obtain the WLSE for the parameters.

$$\frac{\partial S}{\partial \alpha} = 0 \Rightarrow -2\sum_{i=1}^{n} w_i\left(x_i - f(x_i; \alpha, \beta, \lambda)\right) {}^{1}/_{\beta}\left(1 + \left(\frac{x_i - \lambda}{\beta}\right)^{\alpha+1}\right)^{-2} log\left(\frac{x_i - \lambda}{\beta}\right) \tag{31}$$

$$\frac{\partial S}{\partial \beta} = 0 \Rightarrow 2\sum_{i=1}^{n} w_i\left(x_i - f(x_i; \alpha, \beta, \lambda)\right) {}^{\alpha}/_{\beta^2}\left(1 + \left(\frac{x_i - \lambda}{\beta}\right)^{\alpha+1}\right)^{-2}\left(\alpha + 1 - \left(\frac{x_i - \lambda}{\beta}\right)^{\alpha+1}\right) \tag{32}$$

$$\frac{\partial S}{\partial \lambda} = 0 \Rightarrow 2\sum_{i=1}^{n} w_i\left(x_i - f(x_i; \alpha, \beta, \lambda)\right) {}^{\alpha(\alpha + 1)}/_{\beta}\left(1 + \left(\frac{x_i - \lambda}{\beta}\right)^{\alpha+1}\right)^{-2}\left(\frac{x_i - \lambda}{\beta}\right)^{\alpha} \tag{33}$$

Given the complexity involved in solving Equations 33-35, numerical optimization techniques are commonly employed to solve the system of equations formed by setting these derivatives to zero. Some of the widely used optimization algorithms include Newton's method, gradient descent, and quasi-Newton methods. In this context, the objective function is the weighted sum of squared differences, and these algorithms iteratively adjust the parameter values to minimize this function.

Numerical Results and Discussion

This section presents the results of analyzing CO_2 emissions in Tanzania using the Generalized Log-Logistic (GLL) distribution. The GLL distribution is applied to model and interpret statistical patterns in CO_2 emissions and is compared with the log-normal distribution. Additionally, the analysis assesses the GLL distribution alongside its two-parameter sub-models, including the Burr XII, log-logistic, and Weibull distributions, to evaluate its robustness in capturing emission trends. Descriptive statistics, model validity and selection criteria, the likelihood ratio test, and Monte Carlo simulations for comparing estimation method results are also presented in this section.

Descriptive Statistics

The results in Table 2 reveals a mean emission value of 4.35 megatons, indicating the average level of CO_2 emissions over the observed period. The median value of 2.43 megatons suggests that half of the observations fall below this level, reflecting a right-skewed distribution, as further confirmed by the positive skewness value of 1.55. The mode of 2.063 megatons shows the most frequently occurring emission level. The high variance (15.10) and standard deviation (3.89) indicate significant variability in emissions, with values ranging from a minimum of 0.703 to a maximum of 15.57 megatons. The positive kurtosis (1.29) implies a leptokurtic distribution, where there are more extreme values compared to a normal distribution. These findings suggest that while there are periods of low emissions, there are also instances of significantly higher emissions, which could indicate periods of increased industrial activity or other pollution sources. This variability and trend towards higher emission values have serious environmental implications, as rising CO_2 levels are directly linked to climate change, increased greenhouse gas effects, and broader impacts on ecosystems and human health.

Table 2: **Summary Statistic of the Carbon dioxide Emissions**

Mean	Median	Mode	Variance	Skewness	Kurtosis	Minimum	Maximum
4.35	2.43	2.063	15.10	1.55	1.29	0.703	15.57

Table 3: Moments, Standard Deviation, Skewness, and Kurtosis for some GLL parameter values

			(β, α, λ)			
Moments	(0.2,0.2,0.2)	(0.5,1.0,1.0)	2.0,2.5,3.0)	(1.5,2.5,2.0)	(1.5,1,5, 2.5)	(4.5, 5.0,2.5)
μ_1'	0.2151	0.3011	0.3503	0.4932	0.3717	0.4501
μ_2'	0.1456	0.2751	0.3217	0.4021	0.3145	0.2018
μ_3'	0.1034	0.2473	0.3005	0.3671	0.2671	0.1726
μ_4'	0.1006	0.2016	0.2715	0.3237	0.2025	0.1501
μ_5'	0.0345	0.1673	0.2201	0.3025	0.1972	0.1027
SD	0.3218	0.3788	0.3921	0.4521	0.6191	0.5281
Skewness	1.0861	0.8755	0.8023	0.7602	0.8160	-0.2671
Kurtosis	4.7832	3.5671	3.3217	3.1230	4.2910	2.8913

The Table 3 presents statistical properties of a Generalized Log-Logistic Distribution (GLL) for different parameter sets. For the first set of parameters (0.2, 0.2, 0.2), the distribution exhibits a strongly positively skewed shape (skewness of 1.0861), indicating a longer tail on the right side, and a kurtosis of 4.7832 suggests heavy tails and potential outliers. The second set (0.5, 1.0, 1.0) shows a somewhat lower skewness (0.8755) and kurtosis (3.5671), still indicating a right-skewed distribution. The third set (2.0, 2.5, 3.0) exhibits further reductions in skewness and kurtosis, indicating a more symmetric and lighter-tailed distribution. The fourth set (1.5, 2.5, 2.0) shows a similar trend toward less skewness and kurtosis. Finally, the fifth set (1.5, 1.5, 2.5) displays a moderate degree of skewness (0.8160) but lower kurtosis (4.2910), indicating a somewhat heaviertailed distribution. These moments, standard deviations, skewness, and kurtosis statistics provide valuable insights into the shape and behavior of the GLL distribution under various parameter conditions.

Model Validity and Selection Criterion

To assess the suitability of the GLL model for the utilized datasets, several metrics were evaluated including the negative log-likelihood value, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Kolmogorov-Smirnov (K-S) distance, and p-value. The AIC and BIC were calculated using the equations:

The AIC and BIC are defined in Equations (36) and (37) respectively;

$$AIC = 2q - 2l \quad (36)$$

$$BIC = qIn(n) - 2l \quad (37)$$

where l represents the log-likelihood function estimated via Maximum Likelihood Estimation (MLE), n denotes the total number of observations, and q indicates the number of model parameters.

In addition to these criteria, we considered the negative log-likelihood value, which indicates model fit with lower values suggesting better fit, and the K-S distance, which measures the deviation between the empirical and theoretical distributions smaller values are preferable. The p-value assesses the significance of the goodness-of-fit measures, with higher values suggesting a better fit to the data. Detailed results, including these metrics, negative log-likelihood values obtained through MLE, and their standard errors (in parentheses), are presented in Table 4. These metrics provide a comprehensive evaluation of the model's performance and its suitability for the datasets.

Table 4: **Information Criterion, MLE estimates, log-likelihood and Goodness of fit**

Model	Estimates (SE)	AIC	BIC	K-S(p-value)	$-l$
GLL (α, β, λ)	α = 0.871(0.021) β = 0.652(0.056) λ =0.341(0.012)	607.081	614.321	0.032 (0.762)	467.326
Burr-XII (α, β)	α =1.342(0.231) β =0.562(0.038)	615.620	623.083	0.058 (0.742)	470.003
Weibull (α, β)	α =0.651(0.012) β =4.831(0.604)	637.042	641.003	0.046 (0.745)	472.302
LN (μ, σ)	μ =-0.321(0.281) σ =2.092(0.002)	609.027	617.033	0.072 (0.759)	469.002
EXP $(\alpha, \beta\tau)$	α =0.521(0.012) β =0.722(0.010) τ =1.621(0.183)	623.208	637.087	0.062 (0.533)	502.203
LL (α, β)	α =0.431(0.051) β =1.016(0.075)	665.091	671.091	0.064 (0.509)	481.017
SLL (α)	α =0.217(0.027)	626.071	638.081	0.075 (0.677)	478.087

In comparing the Generalized Log Logistic Distribution (GLL) with six other distributions; Burr II, Weibull, Log-Normal (LN), Exponential (EXP), Log-Logistic (LL), and Standard Log-Logistic (SLL) the results in Table 4 reveal that GLL achieves the lowest AIC and BIC among the distributions evaluated. This suggests that the GLL model fits the dataset better than the other distributions, indicating its potential effectiveness in accurately representing the underlying data distribution. These findings highlight GLL's capability to capture the data's statistical characteristics more precisely, which is critical in model selection and data analysis. Additional analysis of model performance metrics could further elucidate the most suitable distribution for the dataset.

The parameters of the Generalized Log Logistic Distribution (GLL) have been estimated using three different methods: Maximum Likelihood Estimation (MLE), Least Squares Estimation (LSE), and Weighted Least Squares Estimation (WLSE). As shown in Table 5, the MLE method demonstrates superior performance compared to the other methods, particularly due to its high p-value. This indicates that MLE provides parameter estimates that align most closely with the data distribution. Consequently, MLE emerges as a robust and reliable method for estimating GLL parameters, highlighting its efficacy in capturing the dataset's statistical properties.

Table 5: The Goodness of Fit Statistics for GLL Parameter Estimated under Various Methods

Estimation Method	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	l	$K-S$	$P-value$
MLE	0.871	0.652	0.341	-467.326	0.051	0.762
LSE	1.023	0.8621	0.674	-470.981	0.078	0.678
WLSE	1.008	0.879	0.832	-475.894	0.082	0.476

Likelihood Ratio Test

The Likelihood Ratio Test (LRT) is a statistical method utilized to compare the goodness of fit between two nested models in hypothesis testing. It determines whether adding extra parameters in a more detailed model significantly enhances the fit compared to a simpler model (Kundu and Raqab 2005). In this study, the LRT was employed to evaluate the goodness of fit of the GLL distribution relative to its five sub-models. The primary goal was to test the following hypotheses:

(a) $\mathbf{H_{01}}$: $\lambda^{\alpha} \to 0$; (The sample data follows a Weibull distribution, indicating that the shape parameter α approaches zero.) vs. $\mathbf{H_{a1}}$: $\lambda^{\alpha} not \to 0$; (The sample data follows a GLL distribution, suggesting that the parameter λ does not tend to zero).

(b) $\mathbf{H_{02}}$: $\lambda = \beta = 1$; (The sample data follows a standard log-logistic distribution, where the scale parameter λ and shape parameter β are both equal to one.) vs. $\mathbf{H_{a2}}$: $\lambda \neq \beta, \beta \neq 1$; (The sample data follows a GLL distribution, indicating that the parameters λ and β are not equal, and β is not equal to one).

(c) $\mathbf{H_{03}}$: $\lambda = 0\ and \alpha = 1$; (The sample data follows an exponential distribution, characterized by λ being zero and the shape parameter α being one.) vs. $\mathbf{H_{a3}}$: $\lambda \neq 0\ and \alpha \neq 1$; (The sample data follows a GLL distribution, where λ is not zero and is not one).

(d) $\boldsymbol{H_{04}}: \lambda = \beta$; (The sample data follows a log-logistic distribution, where the scale parameter λ equals the shape parameter β) Vs $H_{a4}: \lambda \neq \beta$ (The sample data follows a GLL distribution, indicating that λ and β are not equal).

(e) $\boldsymbol{H_{05}}: \lambda = \beta\tau^{-(1/\alpha)}, \tau > 0$; (The sample data follows a Burr XII distribution, where, $\lambda = \beta\tau^{-(1/\alpha)}$, with τ being positive) Vs $\boldsymbol{H_{a5}}: \beta\tau^{-(1/\alpha)}, \tau \leq 0$ (The sample data follows a GLL distribution, where τ is less than or equal to zero, altering the relationship between λ and β).

The test statistic, known as the likelihood ratio (LR), is determined by comparing the maximum likelihood estimates of two models: one based on the null hypothesis (H_0) and the other based on the alternative hypothesis (H_a). The likelihood ratio is calculated as follows:

$LR = -2[[loglikelihood\ under\ H0\ - loglikelihood\ under\ HA]$ OR

$$LR = -2\ln\left(\frac{L(\widehat{\omega}^{*};x)}{l(\widehat{\omega};x)}\right) \qquad (38)$$

The likelihood ratio (LR) test statistic in Equation (38) is based on the comparison of the maximum likelihood estimates obtained under the null hypothesis (H_0) with those obtained under the alternative hypothesis (H_a). Specifically, $\widehat{\omega}^{*}$ denotes the restricted maximum likelihood estimates under the null hypothesis, while($\widehat{\omega}$) represents the unrestricted estimates under the alternative hypothesis. Assuming the models are appropriately specified, the test statistic follows a chi-squared (χ^2) distribution with degrees of freedom equal to the difference in the number of parameters between the two models (i.e., df = number of parameters in H_A - number of parameters in H_0). The null hypothesis is rejected if the p-value is less than the predetermined significance level.

The results presented in Table 6 indicate that the p-values for all five sub-models of the Generalized Log Logistic Distribution (GLL) fall below the significance threshold. This strongly suggests that the dataset aligns well with the GLL, supporting its effectiveness in characterizing the underlying data distribution. These findings highlight the robustness of the GLL model in explaining the observed CO_2 emissions data.

Table 6: Likelihood Ratio Test Results

Sub-models Distribution	Hypotheses	LRT	P-value	Decision
Weibull	$H_{01}: \lambda^{\alpha} \to 0; Vs$ $H_{a1}: \lambda^{\alpha} not \to 0$	7.987	0.0071	H_{01} is rejected
Standard log logistic	$H_{02}: \lambda = \beta = 1; Vs$ $H_{a2}: \lambda \neq \beta, \beta \neq 1$	9.986	0.0021	H_{02} is rejected
Exponential	$H_{03}: \lambda = 0\ and \alpha = 1; Vs$ $H_{a3}: \lambda \neq 0\ and \alpha \neq 1$	34.785	0.0030	H_{03} is rejected
Log logistic	$H_{04}: \lambda = \beta; Vs$ $H_{a4}: \lambda \neq \beta$	25.892	0.0031	H_{04} is rejected
Burr XII	$H_{05}: \beta\tau^{-(1/\alpha)}, \tau > 0; Vs$ $H_{a5}: \beta\tau^{-(1/\alpha)}, \tau \leq 0$	11.891	0.0042	H_{05} is rejected

Monte Carlo Simulation for Comparing Estimation Methods

A Monte Carlo simulation study was conducted to evaluate the performance of various parameter estimation methods for the Generalized Log Logistic Distribution (GLL). The study assessed the effectiveness of different estimators using metrics such as mean average bias and mean squared error (MSE). Simulations were performed across various sample sizes and parameter values, with the process repeated 1,000 times for each combination of sample sizes (n = 25, 50, 100, 500, 1,000) and parameter sets $(\alpha, \beta, \lambda) = (1.5, 0.5, 2.0)\ and\ (1.0, 0.3, 1.5)$.The estimated parameters for the GLL model, obtained through Maximum Likelihood Estimation (MLE), Least Squares Estimation (LSE), and Weighted Least Squares Estimation (WLSE), are detailed in Table 7.

The simulation results revealed a notable trend: as sample size increased, both the average bias and MSE consistently decreased. Among the estimation methods, MLE stood out with superior performance, exhibiting significantly lower MSE compared to the other methods. Additionally, the average bias of the estimates improved with larger sample sizes, reflecting greater accuracy in estimation. As the sample size increase, MLE estimates converged closely to the true parameter values, underscoring their reliability. Consequently, MLE estimates and asymptotic results can be confidently used to construct reliable confidence intervals for model parameters, even in cases with smaller sample sizes.

Table 7: Monte Carlo Simulation Results for Comparing Various GLL Estimation Methods

		I			II		
Parameters	n	**MLE**	**LSE**	**WLSE**	**MLE**	**LSE**	**WSLE**
α	25	2.006 (0.062)	3.098 (0.142)	2.871 (0.212)	2.210 (0.071)	3.812 (0.178)	2.876 (0.216)
	50	1.971 (0.057)	3.005 (0.137)	2.408 (0.175)	1.791 (0.066)	3.084 (0.124)	2.564 (0.203)
	100	1.863 (0.024)	2.891 (0.085)	2.091 (0.143)	1.508 (0.042)	2.781 (0.116)	1.761 (0.184)
	500	1.686 (0.021)	2.536 (0.065)	1.783 (0.125)	1.293 (0.035)	2.007 (0.108)	1.603 (0.159)
	1000	1.502 (0.016)	2.093 (0.046)	1.706 (0.107)	1.006 (0.021)	1.651 (0.092)	1.452 (0.125)
β	25	1.082 (0.054)	2.087 (0.452)	2.451 (0.571)	0.722 (0.045)	1.819 (0.517)	1.605 (1.175)
	50	0.975 (0.051)	1.975 (0.405)	2.091 (0.432)	0.593 (0.036)	1.561 (0.453)	1.106 (0.836)
	100	0.787 (0.035)	1.762 (0.276)	1.672 (0.331)	0.462 (0.031)	1.086 (0.326)	0.784 (0.657)
	500	0.672 (0.021)	1.207 (0.182)	1.577 (0.264)	0.405 (0.026)	0.879 (0.295)	0.592 (0.451)
	1000	0.503 (0.017)	1.036 (0.103)	1.328 (0.205)	0.302 (0.020)	0.698 (0.232)	0.482 (0.354)
λ	25	3.781 (0.082)	4.891 (0.381)	3.983 (0.372)	1.976 (0.033)	2.654 (0.176)	2.781 (0.783)
	50	3.500 (0.063)	3.872 (0.327)	3.562 (0.304)	1.721 (0.030)	2.235 (0.152)	2.567 (0.562)
	100	2.686 (0.041)	3.103 (0.203)	3.005 (0.276)	1.635 (0.021)	2.067 (0.132)	1.971 (0.376)
	500	2.409 (0.034)	2.872 (0.142)	2.651 (0.242)	1.589 (0.018)	1.982 (0.108)	1.784 (0.324)
	1000	2.005 (0.015)	2.784 (0.113)	2.506 (0.193)	1.506 (0.007)	1.819 (0.095)	1.676 (0.291)

Conclusion

The Generalized Log-Logistic (GLL) model demonstrates its suitability for modeling CO_2 emissions data in Tanzania from 1961 to 2022. The GLL model effectively captures the patterns and statistical characteristics of emissions data, showing a better fit than its sub-models (Weibull, Standard log logistic, exponential, Log logistic and Burr XII) and log-normal distribution. Three estimation methods were used in this study: maximum likelihood, least squares, and weighted least squares. Among these, maximum likelihood proved to be the most reliable, demonstrating significantly lower mean squared error (MSE), making it particularly advantageous for larger and small datasets. Additionally, the likelihood ratio test (LRT) was used to evaluate the goodness of fit of the GLL model relative to its five sub-models. The LRT helps determine whether the GLL model provides a significantly better fit to the data compared to its sub-models, ensuring the model's robustness and its ability to accurately represent the underlying CO_2 emissions data. The results of the LRT further supported the effectiveness of the GLL model in characterizing the distribution of emissions.

These findings establish the GLL model as a powerful tool for accurately tracking and analyzing CO_2 emission patterns in Tanzania, which is essential for understanding and managing rising emission levels. The model enhances understanding of emission characteristics, helping policymakers evaluate the scale and sources of emissions. This can inform strategies for transitioning to cleaner energy and mitigating environmental impacts. Further research could expand the use of the GLL model to other pollutants and regions, increasing its applicability in various environmental settings.

Conflict of Interest

Author declares that no conflicts of interest.

Acknowledgements

The author appreciates the support from the Department of Statistics, University of Dar es Salaam.

References

Adeyinka FS 2019 On transmuted four parameters generalized log-logistic distribution. Int. J. *Stat. Distrib. Appl.* 5(2): 32

Aldahlan MA 2020 Alpha power transformed log-logistic distribution with application to breaking stress data. *Adv. Math. Phys.* 2020: Article ID 2193787

Alkasasbeh MR and Raqab MZ 2009 Estimation of the generalized logistic distribution parameters: Comparative study. *Stat. Methodol.* 6(3): 262–279

Aryal GR 2013 Transmuted log-logistic distribution. *J. Stat. Appl. Probab*. 2(1): 11–20

Berger A, Melice J L, and Demuth C 1982 Statistical distributions of daily and high atmospheric SO_2 concentrations. *Atmos. Environ*. 16(12): 2863–2877

Dey S, Dey T, Ali S, and Mulekar MS 2016 Two-parameter Maxwell distribution: Properties and different methods of estimation. *J. Stat. Theory Pract.* 10: 291–310

do Espirito Santo APJ and Mazucheli J 2015 Comparison of estimation methods for the Marshall–Olkin extended Lindley distribution. *J. Stat. Comput. Simul*. 85(17): 3437–3450

Georgopoulos P G and Seinfeld J H 1982 Statistical distributions of air pollution concentrations. *Environ. Sci. Technol.* 16(7): 401A–415

Global Carbon Budget, 2023. *Available at: https://www.globalcarbonproject.org/carbonbudget*

Hafner M, Tagliapietra S, Falchetta G, Occhiali G, Hafner M, Tagliapietra S, and Occhiali G 2019 East Africa: Regional Energy Outlook. *Renew. Energy Access Sustain. Dev. East Afr*. 9–17

Jakeman AJ and Taylor JA 1985 A hybrid ATDL-gamma distribution model for predicting urban area source acid gas concentrations. *Atmos. Environ*. 19: 1959–1967

Kundu D and Raqab MZ 2005 Generalized Rayleigh distribution: different methods of estimations. *Comput. Stat. Data Anal*. 49(1): 187–200

Larsen RI 1974 An air quality data analysis system for interrelating effects, standards, and needed source reductions-Part 2. *J. Air Pollut. Control Assoc*. 24(6): 551–558

Lima SR and Cordeiro GM 2017 The extended log-logistic distribution: properties and application. *Anais Acad. Bras. Ciências* 89(1): 3–17

Mage DT and Ott WR 1978 Refinements of the log-normal probability model for analysis of aerometric data. *J. Air Pollut. Control Assoc.* 28(8): 796–798

Malik AS and Ahmad SP 2020 An extension of log-logistic distribution for analyzing survival data. Pak. *J. Stat. Oper. Res.* 16(4): 789–801

Mazucheli J, Louzada F, and Ghitany ME 2013 Comparison of estimation methods for the parameters of the weighted Lindley distribution. *Appl. Math. Comput*. 220: 463–471

Midhu NN, Sankaran PG, and Nair NU 2013 A class of distributions with the linear mean residual quantile function and its generalizations. *Stat. Methodol.* 15: 1–24

Newman MC, Dixon PM, Looney BB, and Pinder JE 1989 Estimating mean and variance for environmental samples with below detection limit observations. *Water Resour. Bull.* 25(4): 906–916

Singh KP 1989 A generalized log logistic regression model for survival analysis. Hazard rate characteristics. Bio-Prax 29: 49–62

Singh KP, Bartolucci AA, and Bae S 2001 Mathematical modeling of environmental data. *Math. Comput. Model.* 33(6-7): 793–800

Singh KP, Bartolucci AA, and Burgard SL 1994 Two-step for survival data. *Biomet. Praximet.* 34: 1–12

Swain JJ, Venkatraman S, and Wilson JR 1988 Least-squares estimation of distribution functions in Johnson's translation system. *J. Stat. Comput. Simul.* 29(4): 271–297