

ORIGINAL RESEARCH OPEN ACCESS

Ordinary Differential Equation Modeling and a Novel Zubair-Fréchet Distribution for Cholera in Ghana With Control Strategies

Kofi Afriyie Nyamekye¹  | Eric Sampane-Donkor² | Joseph Quarshie¹ | Franklin Asiedu-Bekoe³ | Azuma Abdul-Tawab³ | Justice Moses K. Aheto⁴ 

¹University of Professional Studies, Accra, Ghana | ²University of Ghana Medical School, Accra, Ghana | ³Public Health Division, Accra, Ghana | ⁴University of Ghana, Accra, Ghana

Correspondence: Kofi Afriyie Nyamekye (kofi.anyamekye@upsamail.edu.gh)

Received: 10 April 2025 | **Revised:** 21 January 2026 | **Accepted:** 8 April 2026

Funding: Fogarty International Centre of the National Institutes of Health, Grant/Award Number: UE5TW012566

Keywords: Cholera | Fréchet Distribution | Ordinary differential Equations | SIR model | Zubair - Fréchet (ZF) Distribution

ABSTRACT

Background and Aims: In this paper, we present a valuable and novel interdisciplinary study that applies a classic epidemiological model, Susceptible-Infected-Recovered (SIR) to capture the mean disease dynamics, with a novel Zubair Fréchet (ZF) distribution to model the extreme, heavy-tailed nature of cholera outbreak sizes. Furthermore, the proposed model is linked to both public health and actuarial risk. We demonstrate that the ZF provides a superior fit to cholera case data compared to traditional distributions.

Methods: Data was collected from both primary sources (such as the Ghana Health Service) and secondary sources. The SIR model employs differential equations in modeling. The results of the simulation study showed that as the sample size was increased, the standard error and bias created by the novel Zubair-Fréchet distribution's parameters decreased while depicting a cholera outbreak. The paper finds that the Zubair-Fréchet distribution is more robust compared to some existing distributions, such as the Fréchet and Weibull distributions. Additionally, a number of actuarial properties were determined, and the maximum likelihood technique was used to estimate model parameters.

Conclusion: Findings from the study show that parameter values of infectious contact rate ($\beta = 0.004$) and recovery rate ($\gamma = 0.999$) make the control strategy effective. ZF fits better than the Fréchet/Weibull distribution, as evidenced from the goodness-of-fit table. Furthermore, vaccination reduces transmission evidenced from the data showing less to no cases after vaccination. The results from this simulation could help guide the Ministry of Health and the Ghana Health Service's response to be able to deploy the necessary logistics, update preparedness, and response plans for the country as the rainy season begins. Industry practitioners may apply this family of distributions to microbiology, public health, and targeted interventions to reduce case fatality rates. Additionally, professionals in the insurance industry can apply our model in insurance value-at-risk and tail value-at-risk for fair premium pricing.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2026 The Author(s). *Health Science Reports* published by Wiley Periodicals LLC.

1 | Introduction

1.1 | Background

Statistical distributions are fundamental functions in terms of probabilities of events for modeling a phenomenon of random variables in a host of applications in microbiology, actuarial science, risk and insurance, survival analysis, among many others. In many different domains, a large number of generalized families of continuous distributions have been proposed, investigated, and used to model life-time data sets [1]. However, there are numerous more real-world scenarios where the existing distributions fall short of offering an appropriate fit and flexibility [2]. An example of such data is data on foodborne-related diseases. This calls for an urgent need for new statistical models and distributions with greater versatility and capability to be able to accurately characterize available data, especially in foodborne-related diseases [3].

Mathematical modeling for infectious diseases dates back to Bernoulli in the 18th century [4]. Several mathematical models of foodborne-related diseases have been recently proposed in the literature in response to such outbreaks in developing countries [5]. Many infectious diseases are as a direct result of poor hygienic conditions and contact between an infectious person and a susceptible individual. Cholera is an acute intestinal infectious disease caused by the bacterium *Vibrio cholerae*, characterized by extreme diarrhea and vomiting. It is an infectious disease where an outbreak with no proper intervention and control might lead to death. This is because of the relatively short incubation period of the disease, which usually takes 2 to 5 h. Most cases of cholera currently occur in developing countries in Africa. Cholera was first reported in Africa in 1836 when outbreaks did occur on the Indian Ocean coast, killing as many as 20,000 people in Zanzibar [6]. Within barely half a century, cholera has now been rated an endemic disease, changing from an imported disease in most African countries.

According to the European Center for Disease Prevention and Control (2024), Ghana has a rising number of reported cases of cholera from year 2023 to 2024. In comparison, from 1 January 2023 to 23 October 2023, there were no cases reported. From 31 August 2024 to 18 October 2024, 78 new cases, including one (1) death have been reported. In the same vein, since 1 January 2024 to 18 October 2024; 102 cases, including one (1) death has been reported [7].

The classical approach to disease modeling is to use a type of mathematical model known as a compartmental model in complex ways. This approach is to assign each individual in the population to one of several compartments, based on their infection status; again, define the rates at which individuals move between compartments as their status begins to update. Finally, use this model to define differential equations that can predict the cause of the epidemic.

Using the epidemic modeling approaches, we shall consider the SI model in this study, which is the simplest model used to model infectious diseases in a given population of interest. Thereafter, we shall consider the SIR model, which is an

extension of the SI model. It is also a simple model used to model infectious diseases in a given population of interest. It breaks the population into three groups: susceptible, infected, and recovered. It is a population-based model in the sense that each compartment models the behavior of an average individual in that compartment. Of course, it is an oversimplification of the real-life scenario, but it is actually possible to build more complex models based on this. It was proposed to explain the rapid rise and fall in the number of infected patients observed in epidemics such as the plague and cholera. Diseases represented by this model are diseases that people obtain a lifetime immunity. The SIR model can be extended to other models (e.g., SEIR, SIRS, etc.) depending on the need. Example, for SIRS model, we assume that recovery implies temporary acquired immunity. There is a return of the removed/recovered individuals into the susceptible compartment or group. The underlying assumptions for the SIR models are that firstly, there is an existence of a homogenous population, where people in the same compartment are subject to the same hazards. Moreover, all susceptible individuals have the same risk of getting infected, which implies that there exists a well-mixed population. Lastly, those who recovered from the disease are immune forever [8].

In this paper, it is the authors aim to give a better understanding of the effects of control measures coupled with the transmission dynamics of cholera. This will, in diverse ways, help gain practicable and efficient preventive strategies for the control of the cholera pandemic in Ghana. The paper responds to the need for a generalized distribution to model epidemiology data sets like cholera. Hence, developing a compound distribution, i.e., the Zubair-Fréchet distribution. Furthermore, the paper develops maximum likelihood estimators for the parameters of the Zubair-Fréchet distribution and assesses the performance of its parameters using simulations. Additionally, the paper presents the actuarial properties of the Zubair-Fréchet and the goodness-of-fit for distribution within the food infections data sets. This is important because flexibility in distribution models ensures better conclusions and decisions to be made on specific problem cases [9].

The SIR model was used to frame the disease dynamics, bereft of the likelihood of a specific outcome. Meanwhile, the ZF distribution provides a superior fit to cholera case data compared to traditional distributions, a necessary framework to account for the inherent randomness and uncertainty in extreme, heavy-tailed cholera spread, which makes the predictions more robust and reflective of the stochastic nature of cholera. Thus, making them complementary tools for a more comprehensive and realistic study of epidemics. These allow for quantifying the probability of different outbreak scenarios, which is essential for public health planning and resource allocation.

1.2 | Research Problem

In real-life data analysis, a number of practitioners work with the Fréchet class of distribution. However, the distribution lacks

flexibility since it only has one shape parameter. The aforementioned limitation and also the frequent application of the Fréchet distribution give rise to the research question:

1. How statistically robust does the Fréchet distribution deal with heavy tail distributions with a single shape parameter?
2. Can the single shape parameter perform the functions of kurtosis and skewness efficiently at the same time?

To overcome this drawback, there is a need to introduce an additional scale parameter. This addresses the problem of the lack of scale and shape parameters in the distribution. The Zubair-Fréchet distribution hence has the ability to introduce both scale and shape parameters to an existing Fréchet statistical distribution to make it more flexible.

Nasiru [9] posits that to model data with monotonic and non-monotonic failure rates, as well as produce heavy tail distributions that will control both skewness and kurtosis at the same time, a minimum of two shape/scale parameters are needed. As a result, an extra shape parameter needs to be introduced to make the Fréchet distribution have a more desirable property. Hence, an extension of the Fréchet distribution into what we propose as the Zubair-Fréchet distribution is demonstrated in this paper. This research, therefore, seeks to develop an extended form for the Fréchet distribution by adding an extra parameter to the existing family of distributions. Periodically occurring cholera outbreaks in Ghana put a burden on healthcare systems, reduce economic output, and lead to significant death tolls. The proposed Z-F distribution not only contributes to the body of work on distributions but can also save lives and money when applied to real-world data sets.

1.3 | Objectives

The main objective of this paper is to develop an appropriate model for Ghana's cholera cases. The specific objectives are to:

1. compute the parameters for Ghana cholera SIR model,
2. develop the novel Zubair-Fréchet Distribution for cholera cases,
3. derive the Actuarial Properties of the Zubair-Fréchet Distribution,
4. compare the goodness-of-fit to model Cholera cases.

1.4 | Justification for Selecting ZF Distribution

Introducing additional parameters will generalize the existing distributions. With that, the new distribution will have flexibility and a closed form of cumulative distributive functions, survival functions, as well as hazard rate functions. Moreover, this distribution would add only one parameter instead of adding two or more, which will complicate the existing distributions for its analysis. In the same vein, provision is made for better fits other than modified models. Lastly, it proffers a best

fit to heavy-tailed microbiology data sets for effectively modeling shortfalls [10].

The rest of the paper is organized as follows: literature review, methodology, discussion, conclusion, and recommendations.

2 | Theoretical Literature Review

2.1 | Application of Actuarial Modeling

Actuarial modeling for infectious diseases dates back to Daniel Bernoulli in the 18th century [4]. Several mathematical models of foodborne-related diseases have recently been proposed in the literature in response to outbreaks in developing countries. Recent literature from the year 2024 employs ordinary, partial, fractional and stochastic differential equations to model cholera epidemiology [11]. These models attempt to estimate the dynamics of foodborne transmission and their impact of interventions, with the aim of providing some guidance to policy-makers with decisions such as vaccination, antibiotics, and provision of clean and hygienic water, especially for cholera cases [12]. These resulting compound distributions would be helpful for modeling real-world data sets in the fields of epidemiology, actuarial science, insurance, and other closely related disciplines, as well as distribution theory in statistics [13]. The resulting compound distributions might also be helpful in representing actual events, where the failure rate is not constant due to factors including risk and force of mortality. Ahmad et al. [14] found that the Z family of distributions has the capacity to suitably model data sets that have fat tails for a well-defined risk management system.

This section presents a review of the literature on some probability distribution generators. These generators have been used over the years to modify existing distributions with the purpose of making them flexible in modeling real-life data [15]. Further to this review, the shortcomings of such generators that control skewness, kurtosis and variability are discussed [16]. The section ends with the introduction of the need for the Zubair-Fréchet distribution, which contains both scale parameter and shape parameters [17].

2.2 | Zubair Generated Family of Distributions

Ahmad [10] developed the Zubair-G family of distributions. This generator only has a scale parameter; no shape parameter is present. As a result, since the Zubair family lacks a shape parameter, it is inappropriate to combine it with any distribution that lacks a shape generator as well [18].

2.3 | Empirical Literature Review

According to empirical research, cholera epidemics in Ghana are seasonal, usually reaching their peak during the rainy season owing to sanitation challenges [19]. According to the Ghana Health Service (GHS), flooding caused by heavy rainfall contaminates drinking water sources and raises the risk of cholera in many regions. Seasonal rainfall and outbreak increases are

strongly correlated, according to an empirical study by Addy et al. that examined cholera cases over a ten-year period. Unusually high rainfall levels during the most severe outbreaks, such as those in 2011, 2012, and 2014, exacerbated the unhygienic conditions in urban areas, especially in slums with a high population density.

In many places, cholera is still endemic, especially in cities like Accra and Kumasi, where infrastructure cannot keep up with population expansion. Apenteng et al. [20] examined case data from 2000 to 2018 and found patterns in recurring outbreaks that typically follow a similar geographic and seasonal pattern, indicating a recurring problem with public health preparation and urban management. The prediction of epidemics has long been the focus of attention in epidemiology and mathematical biology. Previous studies have used a variety of approaches to model influenza cases. Goldstein et al. [21] proposed a generalized linear model demonstrating how routine virologic and influenza-like illness (ILI) surveillance data can be used to quantify the dynamics of co-circulating influenza strains and generate short-term predictions of the relative epidemic sizes of each strain during the course of an influenza season in the United States [22]. Soebiyanto et al. [23] employed an Autoregressive Integrated Moving Average (ARIMA) model along with climatic parameters to forecast the next influenza season. Mugglin et al. [24] used a Bayesian hierarchical approach for modeling of influenza epidemic dynamics in both time and space. Kim et al. [25] unexpectedly used Twitter, a free social networking service, to improve the accuracy of forecasting models by providing early warnings of influenza outbreaks.

3 | Materials and Methods

The authors firstly propose a mathematical model which uses differential equations based on the epidemiological compartment modeling with an intervention or control strategy such as vaccination. The computer software package that is used to solve the differential equation model numerically is deSolve in R. Secondly, the Zubair-Fréchet model is also proposed.

3.1 | Representation of Sir Model

The SIR model is represented in threefold. That is by verbal description, graphical description and mathematical description. For a verbal description, we consider S susceptible, I infectious, and R recovered. For a mathematical description using the differential equations formulation is as follows:

$$\frac{dS}{dt} = -\beta IS \quad (1)$$

$$\frac{dI}{dt} = (\beta IS) - (\gamma I) \quad (2)$$

$$\frac{dR}{dt} = \gamma I \quad (3)$$

From the above Equations (1) to (2), the basic reproductive ratio RO is presented in Equation (4) below:

$$RO = \frac{\beta N}{\gamma} \quad (4)$$

Where, $N = S + I + R$ represents the total population, with RO estimating the number of people who will be infected by the average infectious individuals. If RO is less than one (1) it implies an outbreak is likely to be contained, whereas, where is greater than one (1), then an outbreak is likely to be epidemic. Equation (5) below presents the effective reproduction number, where the population of susceptible varies with time [26].

$$eff = RO \left(\frac{S}{N} \right) \quad (5)$$

4 | Graphical Representation of the Sir Model

The Figure 1 above shows a demonstration of the SIR Model as captured above.

4.1 | Comparing a Model's Predictions With Data

Primary data from the Ghana Public Health Service for cholera cases in 2024 will be compared to the model prediction for selected parameters. We plot the points of the cholera data set and visually compare the model's predictions and the data points. If the model does not fit the observed data well, we then train the model on the data by simply changing the gamma and beta parameters. We may start by first changing the beta parameter.

4.2 | The Zubair Fréchet Distribution: Theory and Actuarial Properties

4.2.1 | Generating of Skewed Distribution Techniques

The skewed normal family of distributions was created by Azzalini [27] by combining two separate symmetric distributions, building on the concepts of O'Hagan and Leonard [28]. Azzalini [27] in the existing literature, various distributions based on the skewed family have been created and investigated. By using inverse scaling of the pdf on both sides of the mode, Fernandez and Steel [29], for instance, established and studied a technique for introducing skewness into any given continuous distribution [30]. This makes it possible to create adjustable distribution shapes using just one shape parameter [31]. The epsilon-skew normal family of distributions was also created by Mudholkar and Hutson [32] with an additional parameter that allows it to regulate the degree of skewness. Based on the idea of an inverse probability integral transformation, Azzalini [27] also developed a broad framework that assisted in introducing skewness into symmetric distributions. Azzalini's methodology was recognized as a



FIGURE 1 | Graphical representation of SIR Model. Source: Authors' design for a graphical representation of SIR Model.

special case of Chang and Genton's [33] weighted approach for creating skewed distributions. By creating a larger family of skewed distributions that incorporated the skew normal family [27] and the epsilon-skew normal [32], significantly outperformed the majority of skewed distributions in terms of flexibility.

The second objective of this paper is to develop “Zubair-Fréchet” distribution. The purpose of this section is to create a new type of distribution known as the Zubair-Fréchet that may be used to improve the efficiency of standard distributions by adding skewness to them [34]. The analysis in this research was carried out using R software.

4.2.2 | The Fréchet Distribution

Let a random variable X have the Fréchet Distribution with parameter θ, β where $\theta > 0$ and $\beta > 0$ whose cdf and pdf are defined by $G(x)$ and $g(x)$, respectively. Following Maurice René Fréchet [35], the cdf and pdf of the Fréchet Distribution are given respectively as:

$$G(x) = e^{-\left(\frac{\beta}{x}\right)^\theta} \quad (6)$$

and

$$g(x) = \theta\beta^\theta x^{-(\theta+1)} e^{-\left(\frac{\beta}{x}\right)^\theta} \quad (7)$$

4.2.3 | The Zubair-G Family

Let X denote a non-negative random variable whose cdf and pdf are expressed as $T(x)$ and $t(x)$ respectively, following Ahmad et al. [14], the cdf and pdf are given respectively as:

$$T(x; \phi) = \frac{e^{\alpha G(x; \phi)^2 - 1}}{e^\alpha - 1} \quad \alpha, \phi > 0, x \in \mathfrak{R} \quad (8)$$

$$t(x; \phi) = \frac{2\alpha g(x; \phi) G(x; \phi) e^{\alpha G(x; \phi)^2}}{e^\alpha - 1} \quad (9)$$

4.2.4 | The Zubair-Fréchet (ZF) Distribution

To achieve objective two (2) of this paper, the Zubair-Fréchet (ZF) distribution is developed in this section.

To derive the expressions for the cdf of the Zubair- Fréchet, insert equation (6) as the cdf of the parent distribution into the expression in equation (8), following the distribution we have proposed, the resulting compound distribution gives the cdf of the Zubair- Fréchet distribution as:

$$T(x) = \frac{e^{\alpha \left[e^{-\left(\frac{\beta}{x}\right)^\theta} \right]^2} - 1}{e^\alpha - 1}, \alpha, \beta, \theta > 0, x \in \mathfrak{R}, \quad (10)$$

Now, to derive the pdf of the Zubair-Fréchet distribution, we differentiate the cdf of the distribution. Again, we could also

insert the expression in equation (7) into equation (9), the resulting compound distribution gives the pdf of the Zubair-Fréchet distribution as:

$$t(x) = \frac{2\alpha\theta\beta^\theta x^{-(\theta+1)} e^{-2\left(\frac{\beta}{x}\right)^\theta} e^{\alpha \left[e^{-\left(\frac{\beta}{x}\right)^\theta} \right]^2}}{e^\alpha - 1} \quad (11)$$

Where $\alpha, \beta, \theta > 0, x \in \mathfrak{R}$ and $\phi = (\phi_k) = (\phi_1, \phi_2, \dots)$ is a parameter vector. Hence $T(x; \phi) = T(x)$ and $t(x; \phi) = t(x)$

Proof of ZF distribution in the discrete case

Firstly,

$$\begin{aligned} T(x) &= \frac{e^{\alpha \left[e^{-\left(\frac{\beta}{x}\right)^\theta} \right]^2} - 1}{e^\alpha - 1} \\ \text{When; } x < 0, T(x) &= 0 \\ \text{When; } 0 \leq x < \infty \\ 0 \leq \left[e^{-\left(\frac{\beta}{x}\right)^\theta} \right]^2 &< 1 \\ e^{\alpha \left[e^{-\left(\frac{\beta}{x}\right)^\theta} \right]^2} &\leq e^\alpha \\ 0 \leq T(x) &\leq 1 \end{aligned}$$

For all values of x .

Secondly,

$$\begin{aligned} x &\leq y \\ T(y) - T(x) &= \frac{e^{\alpha \left[e^{-\left(\frac{\beta}{y}\right)^\theta} \right]^2} - 1}{e^\alpha - 1} - \frac{e^{\alpha \left[e^{-\left(\frac{\beta}{x}\right)^\theta} \right]^2} - 1}{e^\alpha - 1} \\ &= \frac{e^{\alpha \left[e^{-\left(\frac{\beta}{y}\right)^\theta} \right]^2}}{e^\alpha - 1} - \frac{e^{\alpha \left[e^{-\left(\frac{\beta}{x}\right)^\theta} \right]^2}}{e^\alpha - 1} \\ &= \frac{e^{\alpha \left[e^{-\left(\frac{\beta}{y}\right)^\theta} \right]^2}}{e^\alpha - 1} \leq \frac{e^{\alpha \left[e^{-\left(\frac{\beta}{x}\right)^\theta} \right]^2}}{e^\alpha - 1} \\ x \leq y, T(y) - T(x) &\geq 0 \\ T(x) &\leq T(y) \end{aligned}$$

Thirdly,

$$\begin{aligned} \lim_{x \rightarrow -\infty} T(x) &= \lim_{x \rightarrow 0} T(x) = \frac{e^0 - 1}{e^\alpha - 1} = 0 \\ \lim_{x \rightarrow \infty} T(x) &= \frac{e^\alpha - 1}{e^\alpha - 1} = 1 \end{aligned}$$

4.3 | Deriving the Various Statistical Properties of the Proposed Zubair-Fréchet

Objective number three (3) was achieved by taking into account statistical properties such as asymptotic behavior, survival function, force of mortality and quantile function [36].

4.3.1 | Survival Function

This defines the probability of an individual surviving beyond a predetermined and specified time in the future. It is the likelihood of a system not failing prior to sometime in the future.

$$S(x) = 1 - F(x) = 1 - \frac{e^{\alpha \left(e^{-\left(\frac{\beta}{x} \right)^{\theta}} \right)^2} - 1}{e^{\alpha} - 1} \quad (12)$$

4.3.2 | Hazard Rate Function

The pdf and survival functions are compounded to give us the hazard rate function.

$$\frac{\left[\frac{2\alpha\theta\beta^{\theta}x^{-(\theta+1)}e^{-2\left(\frac{\beta}{x}\right)^{\theta}}e^{\alpha\left(e^{-2\left(\frac{\beta}{x}\right)^{\theta}}\right)} \right]}{e^{\alpha} - 1} \quad (13)$$

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

4.3.3 | Legitimacy of the Proposed Distribution

Proof: The Zubair Fréchet distribution is valid and legitimate for

$$\int_0^{\infty} t(x)dx = 1$$

$$t(x) = \frac{2\alpha\theta\beta^{\theta}}{e^{\alpha} - 1}x^{-(\theta+1)}e^{-2\left(\frac{\beta}{x}\right)^{\theta}}e^{\alpha\left(e^{-2\left(\frac{\beta}{x}\right)^{\theta}}\right)}$$

$$t(x) = \frac{2\alpha\theta\beta^{\theta}}{e^{\alpha} - 1}\int_0^{\infty}x^{-(\theta+1)}e^{-2\left(\frac{\beta}{x}\right)^{\theta}}e^{\alpha\left(e^{-2\left(\frac{\beta}{x}\right)^{\theta}}\right)}du$$

$$u = e^{-2\left(\frac{\beta}{x}\right)^{\theta}}; \frac{du}{dx} = 2\theta\beta^{\theta}x^{-(\theta+1)}e^{-2\left(\frac{\beta}{x}\right)^{\theta}}$$

$$x = 0, u = 0; x = \infty, u = 1$$

$$\frac{2\alpha\theta\beta^{\theta}}{e^{\alpha} - 1}\int_0^1x^{-(\theta+1)}e^{-2\left(\frac{\beta}{x}\right)^{\theta}}e^{\alpha u} \frac{du}{2\theta\beta^{\theta}x^{-(\theta+1)}e^{-2\left(\frac{\beta}{x}\right)^{\theta}}}$$

$$= \frac{\alpha}{e^{\alpha} - 1}\int_0^1e^{\alpha u}du$$

$$= \frac{\alpha}{e^{\alpha} - 1}\left[\frac{1}{\alpha}e^{\alpha u}\right]_0^1$$

$$= \frac{\alpha}{e^{\alpha} - 1}\left[\frac{1}{\alpha}(e^{\alpha} - 1)\right] = 1$$

4.3.4 | Moment Generating Function

The moment generating function (MGF) helps in finding the moments of a random variable of a given distribution. Generally, the MGF is given as:

$$M_x(t) = E(e^{tx})$$

$$E(x) = \frac{2\alpha\theta\beta^{\theta}}{e^{\alpha} - 1}\int_0^{\infty}x^{-(\theta+1)}e^{-2\left(\frac{\beta}{x}\right)^{\theta}}e^{\alpha\left(e^{-2\left(\frac{\beta}{x}\right)^{\theta}}\right)}du$$

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

$$E(x^n) = \frac{\alpha}{e^{\alpha} - 1}\int_0^1\beta^n\left(-\frac{1}{2}\ln u\right)^{\frac{-n}{\theta}}e^{\alpha u}du$$

Maclaurin;

$$e^{\alpha u} \approx 1 + \alpha u$$

$$E(x^n) \approx \frac{\alpha\beta^n}{e^{\alpha} - 1}\int_0^1\left(-\frac{1}{2}\ln u\right)^{\frac{-n}{\theta}}(1 + \alpha u)du + 0(1)n$$

$$= \frac{\alpha\beta^n}{e^{\alpha} - 1}\left\{2^{\frac{n}{\theta}-1}\left(2^{\frac{n}{\theta}} + 2\right)\Gamma\left(1 - \frac{n}{\theta}\right)\right\} + 0(1); \frac{n}{\theta} < 1$$

$$n = 1;$$

$$E(x) \approx \frac{\alpha\beta}{e^{\alpha} - 1}\left\{2^{\frac{1}{\theta}-1}\left(2^{\frac{1}{\theta}} + 2\right)\Gamma\left(1 - \frac{1}{\theta}\right)\right\} + 0(1); \frac{1}{\theta} < 1$$

$$n = 2$$

$$E(x^2) \approx \frac{\alpha\beta^2}{e^{\alpha} - 1}\left\{2^{\frac{2}{\theta}-1}\left(2^{\frac{2}{\theta}} + 2\right)\Gamma\left(1 - \frac{2}{\theta}\right)\right\} + 0(1); \frac{2}{\theta} < 1$$

4.3.5 | Likelihood Function of Z-F Distribution

θ for a pdf $f(x; \theta)$ is an estimator that maximizes the likelihood function $L(x_1, x_2, \dots, x_n; \theta)$ as a function of θ :

$$L(x_1, x_2, \dots, x_n; \theta) = f(x_1; \theta) \times f(x_2; \theta)$$

$$L\left(x_1, x_2, \dots, x_n; \theta = \prod_{i=1}^n [f(x_i; \theta)]\right)$$

$$L(\alpha, \theta, \beta) = \prod_{i=1}^n \frac{2\alpha\theta\beta^{\theta}}{e^{\alpha} - 1}x^{-(\theta+1)}e^{-2\left(\frac{\beta}{x}\right)^{\theta}}e^{\alpha\left(e^{-2\left(\frac{\beta}{x}\right)^{\theta}}\right)}$$

$$= \frac{2^n\alpha^n\theta^n\beta^{n\theta}}{(e^{\alpha} - 1)^n} \times \prod_{i=1}^n x_i^{-(\theta+1)} \times e^{-2\sum_{i=1}^n \left(\frac{\beta}{x_i}\right)^{\theta}} \times e^{\alpha\sum_{i=1}^n \left[e^{-2\left(\frac{\beta}{x_i}\right)^{\theta}}\right]}$$

4.4 | Methods of Estimation

Objective three is to develop maximum likelihood estimators for parameters of the Zubair- Fréchet distribution. There are several methods of estimation.

4.5 | Monte Carlo Simulation Study

This part provides a thorough simulation study to evaluate the MLEs' behavior. It is simple to imitate the Z-F family by inverting (10) as shown below: If the distribution of U is uniform (0, 1), the nonlinear equation is:

$$\beta \left[\left(\ln \left(\frac{1}{\alpha} \ln(u(e^\alpha - 1) + 1) \right)^{-\frac{1}{2}} \right)^{\frac{1}{\beta}} \right]^{-1} = 0 \quad (14)$$

The purpose of the simulation study is to determine whether:

- The MLEs converge to the true value in every case as the sample size increases,
- Whether there is a small or significant difference between the average estimates and true values, and
- As the sample size grows, the standard errors of the MLEs would decrease. The accuracy of the approximation of the standard error of the MLEs is assessed by the following criteria for evaluating the performance of the estimators.

5 | Epidemiological Results vs Statistical Results

This section provides a detailed analysis of the development of the suggested distribution.

5.1 | Data Source

Primary data were sourced from the Office of the Director of Public Health (Ghana Health Service, Headquarters, Accra). The time series data received was analyzed from 1st October, 2024, to 1st March, 2025. The frequency was on a daily basis from the onset.

The parameters beta and gamma were traced from fitting to visually compare the model's predictions and the data points, where the model's predictions appeared closer to the data points. The parameters beta and gamma, which are respectively the infectious contact rate (i.e., the average number of secondary infections per unit time, as averagely 0.004 means a lower rate of contracting the disease) and the recovery rate; gamma of 0.999 indicates a faster rate of recovery, almost getting to one (1) making it very significant. These parameters are per-day rates that relate positively to the observed dynamics. These immediate containment measures focus on rapid patient treatment and interrupting transmission through water, sanitation, and hygiene (WASH), as well as a control strategy of community engagement and vaccination.

Quantitative measures such as the root mean square error (RMSE) had lower values of 5.689045, 5.68904 and 5.688973; mean square error (MSE): 32.36523, 32.36518, 32.36442 and residual sum of squares (RSS): 2297.931, 2297.928, 2297.874 for selected sample sizes with parameter values $N=10, 1000, 10,000$ and beta = 0.004, theta = 0.999, respectively. A sensitivity analysis is performed where beta and theta values are changed to 0.006 and 0.995, respectively. RMSE, MSE and RSS values are the latter shown (see Table B1 in Appendix B).

5.1.1 | Sir Model

As shown in Figure 2, parameter values of beta = 0.004, gamma = 0.999, showing the respective risk and probability rate

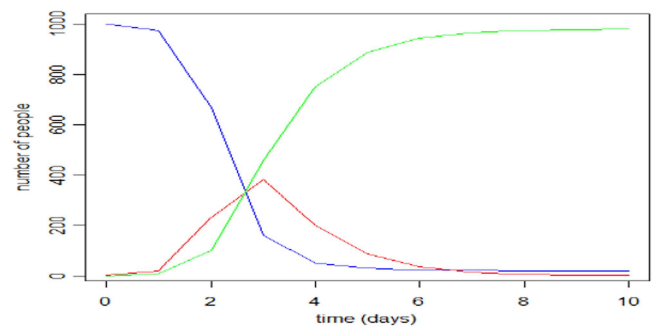


FIGURE 2 | SIR model. Source: Authors' result of graph of SIR model for cholera cases in Ghana, 2024.

TABLE 1 | SIR simulator, ghana cholera cases.

	Time	S	I	R
1	0	999.00	1.00	0.00
2	1	974.27	19.46	6.26
3	2	667.18	231.99	100.81
4	3	159.01	382.00	458.97
5	4	49.56	200.30	750.13
6	5	28.79	85.46	885.73
7	6	22.97	34.83	942.19
8	7	20.96	13.99	965.04
9	8	20.20	5.59	974.19
10	9	19.91	2.23	977.85
11	10	19.79	0.88	979.31

Source: Authors' compilation of SIR Simulator for Ghana cholera cases, 2024.

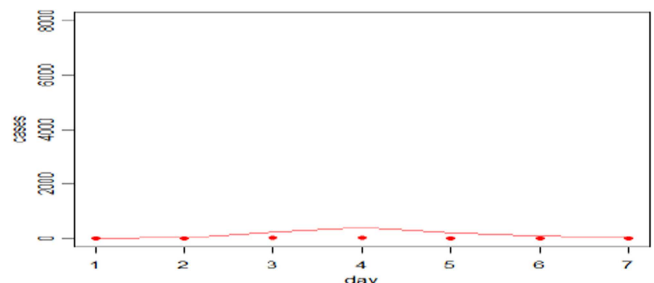


FIGURE 3 | Last week cholera graph, 2024. Source: Authors' results for graph for the last week of cholera cases, 2024 with given parameters.

for the midweek cholera cases from the onset in October, the data set is divided into three parts, with cholera peaks taken into recognition.

As shown in Table 1, the simulated results show susceptible from 999 to 19.79 as well as recovery from 0.00 to 979.31. This gives values close to the population as earlier specified.

As shown in Figure 3, the parameter values of beta = 0.004, gamma = 0.999, showing the respective risk and probability rate for the last week's cholera cases from the onset in October, the data set is divided into three parts, with cholera peaks taken into recognition. The graph shows a weekly trend that starts

from a lower number, rises and drops eventually. The peak hitting around the mid of that week in the year 2024.

As shown in Figure 4, the parameter values of $\beta = 0.004$, $\gamma = 0.999$, showing the respective risk and probability rate for the first week cholera cases from the onset in October, the data set is divided into three parts, with cholera peaks taken into recognition. The graph shows a weekly trend that starts from a lower number, rises and drops eventually. The peak hitting around the mid of that week, but closer to the end of same.

As shown in Figure 5, the parameter values of $\beta = 0.004$, $\gamma = 0.999$, showing the respective risk and probability rate for the mid-week cholera cases from the onset in October, the data set is divided into three parts, with cholera peaks taken into recognition. The graph shows a weekly trend that starts from a lower number, rises and drops eventually. The peak hitting around the mid of that week, but closer to the start of same.

5.2 | Intervention of Vaccination as a Control Strategy

Ultimately, the intention of vaccination has helped to alter the course of the epidemic by reducing the transmission coefficient, β . We see that after the implementation of vaccination from

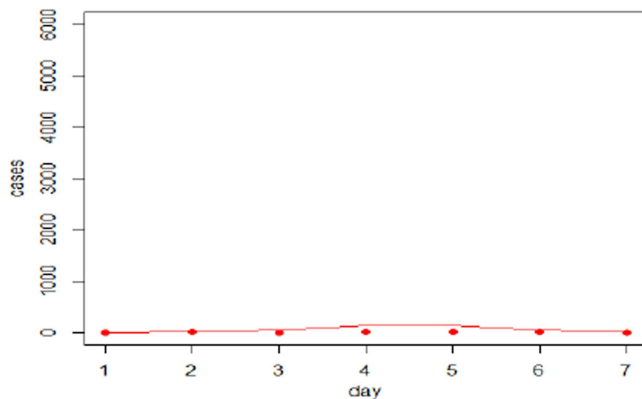


FIGURE 4 | First week cholera graph, 2024. Source: Authors' results of graph of a representation of the first week of cholera cases, 2024 with selected parameters.

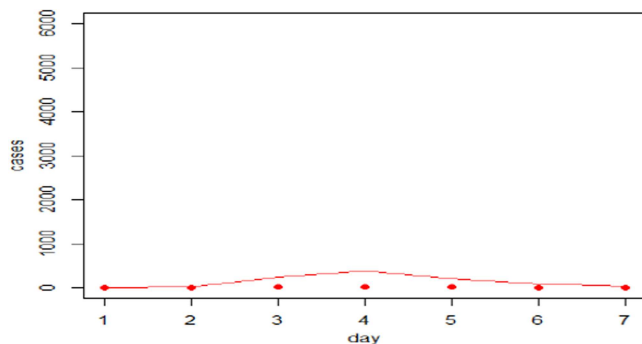


FIGURE 5 | Midweek graph for cholera cases, 2024. Source: Authors' results of graph of the midweek, 2024 of cholera cases with given parameters.

27th November, 2024, there is a huge drop in the transmission and infection rates. The only change is the extra `ifelse()` statement to adjust the value of β to 0.1 during vaccination and β to 0.5 before and after the vaccination. Thus, if before or after the vaccination, give beta value to be 0.5 (i.e., increased transmission rate), but if within the vaccination periods, assign beta value to be 0.1 (i.e., reduced transmission rate). These values are selected as an illustrative case. You need to pass two new parameters to the model. These are the start and end times of the vaccination periods.

Vaccinated individuals acquire immunity. A group of people in a given jurisdiction is said to have herd immunity if enough people get immune, hence curtailing the spread. Individuals then move from the $S(t)$ compartment to the $R(t)$ compartment, where mortality is assumed to be zero (0) in the immunized population.

Let p represent the proportion vaccinated of the total population N and $1 - p$ the proportion unvaccinated, where $0 < p < 1$. Hence, the total population that are vaccinated is $aN(t)p$, which is added to the recovery compartment $R(t)$. The unvaccinated population is $aN(t)(1 - p)$, which will still remain in the susceptible group $S(t)$. Therefore, after vaccination, the equations of the SIR model can be written as:

$$\frac{dS}{dt} = -\beta IS(aN(t)(1 - p)) \quad (15)$$

$$\frac{dI}{dt} = (\beta IS) - (\gamma I) \quad (16)$$

$$\frac{dR}{dt} = \gamma I + aN(t)p \quad (17)$$

5.3 | Immunity Loss of Vaccine

In the real field of epidemiology, no vaccine gives 100% protection to cholera. Hence, the vaccinated individuals are not perfectly immunized. The immunization over time decreases with respect to the vaccination. The population now becomes susceptible to cholera.

Let $e > 0$ represent the rate of immunity loss, then in $R(t)$ compartment, the population in $eR(t)$ reduces as a result of the loss of immunity; then the individuals $eR(t)$ adds to $S(t)$ compartment. The equation of SIR model can be written as:

$$\frac{dS}{dt} = -\beta IS(aN(t)(1 - p)) + eR(t) \quad (18)$$

$$\frac{dI}{dt} = (\beta IS) - (\gamma I) \quad (19)$$

$$\frac{dR}{dt} = \gamma I + aN(t)p - eR(t) \quad (20)$$

5.3.1 | Stochastic or Delay Effects

Further research is encouraged to use mathematical, stochastic or delay effects as modeling approaches applied in the

optimization of vaccination strategies to minimize the impact of the spread of cholera. The first strategy is the randomness in cholera transmission with the interactions in the population, where such random events alter the outcomes. The time delay effects account for the lag time in vaccine deployment. In essence, early vaccination, where the optimal strategy is to concentrate available vaccines in cities from which cholera originated, and delayed vaccination with an equal allocation of vaccines in cities where cholera already spreads. Incorporating these models will provide a more realistic framework as a control strategy.

5.4 | PDF of the Zubair-Fréchet Distribution

Suppose that X has a continuous distribution on R with probability density function f and distribution function F . Then the pdf plot of the ZF distribution as shown in Figure 6 below, can be left-skewed or right-skewed with small and large values of skewness and kurtosis for different values of parameters [37].

As shown in Figure 6, the pdf of the Z-F distribution shows for different parameters the flexibility of the distribution to handle

values in such a domain and class, especially in the handling of cholera disease data points.

5.5 | Cdf of the Zubair- Fréchet Distribution

The graphical representation of the CDF of ZF distribution is shown below:

As shown in Figure 7, the cdf of the Z-F distribution shows for different parameters the flexibility of the distribution to handle values in such a domain and class, especially in the handling of cholera disease data points. The properties of the CDF are clearly shown from the curve.

5.6 | Survival Function of Zf Distribution

The Graphical representation of the various survival functions of the ZF distribution is shown in the figure below.

As shown in Figure 8, the Survival of the Z-F distribution shows for different parameters the flexibility of the distribution to handle values in such a domain and class,

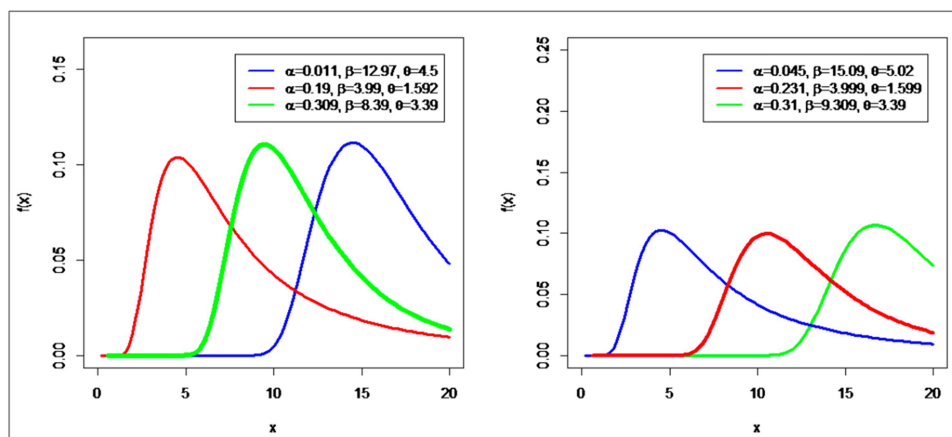


FIGURE 6 | PDF plots for ZF with different parameters. *Source:* Authors' compilation of graph of pdf plots of ZF distribution with different values and parameters.

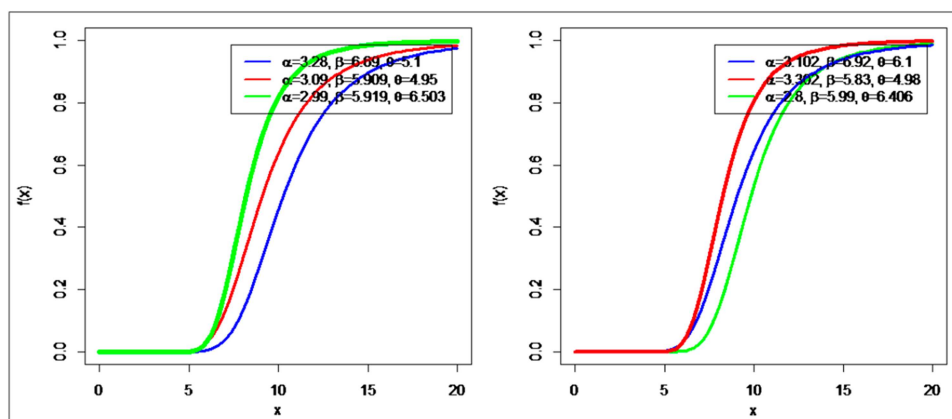


FIGURE 7 | CDF plots of ZF graph for different parameter values. *Source:* Authors' compilation of results of graph of CDF of ZF distribution with given parameters.

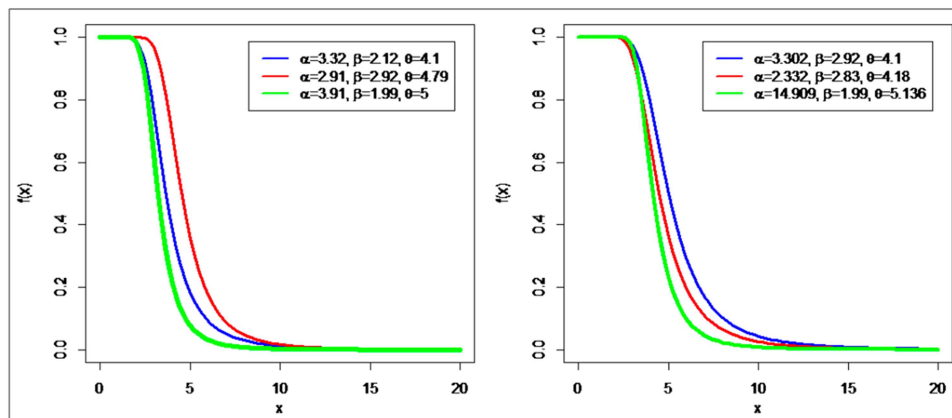


FIGURE 8 | Survival plots for ZF with different parameter values. *Source:* Authors' compilation of results of graph of the survival function graph with selected parameters.

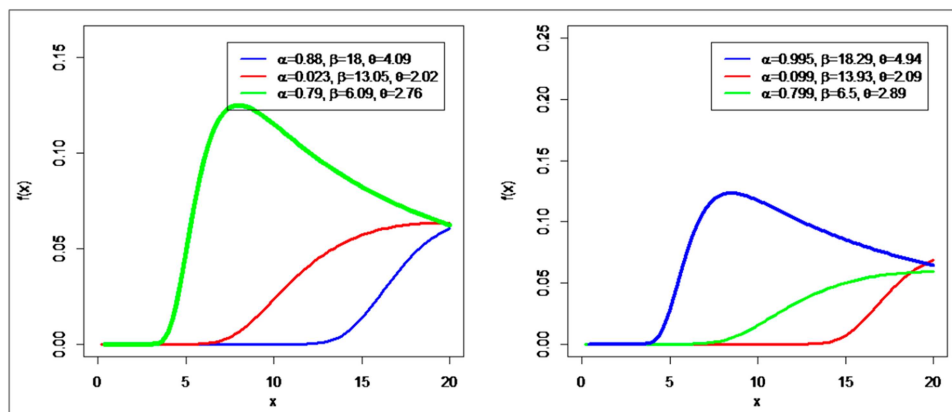


FIGURE 9 | Hazard plot for ZF. *Source:* Authors' compilation of results for graph of hazard rate function curve with selected parameters.

especially in the handling of cholera disease data points. The properties of the Survival are clearly shown from the curve.

5.7 | Hazard Rate Function of the Zf Distribution

The graphical representation of the various hazard rate functions are given below:

As depicted in the Figure 9, the hazard rate function (ZF distribution), it can be inferred that the hazard rate function has many shapes, such as a bathtub, and either monotonically increases or decreases depending on the shape. This demonstrates the ZF distribution's flexibility even more.

5.8 | Asymptotic Behavior

The asymptotic properties or limiting behavior involve the study of the distribution at $x = 0$ and as $x \rightarrow \infty$. In this paper, $\lim_{x \rightarrow \delta} T(x)$ as $\delta \rightarrow \infty$ was considered. It is expected that $\lim_{x \rightarrow \delta} T(x)$ will equal to or approach one.

Let X denote a non-negative random variable from an arbitrary distribution whose pdf and cdf are defined by

$t(x)$ and $T(x)$, respectively. A proof is given at the Appendix A.

5.9 | Monte Carlo Simulation Studies of the ZF Distribution

The properties of the maximum likelihood estimators of the ZF distribution were examined using the Monte Carlo simulation technique. This section responds to the third objective of the paper. The average bias (Abias) and the root mean square error (RMSE) are presented in Table 1. The quantile function of the Z-F distribution is used to generate the random samples of the distribution. This simulation experiment has been repeated for $N = 1000$ times for each and every model with various sample sizes of $n = 25, 50, 75, 100$ and different parameter values. Table 1 shows the Abias and RMSE of the maximum likelihood estimation of (α, θ, β) for $n = 25, 50, 75, 100$ for the ZF distribution. As sample sizes increase, the estimators of the parameters of RMSE and Abias both decline.

As shown in Table 2, (Abias) and (RMSE) of the Z-F distribution show for different parameters the flexibility of the distribution to handle values in such a domain and class, especially in the handling of cholera disease data points. The properties of these are clearly shown from the table. We see that as the

TABLE 2 | The average bias and root mean square of MLE(ZF).

(α, θ, β)	Sample Size (<i>n</i>)	ABias of estimated parameters			RMSE of estimated parameters		
		$(\hat{\alpha})$	$(\hat{\theta})$	$(\hat{\beta})$	$(\hat{\alpha})$	$(\hat{\theta})$	$(\hat{\beta})$
(0.02, 0.0880, 0.001)	25	7.7223	−0.0842	1.9121	7.7382	0.0847	2.7114
	50	7.5673	−0.0845	1.8809	7.5853	0.0846	2.6801
	75	7.4278	−0.0846	1.8233	7.4497	0.0845	2.5766
100	7.1071	−0.0847	1.8346	7.1328	0.0842	2.3475	
(0.02, 0.089, 0.001)	25	7.693	−0.0842	2.0168	7.7092	0.0842	2.766
	50	7.6021	−0.0845	1.7754	7.6197	0.0845	2.7629
	75	7.4626	−0.0846	1.8298	7.4844	0.0846	2.5219
100	7.124	−0.0846	1.8071	7.1527	0.0846	2.5143	
(0.2, 0.1, 0.3)	25	7.3262	−0.0944	1.9805	6.7496	0.0944	2.6773
	50	7.2141	−0.095	0.9433	7.0852	0.095	1.9805
	75	7.0633	−0.0953	0.4203	7.2319	0.0953	1.4307
100	6.7224	−0.0954	0.2124	7.3418	0.0954	0.8572	
(0.2, 0.0879, 0.001)	25	7.6963	−0.0841	2.035	7.7118	0.0845	2.8161
	50	7.5789	−0.0844	1.8825	7.5967	0.0834	2.7944
	75	7.4167	−0.0845	1.7983	7.4366	0.0725	2.5911
100	7.1217	−0.0845	1.7655	7.1472	0.0645	2.5465	
(0.2, 0.0881, 0.001)	25	7.719	−0.0843	1.9759	7.7355	0.0847	2.6535
	50	7.6169	−0.0845	1.9223	7.6353	0.0845	2.6405
	75	7.4062	−0.0847	1.9096	7.4274	0.0837	2.6299
100	7.1154	−0.0847	1.8209	7.1429	0.0821	2.7273	

Source: Author's compilation of the average bias (Abias) of the MLE of the Z-F distribution and the Root Mean Square Error (RMSE) of the MLE of the Z-F distribution.

TABLE 3 | Goodness of fits of comparative models.

Data VI				
D	AIC	BIC	CAIC	HQIC
Zubair-Fréchet	−27.61176	−22.12584	−27.04034	−25.5567
Fréchet	213.3038	216.9611	213.5829	214.6738
Weibull	219.0484	222.7057	219.3275	220.4184

Source: Authors' computation of goodness-of-fits. cholera, 2024.

sample size increases, then the estimators of the parameters of RMSE and Abias both decline.

those of the Fréchet (213.3) and Weibull (219.0), indicating a decisively better fit to the cholera case data.

5.10 | Analytical Measures and Model Selection

The analytical measures used in this paper include the Akaike information criterion (AIC), Bayesian information criterion (BIC), Consistent Akaike information criterion (CAIC), and Hannan-Quinn information criterion (HQIC). The different tests are determined by the lowest values of these discrimination measures.

Evidently, the results above show that the Z-F distribution provides a better fit than the other distributions, such as the Weibull distribution and the Fréchet distribution in heavy-tailed distribution comparison. For instance, from Table 3, the Z-F distribution AIC value of −27.6 is substantially lower than

5.11 | Risk Measures in Health Resource Allocation and Outbreak Severity

Value at Risk (VaR)—a concept commonly used in finance to quantify financial losses—is applied to outbreak severity of cholera, especially with health resource allocation, primarily by using Tail Value at Risk (TVaR) to address extreme events and optimize resource distribution under uncertain events. This will help health professionals to analyze the worst-case scenario within a given time frame and confidence level.

Results for the health resource allocation and outbreak severity by the use of VaR and TVaR are shown in the table below.

TABLE 4 | Numerical study using risk metrics.

Normal distribution using risk metrics (ghana)				
Probabilities	VaR	TVaR/ES	VaR	TVaR/ES
0.950	8.117738	10.17997	12.20702	15.30810
0.990	11.481072	13.15346	17.26462	19.77947
0.999	15.251021	16.61738	22.93367	24.98833

Source: Authors' compilation of Numerical Study of VaR and TVaR for Ghana cholera cases, 2024.

TABLE 5 | Numerical study of POT and GPD.

Peaks over threshold at threshold of 0.025 (Ghana)			Generalized Pareto Distribution with a threshold of 0.025 (Ghana)		
Probabilities	VaR	TVaR/ES	Probability	VaR	TVaR
0.950	11.06480	13.05777	0.950	11.06541	13.05808
0.990	14.36534	15.70347	0.990	14.36548	15.70324
0.999	17.29532	18.05213	0.999	17.29458	18.05104

Peaks over threshold at threshold of 0.025 (Africa)			Generalized Pareto Distribution with a threshold of 0.025 (Africa)		
Probabilities	VaR	TVaR/ES	Probability	VaR	TVaR
0.950	10.10064	10.12343	0.950	10.11214	10.12776
0.990	10.13382	10.13490	0.990	10.13458	10.13518
0.999	10.13545	10.13546	0.999	10.13546	10.13547

Source: Authors' compilation of Numerical Study of VaR and TVaR for Ghana and Africa cholera cases, 2024.

The values from Table 4 show a wider gap in value. The consistency can be questioned.

As shown in Table 5, the POT and GPD are compared for Ghana and Africa. The results show much consistency (little numerical gaps between VaR and TVaR for different probability values) with the use of the POT for the cholera data. The results above quantify the maximum potential loss Ghana spends over a specific time period and at the given confidence level. A VaR of GH¢ 1 million at a 95% confidence level over 1 day means there is 95% probability the loss will not exceed GH¢ 1 million in a single day. VaR in this instance measures potential losses in terms of public health outcomes. The change in financial VaR values captured above, indirectly, is a reflection of the perceived economic severity during an outbreak. An increase in VaR values indicates that the public health situation has worsened. A low VaR value suggests that potential financial losses are minimal, hence the economy of Ghana becomes more stable and hence less affected by the outbreak at such a specific time horizon. For instance, from Table 4, taking a “95% VaR” that an outbreak will not result in more than eight (8) deaths or severe cases in Ghana. Key measures of the severity of the cholera outbreak come in the form of infection fatality risk, attack rate, case fatality risk and hospitalization risk. The TVaR focuses on the tail of the distribution and targets the worst-case scenarios. TVaR is used to develop models for risk-aware resource allocation, waiting times of patients, especially for improving equity outcomes in health systems as captured in this paper (Risk Measures for Zubair Fréchet), for the financial impact of an outbreak.

Value at risk in health estimates the maximum financial cost of a health emergency within a specific confidence level [38].

Linking financial risk measures to public health resource allocation involves the medical impoverishment, catastrophic health expenditure and fiscal space for health, which involves the assessment of the government's capacity to provide additional budgetary resources for cholera treatment [39].

5.12 | Comparison in Risk Measures

The value at risk and Tail value at risk are also compared for the Z-F, Weibull and Fréchet distributions. The Z-F distribution gives a better selection than the others. VaR indicates heavier-tailed outbreaks, hence a greater preparedness cost is needed [40]. A recommendation to consider Z-F distribution in the modeling of cholera cases, for further research.

As shown in Table 6, the Value at risk and expected value at risk are also compared for the Z-F and Fréchet distribution. The Z-F distribution gives a better selection than the others, as shown from Table 6 by a simulation of over 10,000 iterations. This does not give the VaR and TVaR of the cholera data. It generalizes the comparison of the classic traditional distribution Fréchet and novel distribution Zubair Fréchet by simulation, accounting for higher values than for a cholera case data.

The Value at risk and expected value at risk are also compared for the Z-F and Fréchet distribution. The Z-F distribution gives a better selection than the others, as shown from Figure 10.

The Value at risk and expected value at risk are also compared for the Z-F and Weibull distribution. The Z-F distribution gives a better selection than the others, as shown from Figure 11.

TABLE 6 | Risk measure comparison.

Z-Fréchet					Fréchet			
Level of Significance	0.7	0.8	0.9	0.95	0.7	0.8	0.9	0.95
Parameters	Alpha = 0.7	Beta = 4.4	Theta = 0.32	—	—	Beta = 4.4	Theta = 0.32	
VaR	2357.5	11000.2	119485.7	1237899	111.7	474.2	5069.5	46081.5
TVaR	7.7E + 13	4.3E + 11	1.4E + 17	6.8E + 16	3.4E + 11	6.19E + 11	1.16E + 18	1.11E + 12

Source: Authors' compilation of Numerical Study of VaR and TVaR for a different set of parameters.

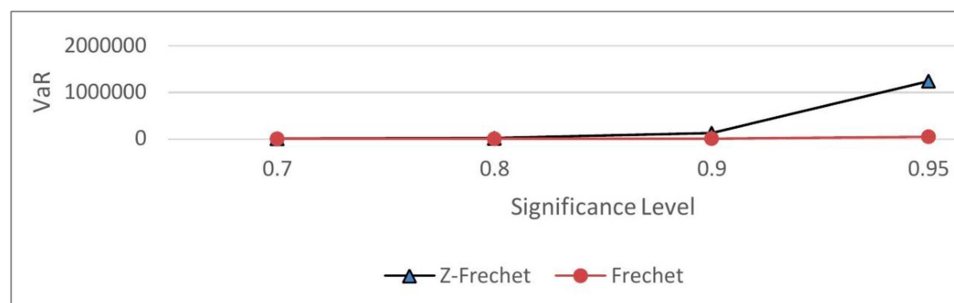


FIGURE 10 | Graph of ZF and Fréchet compared. Source: Authors' computation and graph of comparison of VaR of ZF and Fréchet distribution.

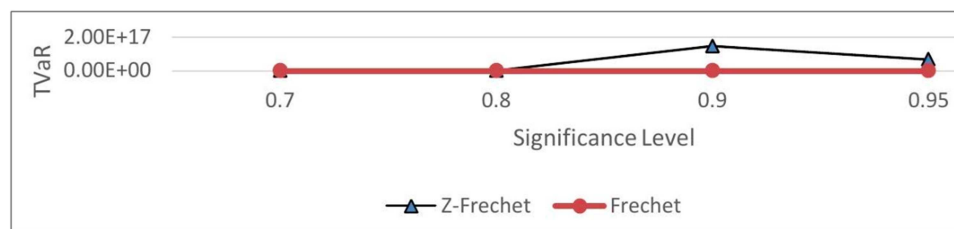


FIGURE 11 | Graph of comparison of ZF and Weibull. Source: Authors' computation of a comparative analysis and graph of TVaR for ZF and Weibull distributions.

6 | Conclusion and Recommendation

This paper proposes an enhanced model for cholera disease insurance by incorporating disease-specific dynamics for improved financial risk assessment through a robust model for quantifying financial effects on policyholders and insurers of epidemics.

6.1 | Summary of Findings

The Z-F generator was developed. The Zubair generator has one shape parameter, whereas the Fréchet distribution has one shape parameter and one scale parameter. The pdf, cdf and the hazard functions were primarily derived. The new generator possesses three parameters, making it more flexible and sufficient to display symmetrical, left-skewed and right-skewed data illustrated by the pdf plots. By virtue of the hazard rate function plots, the model also exhibited a bathtub shape, observed as upside-down bathtub shapes. The actuarial properties of the Z-F distribution are derived. These are the quantile function, force of mortality, cumulative hazard function, distribution of order statistics, survival analysis, value at risk and expected shortfall. The maximum likelihood technique was used to derive the estimators of the parameters.

A study on simulation was conducted to check the behavior of the parameter estimators. Sample size did increase from

25, 50, 75 to 100, as there was a corresponding responsive decrease of the root mean square error (RMSE) and average bias (ABias). This then satisfies the property of consistency for the maximum likelihood estimator of the Z-F distribution. The Z-Fréchet, Fréchet and Weibull distributions are fitted by maximum likelihood estimation to numerical simulation and cholera data from Ghana and Africa. The application results show that the Z-F distribution provides a better fit. A model with a higher risk measure connotes a heavier tail. The results of the simulation in Table 2 above suggest that the proposed Z-F distribution has higher values of the risk measures in comparison with the traditional Fréchet distribution. Hence, the proposed Zubair-Fréchet distribution has heavier tails than the Fréchet distribution. Parameters for the SIR model for the cholera case in Ghana for 2024 have been determined for further research.

6.2 | Contribution to Knowledge

The paper has led to the following contributions. Actuarial properties have been developed from the Z-F, which could be of use for key stakeholders such as researchers and students. In practice, the risk measures help in setting premiums in property and casualty insurance. There is also the empirical development of an extension of the Fréchet distribution.

6.3 | Recommendation and Further Research

There should be the consideration of application of more link functions, which result in generalized families of distributions with induced skewness having additional parameters, such as shape, scale or location are developed for fat-tailed distributions. There should be other analytical research performed with Z-F to ascertain the best-fitting model for several heavy-tailed real-life data sets. This is crucial because there are few or no techniques that can handle the extreme value analysis of multiplex data.

In Ghana, cholera is a public health concern that needs more focus. Every year since the initial outbreak in 1970, a number of persons are impacted. Additionally, cholera claims the lives of numerous people every year. These suggestions could reduce, if not completely eliminated. Mathematical models should be used to help in the prediction of such diseases, as well as the intensity for commensurate mitigation of the same.

The results from this simulation could help guide the Ministry of Health and the Ghana Health Service Response to be able to preposition the necessary logistics and update preparedness and response plans for the country as the rainy season begins, especially in the Greater Accra Region, due to overcrowding and contaminated water and poor sanitation. The country has started various preparedness activities in the most affected areas, and evidence from this could help tailor various training activities, request for vaccines, procurement of RDTs, aqua tabs, provision of WASH facilities, medications and logistics.

Ghana has initiated processes to update the National Action Plan and guidelines on cholera, and this information will help influence what goes into the two documents as lessons learnt from this outbreak will guide the process.

7 | Limitations and Suggested Further Research

The reproduction and effective reproduction of the model can be analyzed in the light of simulated results. Again, the SIR framework lacks ecological fidelity for *V. cholerae* transmission, hence future research could capture and analyze the presence of an environmental reservoir compartment. There should be an extension to an SIR-B or SIRC model.

This is a result of data limitations with an assumption of simple SIR modeling. Further studies are recommended on the ZF VaR and TVaR comparison of risk measures against traditional models.

The threshold selection for POT has been a challenge in the literature [41]. This has been a shortcoming that should be further looked into for other preferred selection processes.

Author Contributions

The study was designed by Kofi Afriyie Nyamekye, the corresponding author. Material preparation and data collection were performed by Kofi Afriyie Nyamekye, Franklin Asiedu-Bekoe, and Joseph Quarshie. The data analysis was performed by Prof. Justice Moses K. Aheto, Prof. Eric Sampene-Donkor, and Kofi Afriyie Nyamekye. The first draft of the

manuscript was written by Kofi Afriyie Nyamekye and Justice Moses K. Aheto, and all authors commented on the previous versions of the manuscript. All authors read and approved the final manuscript.

Acknowledgments

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article. This paper was supported by the Fogarty International Centre of the National Institutes of Health through the Application of Data Science to Build Research Capacity in Zoonoses and Foodborne Infections in West Africa Training Program hosted at the Department of Medical Microbiology, University of Ghana Medical School (Award Number: UE5TW012566). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

Justice Moses K. Aheto is an Editorial Board member of Health Science Reports and a co-author of this article. To minimize bias, he has been excluded from all editorial decision-making related to the acceptance of this article for publication.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Transparency Statement

The lead author Kofi Afriyie Nyamekye affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

References

1. M. Bourguignon, R. B. Silva, and G. M. Cordeiro, "The Weibull-G Family of Probability Distributions," *Journal of Data Science* 12, no. 1 (2014): 53–68, [https://doi.org/10.6339/JDS.2014.12\(1\).1210](https://doi.org/10.6339/JDS.2014.12(1).1210).
2. S. Huang and B. Oluyede, "Exponentiated Kumaraswamy–Dagum Distribution With Applications to Income Life Data," *Journal of Statistical Distributions and Applications* 1, no. 8 (2014): 1–20, <https://doi.org/10.1186/2195-5832-1-8>.
3. A. Al-Babtain, A. A. Fattah, A.-H. N. Ahmed, and F. Merovci, "The Kumaraswamy-Transmuted Exponentiated Modified Weibull Distribution," *Communications in Statistics - Simulation and Computation* 46, no. 5 (2017): 3812–3832.
4. D. Bernoulli, *Essai d'une nouvelle analyse de la mortalité causée par la petite vérole et des avantages de l'inoculation pour la prévenir* (Histoire de l'Académie Royale des Sciences, 1766), 1–45.
5. W. Shaw and I. Buckley (2007). The Alchemy of Probability Distributions: Beyond Gram-Charlier Expansions and a Skew-Kurtotic-Normal Distribution From a Rank Transmutation Map, Research Report.
6. D. Olago, M. Marshall, S. O. Wandiga, et al., "Climatic, Socio-Economic, and Health Factors Affecting Human Vulnerability to Cholera in the Lake Victoria Basin, East Africa," *Ambio* 36, no. 4 (2007): 350–358, [https://doi.org/10.1579/0044-7447\(2007\)36\[350:csahfa\]2.0.co;2](https://doi.org/10.1579/0044-7447(2007)36[350:csahfa]2.0.co;2).
7. European Centre for Disease Prevention and Control (ECDC), *Cholera Worldwide Overview* (ECDC, 2024), <https://www.ecdc.europa.eu/en/all-topics-z/cholera/surveillance-and-disease-data/cholera-monthly>.
8. O. El Deeb and A. Matar (2024). Impact of Segregation and Integration on Infection Dynamics: Analysis of a Cholera Model in a Two-Population System, IOP Publishing Ltd, Phys. Scr.99 105016.

9. S. Nasiru. A New Generalisation of Transformed Transformer Family of Distributions. Doctoral Thesis, Pan African University, Institute for Basic Sciences, Technology and Innovation. (2018).
10. Z. Ahmad, "The Zubair-G Family of Distributions: Properties and Applications," *Annals of Data Science* 7, no. 2 (2020): 195–208, <https://doi.org/10.1007/s40745-018-0169-9>.
11. M. Awadalla, T. N. Cheema, A. Raza, et al., "Intelligent Optimization Analysis of the Cholera Epidemic Model," *Discrete Dynamics in Nature and Society* 2024 (2024): 1–16, <https://doi.org/10.1155/ddns/9677485>.
12. Y. H. Grad, J. C. Miller, and M. Lipsitch, "Cholera Modeling: Challenges to Quantitative Analysis and Predicting the Impact of Interventions," *Epidemiology* 23, no. 4 (2012): 523–530, <https://doi.org/10.1097/EDE.0b013e3182572581>.
13. F. Merovci, "Transmuted Exponentiated Exponential Distributions," *Mathematical Sciences and Applications E-Notes* 1, no. 2 (2013a): 112–122.
14. Z. Ahmad, E. Mahmoudi, G. G. Hamedani, and O. Kharazmi, "New Methods to Define Heavy-Tailed Distributions With Applications to Insurance Data," *Journal of Taibah University for Science* 14, no. 1 (2020): 359–382, <https://doi.org/10.1080/16583655.2020.1741942>.
15. A. Mahdavi and D. Kundu, "A New Method for Generating Distributions With an Application to Exponential Distribution," *Communications in Statistics - Theory and Methods* 46, no. 13 (2017): 6543–6557, <https://doi.org/10.1080/03610926.2015.1130839>.
16. D. G. G. Hamedani, H. M. Yousof, M. Rasekhi, M. Alizadeh, and S. M. Najibi, "Type I General Exponential Class of Distributions," *Pakistan Journal of Statistics and Operation Research* 14, no. 1 (2018): 39–55, <https://doi.org/10.18187/pjsor.v14i1.2193>.
17. K. Pearson, "Contributions to the Mathematical Theory of Evolution II: Skew Variation in Homogeneous Material," *Philosophical Transactions of the Royal Society A: Containing Papers of a Mathematical or Physical Character* 186, no. 186 (1895): 343–414, <https://doi.org/10.1098/rsta.1895.0010>.
18. P. E. Oguntunde, A. O. Adejumo, and K. A. Adepoju, "Assessing the Flexibility of the Exponentiated Generalised Exponential Distribution," *Pacific Journal of Science and Technology* 17, no. 1 (2016b): 49–57.
19. F. Kungu, S. N. A. Yartey, A. A. Asantewaa, and E. S. Donkor, "Cholera Burden in Ghana: A Systematic Review and Meta-Analysis of Prevalence, Antimicrobial Resistance and Risk Factors," *International Health* 17 (2025): 869–880, <https://doi.org/10.1093/inthealth/ihaf069>.
20. P. Apenteng, N. Maimaiti, and C. A. G. Donkor, "Outbreak of Cholera in Ghana: A Systematic Review From 2010 to 2020," *Asian Journal of Research in Infectious Diseases* 13, no. 2 (2023): 26–33, <https://doi.org/10.9734/ajrid/2023/v13i2261>.
21. E. Goldstein, S. Cobey, S. Takahashi, J. C. Miller, and M. Lipsitch, "Predicting the Epidemic Sizes of Influenza A/H1N1, A/H3N2, and B: A Statistical Method," *PLOS Medicine* 8, no. 7 (2011): e1001051, <https://doi.org/10.1371/journal.pmed.1001051>.
22. S. Nadarajah and S. Kotz, "The Exponentiated Type Distributions," *Acta Applicandae Mathematicae* 92, no. 2 (2006): 97–111, <https://doi.org/10.1007/s10440-006-9055-0>.
23. R. P. Soebiyanto, F. Adimi, and R. K. Kiang, "Modeling and Predicting Seasonal Influenza Transmission in Warm Regions Using Climatological Parameters," *PLOS ONE* 5, no. 3 (2010): e9450, <https://doi.org/10.1371/journal.pone.0009450>.
24. A. S. Mugglin, N. Cressie, and I. Gemm, "Hierarchical Statistical Modelling of Influenza Epidemic Dynamics in Space and Time," *Statistics in Medicine* 21, no. 18 (2002): 2703–2721, <https://doi.org/10.1002/sim.1217>.
25. E. K. Kim, J. H. Seok, J. S. Oh, H. W. Lee, and K. H. Kim, "Use of Hangeul Twitter to Track and Predict Human Influenza Infection," *PLOS ONE* 8, no. 7 (2013): e69305, <https://doi.org/10.1371/journal.pone.0069305>.
26. J. C. Miller, "Mathematical Models of Sir Disease Spread With Combined Non-Sexual and Sexual Transmission Routes," *Infectious Disease Modelling* 2, no. 1 (2017): 35–55, <https://doi.org/10.1016/j.idm.2016.12.003>.
27. A. Azzalini, "A Class of Distributions Which Includes the Normal Ones," *Scandinavian Journal of Statistics* 12, no. 2 (1985): 171–178.
28. A. O'hagan and T. Leonard, "Bayes Estimation Subject to Uncertainty About Parameter Constraints," *Biometrika* 63, no. 1 (1976): 201–203, <https://doi.org/10.1093/biomet/63.1.201>.
29. C. Fernández and M. F. J. Steel, "On Bayesian Modeling of Fat Tails and Skewness," *Journal of the American Statistical Association* 93, no. 441 (1998): 359–371, <https://doi.org/10.2307/2669632>.
30. C. Lee, F. Famoye, and A. Y. Alzaatreh, "Methods for Generating Families of Univariate Continuous Distributions in the Recent Decades," *WIREs Computational Statistics* 5, no. 3 (2013): 219–238, <https://doi.org/10.1002/wics.1255>.
31. A. Z. Afify, G. G. Hamedani, I. Ghosh, and M. E. Mead, "The Transmuted Marshall-Olkin Fréchet Distribution: Properties and Applications," *International Journal of Statistics and Probability* 4, no. 4 (2015): 132–148, <https://doi.org/10.5539/ijsp.v4n4p132>.
32. G. S. Mudholkar and A. D. Hutson, "The Epsilon-Skew-Normal Distribution for Analyzing Near-Normal Data," *Journal of Statistical Planning and Inference* 83, no. 2 (2000): 291–309, [https://doi.org/10.1016/S0378-3758\(99\)00096-8](https://doi.org/10.1016/S0378-3758(99)00096-8).
33. S. M. Chang and M. G. Genton, "Extreme Value Distributions for the Skew-Symmetric Family of Distributions," *Communications in Statistics - Theory and Methods* 36 (2007): 1705–1717.
34. A. M. Abd-Elfattah, S. M. Assar, and H. I. Abd-Elghaffar, "Exponentiated Generalised Fréchet Distribution," *International Journal of Mathematical Analysis and Applications* 3, no. 5 (2016): 39–48.
35. M. Fréchet, "Sur la loi de Probabilité de l'écart Maximum [On the Probability Law of the Maximum Deviation]," *Annales de la Société Polonaise de Mathématique* 6 (1927): 93–116.
36. S. Rezaei, A. K. Marvasty, S. Nadarajah, and M. Alizadeh, "A New Exponentiated Class of Distributions: Properties and Applications," *Communications in Statistics - Theory and Methods* 46, no. 12 (2017): 6054–6073, <https://doi.org/10.1080/03610926.2015.1116579>.
37. N. L. Johnson, "Systems of Frequency Curves Generated by Methods of Translation," *Biometrika* 36, no. 2 (1949): 149–176, <https://doi.org/10.1093/biomet/36.1-2.149>.
38. P. Abad and S. Benito, "A Detailed Comparison of Value at Risk Estimates," *Mathematics and Computers in Simulation* 94 (2013): 258–276, <https://doi.org/10.1016/j.matcom.2012.05.011>.
39. S. Coles, *An Introduction to Statistical Modeling of Extreme Values* (Springer, 2001), <https://doi.org/10.1007/978-1-4471-3675-0>.
40. D. Cummins, "Insurer's Risk: A Restatement," *Journal of Risk and Insurance* 4, no. 1 (2014): 147–157.
41. R. S. Tsay, *Analysis of Financial Time Series* (3rd Edition. John Wiley & Sons, <https://doi.org/10.1002/9780470644560>.

Appendix A

Steps for anti-derivative of Z-F pdf

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

$$\int t(x) dx = T(x)$$

From the mutually computed anti-derivative above, it is assumed that,

$$-c - 1 \neq 0, c \neq 0$$

Hence,

$$\int \frac{2\alpha\beta^\theta \theta x^{-\theta-1} e^{\alpha e^{-2\left(\frac{\beta}{x}\right)^\theta} - 2\left(\frac{\beta}{x}\right)^\theta}}{e^\alpha - 1} dx$$

$$\frac{2\alpha\beta^\theta \theta}{e^\alpha - 1} x^{-\theta-1} e^{\alpha e^{-\frac{2\beta^\theta}{x^\theta} - \frac{2\beta^\theta}{x^\theta}}} dx$$

Applying linearity,

Now solving,

$$\int x^{-\theta-1} e^{\alpha e^{-\frac{2\beta^\theta}{x^\theta} - \frac{2\beta^\theta}{x^\theta}}} dx$$

$$u = e^{-\frac{2\beta^\theta}{x^\theta}} \rightarrow \partial u = 2\beta^\theta \theta x^{-\theta-1} e^{-\frac{2\beta^\theta}{x^\theta}} dx$$

$$x^{-\theta-1} = \frac{\beta^{-\theta-1} \cdot 2^{-\frac{\theta-1}{\theta}}}{(-\ln(u))^{\frac{-\theta-1}{\theta}}}$$

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

$$= \frac{1}{2\theta\beta^\theta} \int \left(e^{-\frac{(e-1)(\ln(u) + \alpha u)}{u}} \right) \partial u$$

Substitute

$$v = \frac{e^{-\frac{(-\theta-1)(\ln(u) + \alpha u)}{\theta+1}}}{u} \rightarrow \partial v$$

$$= \left(-\frac{\left(-\theta - 1 \left(\frac{1}{u} + \alpha \right) e^{-\frac{(-\theta-1)(\ln(u) + \alpha u)}{e+1}} \right)}{(\theta+1)u} \right)$$

$$- \frac{e^{-\frac{(-\theta-1)(\ln(u) + \alpha u)}{e+1}}}{u^2} \partial u$$

$$\int \left(\frac{u}{v \cdot \left(\frac{\alpha}{\ln(v)} + \alpha \right) - \frac{\alpha v}{\ln(v)}} \right) \partial v$$

Since,

$$\frac{1}{\alpha} \int 1 \partial v$$

Undo substitution

$$v = \frac{e^{-\frac{(-\theta-1)(\ln(u) + \alpha u)}{\theta+1}}}{u}$$

$$\frac{1}{2\theta\beta^\theta} \int \left(\frac{e^{-\frac{(-\theta-1)(\ln(u) + \alpha u)}{\theta+1}}}{u} \right) \partial v$$

$$\int \left(\frac{e^{-\frac{(-\theta-1)(\ln(u) + \alpha u)}{\theta+1}}}{2\alpha\theta\beta^\theta u} \right)$$

$$u = e^{-\frac{2\beta^\theta}{x^\theta}}$$

After undoing substitution,

$$\ln \left(e^{-\frac{2\beta^\theta}{x^\theta}} \right) = -\frac{2\beta^\theta}{x^\theta}$$

$$= \frac{2\beta^\theta}{e^{\frac{2\beta^\theta}{x^\theta} - \frac{(-\theta-1)\alpha e^{-\frac{2\beta^\theta}{x^\theta} - \frac{2\beta^\theta}{x^\theta}}}{\theta+1}}}$$

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

$$= \frac{2\beta^\theta}{e^{\frac{2\beta^\theta}{x^\theta} - \frac{(-\theta-1)\alpha e^{-\frac{2\beta^\theta}{x^\theta} - \frac{2\beta^\theta}{x^\theta}}}{\theta+1}}} + C$$

$$= \frac{e^{\alpha e^{-2\beta^\theta}}}{e^\alpha - 1} + C$$

Appendix B

TABLE B1 | Quantitative measures (SIR model).

Sample size, parameters of quantitative measures	RMSE	MSE	RSS
N = 10; (beta = 0.004, gamma = 0.999)	5.689045	32.36523	2297.931
N = 1000; (beta = 0.004, gamma = 0.999)	5.68904	32.36518	2297.928
N = 10,000; (beta = 0.006, gamma = 0.995)	5.688973	32.36442	2297.874

Source: Authors' compilation of Quantitative Measures for a different set of parameters.