Research Article

Statistical Analysis of Water Purification (Using Vinyl Chloride) Data

Aqsa Rafique, 1 Naz Saud, 1 Naila Amjad, 1 Muhammad Ijaz, 2 Fastel Chipepa, 3 and Mahmoud El-Morshedy 4, 5

1 Department of Statistics, Lahore College for Women University, Lahore, Pakistan
2 Department of Mathematics and Statistics, University of Haripur, Haripur, KPK, Pakistan
3 Department of Mathematics and Statistical Sciences, Botswana International University of Science and Technology, Gaborone, Botswana
4 Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
5 Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt

Correspondence should be addressed to Mahmoud El-Morshedy; mah_elmorshedy@mans.edu.eg

Received 28 October 2021; Revised 9 May 2022; Accepted 11 May 2022; Published 10 June 2022

Academic Editor: A. M. Bastos Pereira

Copyright © 2022 Aqsa Rafique et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Water is important to live because of its impacts on food supply and the natural environment for all living things. Approximately 0.3 percent of water resources are usable in the world. Groundwater is the principal source of drinking water but the air, sea, land, rivers, lakes, ocean, and wells are also the essential sources of water. Many statistical models have been used to analyze the water data set so that to make predictions in a better way with limited resources of water. In this paper, a new approach is offered to model the data of water purification using vinyl chloride data. Various statistical properties of the proposed model have been derived. Maximum likelihood estimation is used to estimate the model parameters. Monte Carlo simulations are used to show the consistency of parameters. Using water purification using vinyl chloride, the suggested model application is studied and compared with that of other existing models such as EGF, GF, and F. The results showed that our model provides a much better fit while modeling this data set rather than EGF, GF, and F distributions.

1. Introduction

In 21st century, one of the main issues of water resources is the preservation and rational utilization. Water is a valuable resource in terms of its quality, quantity, and location, with quality referring to the state of the water system as measured by chemical, biological indicators, and physical-chemical. Economic and geographic influence and influence on ecosystem and human health are all key factors to consider when assessing the top-priority issues of water quality [1]. Changes in the physico-chemical features of water quality are determined by anthropogenic factors and also by natural processes such as lithology and topography, hydrological conditions, precipitation inputs, climate, edaphic and tectonic factors, and catchment area erosion, in combination with environmental influences [2]. The greatest threat to water quality is posed by point sources of industries and municipalities. In terms of water quality, anthropogenic influences can have negative consequences in a short time period [3]. The organic solid load and the dynamics of its degradation are very good indicators of anthropogenic impact on water [4].

Water pollution threatens the environment. A water resource policy is needed to resolve this problem. The causes of the death and diseases ratio are increased and about 14000 people dead cases are recorded every day due to water pollution. Many factors are considered that cause the water pollution nowadays such as industrial wastes, atmospheric pollutants, pesticides, and herbicides [5].

Water is purified by biological and physico-chemical processes that occur at the air-water interface and within the bulk water, over the surface of the sand and within the bed of
sand [6]. Because of their wide surface area, excellent mechanical strength, high chemical reactivity, and low cost, nanoscale composite materials have enormous potential to filter water in a variety of ways. Metal nanocomposite, carbon nanocomposite, metal oxide nanocomposite, and polymer nanocomposite plays an active role in water purification.

In the literature, there exist a lot of distributions that can be used to model the real data sets but these distributions do not increase the goodness of fit and also do not capture the non-monotonicity of the data. For example, the Weibull, Gamma, and Exponential distributions fail to model the data with a nonmonotonic hazard rate. To overcome the limitations of these distributions, researchers are working to define new models. The detail is given under the material and methodology section.

The aim of this paper is to define a new probability model and test distributions with application to the data set of water purification using vinyl chloride. The following points are of interest:

1. To propose a model that can model the data sets with both monotonic and nonmonotonic hazard rate functions
2. To increase the goodness of fit as compared with other models

The article is arranged as follows. Materials and Methods are discussed in Section 2. Section 3 introduces the FFP lifetime distribution and the expressions for density function, distribution function, survival function, and hazard rate functions are presented. In this Section 3, two propositions are discussed. In Section 4, mathematical properties of the proposed distribution are discussed. Parameter Estimation is done in Section 5. A simulation study is conducted in Section 6. In Section 7, an application study is carried out to show the flexibility of the FFP model. We present concluding remarks in Section 8.

2. Material and Methodology

Statistical distributions have become propitious in describing and predicting real-world applications. The basic idea behind developing these distributions is that the lifetime of a system with Y number of elements (discrete random variable) and a positive continuous random variable X, which denotes the lifetime of the i^{th} element, can be represented by nonnegative random variable \( T = \max(X_1, X_2, \ldots, X_Y) \) if the elements are in a parallel or \( T = \min(X_1, X_2, \ldots, X_Y) \) if the elements are in a series. Recently, various distributions to model lifetime data have been introduced; for example, Tahmasebi and Jafari [7] proposed a new distribution named Generalized Gompertz-Power Series Distributions. Rafique and Saud [8] proposed lifetime distribution with modified upside-down bathtub, increasing, decreasing, and reverse J-shaped hazard rate by compounding exponentiated generalized Frechet with geometric distribution. Another compounded inverse Weibull distribution with monotone and nonmonotone failure rates was proposed by Chakrabarty and Chowdhury [9]. Later Joshi and Dhungana [10], Louzada et al. [11], and Ibrahim [12] have come up with similar studies with Exponentiated Rayleigh poisson, Quasi Xgamma-Poisson, and Poisson Rayleigh Burr XII distributions.

The Exponentiated Generalized Frechet (EGF) distribution was introduced by Cordeiro et al. [13]. Various fields of engineering applications need to model the statistical performance of material properties and also be used to determine different failure characteristics as useful life, infant mortality, and wear-out periods. It is the appropriate model of extreme value continuous distribution. Extreme value distribution is required for a statistical model to maximize or minimize the collection of random observations. Extreme value theory (EVT) has been used to predict probabilities for maximum or minimum of the extreme values and measure events occurring with a small probability.

Let \( X \) be a continuous random variable following EGF distribution with \( \kappa, \delta, \xi, \phi \) and its probability density function (PDF) and cumulative distribution function (CDF) are, respectively, defined by

\[
g(x) = \kappa \delta \phi \xi x^{-\phi+1} e^{-\xi(t)x^\delta}(1 - e^{-\xi(t)x^\delta})^{\kappa-1} \\
G(x) = \left[ 1 - \left( 1 - e^{-\xi(t)x^\delta} \right)^{\kappa-1} \right]^{\delta}, \quad x > 0, \kappa, \delta, \phi, \xi > 0, \quad (1)
\]

where \( \kappa, \delta > 0 \) are the shape parameters and \( \xi > 0 \) is the scale parameter.

Let \( Y \) be a discrete random variable following zero a truncated Poisson distribution with parameter \( \beta > 0 \). The probability mass function of “\( Y^* \)” is given as

\[
P(y; \beta) = \frac{e^{-\beta} \beta^y}{\Gamma(y + 1)(1 - e^{-\beta})}, \quad y \in \mathbb{N}, \beta > 0, \quad (2)
\]

Now, assume \( X_1, X_2, \ldots \) a sequence of iid random variables extracted from \( X \) and also independent from \( Y \). We define \( T = \min\{X_1, X_2, \ldots, X_Y\} \).

Then, we get

\[
f_{\eta Y}(t; \phi, \kappa, \delta, \phi, \xi) = \frac{\eta \delta \phi \xi e^{-\phi t} e^{-\xi(t)t^\delta} \left( 1 - e^{-\xi(t)t^\delta} \right)^{\kappa-1} \left[ 1 - \left( 1 - e^{-\xi(t)t^\delta} \right)^{\kappa-1} \right]^{\delta-1}}{\left( 1 - \left( 1 - e^{-\xi(t)t^\delta} \right)^{\kappa-1} \right)^{\delta-1}}, \quad (3)
\]
The joint probability density function of \((T, Y)\) is

\[
f_{t,y}(t, y; \kappa, \delta, \phi, \xi) = y \kappa \delta \phi \xi t^{-(\phi+1)} e^{-(\xi y)t} \left(1 - e^{-(\xi y)t}\right)^{\kappa-1} \left[1 - \left(1 - e^{-(\xi y)t}\right)^{\delta}\right]^{-1}\left(1 - \left(1 - e^{-(\xi y)t}\right)^{\delta}\right)^{1-\phi - 1} e^{-\beta \phi y} \frac{y-1}{\Gamma(y+1)(1-e^{-\beta})}.
\]

\[
(4)
\]

3. The FFP Distribution

In this section, we derive a novel probability model using (4) called Flexible Frechet Poisson (FFP) distribution as a marginal density of \(T\) with the following PDF and CDF:

\[
f(t; \hat{\theta}) = \frac{\beta \kappa \delta \phi \xi t^{-(\phi+1)} e^{-(\xi y)t} \left(1 - e^{-(\xi y)t}\right)^{\kappa-1} \left[1 - \left(1 - e^{-(\xi y)t}\right)^{\delta}\right]^{-1} \exp\left(\beta \left(1 - \left(1 - e^{-(\xi y)t}\right)^{\delta}\right)\right)}{e^\delta - 1},
\]

\[
F(t; \hat{\theta}) = 1 - \frac{\exp\left(\beta \left(1 - \left(1 - e^{-(\xi y)t}\right)^{\delta}\right)\right) - 1}{e^\delta - 1}, \quad t > 0, \beta, \xi, \kappa, \delta, \phi > 0,
\]

where \(\hat{\theta} = /{\kappa, \delta, \phi, \xi, \beta, \beta}\)

The survival function and hazard function of \(T\) are given by

\[
S(t) = \frac{\exp\left(\beta \left(1 - \left(1 - e^{-(\xi y)t}\right)^{\delta}\right)\right) - 1}{e^\delta - 1},
\]

\[
h(t) = \frac{\beta \kappa \delta \phi \xi t^{-(\phi+1)} e^{-(\xi y)t} \left(1 - e^{-(\xi y)t}\right)^{\kappa-1} \left[1 - \left(1 - e^{-(\xi y)t}\right)^{\delta}\right]^{-1} \exp\left(\beta \left(1 - \left(1 - e^{-(\xi y)t}\right)^{\delta}\right)\right)}{\exp\left(\beta \left(1 - \left(1 - e^{-(\xi y)t}\right)^{\delta}\right)\right) - 1}.
\]

Proposition 1. The density function of the FFP model can be shown as infinite mixture of EGF density with parameters \(\xi, \phi, \kappa,\) and \(\delta\) using PDF form (5).

The Taylor series formula for real \(\tau\) is given by

\[
e^\tau = \sum_{\tau} \frac{\tau^e}{e!}
\]

The binomial expansion for a real, noninteger, \(s > 0\) is

\[
(1 - m)^{s-1} = \sum_{j=0}^{\infty} \binom{s-1}{j} (-1)^j m^j.
\]

Applying (8) and (9) into (5) and after simplification, we have the expansion of the PDF as
Figure 1: The PDF plots of the FFP distribution.

Figure 2: The HRF plots for the FFP distribution.
The EGF distribution with parameters $\xi, \phi, \kappa,$ and $\delta$ is a limiting form of the FFP class of distributions when $\beta^* \to 0$.

\[
\lim_{\beta \to \beta^*} F(t, \beta) = \lim_{\beta \to \beta^*} 1 - \frac{\exp\left(\left(1 - \left[1 - \left(1 - e^{-(\xi t)^\phi}\right)^\delta\right]\right)\right)}{e^\delta - 1} - 1 = \left[1 - \left(1 - e^{-(\xi t)^\phi}\right)^\delta\right],
\]

which is the CDF of EGF distribution with parameter $\xi, \phi, \kappa,$ and $\delta$.

4. Statistical Properties

In this section, we obtained the expressions of the statistical measures such as moment, quantile function, Lorenz, and Bonferroni curve.

4.1. Quantile Function and Random Number Generator. Quantile function is used to obtain the value of the random variable in the given probability distribution at a specific probability level. The $d^{th}$ quantile of any probability distribution can be attained by solving

\[
F(t_d) = d,
\]

where $0 \leq d \leq 1$.

The random number generator of the FFP distribution is given as

\[
T = \xi\left(1 - \left[1 - \left(1 - e^{-(\xi t)^\phi}\right)^\delta\right]\right)^{1/\phi} - 1/\phi.
\]

where $R$ represents random number that follows uniform distribution $U(0, 1)$.

For illustration purpose, we considered different set of parameters value (hypothetical) subject to the conditions that the parameter value must be positive. The quantile function of FFP distribution using different sets of parameters is shown in Table 1.

4.2. Moments. The moments of any probability distribution have major importance in statistical analysis and also in the application of real data sets. The $r^{th}$ moment is obtained as follows:

\[
\mu'_r = \int_0^\infty t^r f(t)dt,
\]

\[
\mu'_r = \int_0^\infty L_{b_{max}} \kappa (a\delta) \phi^x t^{-(\phi+1)+r} \exp\left(\left(-(b + 1)\left(\frac{\xi}{\delta}\right)\right)^{\phi}\right) dt,
\]

then

\[
\mu'_r = L_{b_{max}} (b + 1)^{r/\phi - 1} \kappa \phi \Gamma\left(1 - \frac{r}{\phi}\right).
\]
4.3. Order Statistics. Let \( T_{(1)}, T_{(2)}, T_{(3)}, \ldots, T_{(n)} \) be the order statistics of sample size \( n \) from FFP. Then, the PDF of the \( w \)th order statistic is given by

\[
 f_{1:n}(t) = n \left[ \frac{\exp(\beta \left( 1 - \left( 1 - e^{-\left(\frac{t}{\xi}\right)^\delta}\right)^{\delta}\right)}{e^\beta - 1} \right]^{n-1} \frac{n!}{r!(n-r)!},
\]

\[
 f_{n:n}(t) = n \left[ \frac{\exp(\beta \left( 1 - \left( 1 - e^{-\left(\frac{t}{\xi}\right)^\delta}\right)^{\delta}\right)}{e^\beta - 1} \right]^{n-1} \frac{n!}{r!(n-r)!},
\]

Using (5) and (6), the smallest and largest order statistics of FFP can be obtained, respectively, by using \( w = 1 \) and \( w = n \) as

\[
 E(T) = L_{\text{max}}(b + 1)^{(1/\phi) - 1} \kappa a \delta \Gamma(1 - (1/\phi)),
\]

\[
 V(T) = \xi^2 \left[ L_{\text{max}}(b + 1)^{(2/\phi - 1)} \kappa a \delta \Gamma^2 \left( \frac{-2}{\phi} + 1 \right) - \left( L_{\text{max}}(b + 1)^{(1/\phi - 1)} \kappa a \delta \Gamma \left( \frac{-1}{\phi} + 1 \right) \right)^2 \right].
\]

The descriptive measures for different parameter values for the FFP distribution are given in Table 2.

### Table 2: Quantiles of FFP distribution for different choices of parameters.

<table>
<thead>
<tr>
<th>(d)</th>
<th>((0.3, 1, 1.5, 1.2, 0.5))</th>
<th>((0.9, 6.7, 4.5, 1.4, 0.3))</th>
<th>((3.1, 4.5, 3.5, 0.6, 5))</th>
<th>((1.2, 2.5, 5.5, 0.2, 2))</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.711</td>
<td>0.086</td>
<td>0.061</td>
<td>2.699</td>
</tr>
<tr>
<td>0.2</td>
<td>0.870</td>
<td>0.157</td>
<td>0.088</td>
<td>2.782</td>
</tr>
<tr>
<td>0.3</td>
<td>1.016</td>
<td>0.254</td>
<td>0.118</td>
<td>2.848</td>
</tr>
<tr>
<td>0.4</td>
<td>1.172</td>
<td>0.401</td>
<td>0.152</td>
<td>2.908</td>
</tr>
<tr>
<td>0.5</td>
<td>1.351</td>
<td>0.640</td>
<td>0.195</td>
<td>2.968</td>
</tr>
<tr>
<td>0.6</td>
<td>1.574</td>
<td>1.076</td>
<td>0.253</td>
<td>3.033</td>
</tr>
<tr>
<td>0.7</td>
<td>1.879</td>
<td>2.005</td>
<td>0.342</td>
<td>3.108</td>
</tr>
<tr>
<td>0.8</td>
<td>2.363</td>
<td>4.620</td>
<td>0.502</td>
<td>3.206</td>
</tr>
<tr>
<td>0.9</td>
<td>3.390</td>
<td>18.501</td>
<td>0.907</td>
<td>3.360</td>
</tr>
</tbody>
</table>

4.4. Mean Deviation. In statistics, we measure the strength of scatter points in population by mean deviation about the mean and median. Let \( T \) be a random variable coming from FFP distribution with CDF \( F(t) \), mean \( \mu = E(T) \), and \( M = \text{Med} \), then mean deviation about mean and mean deviation about median are given by

\[
 \lambda_1(t) = \int_0^\infty |t - \mu| f(t) dt = 2(\mu F(\mu) - I(\mu)),
\]

\[
 \lambda_2(t) = \int_0^\infty |t - M| f(t) dt = 2MF(M) - 2I(M) - M + \mu,
\]

respectively, where
\begin{equation}
I(z) = \int_0^z t f(t) dt = L_{\text{max}} \kappa \alpha \delta \xi (b + 1)^{(1/\phi) - 1} \Gamma \left( \left( \frac{-1}{\phi} + 1 \right), (b + 1) \left( \frac{\xi}{z} \right)^{\phi} \right). \tag{25}
\end{equation}

where \( \Gamma((1/\phi) + 1), (b + 1)(\xi/z)^{\phi} \) is upper incomplete gamma function [14].

\begin{align}
L(t) &= \frac{1}{\mu} \int_0^q t f(t) dt \\
&= \frac{1}{\mu} \left[ L_{\text{max}} \kappa \alpha \delta \xi (b + 1)^{(1/\phi) - 1} \Gamma \left( \left( \frac{-1}{\phi} + 1 \right), (b + 1) \left( \frac{\xi}{z} \right)^{\phi} \right) \right]. \tag{26}
\end{align}

\begin{align}
B(t) &= \frac{1}{\mu F(t)} \int_0^q t f(t) dt \\
&= \frac{1}{\mu F(t)} \left[ L_{\text{max}} \kappa \alpha \delta \xi (b + 1)^{(1/\phi) - 1} \Gamma \left( \left( \frac{-1}{\phi} + 1 \right), (b + 1) \left( \frac{\xi}{q} \right)^{\phi} \right) \right].
\end{align}

where \( q = F^{-1}(t) \).
5. Parameter Estimation

Let \( T_1, T_2, \ldots, T_n \) be a random sample of size \( n \) from FFP distribution with parameters \( (\xi, \kappa, \delta, \beta, \phi) \). The log-likelihood function, say \( l \), can be written as

\[
\ell(t, \theta) = n \ln \beta + n \ln \kappa + n \ln \delta + n \ln \phi - (\phi + 1) \sum_{i=1}^{n} \ln t_i + n \phi \ln \xi - \sum_{i=1}^{n} \left( \frac{\xi}{t_i} \right)^{\phi} + (\kappa - 1) \sum_{i=1}^{n} \ln \left( 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right) - n \ln(\delta - 1) + (\delta - 1) \sum_{i=1}^{n} \ln \left[ 1 - \left( 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right) \right] + \beta \sum_{i=1}^{n} \left[ 1 - \left( 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right) \right]^{\delta}.
\]

To obtain MLEs, we will take partial derivative with respect to \( \xi, \kappa, \delta, \beta, \) and \( \phi \), respectively, and equating to zero. The associated components of the score function are as follows:

\[
\frac{\partial \ell}{\partial \xi} = \frac{n \phi}{\xi} - \phi \sum_{i=1}^{n} \left( \frac{\xi}{t_i} \right)^{\phi-1} \left( \frac{1}{t_i} \right) + \phi (\kappa - 1) \sum_{i=1}^{n} \frac{e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}}}{\left( 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right)} \left( \frac{1}{t_i} \right) - \phi \kappa (\delta - 1) \sum_{i=1}^{n} \frac{\left( 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right)^{\kappa-1}}{\left[ 1 - \left( 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right) \right]^{\delta}}
\]

\[
e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \left( \frac{\xi}{t_i} \right)^{\phi-1} \left( \frac{1}{t_i} \right) + \phi \beta \delta \kappa \sum_{i=1}^{n} \left[ 1 - \left( 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right) \right] \left( 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right)^{\kappa-1} \left( \frac{\xi}{t_i} \right)^{\phi-1} \left( \frac{1}{t_i} \right) = 0
\]

\[
\frac{\partial \ell}{\partial \kappa} = \frac{n}{\kappa} + \sum_{i=1}^{n} \ln \left[ 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right] + (\delta - 1) \sum_{i=1}^{n} \frac{n \left( 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right)^{\kappa} \ln \left[ 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right]}{\left[ 1 - \left( 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right) \right]^{\delta-1}}
\]

\[
\cdot \left( 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right)^{\kappa} \ln \left[ 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right] = 0
\]

\[
\frac{\partial \ell}{\partial \delta} = \frac{n}{\delta} + \sum_{i=1}^{n} \ln \left[ 1 - \left( 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right) \right] - \beta \sum_{i=1}^{n} \left[ 1 - \left( 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right) \right]^{\kappa} \delta \ln \left[ 1 - \left( 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right) \right] = 0
\]

\[
\frac{\partial \ell}{\partial \beta} = \frac{n}{\beta} - \frac{n}{\beta e^{\beta-1}} + \sum_{i=1}^{n} \left[ 1 - \left( 1 - \exp \left\{ -\left(\frac{\xi}{t_i}\right)^{\phi} \right\} \right) \right]^{\kappa \delta}
\]

\[
= 0 \frac{\partial \ell}{\partial \phi} = \frac{n}{\phi} + n \ln \xi - \sum_{i=1}^{n} \ln t_i - \sum_{i=1}^{n} \left( \frac{\xi}{t_i} \right)^{\phi} \ln \left( \frac{\xi}{t_i} \right) + (\kappa - 1) \sum_{i=1}^{n} \frac{e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}}}{\left( 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right)} \left( \frac{\xi}{t_i} \right)^{\phi} \ln \left( \frac{\xi}{t_i} \right)
\]

\[
- \kappa (\delta - 1) \sum_{i=1}^{n} \frac{\left( 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right)^{\kappa-1}}{\left[ 1 - \left( 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right) \right]^{\delta-1}} \left( 1 - e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \right)^{\kappa-1}
\]

\[
e^{-\left(\frac{t_i}{\xi}\right)^{\kappa}} \left( \frac{\xi}{t_i} \right)^{\phi} \ln \left( \frac{\xi}{t_i} \right) = 0.
\]
The maximum likelihood estimate (MLE) of $\theta$, say $\hat{\theta}$, is obtained by solving the nonlinear system

$$U_n(\theta) = [ (\partial \ell / \partial \kappa), (\partial \ell / \partial \delta), (\partial \ell / \partial \phi), (\partial \ell / \partial \xi), (\partial \ell / \partial \beta) ]^T = 0.$$  

These equations cannot be solved analytically, and statistical software can be used to solve them numerically via iterative methods. We can use iterative techniques such as a Newton–Raphson type algorithm to obtain the estimate $\hat{\theta}$.

### 6. Simulation Study

We performed a Monte Carlo simulation study to demonstrate the performance of the ML estimates for FFP distribution for different sample sizes. R statistical software is used to obtain the maximum likelihood estimates (MLE’s), their standard errors, bias, and root mean square error (RMSE). The experiment is repeated 1000 times with sample size $n = 30, 50, 75, 100, 300, \text{and} 500$. In each trial, the estimates are obtained through maximum likelihood method of estimation using the Newton–Raphson method for obtaining approximate solution. For this purpose, we considered 0.2, 0.5, 1.5, 0.2, and 0.12 for parameters $(\kappa, \delta, \phi, \xi, \beta)$ as initial guess (assumed true values) subject to the condition that these values must be positive. The means of ML estimates along with standard error, bias, and root mean square error are computed.

From all results of Table 3, it can be noticed that

1. RMSE decreases as sample size increases
2. Biases decreases as sample size increase
3. Estimates of the unknown parameters are closer to true values as sample size increases

### 7. Application

To demonstrate the superiority and applicability of the FFP distribution as compared with the other distributions, namely, Exponentiated Generalized Frechet (EGF), Generalized Frechet (GF), and Frechet (F) distribution, an application of water purification data is used. The ML estimates, standard errors, and goodness of fit measures, i.e., AIC, BIC, Kolmogorov–Smirnov (KS), Cramer Von Mises ($W^*$), and Anderson-Darling ($A^*$) are computed for the four models.

#### 7.1. Vinyl Chloride Data

The data consist of 34 observations and were analyzed by Bhauamik et al. [16] for testing the parameters of Gamma distribution. The data are about vinyl chloride (in mg/l) which was obtained from clean-up gradient monitoring wells. The observations are given in Table 4. The maximum likelihood estimates of the model parameters, test statistics, and corresponding p-values are presented in Tables 5 and 6. The suggested model has the

<table>
<thead>
<tr>
<th>Sample size</th>
<th>Mean</th>
<th>Standard error</th>
<th>Bias</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa$</td>
<td>0.2931</td>
<td>0.5297</td>
<td>0.0931</td>
<td>0.1658</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.7003</td>
<td>0.8868</td>
<td>0.2003</td>
<td>0.2945</td>
</tr>
<tr>
<td>$\phi$</td>
<td>1.6472</td>
<td>0.4289</td>
<td>0.14728</td>
<td>0.3011</td>
</tr>
<tr>
<td>$\xi$</td>
<td>0.5763</td>
<td>1.4711</td>
<td>0.3763</td>
<td>0.4367</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.2423</td>
<td>0.6480</td>
<td>0.1214</td>
<td>0.2458</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.2585</td>
<td>0.2122</td>
<td>0.0585</td>
<td>0.1160</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.6515</td>
<td>0.3682</td>
<td>0.1515</td>
<td>0.2193</td>
</tr>
<tr>
<td>$\phi$</td>
<td>1.5672</td>
<td>0.3074</td>
<td>0.0673</td>
<td>0.2056</td>
</tr>
<tr>
<td>$\xi$</td>
<td>0.4592</td>
<td>0.5871</td>
<td>0.2592</td>
<td>0.3057</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.2140</td>
<td>0.6954</td>
<td>0.0930</td>
<td>0.2124</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.2579</td>
<td>0.1996</td>
<td>0.0579</td>
<td>0.1033</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.6423</td>
<td>0.2707</td>
<td>0.1423</td>
<td>0.1906</td>
</tr>
<tr>
<td>$\phi$</td>
<td>1.5503</td>
<td>0.2352</td>
<td>0.0503</td>
<td>0.1739</td>
</tr>
<tr>
<td>$\xi$</td>
<td>0.4120</td>
<td>0.4144</td>
<td>0.2120</td>
<td>0.2600</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.1775</td>
<td>0.4606</td>
<td>0.0565</td>
<td>0.1763</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.2425</td>
<td>0.1842</td>
<td>0.0426</td>
<td>0.0901</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.6154</td>
<td>0.2228</td>
<td>0.1154</td>
<td>0.1599</td>
</tr>
<tr>
<td>$\phi$</td>
<td>1.5246</td>
<td>0.2128</td>
<td>0.0247</td>
<td>0.1475</td>
</tr>
<tr>
<td>$\xi$</td>
<td>0.3996</td>
<td>0.3715</td>
<td>0.1995</td>
<td>0.2357</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.1689</td>
<td>0.3997</td>
<td>0.0479</td>
<td>0.1511</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.2263</td>
<td>0.1422</td>
<td>0.0263</td>
<td>0.0575</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.5792</td>
<td>0.1539</td>
<td>0.0782</td>
<td>0.1023</td>
</tr>
<tr>
<td>$\phi$</td>
<td>1.4780</td>
<td>0.1282</td>
<td>-0.0220</td>
<td>0.0807</td>
</tr>
<tr>
<td>$\xi$</td>
<td>0.3548</td>
<td>0.3357</td>
<td>0.1548</td>
<td>0.1745</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.1388</td>
<td>0.2562</td>
<td>0.0178</td>
<td>0.0988</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.2138</td>
<td>0.1096</td>
<td>0.0138</td>
<td>0.0413</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.5652</td>
<td>0.1237</td>
<td>0.0652</td>
<td>0.0762</td>
</tr>
<tr>
<td>$\phi$</td>
<td>1.4723</td>
<td>0.0922</td>
<td>-0.0277</td>
<td>0.0553</td>
</tr>
<tr>
<td>$\xi$</td>
<td>0.3290</td>
<td>0.3143</td>
<td>0.1290</td>
<td>0.1414</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.1102</td>
<td>0.1598</td>
<td>-0.0108</td>
<td>0.0605</td>
</tr>
</tbody>
</table>
The lowest values of the test statistics which shows that the FFP distribution is more suitable to model the water purification data. Figure 3 describes the theoretical and empirical PDF, CDF, and TTT plots of the data. Figure 3 represents the graphs of estimated PDF and CDF for these data. Figure 4 presents the TTT plot for these data showing a decreasing tendency of the hazard function. Thus, according to all these graphs and tests, the FFP model is a good fit for this data set.

The observations are given in Table 4. The maximum likelihood estimates of the model parameters, test statistics, and corresponding p-values are presented in Tables 5 and 6. The suggested model has the lowest values of the test statistics showing that the FFP distribution is more suitable to model the water purification data. Figure 3 describes the theoretical and empirical PDF, CDF, and TTT plots of the data. Figure 3 represents the graphs of estimated PDF and CDF for these data. Figure 4 presents the TTT plot for these data showing a decreasing tendency of hazard function. Thus, according to all these graphs and tests, the FFP model is good fit for this data set.

---

### Table 4: Data set of vinyl chloride.

<table>
<thead>
<tr>
<th>Data set</th>
<th>8.0</th>
<th>6.8</th>
<th>5.3</th>
<th>5.1</th>
<th>4.0</th>
<th>3.2</th>
<th>2.9</th>
<th>2.7</th>
<th>2.5</th>
<th>2.4</th>
<th>2.3</th>
<th>2.0</th>
<th>2.0</th>
<th>1.8</th>
<th>1.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2</td>
<td>1.2</td>
<td>1.1</td>
<td>1.0</td>
<td>0.9</td>
<td>0.9</td>
<td>0.8</td>
<td>0.6</td>
<td>0.6</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.1</td>
</tr>
<tr>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

---

### Table 5: Estimates and standard error for vinyl chloride.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>$\kappa$</th>
<th>$\xi$</th>
<th>$\delta$</th>
<th>$\phi$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFP</td>
<td>47.4086 (0.3821)</td>
<td>77.6133 (0.0584)</td>
<td>0.1523 (0.0708)</td>
<td>0.509 (0.0503)</td>
<td>0.0202 (1.2049)</td>
</tr>
<tr>
<td>EGF</td>
<td>4.2204 (2.8717)</td>
<td>7.4481 (8.7288)</td>
<td>0.3552 (0.493)</td>
<td>0.6107 (0.2062)</td>
<td>—</td>
</tr>
<tr>
<td>GF</td>
<td>—</td>
<td>2.3347 (27.533)</td>
<td>0.3099 (3.2183)</td>
<td>0.8804 (0.1093)</td>
<td>—</td>
</tr>
<tr>
<td>F</td>
<td>—</td>
<td>0.6174 (0.278)</td>
<td>—</td>
<td>0.8803 (0.1093)</td>
<td>—</td>
</tr>
</tbody>
</table>

---

### Table 6: $(\tilde{l}), \text{AIC}, \text{BIC}, W^*, A^*, \text{KS},$ and $p$-value for the vinyl chloride data.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>$\tilde{l}$</th>
<th>AIC</th>
<th>BIC</th>
<th>$A^*$</th>
<th>$W^*$</th>
<th>KS</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFP</td>
<td>-54.3428</td>
<td>118.6856</td>
<td>123.3174</td>
<td>0.1687</td>
<td>0.0223</td>
<td>0.07838</td>
<td>0.9851</td>
</tr>
<tr>
<td>EGF</td>
<td>-55.64101</td>
<td>119.2820</td>
<td>124.3059</td>
<td>0.3291</td>
<td>0.0439</td>
<td>0.0919</td>
<td>0.9360</td>
</tr>
<tr>
<td>GF</td>
<td>-58.6265</td>
<td>123.2532</td>
<td>127.8323</td>
<td>0.7721</td>
<td>0.1027</td>
<td>0.1134</td>
<td>0.7749</td>
</tr>
<tr>
<td>Frechet</td>
<td>-58.6265</td>
<td>121.2532</td>
<td>124.3059</td>
<td>0.7716</td>
<td>0.1026</td>
<td>0.1135</td>
<td>0.7749</td>
</tr>
</tbody>
</table>
8. Conclusion

This paper introduced a novel probability model for modeling the water purification data using vinyl chloride. The proposed model is called Flexible Frechet Poisson (FFP) distribution. Various statistical properties are derived such as quantile function, survival function, hazard rate function moments, and mean deviation. The maximum likelihood procedure was applied to evaluate the model parameters, with precision supported by a simulation study. Using the application of water data, the suggested model is compared with three other existing models such as EGF, GF, and F and test for goodness of fit. The results showed that our model provides a better fit as compared with other existing models.

Data Availability

The data set is taken from the literature and is given in a paper with the reference at the end.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References