

## RESEARCH ARTICLE

# A modified truncated distribution for modeling the heavy tail, engineering and environmental sciences data

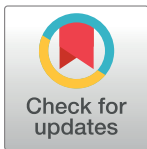
Ahtasham Gul<sup>1,2,3\*</sup>, Muhammad Mohsin<sup>1</sup>, Muhammad Adil<sup>2</sup>, Mansoor Ali<sup>3</sup>

**1** Department of Statistics, COMSATS University Islamabad, Lahore Campus, Lahore, Pakistan, **2** Pakistan Bureau of Statistics, Islamabad, Pakistan, **3** Ministry of PD&SI, Islamabad, Pakistan

\* [ahtashamgul@gmail.com](mailto:ahtashamgul@gmail.com)

## Abstract

Truncated models are imperative to efficiently analyze the finite data that we observe in almost all the real life situations. In this paper, a new truncated distribution having four parameters named Weibull-Truncated Exponential Distribution (W-TEXPD) is developed. The proposed model can be used as an alternative to the Exponential, standard Weibull and shifted Gamma-Weibull and three parameter Weibull distributions. The statistical characteristics including cumulative distribution function, hazard function, cumulative hazard function, central moments, skewness, kurtosis, percentile and entropy of the proposed model are derived. The maximum likelihood estimation method is employed to evaluate the unknown parameters of the W-TEXPD. A simulation study is also carried out to assess the performance of the model parameters. The proposed probability distribution is fitted on five data sets from different fields to demonstrate its vast application. A comparison of the proposed model with some extant models is given to justify the performance of the W-TEXPD.



## OPEN ACCESS

**Citation:** Gul A, Mohsin M, Adil M, Ali M (2021) A modified truncated distribution for modeling the heavy tail, engineering and environmental sciences data. PLoS ONE 16(4): e0249001. <https://doi.org/10.1371/journal.pone.0249001>

**Editor:** Feng Chen, Tongji University, CHINA

**Received:** December 23, 2020

**Accepted:** March 9, 2021

**Published:** April 6, 2021

**Copyright:** © 2021 Gul et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper.

**Funding:** The author(s) received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

## 1 Introduction

Truncated probability models are efficiently used when stochastic variable is confined in some domains. They are required in almost every field like astronomy, epidemiology, biometry, engineering and economy. For instance, the government is interested to know the population of families who are living in New York city having monthly income more than 50,000 US dollars. Another example is recruitment of police officials who meet the minimum prerequisite academic qualification. In engineering, the measurements are taken by using a detector which detects the signals above a specific limit and the weak signals are not taken into account. In all the above situations we need truncated probability distributions to model them.

Weibull and Exponential distributions are immensely utilized in reliability and lifetime analysis due to their simplicity and easy mathematical manipulations. The Weibull distribution is generated by the Swedish physicist [1]. It is commonly used for modeling reliability, electrical engineering [2], mechanical engineering [3], life time and environmental sciences data due to wide-variety of shapes. A considerable literature discussing the methods of estimation of Weibull parameters is given by [4, 5]. [6] stating that Weibull distribution becomes

reversed J-shaped, exponential and bell shaped for the shape parameter  $<$ ,  $=$  and  $>$  0 respectively. A comprehensive account on truncated Weibull distribution is given by [7, 8]. [9] fits the truncated Weibull distribution in different areas like to analyse the diameter of trees by truncating data at a specific threshold level and to infer the height of small trees. [10] studies the method of moments to compute the moment expression for two parameters, three-parameters and truncated (left, right and doubly) Weibull distributions. Exponential distribution is also famous for modeling the data due to availability of good estimators, and its nice mathematical properties (e.g. being memory less). [11] defines the maximum likelihood estimator of scale parameter for Exponential distribution.

[12] computes the parameter estimates of truncated Gamma probability density function (pdf). [13] distinguishes the worth of truncated probability density function in hydrology by computing the truncated moment expressions (TMEs) as well as complete moments of different densities and notify that complete moments are regarded as a special case of truncated moment expressions. [14] utilizes both skew-Cauchy (SK-CD) and truncated skew-Cauchy (TSK-CD) probability functions for modeling the exchange rate between the U.K pound sterling and the U.S dollar from 1800 to 2003 and verdicts that TSK-CD is a better probability function to model the data set contrary to SK-CD. [15] studies the truncated version of the Birnbaum-Saunders (BS) distribution to enhance a forecasting of actuarial model, specifically for modeling data regarding insurance payments that establish a deduction.

In the field of hydrology, [16] uses the generalized exponential (GE) distribution to study the flood frequency for Polish Rivers. [17] employs Weibull density function to study accrual failure detector and calls it “Weibull Distribution Failure Detector for Cloud Computing”. [18] introduces a new generalized form of Weibull probability model i.e. “Alpha logarithmic transformed Weibull distribution” (ALTW) to model the failure time of turbocharger of engine.

In engineering, [19] fitted the micro-level spatial joint and macro-level model with conditional autoregressive (CAR) to analyze the zonal crash using three years urban highway data in the USA. They conclude that micro-level model better fits the data. [20] proposes a new Bayesian Spatio-temporal model to study the association between frequency of free way incidence and other risk factors. [21] considers the mixed logit model to identify the main factors of single and multiple vehicle accidents. Similarly, in another study [22] applies the mixed logit model to analyze the significant factors of single vehicle (SV) and multiple vehicle (MV) accidents by using the 10 years truck drivers data at rural highway of the USA. [23] develops Weibull-Lindely distribution by compounding of two distributions and highlights its worth by fitting it on three medica data sets. [24] introduces U-statistics for Weibull distribution parameters and compares it with nine parameter estimation techniques.

Some distinct characteristics motivated us to demonstrate the W-TEXPD like: (i) it is distinctive by the induction of a new scale parameter obtained from the new truncated transformed distribution along with the usual induction of location parameter; (ii) the W-TEXPD shows monotonic, non-monotonic and bathtub shaped hazard rates which make the W-TEXPD a better model than those lifetime models that only demonstrate constant or monotonically increasing/decreasing hazard rates; (iii) it can be viewed that various known lifetime classical models are the special cases of W-TEXPD; (iv) it can be observed that W-TEXPD is appropriate for fitting the scattered, skewed (spread) and/or heavy tailed (flat curved) data which may not be appropriately fitted by other typical probability density functions; and (v) the results achieved by Monte Carlo simulation study for different sample sizes reveals the stability of the model parameters. Finally, we intend to find that how well W-TEXPD performs as compared to several renowned classical lifetime models by using five data sets having skewed and heavy tailed data.

The manuscript is sorted as: In Section 2, W-TEXPD is described and its characteristics such as hazard function, cumulative hazard function, raw and central moments, skewness, kurtosis, Shannon’s entropy and order statistics are derived. In Section 3, Maximum likelihood estimates of the model parameters are obtained. In Section 4, Monte Carlo simulation study is performed to examine the performance of W-TEXPD for different choices of the model parameters. In Section 5, the feasibility of the proposed model is studied by fitting it to the real data sets and comparing with some baseline models. Some concluding remarks are recorded in Section 6.

## 2 Weibull-Truncated Exponential distribution (W-TEXPD)

[25] suggests a new method for generating a family of truncated distributions called T- $X_T$  family of distributions by using a new function given as

$$W(F(x_T)) = -\log \{1 - F(x|x > \tau)\}. \tag{1}$$

Let X be a non-negative random variable truncated on left having probability density function (pdf)  $f(x_T)$  and distribution function (cdf)  $F(x_T)$  on domain  $[\tau, \infty)$ . Also let T be a random variable with pdf  $r(t)$  and cdf  $R(t)$  on interval  $[-\infty \leq a \leq t \leq b \leq \infty)$ .

Then the cdf of T- $X_T$  family of distributions is

$$G(x_T) = \int_{\tau}^{-\log \{1 - F(x|x > \tau)\}} r(t) dt, \tag{2}$$

$$G(x_T) = R[-\log \{1 - F(x_T)\}], \tag{3}$$

where  $R(t)$  is the cdf of random variable T, while the corresponding pdf of T- $X_T$  family of distributions is

$$g(x_T) = r[-\log \{1 - F(x|x > \tau)\}] \frac{f(x|x > \tau)}{1 - F(x|x > \tau)}, \quad x > \tau. \tag{4}$$

$$g(x_T) = r[-\log \{1 - F(x_T)\}] \left\{ \frac{f(x_T)}{1 - F(x_T)} \right\}, \tag{5}$$

$$g(x_T) = r[H(x_T)]h[x_T]. \tag{6}$$

The idea presented in Eq (2) is extended by the method of generating a new family of distributions called T-X family of distributions proposed by [26] which is the extension of Beta Generated distributions originally introduced by [27].

Suppose X be an exponential random variable having density function

$$f(x) = \theta e^{-\theta x}, \tag{7}$$

with corresponding cdf

$$F(x) = 1 - \theta e^{-\theta x}. \tag{8}$$

The T-Truncated Exponential distribution defined by [25] is expressed as:

$$g(x) = \theta r\{\theta(x - a)\}. \tag{9}$$

Let T be the Weibull random variable having pdf

$$r(t) = \frac{\beta}{\alpha^\beta} t^{\beta-1} e^{-(t/\alpha)^\beta}, \quad 0 < t < \infty. \tag{10}$$

The Weibull-Truncated exponential distribution (W-TEXPD) is defined by using Eq (9) as

$$g(x) = \left(\frac{\theta}{\alpha}\right)^\beta \beta (x - \tau)^{\beta-1} e^{-\left\{\frac{\theta(x-\tau)}{\alpha}\right\}^\beta} \quad \tau < x < \infty. \tag{11}$$

Where

$\tau$  ———> Location parameter.

$\theta, \alpha$  ———> Scale parameter.

$\beta$  ———> Shape parameter.

**Some special cases of W-TEXPD**

- W-TEXPD reduces to Exponential distribution for  $\tau = 0$  and  $\theta, \beta = 1$ .
- W-TEXPD reduces to Weibull for  $\tau = 0, \theta = 1$ .
- W-TEXPD reduces to Shifted Gamma-Weibull [28] and three parameter Weibull [29] distribution for  $\theta = 1$ .

Fig 1 displays different shapes of W-TEXPD for different values of the parameters.

**Some characteristics of W-TEXPD are:**

**Lemma 2.0.1.** *The hazard function of W-TEXPD is*

$$h(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{R(t)}.$$

$$h(t) = \beta \left(\frac{\theta}{\alpha}\right)^\beta (t - \tau)^{\beta-1}. \tag{12}$$

Fig 2 highlights that the W-TEXPD can model both monotonically and non-monotonically hazard rate shapes with different values of the parameters.

**Lemma 2.0.2.** *The cumulative hazard function of W-TEXPD is computed as*

$$H(t) = \int_0^x h(t) dt.$$

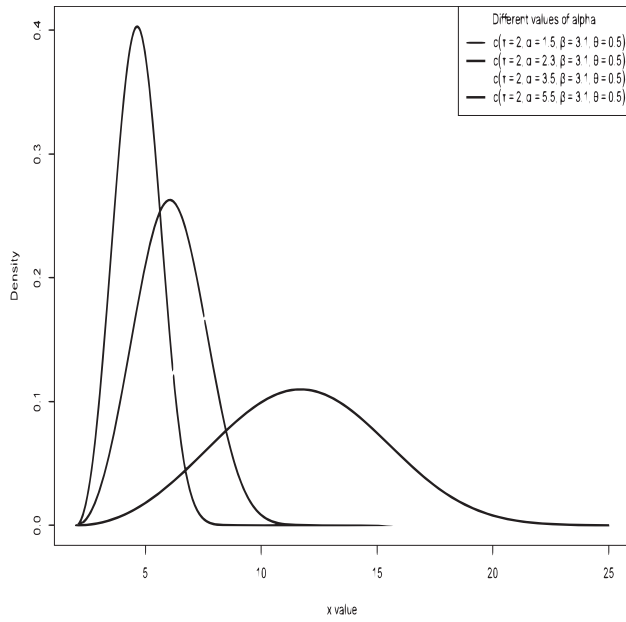
$$H(t) = \left(\frac{\theta}{\alpha}\right)^\beta (x - \tau)^\beta, \quad \beta > 0. \tag{13}$$

**Lemma 2.0.3.** *The  $p^{th}$  percentile of W-TEXPD is given by*

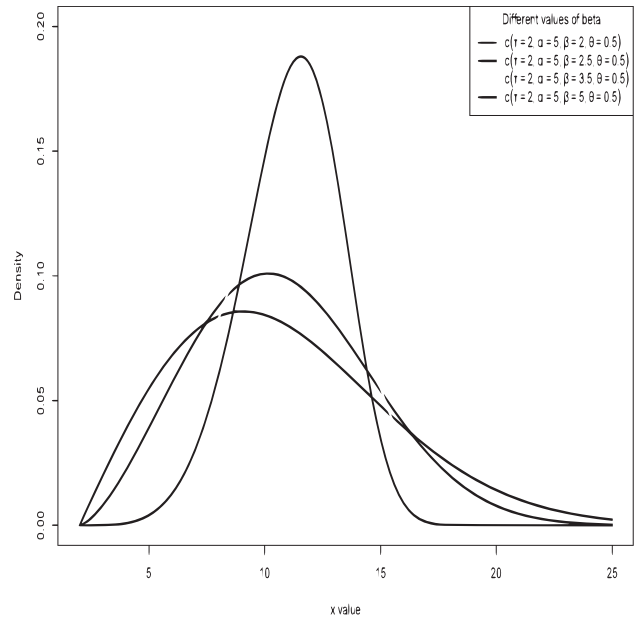
$$G(x) = P.$$

$$1 - e^{-\left\{\frac{\theta(x-\tau)}{\alpha}\right\}^\beta} = P,$$

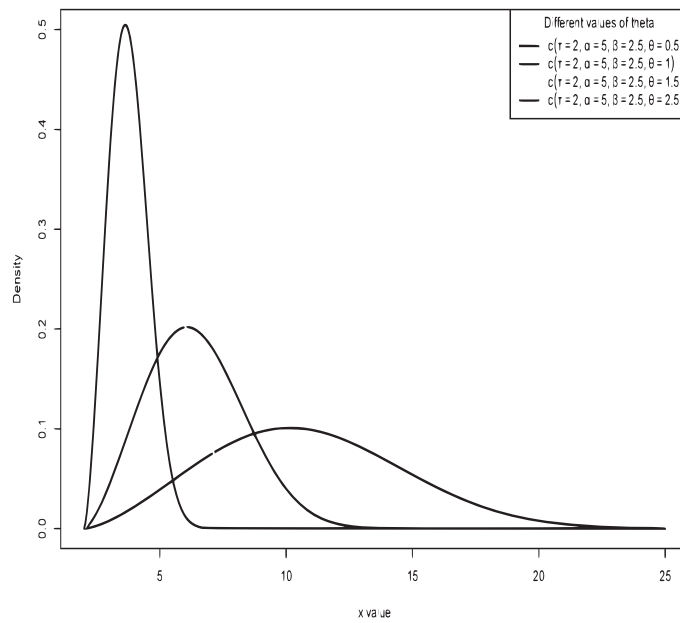
$$x = \tau + \frac{\alpha}{\theta} \log(1 - P). \tag{14}$$



(a) different values of  $\alpha$



(b) different values of  $\beta$



(c) different values of  $\theta$

**Fig 1. PDF of W-TEXPD for different values of  $\alpha$ ,  $\beta$  and  $\theta$ .**

<https://doi.org/10.1371/journal.pone.0249001.g001>

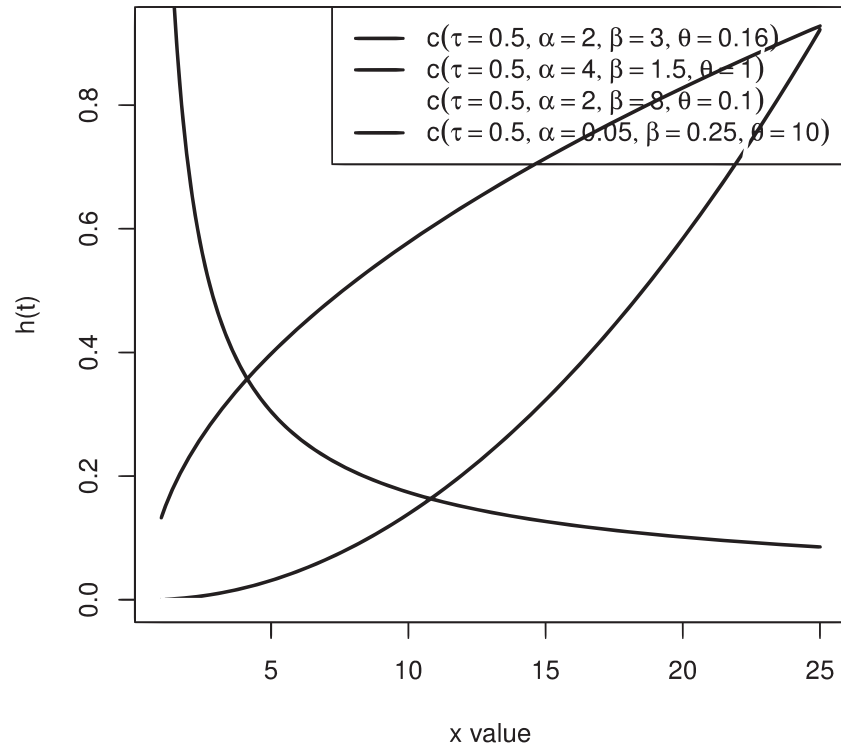


Fig 2.  $h(t)$  of W-TEXPD for different values of  $\alpha, \beta$  and  $\theta$ .

<https://doi.org/10.1371/journal.pone.0249001.g002>

**Lemma 2.0.4.** *The first four raw moments of W-TEXPD are given by*

$$E(X) = \int_{\tau}^{\infty} xg(x)dx, \tag{15}$$

$$E(X) = \tau + \frac{\alpha}{\theta} \Gamma\left(\frac{1}{\beta} + 1\right).$$

$$E(X^2) = \beta \left(\frac{\theta}{\alpha}\right)^{\beta} \int_{\tau}^{\infty} x^2(x - \tau)^{\beta-1} e^{-\left\{\frac{\theta(x-\tau)}{\alpha}\right\}^{\beta}} dx, \tag{16}$$

$$E(X^2) = \tau^2 + \left(\frac{\alpha}{\theta}\right)^2 \Gamma\left(\frac{2}{\beta} + 1\right) + 2\frac{\tau\alpha}{\theta} \Gamma\left(\frac{1}{\beta} + 1\right).$$

$$E(X^3) = \int_{\tau}^{\infty} x^3g(x)dx, \tag{17}$$

$$E(X^3) = \tau^3 + \left(\frac{\alpha}{\theta}\right)^3 \Gamma\left(\frac{3}{\beta} + 1\right) + 3\tau\left(\frac{\alpha}{\theta}\right)^2 \Gamma\left(\frac{2}{\beta} + 1\right) + 3\tau^2\left(\frac{\alpha}{\theta}\right) \Gamma\left(\frac{1}{\beta} + 1\right).$$

$$E(X^4) = \tau^4 + 4\tau^3\left(\frac{\alpha}{\theta}\right) \Gamma\left(\frac{1}{\beta} + 1\right) + 6\tau^2\left(\frac{\alpha}{\theta}\right)^2 \Gamma\left(\frac{2}{\beta} + 1\right) + 4\tau\left(\frac{\alpha}{\theta}\right)^3 \Gamma\left(\frac{3}{\beta} + 1\right) + \left(\frac{\alpha}{\theta}\right)^4 \Gamma\left(\frac{4}{\beta} + 1\right). \tag{18}$$

**Lemma 2.0.5.** *The first four central moments of W-TEXPD are given by*

$$\mu_1 = 0. \tag{19}$$

$$\begin{aligned} \mu_2 &= \text{Var}(x) = E(x^2) - \{E(x)\}^2, \\ \text{Var}(x) &= \left(\frac{\alpha}{\theta}\right)^2 \left[ \Gamma\left(\frac{2}{\beta} + 1\right) - \left\{ \Gamma\left(\frac{1}{\beta} + 1\right) \right\}^2 \right]. \end{aligned} \tag{20}$$

$$\begin{aligned} \mu_3 &= \mu_3' - 3\mu_1'\mu_2' + 2(\mu_1')^3, \\ \mu_3 &= \left(\frac{\alpha}{\theta}\right)^3 \Gamma\left(\frac{3}{\beta} + 1\right) + \Gamma\left(\frac{1}{\beta} + 1\right) \left[ 6\tau^2 \left(\frac{\theta}{\alpha} - \frac{\alpha}{\theta}\right) - 3\tau^2 \left(\frac{\alpha}{\theta}\right) + 2\left(\frac{\theta}{\alpha}\right)^2 + 6\tau \left(\frac{\theta}{\alpha}\right) \right]. \end{aligned} \tag{21}$$

$$\begin{aligned} \mu_4 &= \mu_4' - 4\mu_1'\mu_3' + 6(\mu_1')^2\mu_2' - 3(\mu_1')^4, \\ \mu_4 &= \left[ \begin{aligned} &4\tau^3 \left(\frac{\theta}{\alpha}\right) \Gamma\left(\frac{1}{\beta} + 1\right) - 6\tau^2 \left(\frac{\theta}{\alpha}\right)^2 \Gamma\left(\frac{2}{\beta} + 1\right) - 12\tau^2 \left(\frac{\theta}{\alpha}\right) \Gamma\left(\frac{1}{\beta} + 1\right) - 4\tau \Gamma\left(\frac{3}{\beta} + 1\right) \\ &\quad + 12\tau \left(\frac{\alpha}{\theta}\right)^2 \left\{ \Gamma\left(\frac{1}{\beta} + 1\right) \right\}^3 \\ &- 6\tau^2 \left(\frac{\alpha}{\theta}\right)^2 \left\{ \Gamma\left(\frac{2}{\beta} + 1\right) + \left\{ \Gamma\left(\frac{1}{\beta} + 1\right) \right\}^2 - \left\{ \Gamma\left(\frac{1}{\beta} + 1\right) \right\}^2 \Gamma\left(\frac{2}{\beta} + 1\right) \right\} \\ &+ 8\tau^3 \left(\frac{\alpha}{\theta}\right) \Gamma\left(\frac{1}{\beta} + 1\right) - \left(\frac{\alpha}{\theta}\right)^4 \Gamma\left(\frac{1}{\beta} + 1\right) \left[ 4\Gamma\left(\frac{3}{\beta} + 1\right) + 3\left\{ \Gamma\left(\frac{1}{\beta} + 1\right) \right\}^3 \right] \end{aligned} \right]. \end{aligned} \tag{22}$$

**Lemma 2.0.6.** *The skewness and kurtosis of W-TEXPD are defined as*

$$\begin{aligned} \text{Skewness} &= \frac{\mu_3}{(\mu_2)^{3/2}}, \\ \text{Skewness} &= \frac{\left(\frac{\alpha}{\theta}\right)^3 \Gamma\left(\frac{3}{\beta} + 1\right) + \Gamma\left(\frac{1}{\beta} + 1\right) \left[ 6\tau^2 \left(\frac{\theta}{\alpha} - \frac{\alpha}{\theta}\right) - 3\tau^2 \left(\frac{\alpha}{\theta}\right) + 2\left(\frac{\theta}{\alpha}\right)^2 + 6\tau \left(\frac{\theta}{\alpha}\right) \right]}{\left(\frac{\alpha}{\theta}\right)^3 \left[ \Gamma\left(\frac{2}{\beta} + 1\right) - \left\{ \Gamma\left(\frac{1}{\beta} + 1\right) \right\}^2 \right]^{3/2}}. \end{aligned} \tag{23}$$

$$\begin{aligned} \text{Kurtosis} &= \frac{\mu_4}{(\mu_2)^2}, \\ \text{Kurtosis} &= \frac{\left[ \begin{aligned} &4\tau^3 \left(\frac{\theta}{\alpha}\right) \Gamma\left(\frac{1}{\beta} + 1\right) - 6\tau^2 \left(\frac{\theta}{\alpha}\right)^2 \Gamma\left(\frac{2}{\beta} + 1\right) - 12\tau^2 \left(\frac{\theta}{\alpha}\right) \Gamma\left(\frac{1}{\beta} + 1\right) \\ &- 4\tau \Gamma\left(\frac{3}{\beta} + 1\right) + 12\tau \left(\frac{\alpha}{\theta}\right)^2 \left\{ \Gamma\left(\frac{1}{\beta} + 1\right) \right\}^3 - 6\tau^2 \left(\frac{\alpha}{\theta}\right)^2 \Gamma\left(\frac{2}{\beta} + 1\right) \\ &+ \left\{ \Gamma\left(\frac{1}{\beta} + 1\right) \right\}^2 - \left\{ \Gamma\left(\frac{1}{\beta} + 1\right) \right\}^2 \Gamma\left(\frac{2}{\beta} + 1\right) + 8\tau^3 \left(\frac{\alpha}{\theta}\right) \Gamma\left(\frac{1}{\beta} + 1\right) \end{aligned} \right]}{\left(\frac{\alpha}{\theta}\right)^4 \left[ \Gamma\left(\frac{2}{\beta} + 1\right) - \left\{ \Gamma\left(\frac{1}{\beta} + 1\right) \right\}^2 \right]^2}. \end{aligned} \tag{24}$$

**Theorem 2.1.** *Let  $X_T$  be a stochastic variable following W-TEXPD, then Shannon entropy is given by*

$$\eta_{x_T} = 1 + \frac{(\beta - 1)}{\beta} \left\{ C + \beta \ln \left( \frac{\theta}{\alpha} \right) \right\} - \ln \left\{ \beta \left( \frac{\theta}{\alpha} \right)^\beta \right\}.$$

*Proof.* The Shannon entropy of a random variable  $X_T$  is a measure of uncertainty given as

$$\eta_{x_T} = E\{\ln g(x_T)\} = \int_{\tau}^{\infty} \{-\ln g(x_T)\} g(x_T) dx_T,$$

$$\eta_{x_T} = - \left[ \begin{aligned} &\ln \left\{ \beta \left( \frac{\theta}{\alpha} \right)^{\beta} \right\} \int_{\tau}^{\infty} g(x_T) dx_T + \int_{\tau}^{\infty} (\beta - 1) \ln(x_T - \tau) g(x_T) dx_T \\ &- \left( \frac{\theta}{\alpha} \right)^{\beta} \int_{\tau}^{\infty} (x_T - \tau)^{\beta} g(x_T) dx_T \end{aligned} \right],$$

$$\eta_{x_T} = \left[ \begin{aligned} &\left( \frac{\theta}{\alpha} \right)^{\beta} \int_{\tau}^{\infty} (x_T - \tau)^{\beta} g(x_T) dx_T - (\beta - 1) \int_{\tau}^{\infty} \ln(x_T - a) g(x_T) dx_T \\ &- \ln \left\{ \beta \left( \frac{\theta}{\alpha} \right)^{\beta} \right\} \int_{\tau}^{\infty} g(x_T) dx_T \end{aligned} \right],$$

$$\eta_{x_T} = \left( \frac{\theta}{\alpha} \right)^{\beta} I_1 - (\beta - 1) I_2 - \ln \left\{ \beta \left( \frac{\theta}{\alpha} \right)^{\beta} \right\}. \tag{25}$$

$$I_1 = \beta \left( \frac{\theta}{\alpha} \right)^{\beta} \int_a^{\infty} (x_T - a)^{2\beta-1} e^{-\left\{ \frac{\theta(x_T-a)}{\alpha} \right\}^{\beta}} dx_T,$$

let

$$\left\{ \frac{\theta(x_T - a)}{\alpha} \right\}^{\beta} = u,$$

$$I_1 = \left( \frac{\theta}{\alpha} \right)^{-\beta} \int_0^{\infty} u e^{-u} du,$$

$$I_1 = \left( \frac{\alpha}{\theta} \right)^{\beta},$$

$$I_2 = \int_a^{\infty} \ln(x_T - a) g(x_T) dx_T,$$

$$I_2 = \beta \left( \frac{\theta}{\alpha} \right)^{\beta} \int_a^{\infty} \{ \ln(x_T - a) \} (x_T - a)^{\beta-1} e^{-\left( \frac{\theta}{\alpha} \right)^{\beta} (x_T - a)^{\beta}} dx_T,$$

$$(x_T - a)^{\beta} = u,$$



$$\beta \left(\frac{\theta}{\alpha}\right)^\beta \int_a^\infty \{\ln(x - \tau)\} (x - \tau)^{\beta-1} e^{-(\theta/x)^\beta (x-\tau)^\beta} dx = \frac{1}{\beta} \left(\frac{\theta}{\alpha}\right)^\beta \int_0^\infty \ln ue^{-(\theta/x)^\beta u} du,$$

$$I_2 = \frac{1}{\beta} \left(\frac{\theta}{\alpha}\right)^\beta \int_0^\infty \ln ue^{-(\theta/x)^\beta u} du.$$

To solve the above integral, we use the following combinations of logarithms and exponential given by [30] (Jeffrey and Zwillinger, 2007, 7th edition, Eq. (4.331.1), p. 571).

$$\int_0^\infty \ln ue^{-\mu u} du = -\frac{1}{\mu} (C + \ln \mu) \quad [\text{Re } \mu > 0],$$

$$I_2 = -\frac{1}{\beta} \left\{ C + \beta \ln \left(\frac{\theta}{\alpha}\right) \right\}.$$

By using  $I_1$  and  $I_2$ , Eq (25) becomes

$$\eta_{x_T} = 1 + \frac{(\beta - 1)}{\beta} \left\{ C + \beta \ln \left(\frac{\theta}{\alpha}\right) \right\} - \ln \left\{ \beta \left(\frac{\theta}{\alpha}\right)^\beta \right\}.$$

**Theorem 2.2.** The  $r^{th}$  order statistics  $f_{r,n}(x)$  of a random sample of size  $n$  for the W-TEXPD distribution is given by

$$f_{r,n}(x) = \beta \left(\frac{\theta}{\alpha}\right)^\beta (x - \tau)^{\beta-1} \{F(x)\}^{r-1} \{1 - F(x)\}^{n-r+1}.$$

*Proof.* By definition

$$f_{r,n}(x) = C_{r,n} g(x) \{G(x)\}^{r-1} \{1 - G(x)\}^{n-r},$$

or

$$f_{r,n}(x) = C_{r,n} f(x) \{F(x)\}^{r-1} \{1 - F(x)\}^{n-r} \tag{26}$$

where

$$f(x) = \left(\frac{\theta}{\alpha}\right)^\beta \beta (x - a)^{\beta-1} e^{-\left(\frac{\theta(x-a)}{\alpha}\right)^\beta},$$

$$F(x) = 1 - e^{-\left(\frac{\theta(x-a)}{\alpha}\right)^\beta} \tag{27}$$

$$f_{r,n}(x) = \beta \left(\frac{\theta}{\alpha}\right)^\beta (x - \tau)^{\beta-1} \{F(x)\}^{r-1} \{1 - F(x)\}^{n-r+1}.$$

### 3 Estimation of model parameters by using Maximum Likelihood (ML) method

In this Section, we estimate the unknown parameters of W-TEXPD by applying maximum likelihood estimation method as defined by [31]. The log-likelihood function of W-TEXPD is given by:

$$\ln L(\tau, \theta, \alpha, \beta; x) = \ln \prod_{i=1}^n \left[ \beta \left( \frac{\theta}{\alpha} \right)^\beta (x_i - \tau)^{\beta-1} e^{-\left\{ \frac{\theta(x_i - \tau)}{\alpha} \right\}^\beta} \right],$$

$$\ln L(\tau, \theta, \alpha, \beta; x) = \left[ \begin{aligned} n \ln \beta + n \ln \theta - n \beta \ln \alpha + \beta \sum_{i=1}^n \ln (x_i - \tau) - \sum_{i=1}^n \ln (x_i - \tau) \\ - \left( \frac{\theta}{\alpha} \right)^\beta \sum_{i=1}^n (x_i - \tau)^\beta \end{aligned} \right]. \tag{28}$$

Now computing the first partial derivatives of (28) with respect to  $\tau, \theta, \alpha, \beta$  and equating the results to zero, we have

$$\frac{\partial \ln L(\tau, \theta, \alpha, \beta; x)}{\partial \tau} = \text{Min}[x_i], \quad i = 0, 1, 2, \dots, n, \tag{29}$$

$$\frac{\partial \ln L(\tau, \theta, \alpha, \beta; x)}{\partial \theta} = \frac{n\beta}{\theta} - \frac{\beta\theta^{\beta-1}}{\alpha^\beta} \sum_{i=1}^n (x_i - \tau)^\beta = 0, \tag{30}$$

$$\frac{\partial \ln L(\tau, \theta, \alpha, \beta; x)}{\partial \alpha} = \beta\theta^\beta \alpha^{-(\beta+1)} \sum_{i=1}^n (x_i - \tau)^\beta - \frac{n\beta}{\alpha} = 0, \tag{31}$$

$$\frac{\partial \ln L(\tau, \theta, \alpha, \beta; x)}{\partial \beta} = \left[ \begin{aligned} \frac{n}{\beta} + n \ln \theta - n \ln \alpha + \sum_{i=1}^n \ln (x_i - \tau) - \left( \frac{\theta}{\alpha} \right)^\beta \ln \left( \frac{\theta}{\alpha} \right) \sum_{i=1}^n (x_i - \tau)^\beta \\ - \left( \frac{\theta}{\alpha} \right)^\beta \sum_{i=1}^n (x_i - \tau)^\beta \ln (x_i - \tau) \end{aligned} \right] = 0, \tag{32}$$

respectively. Since the Eqs (30) to (32) are not in closed form, we use a well-known iterative method i.e. Newton Raphson to obtain the approximate ML estimates for the parameters  $\theta, \alpha$  and  $\beta$ .

#### 3.1 Asymptotic confidence bounds

It is observed that ML estimates of the unknown parameters  $\theta, \alpha, \beta$  of W-TEXPD are not in closed forms. In this situation, we compute the asymptotic confidence bounds of W-TEXPD based on the asymptotic distribution of the MLE.

The Fisher Information matrix can be used for interval estimation and hypothesis testing. For W-TEXPD, Information matrix is obtained by computing the second partial derivatives of

the Eqs (30) to (32) as:

$$I_n = \begin{pmatrix} I_{\alpha\alpha} & I_{\alpha\beta} & I_{\alpha\theta} \\ I_{\beta\alpha} & I_{\beta\beta} & I_{\beta\theta} \\ I_{\theta\alpha} & I_{\theta\beta} & I_{\theta\theta} \end{pmatrix},$$

the entries of Fisher Information matrix of W-TEXPD are:

$$I_{\alpha\alpha} = \frac{\partial^2 \ln L(\tau, \theta, \alpha, \beta; x)}{\partial \alpha^2} = \frac{n\beta}{\alpha^2} - \frac{\beta(\beta + 1)\theta^\beta}{\alpha^{\beta+2}} \sum_{i=1}^n (x_i - \tau)^\beta. \tag{33}$$

$$I_{\beta\beta} = \frac{\partial^2 \ln L(\tau, \theta, \alpha, \beta; x)}{\partial \beta^2} = \begin{bmatrix} -\frac{n}{\beta^2} - 2\left(\frac{\theta}{\alpha}\right)^\beta \ln\left(\frac{\theta}{\alpha}\right) \sum_{i=1}^n (x_i - \tau)^\beta \\ -\left(\frac{\theta}{\alpha}\right)^\beta \ln\left(\frac{\theta}{\alpha}\right) \sum_{i=1}^n \{\ln(x_i - \tau)\} (x_i - \tau)^\beta \\ -\left(\frac{\theta}{\alpha}\right)^\beta \ln\left(\frac{\theta}{\alpha}\right) \sum_{i=1}^n \{\ln(x_i - \tau)\} (x_i - \tau)^\beta \\ -2\left(\frac{\theta}{\alpha}\right)^\beta \sum_{i=1}^n \{\ln(x_i - \tau)\} (x_i - \tau)^\beta \end{bmatrix}. \tag{34}$$

$$I_{\theta\theta} = \frac{\partial^2 \ln L(\tau, \theta, \alpha, \beta; x)}{\partial \theta^2} = -\frac{n\beta}{\theta^2} - \frac{\beta(\beta - 1)\theta^{\beta-2}}{\alpha^\beta} \sum_{i=1}^n (x_i - \tau)^\beta. \tag{35}$$

$$I_{\theta\alpha} = \frac{\partial \ln L(\tau, \theta, \alpha, \beta; x)}{\partial \theta \partial \alpha} = \frac{\beta^2 \theta^{\beta-1}}{\alpha^{\beta+1}} \sum_{i=1}^n (x_i - \tau)^\beta. \tag{36}$$

$$I_{\theta\beta} = \frac{\partial \ln L(\tau, \theta, \alpha, \beta; x)}{\partial \theta \partial \beta} = - \begin{bmatrix} \theta^{\beta-1} \alpha^{-\beta} \sum_{i=1}^n (x_i - \tau)^\beta + \beta \theta^{\beta-1} \alpha^{-\beta} \ln \theta \sum_{i=1}^n (x_i - \tau)^\beta \\ + \beta \theta^{\beta-1} \alpha^{-\beta} \ln \alpha \sum_{i=1}^n (x_i - \tau)^\beta \\ + \beta \theta^{\beta-1} \alpha^{-\beta} \sum_{i=1}^n (x_i - \tau)^\beta \sum_{i=1}^n \{\ln(x_i - \tau)\} \end{bmatrix}. \tag{37}$$

$$I_{\alpha\beta} = \frac{\partial \ln L(\tau, \theta, \alpha, \beta; x)}{\partial \alpha \partial \beta} = - \begin{bmatrix} \theta^\beta \alpha^{-(\beta+1)} \sum_{i=1}^n (x_i - \tau)^\beta + \beta \theta^\beta \alpha^{-(\beta+1)} \ln \theta \sum_{i=1}^n (x_i - \tau)^\beta \\ + \beta \theta^\beta \alpha^{-(\beta+1)} \ln \alpha \sum_{i=1}^n (x_i - \tau)^\beta \\ + \beta \theta^\beta \alpha^{-(\beta+1)} \sum_{i=1}^n (x_i - \tau)^\beta \sum_{i=1}^n \{\ln(x_i - \tau)\} - \frac{n}{\alpha} \end{bmatrix} = 0. \tag{38}$$

The asymptotic confidence intervals are obtained by using either the approximate normal distribution or the approximate log-normal distribution of the ML estimates  $\hat{g} = (\hat{\alpha}, \hat{\beta}, \hat{\theta})$ . The estimated standard errors of  $\hat{\alpha}$ ,  $\hat{\beta}$  and  $\hat{\theta}$  are expressed as:

$$\sigma(\hat{\alpha}, \hat{\beta}, \hat{\theta}) = \sqrt{\sum_{jj}}, \quad \text{where } \Sigma = [I_n]^{-1}.$$

For instance, the expressions for  $(1 - \xi)100\%$  confidence interval of  $\alpha$  calculated by using the approximate normal distribution and log-normal distribution are

$$\hat{\alpha} \pm \delta_{\xi/2} \sigma(\hat{\alpha}). \tag{39}$$

and

$$\hat{\alpha} \exp \{ \pm \delta_{\xi/2} \sigma(\hat{\alpha}) / \hat{\alpha} \},$$

or

$$\hat{\alpha} e^{\delta_{\xi/2} \sigma(\hat{\alpha}) / \hat{\alpha}} \leq \alpha \leq \hat{\alpha} e^{-\delta_{\xi/2} \sigma(\hat{\alpha}) / \hat{\alpha}}, \tag{40}$$

respectively, where  $\delta_{\xi/2}$  is the  $1 - \delta_{\xi/2}$  percentile of standard normal distribution. The log-normal approximation works well if the standard error of parameters is greater than half of their point estimate.

### 4 Simulation study

The core feature of probability is randomness and uncertainty. The randomness exists in every field of life. Simulation imitates the realization of a random experiment, so that random values are generated (that are deterministic) by using an appropriate model designed on the basis of random experiment. A simple such model can be a probability distribution that is used to sketch a real mechanism that produces values of some quantity of interest.

Here, we carry out Monte Carlo simulation studies to assess the performance of maximum likelihood estimates (MLEs) using the R programming. The Monte Carlo simulations are run 1000 times and in each replication, random sample of size n is drawn from the W-TEXPD ( $\alpha, \beta, \theta$ ). The model parameters are estimated by maximum likelihood method.

Table 1 presents the average point estimates of three parameters with standard errors (SEs), bias and mean square errors (MSEs) for the sample sizes 20, 50, 100 and 200. A fixed seed is used to generate such random numbers implying that all results of these studies can always be exactly replicated.

The assessment is based on a simulation study by applying following steps:

1. Generate one thousand samples of size n each using Eq (14).
2. Compute the MLEs for the one thousand samples, say  $(\hat{\alpha}_i, \hat{\beta}_i, \hat{\theta}_i)$  for  $i = 1, 2, 3, \dots, 1000$ .
3. Compute the SEs of the MLEs for the one thousand samples.
4. Compute the biases and mean square errors by using

$$bias = \frac{1}{1000} \sum_{i=1}^{1000} \hat{\delta}_i - \delta,$$

and

$$MSE = \frac{1}{1000} \sum_{i=1}^{1000} \hat{\sigma}_x^2 + \{bias(\hat{\delta}_i)\}^2$$

respectively.

Table 1 shows that biases and MSEs vary with respect to n. The biases and MSEs for each parameter approaches to zero as sample size increases.

## 5 Real life application

### 5.1 Application 1: Aeronautical engineering

To demonstrate the strength of W-TEXPD, we use Aeronautical Engineering data set to show that the proposed distribution can be a better model than the base line distributions i.e. Weibull, truncated Exponential (TEXPD), Gamma and Exponential distributions. We re-analyse

Table 1. Average estimated values, corresponding SEs (given in parentheses) bias and MSE of model parameters.

Actual	n	Average Estimate (S.E)			Bias			MSE		
		$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$
$\alpha = 0.1$ $\beta = 0.5$ $\theta = 2.0$	20	0.08583 (0.00021)	0.66002 (0.00208)	1.99978 (0.00003)	-0.01417	0.16001	-0.00001	0.00007	0.02991	0.00001
	50	0.09253 (0.00005)	0.72130 (0.00116)	1.99985 (0.00001)	-0.00747	0.22130	-0.00014	0.00005	0.05033	0.00000
	100	0.09603 (0.00002)	0.73558 (0.00103)	1.99999 (0.00004)	-0.00397	0.23558	-0.00007	0.00001	0.05655	0.00001
	200	0.09793 (0.00001)	0.74491 (0.00110)	1.99991 (0.00001)	-0.00207	0.24492	-0.00004	0.00004	0.06119	0.00001
$\alpha = 0.1$ $\beta = 1.0$ $\theta = 2.0$	20	0.09378 (0.00034)	1.03364 (0.00171)	2.00001 (0.00003)	-0.00622	0.03364	0.00001	0.00016	0.00404	0.00001
	50	0.09842 (0.00011)	1.02537 (0.00159)	2.00006 (0.00009)	-0.00158	0.02537	0.00006	0.00001	0.00317	0.00000
	100	0.09942 (0.00003)	1.02011 (0.00125)	2.00000 (0.00002)	-0.00057	0.02011	0.00000	0.00001	0.00196	0.00000
	200	0.09979 (0.00001)	1.01700 (0.00102)	2.00000 (0.00000)	-0.00002	0.01700	0.00000	0.00000	0.00133	0.00000
$\alpha = 0.1$ $\beta = 1.0$ $\theta = 5.0$	20	0.09399 (0.00032)	1.03973 (0.00196)	5.00016 (0.00001)	-0.00060	0.03973	0.00016	0.00014	0.00542	0.00001
	50	0.09857 (0.00001)	1.02870 (0.00171)	5.00001 (0.00000)	-0.00142	0.02871	0.00001	0.00001	0.00377	0.00000
	100	0.09950 (0.00003)	1.02194 (0.00130)	5.00000 (0.00000)	-0.00005	0.02194	0.00000	0.00001	0.00218	0.00000
	200	0.09980 (0.00001)	1.01827 (0.00105)	5.00000 (0.00000)	-0.00002	0.01827	0.00000	0.00000	0.00144	0.00000
$\alpha = 2.5$ $\beta = 1.0$ $\theta = 2.0$	20	2.50850 (0.00301)	1.00439 (0.00096)	2.00047 (0.00031)	0.00850	0.00439	0.00047	0.00915	0.00095	0.00010
	50	2.50775 (0.00126)	1.01231 (0.00132)	2.00001 (0.00000)	0.00775	0.01231	0.00001	0.00166	0.00189	0.00000
	100	2.50491 (0.00079)	1.01451 (0.00130)	2.00000 (0.00000)	0.00491	0.01451	0.00000	0.00065	0.00191	0.00000
	200	2.50192 (0.00015)	1.01335 (0.00097)	2.00000 (0.00000)	0.00192	0.01335	0.00000	0.00002	0.00113	0.00000

<https://doi.org/10.1371/journal.pone.0249001.t001>

**Table 2. Descriptive statistics of failure times of 31 air conditioners of airplane.**

Min.	Q1	Q2	Mean	Q3	Max.	S.D.	Skewness	Kurtosis
1.00	11.50	21.00	58.03	79.00	261.00	71.21	1.74	5.14

<https://doi.org/10.1371/journal.pone.0249001.t002>

the data extracted from [32] to illustrate our proposed model. The data given below are the failure times of air conditioning system in an airplane. [33] used Exponentiated Exponential distribution to model the same data and estimate the parameter as well. The data set of 31 observations is recorded as: 23, 261, 87, 7, 120, 14, 62, 47, 225, 71, 246, 21, 42, 20, 5, 12, 120, 11, 3, 14, 71, 11, 14, 11, 16, 90, 1, 16, 52 and 95.

Table 2 reveals certain descriptive statistics regarding set of observations under study which tells that the data set is positive skewed and heavy tailed towards right.

Table 3 provides the estimated values along with standard errors of unknown parameters of W-TEXPD, Weibull, Gamma, Exponential and Truncated Exponential (TEXPD) distributions by using ML method. The negative log-likelihood, Akaike information criterion (AIC) and Bayesian information criterion (BIC) are computed to compare the models. A distribution with the highest negative log-likelihood value and the smallest AIC and BIC values indicates the better model contrary to other fitted distributions. The values in Table 3 verdict that the proposed model has the highest value of negative log-likelihood and the lowest values of AIC and BIC supporting the new suggested distribution.

Table 4 provides the values of different test statistics which are used to analyse the goodness of fit for the distributions. The distribution having the smallest value of test statistics fits the best. It is obvious from values in the Table 4 that the W-TEXPD distribution leads to a better fit than the Weibull, Gamma, Exponential and TEXPD distributions.

We graphically study the efficacy of new proposed distribution W-TEXPD by sketching pdf, Q-Q, cdf and P-P plots for the above data set to check the goodness of fit. From Fig 3, it is

**Table 3. Negative log-likelihood values ( $\hat{l}$ ), MLEs of model parameters, the corresponding SEs (given in parentheses) and the statistics AIC and BIC of failure times of 31 air conditioning system of airplane.**

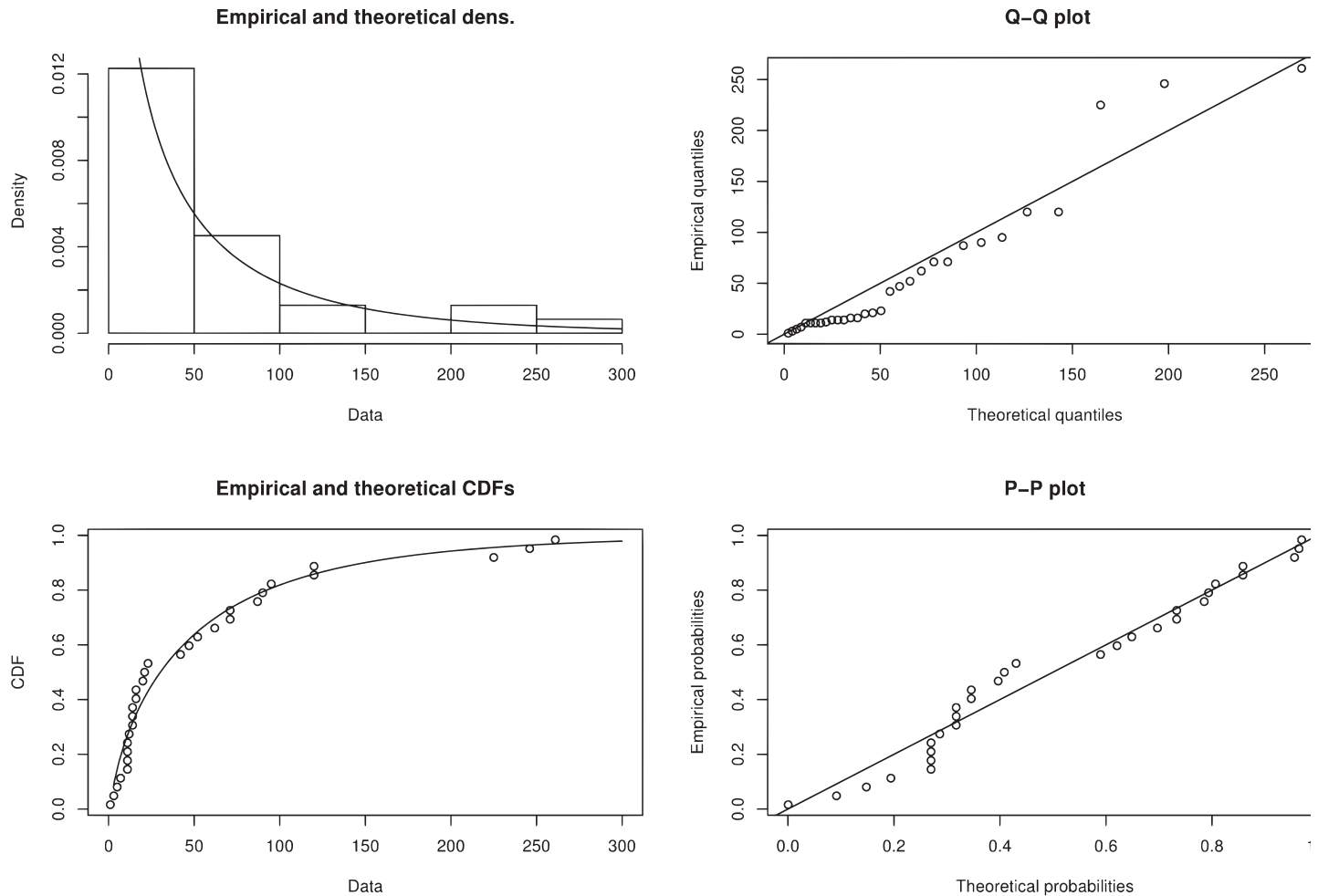
Model	Estimates					Statistic	
	$\hat{l}$	$\hat{a}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$	AIC	BIC
W-TEXPD	-153.67	1.00	0.145 (0.070)	0.738 (0.105)	0.003 (0.001)	313.35	317.65
Weibull	-156.10	—	52.883 (11.853)	0.848 (0.116)	—	316.21	319.079
Gamma	-156.39	—	0.014 (0.004)	0.807 (0.176)	—	316.77	319.64
Exponential	-156.89	—	0.017 (0.003)	—	—	315.78	317.22
TEXPD	-156.353	1.00	0.018 (0.003)	—	—	314.71	316.32

<https://doi.org/10.1371/journal.pone.0249001.t003>

**Table 4. Goodness of fit statistic of failure times of 31 air conditioners of airplane.**

Model	W-TEXPD	Weibull	Gamma	Exponential	TEXPD
K-Smirnov	0.140	0.159	0.176	0.221	0.228
C-Von	0.091	0.118	0.143	0.253	0.284
A-Darling	0.633	0.659	0.783	1.359	1.734

<https://doi.org/10.1371/journal.pone.0249001.t004>



**Fig 3.** The fitted pdf of W-TEXPD on the histogram of failure times of 31 air conditioning system of airplane along with their cdf, Q-Q and probability plots.

<https://doi.org/10.1371/journal.pone.0249001.g003>

evident from pdf plot that theoretical/predicted probabilities are closer to the empirical histogram which highlights that W-TEXPD better fits the data set. The corresponding cdf, Q-Q and P-P plots suggest the same results as well. It is also observed that the W-TEXPD follows the diagonal line more closely than the empirical line.

### 5.2 Application 2: Electrical engineering

In this section, we use a real data set to show that the W-TEXPD distribution can be a better model than Weibull, truncated Exponential (TEXPD), Gamma and Exponential distributions. The following data set is taken from [7] which represents the failure times of 50 components (per 1000h). [34] use McDonald modified Weibull distribution to model the same data set. The data set of 50 observations is recorded as: 0.036, 0.058, 0.061, 0.074, 0.078, 0.086, 0.102, 0.103, 0.114, 0.116, 0.148, 0.183, 0.192, 0.254, 0.262, 0.379, 0.381, 0.538, 0.570, 0.574, 0.590, 0.618, 0.645, 0.961, 1.228, 1.600, 2.006, 2.054, 2.804, 3.058, 3.076, 3.147, 3.625, 3.704, 3.931, 4.073, 4.393, 4.534, 4.893, 6.274, 6.816, 7.896, 7.904, 8.022, 9.337, 10.940, 11.020, 13.880, 14.730 and 15.080.

The descriptive statistics in Table 5 connotes that the data are highly spread, skewed and long right tailed.

**Table 5. Descriptive statistics of failure times of 50 components (per 1000 hours).**

Min.	Q1	Q2	Mean	Q3	Max.	S.D.	Skewness	Kurtosis
0.036	0.207	1.414	3.329	4.499	15.080	4.184	1.423	4.092

<https://doi.org/10.1371/journal.pone.0249001.t005>

**Table 6. Negative log-likelihood values ( $\hat{\ell}$ ), MLEs of model parameters, the corresponding SEs (given in parentheses) and the statistics AIC and BIC of failure time of 50 components (per 1000 hours).**

Model	Estimates					Statistic	
	$\hat{\ell}$	$\hat{a}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$	AIC	BIC
W-TEXPD	<b>-95.86</b>	<b>0.036</b>	<b>-0.279</b> (4.755)	<b>0.583</b> (0.067)	<b>-0.124</b> (2.111)	<b>197.73</b>	<b>203.46</b>
Weibull	-102.07	—	2.514 (0.568)	0.661 (0.074)	—	208.14	211.97
Gamma	-102.18	—	0.163 (0.042)	0.545 (0.091)	—	208.37	212.19
Exponential	-110.13	—	0.300 (0.042)	—	—	222.25	224.17
TEXPD	-109.58	0.036	0.304 (0.042)	—	—	221.17	223.08

<https://doi.org/10.1371/journal.pone.0249001.t006>

Table 6 provides parameter estimates of the fitted distributions along with ( $\hat{\ell}$ ), AIC and BIC values. The above table shows that the suggested model leads to better fit than the baseline distributions i.e. Weibull & TEXPD, Gamma and Exponential distributions for describing the certain data.

The smallest values of goodness of fit statistic, i.e. K-S, C-Von and A-Darling for W-TEXPD in Table 7 prove that the proposed model fits the given data best among the other models.

We graphically studied the performance of W-TEXPD by sketching pdf, Q-Q, cdf and P-P plots. We can observe in Fig 4 that the closer empirical and theoretical lines show that the W-TEXPD better fits the above data.

### 5.3 Application 3: Mechanical engineering

To demonstrate the strength of W-TEXPD, we fit our suggested model on the data extracted from [35]. The following observations are the numbers of revolutions (in millions) before failure of 23 ball bearings in a life testing experiment. [28, 36] use the Gamma-Weibull and extended generalized Gamma distributions to model the same data respectively. The data set of 23 observations is: 17.88, 28.92, 33.00, 41.52, 42.12, 45.6, 48.8, 51.84, 51.96, 54.12, 55.56, 67.8, 68.64, 68.64, 68.88, 84.12, 93.12, 98.64, 105.12, 105.84, 127.92, 128.04, 173.4.

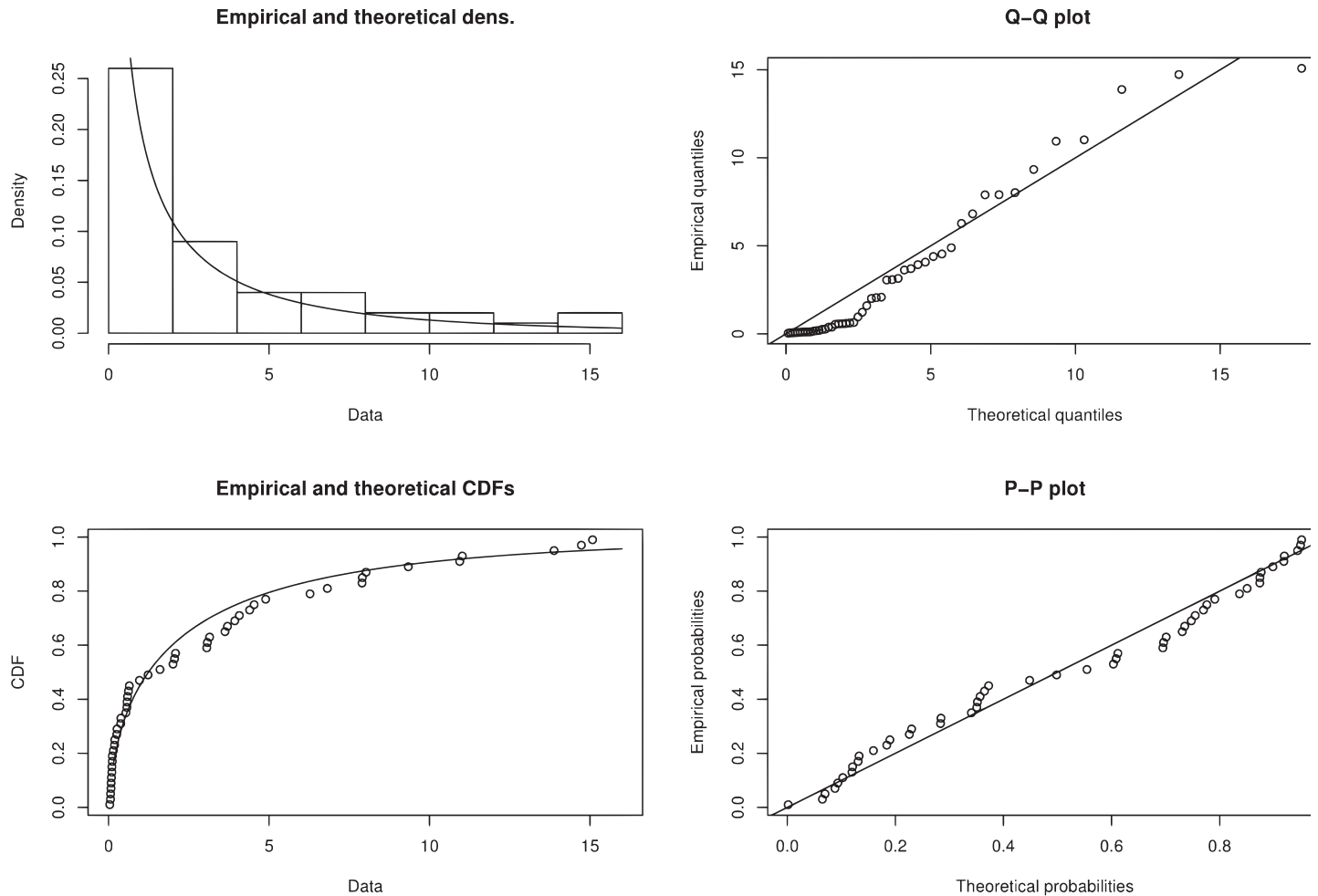
Table 8 reveals certain descriptive statistics of the under study data suggesting that the data are slightly skewed and right tailed.

**Table 7. Goodness of fit statistic of failure time of 50 components (per 1000 hours).**

Model	W-TEXPD	Weibull	Gamma	Exponential	TEXPD
K-Smirnov	<b>0.140</b>	0.159	0.176	0.221	0.228
C-Von	<b>0.091</b>	0.118	0.143	0.253	0.284
A-Darling	<b>0.633</b>	0.659	0.783	1.359	1.734

<https://doi.org/10.1371/journal.pone.0249001.t007>





**Fig 4.** The fitted pdf of W-TEXPD on the histogram of failure times of 50 components (per 1000 hours) along with their cdf, Q-Q and probability plots.

<https://doi.org/10.1371/journal.pone.0249001.g004>

Table 9 provides parameter estimates. The larger value of  $\hat{l}$  and the smaller values of AIC and BIC reflect a better model. In this aspect, it is evident from the statistics that the suggested model provides a better fit than the baseline distributions i.e. Weibull & TEXPD, Gamma and Exponential distributions for the certain data.

Table 10 gives numerical values of goodness of fit tests. We can make an upshot from these statistical values that suggested model better fits to the above data.

The histogram that is superimposed by the empirical pdf in Fig 5 also suggest that W-TEXPD better fits the data. Similarly P-P plot also support the proposed model and gives evidence that it is a pliable probability model for such type of data.

### 5.4 Application 4: Bio-chemical engineering

This example consist of vinyl chloride data which has been taken from the clean up-gradient monitoring wells in  $\mu\text{mg/L}$ . [37] utilized following data set to estimate the upper confidence

**Table 8.** Descriptive statistics of ball bearing data.

Min.	Q1	Q2	Mean	Q3	Max.	S.D.	Skewness	Kurtosis
17.88	47.20	67.80	72.24	95.88	173.40	37.48	0.94	3.49

<https://doi.org/10.1371/journal.pone.0249001.t008>

**Table 9. Negative log-likelihood values, MLEs of the model parameters, the corresponding SEs (in parentheses) along with AIC and BIC values.**

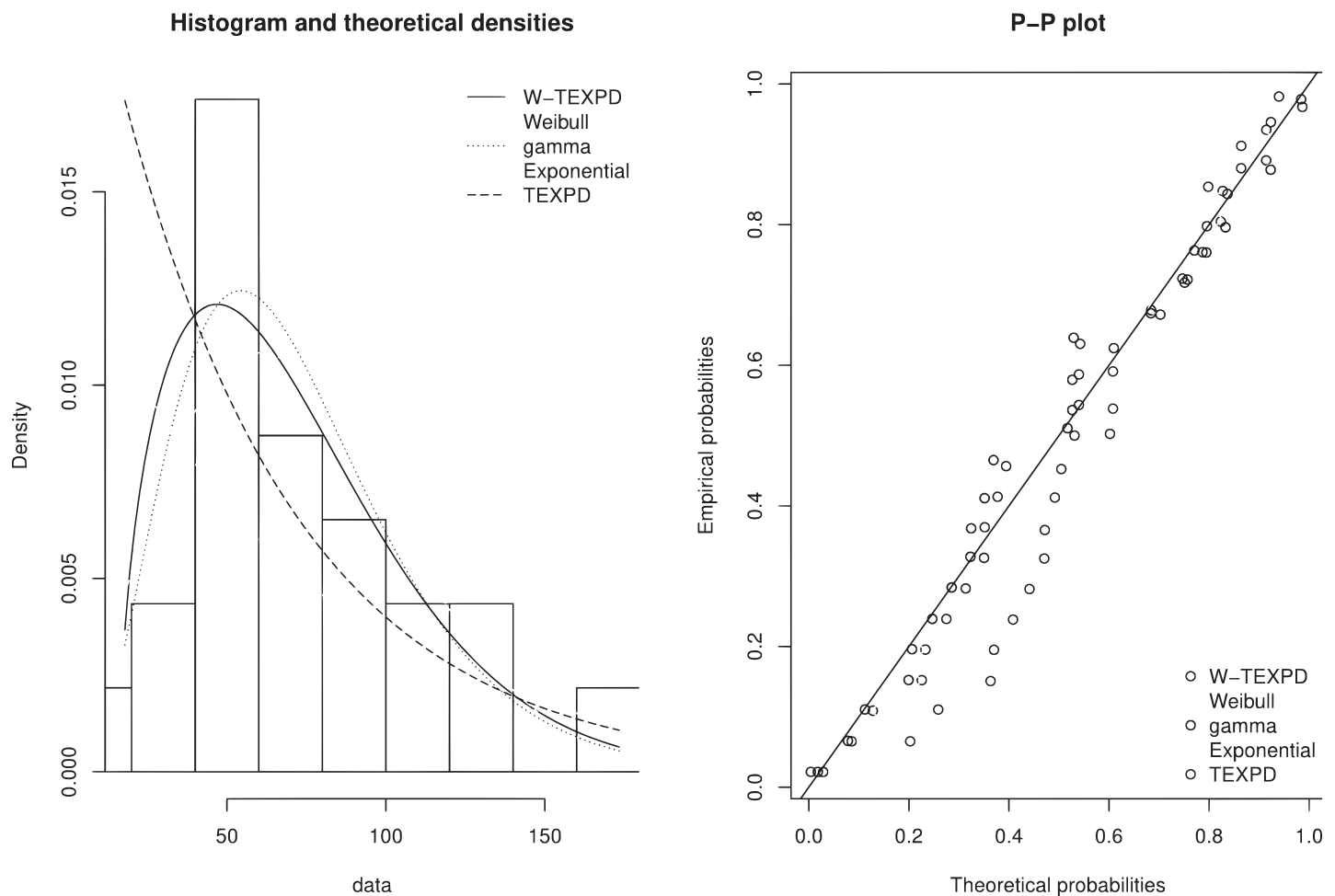
Model	Estimates					Statistics	
	$\hat{l}$	$\hat{a}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$	AIC	BIC
W-TEXPD	-112.91	16.25	1.18 (2.56)	1.52 (0.25)	0.02 (0.04)	231.81	235.22
Weibull	-113.69	—	2.10 (0.33)	81.90 (8.60)	—	231.38	233.64
Gamma	-113.03	—	4.03 (1.14)	0.06 (0.02)	—	230.05	232.32
Exponential	-121.44	—	0.01 (0.003)	—	—	244.88	246.01
TEXPD	-115.58	—	0.02 (0.004)	—	—	233.15	234.29

<https://doi.org/10.1371/journal.pone.0249001.t009>

**Table 10. Goodness of fit statistic.**

Model	W-TEXPD	Weibull	Gamma	Exponential	TEXPD
K-Smirnov	0.109	0.151	0.123	0.307	0.233
C-Von	0.034	0.582	0.039	0.537	0.240
A-Darling	0.240	0.329	0.216	2.814	1.272

<https://doi.org/10.1371/journal.pone.0249001.t010>



**Fig 5. Estimated pdf and P-P plot of W-TEXPD, Weibull, Gamma, Exponential and TEXPD distributions.**

<https://doi.org/10.1371/journal.pone.0249001.g005>

**Table 11. Descriptive statistics of Vinyl chloride data ( $\mu\text{g/L}$ ).**

Min.	Q1	Q2	Mean	Q3	Max.	S.D.	Skewness	Kurtosis
0.10	0.50	1.15	1.88	2.47	8.00	1.95	1.60	5.01

<https://doi.org/10.1371/journal.pone.0249001.t011>

**Table 12. Negative log-likelihood values, MLEs of model parameters, the corresponding SEs (in parentheses) along with the AIC and BIC values.**

Model	Estimates					Statistic	
	$\hat{l}$	$\hat{a}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$	AIC	BIC
W-TEXPD	-53.19	0.09	0.121 (0.93)	0.86 (0.12)	0.07 (0.56)	112.38	116.96
Weibull	-55.45	—	1.01 (0.13)	1.88 (0.34)	—	114.90	117.95
Gamma	-55.41	—	1.06 (0.22)	0.56 (0.15)	—	114.82	117.88
Exponential	-55.45	—	0.53 (0.09)	—	—	112.90	114.43
TEXPD	-53.77	—	0.56 (0.09)	—	—	109.53	111.06

<https://doi.org/10.1371/journal.pone.0249001.t012>

level (UCL) for the mean Vinyl chloride to facilitate the researchers to compare it with the average obtained from down-gradient monitoring well. Similarly, [38] also model the same data by using a New Extended Burr XII distribution. The observations of the Vinyl Chloride data are: 0.2, 2.0, 1.2, 1.3, 0.6, 0.5, 2.4, 0.5, 1.1, 8.0, 0.8, 0.4, 0.6, 0.9, 0.4, 6.8, 1.2, 0.5, 5.3, 3.2, 2.7, 2.9, 2.3, 1.0, 0.2, 2.5, 0.1, 0.1, 1.8, 0.9, 2.0, 4.0, 0.4, 5.1.

Table 11 displays certain descriptive statistics of the under study data reflecting that the data set is skewed and right tailed.

The statistics in Table 12 display that the suggested model W-TEXPD better explains the vinyl chloride data than the classical probability distributions.

Table 13 shows the values of goodness of fit test statistics. It is observed that W-TEXPD has the smallest values which establish that this model fits the best among the rest.

One can clearly observe from histogram and P-P plots of W-TEXPD and the other models given in Fig 6 that the proposed distribution provides the best fit among the competitive models for the Vinyl chloride data.

### 5.5 Application 5: Environmental sciences

The data is taken from a book “Loss Distributions” [39]. In 1977 the following 40 losses, due to wind-related catastrophes, were recorded to the nearest \$ 1,000,000. These data include only those losses of 2,000,000 or more; and, for convenience, they have been ordered and recorded in millions. [40] fits Alpha-Power Pareto distribution and [41] fits Lomax exponential model on the same data. The observations in the data set are: 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 4, 4, 4, 5, 5, 5, 6, 6, 6, 6, 8, 8, 9, 15, 17, 22, 23, 24, 24, 25, 27, 32 and 43.

**Table 13. Goodness of fit statistic.**

Model	W-TEXPD	Weibull	Gamma	Exponential	TEXPD
K-Smirnov	0.088	0.092	0.097	0.089	0.105
C-Von	0.025	0.043	0.051	0.041	0.061
A-Darling	0.220	0.282	0.313	0.272	0.486

<https://doi.org/10.1371/journal.pone.0249001.t013>

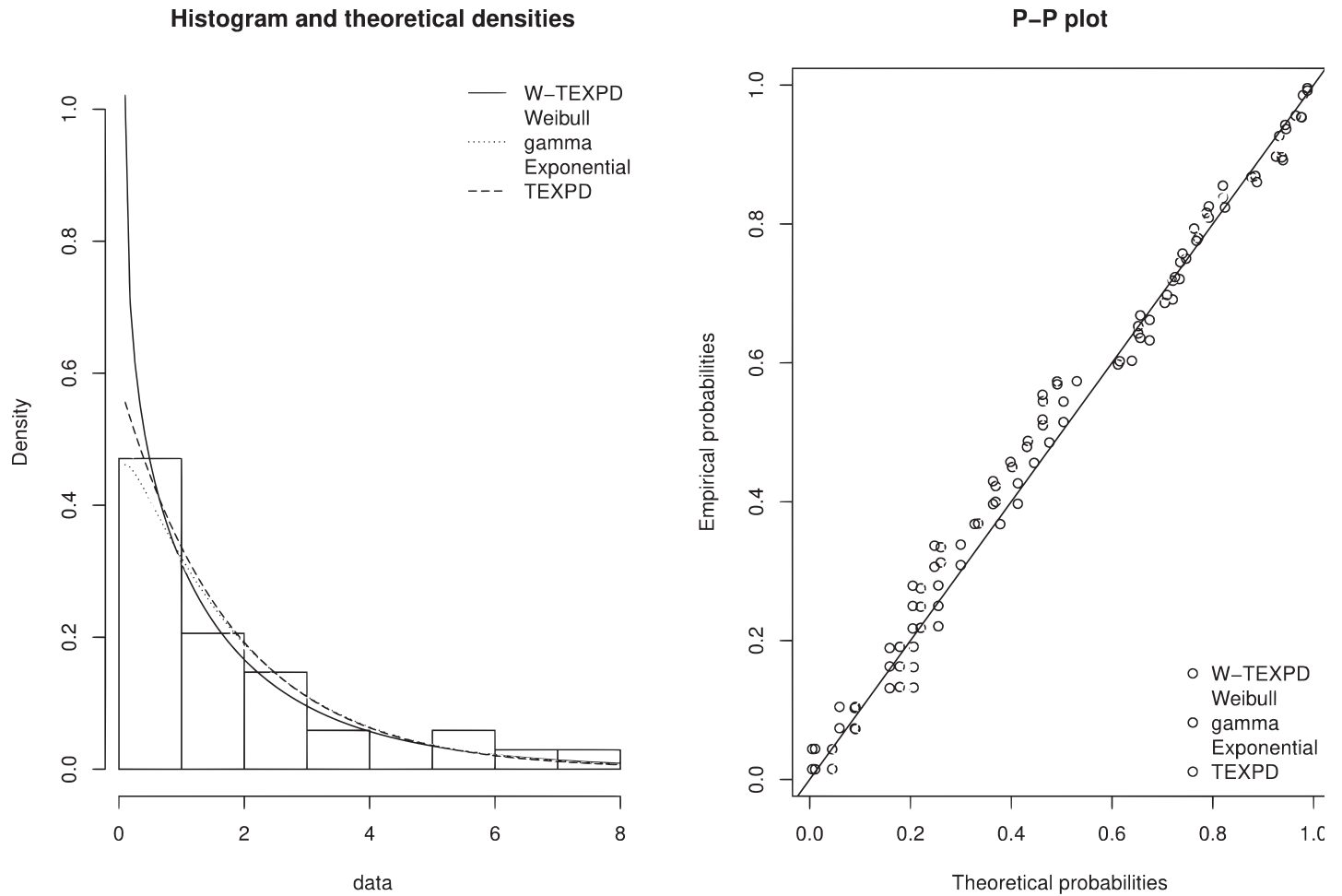


Fig 6. Estimated pdf and P-P plot of W-TEXPD, Weibull, Gamma, Exponential and TEXPD models.

<https://doi.org/10.1371/journal.pone.0249001.g006>

Table 14. Descriptive statistics of 40 losses due to wind-related catastrophes.

Min.	Q1	Q2	Mean	Q3	Max.	S.D.	Skewness	Kurtosis
2.00	2.00	5.00	9.50	13.50	43.00	10.433	1.503	4.369

<https://doi.org/10.1371/journal.pone.0249001.t014>

Table 15. Negative log-likelihood values ( $\hat{\ell}$ ), MLEs of model parameters, the corresponding SEs (given in parentheses) along with the AIC and BIC values for 40 losses due to wind-related catastrophes.

Model	Estimates					Statistic	
	$\hat{\ell}$	$\hat{a}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$	AIC	BIC
W-TEXPD	-73.96	2.00	-0.024 (0.025)	0.359 (0.049)	-0.008 (0.077)	153.92	158.83
Weibull	-123.55	—	9.507 (1.637)	1.001 (0.121)	—	251.10	254.37
Gamma	-123.42	—	0.116 (0.029)	1.108 (0.225)	—	250.85	254.12
Exponential	-123.54	—	0.105 (0.017)	—	—	249.10	250.74
TEXPD	-114.57	2.00	0.133 (0.021)	—	—	231.15	232.79

<https://doi.org/10.1371/journal.pone.0249001.t015>

**Table 16. Goodness of fit statistic of 40 losses due to wind-related catastrophes.**

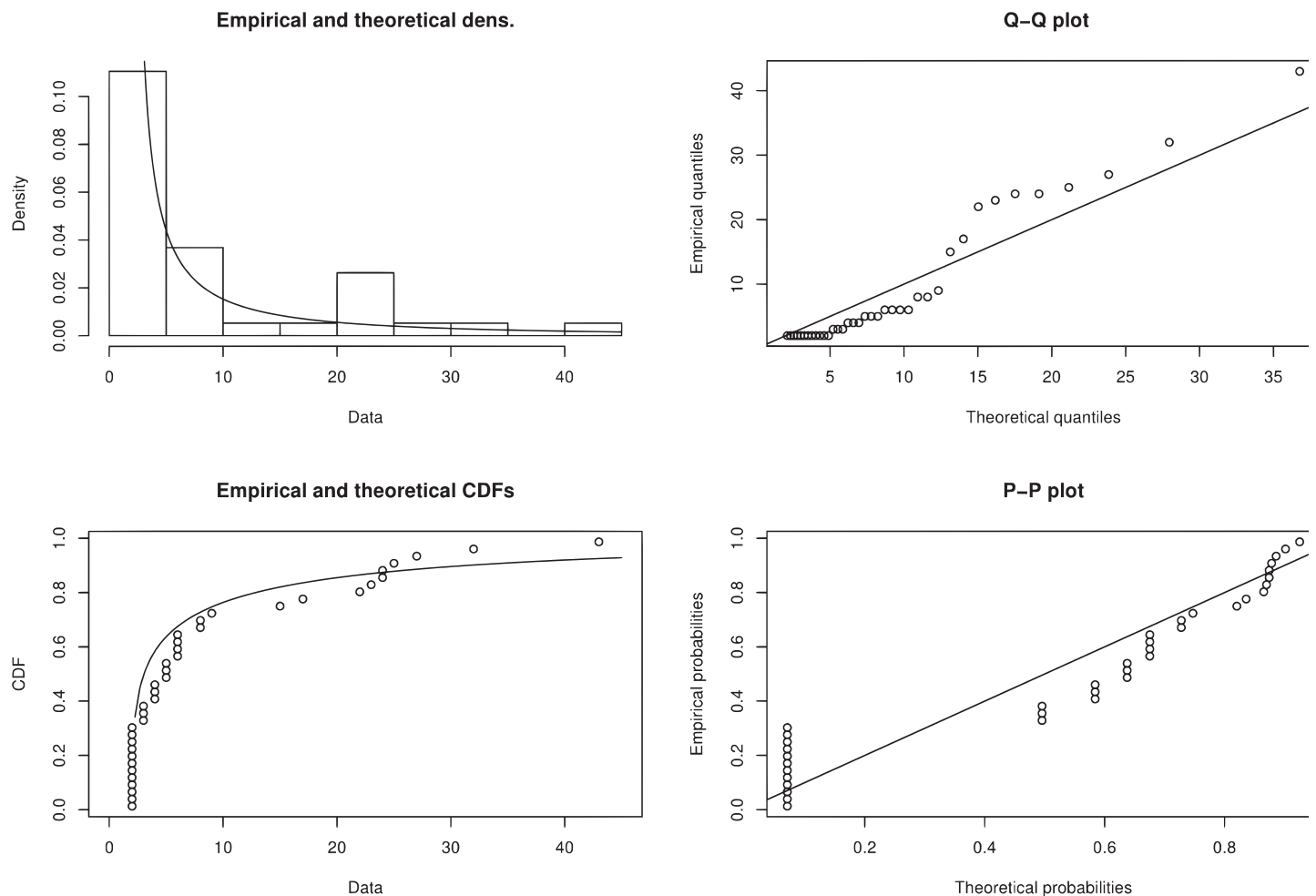
Model	W-TEXPD	Weibull	Gamma	Exponential	TEXPD
K-Smirnov	<b>0.224</b>	0.190	0.205	0.189	0.315
C-Von	<b>0.422</b>	0.343	0.398	0.341	1.041
A-Darling	<b>3.029</b>	2.055	2.313	2.049	2.165

<https://doi.org/10.1371/journal.pone.0249001.t016>

Table 14 shows that there is a large variation in the data. Furthermore, observations under study are positively skewed and flat towards right.

Table 15 provides the estimated values along with standard errors of unknown parameters by using ML method for W-TEXPD and rest of the fitted models. The negative log-likelihood, Akaike information criterion (AIC) and Bayesian information criterion (BIC) are computed to compare the models. The values in Table 15 highlight that the proposed model is statistically better than Weibull, Gamma, Exponential and Truncated Exponential (TEXPD) distributions.

Table 16 provides the values of different test statistics which are used to analyze the goodness of fit of the distributions. The distribution having the smallest value of test statistics fits



**Fig 7. The fitted pdf of W-TEXPD on the histogram of 40 losses due to wind-related catastrophes along with their cdf, Q-Q and probability plots.**

<https://doi.org/10.1371/journal.pone.0249001.g007>

the best. It is obvious from the values in the [Table 16](#) that the W-TEXPD distribution gives a better fit than the Weibull, Gamma, Exponential and TEXPD distributions.

The graphical study reveals the performance of the W-TEXPD by sketching pdf, Q-Q, cdf and P-P plots in terms of the goodness of fit. It is evident from [Fig 7](#) that the observed probabilities plotted against the predicted probabilities are closer and follow the diagonal line. Hence, it is concluded that W-TEXPD is the best choice for modeling the above data.

## 6 Concluding remarks

In this research paper, we introduce a new four parameter left truncated distribution called Weibull-Truncated Exponential distribution (W-TEXPD) by employing a new generator. The objective of the present research is to provide a truncated model for finite data. Besides, the additional scale parameter shows a significant impact on the shape of the distribution. A number of distributions are observed as the special cases of the proposed distribution.

It is demonstrated through real life applications that the truncated distribution can quite effectively be used to model a variety of data sets from different fields. It is also concluded that our proposed model better fits the data comprising extreme and/or scattered values as well as skewed (spread) and heavy tailed (flat curved) data. W-TEXPD is effectively applied in engineering and environmental sciences where such type of truncated data are commonly encountered. In future research, we will make a study to compare these estimators for censored data.

## Acknowledgments

Authors are deeply thankful to the editor in chief and reviewers for their valuable suggestions for improving the quality of this manuscript.

## Author Contributions

**Conceptualization:** Ahtasham Gul, Muhammad Mohsin, Mansoor Ali.

**Data curation:** Mansoor Ali.

**Investigation:** Muhammad Adil.

**Methodology:** Ahtasham Gul, Muhammad Mohsin, Muhammad Adil, Mansoor Ali.

**Project administration:** Mansoor Ali.

**Resources:** Muhammad Adil.

**Software:** Ahtasham Gul, Muhammad Adil.

**Supervision:** Muhammad Mohsin.

**Validation:** Muhammad Adil.

**Visualization:** Muhammad Adil.

**Writing – original draft:** Ahtasham Gul, Mansoor Ali.

**Writing – review & editing:** Muhammad Mohsin, Mansoor Ali.

## References

1. Weibull Waloddi. Wide applicability. *Journal of applied mechanics*. 1951; 730(103):293–297.
2. Cordeiro Gauss M, Silva Rodrigo B. The complementary extended Weibull power series class of distributions. *Ciência e Natura*. Universidade Federal de Santa Maria. 2014; 36:1–13.

3. Dey Sanku, Raheem Enayetur, Mukherjee Saikat. Statistical properties and different methods of estimation of transmuted Rayleigh distribution. *Revista Colombiana de Estadística*. 2017; 40(01):165–203. <https://doi.org/10.15446/rce.v40n1.56153>
4. Teimouri Mahdi, Gupta Arjun K. On the three-parameter Weibull distribution shape parameter estimation. *Journal of Data Science*. 2013; 11(03):403–414.
5. Nadarajah Saralees, Haghghi Firoozeh. An extension of the exponential distribution. *Statistics*. 2011; 45(06):543–558. <https://doi.org/10.1080/02331881003678678>
6. Kantar Yeliz Mert, Şenoğlu Birdal. A comparative study for the location and scale parameters of the Weibull distribution with given shape parameter. *Computers & Geosciences*. 2008; 34(12):1900–1909. <https://doi.org/10.1016/j.cageo.2008.04.004>
7. Murthy DP and Xie M, Jiang R. Weibull models. John Wiley & Sons. 2004; 505:1900–1909.
8. Zhang Tieling, Xie Min. On the upper truncated Weibull distribution and its reliability implications. *Reliability Engineering & System Safety*. 2011; 96(01):194–200. <https://doi.org/10.1016/j.ress.2010.09.004>
9. Nanos Nikos, Sjöstedt de Luna Sara. Fitting diameter distribution models to data from forest inventories with concentric plot design. *Forest Systems*. 2017; 26(02). <https://doi.org/10.5424/fs/2017262-10486>
10. McEwen Robert P, Parresol Bernard R. Moment expressions and summary statistics for the complete and truncated Weibull distribution. *Communications in Statistics-Theory and Methods*. 1991; 20(04):1361–1372. <https://doi.org/10.1080/03610929108830570>
11. Deemer Walter L Jr, Votaw David F Jr. Estimation of parameters of truncated or censored exponential distributions. *The Annals of Mathematical Statistics*. 1955:498–504. <https://doi.org/10.1214/aoms/1177728494>
12. Chapman Douglas G. Estimating the parameters of a truncated gamma distribution. *The Annals of Mathematical Statistics*. 1956:498–506. <https://doi.org/10.1214/aoms/1177728272>
13. Jawitz James W, Moments of truncated continuous univariate distributions. *Advances in water resources*. 2004; 27(3):269–281. <https://doi.org/10.1016/j.advwatres.2003.12.002>
14. Jamalizadeh A, Pourmousa R, Balakrishnan N. Truncated and limited skew-normal and skew-t distributions: properties and an illustration. *Communications in Statistics—Theory and Methods*. 2009; 38(16-17):2653–2668. <https://doi.org/10.1080/03610910902936109>
15. Ahmed S Ejaz, Castro-Kuriss Claudia, Flores Esteban, Leiva Víctor, Sanhueza Antonio. A TRUNCATED VERSION OF THE BIRNBAUM-SAUNDERS DISTRIBUTION WITH AN APPLICATION IN FINANCIAL RISK. *Pakistan Journal of Statistics*. 2010; 26(01).
16. Markiewicz Iwona, Strupczewski Witold, Bogdanowicz Ewa, Kochanek Krzysztof. Generalized exponential distribution in flood frequency analysis for polish rivers. *PloS one*. 2015; 10(12). <https://doi.org/10.1371/journal.pone.0143965> PMID: 26657239
17. Liu Jiayi, Wu Zhibo, Wu Jin, Dong Jian, Zhao Yao, We Dongxin. A Weibull distribution accrual failure detector for cloud computing. *PloS one*. 2017; 12(03).
18. Nassar Manal Mohamed, Radwan Salwa Said, Elmasry Ayat Salah. the Exponential Modified Weibull Logistic Distribution (EMWL). *EPH-International Journal of Mathematics and Statistics (ISSN: 2208-2212)*. 2018; 4(01):22–38.
19. Huang Helai, Song Bo, Xu Pengpeng, Zeng Qiang, Lee Jaeyoung, Abdel-Aty Mohamed. Macro and micro models for zonal crash prediction with application in hot zones identification. *Journal of transport geography*. 2016; 54:248–256. <https://doi.org/10.1016/j.jtrangeo.2016.06.012>
20. Wen Huiying, Zhang Xuan, Zeng Qiang, Sze NN. Bayesian spatial-temporal model for the main and interaction effects of roadway and weather characteristics on freeway crash incidence. *Accident Analysis & Prevention*. 2019; 132:105249. <https://doi.org/10.1016/j.aap.2019.07.025> PMID: 31415995
21. Dong Bowen, Ma Xiaoxiang, Chen Feng, Chen Suren. Investigating the differences of single-vehicle and multivehicle accident probability using mixed logit model. *Journal of Advanced Transportation*. (2018); 2018. <https://doi.org/10.1155/2018/2702360>
22. Chen Feng, Chen Suren. Injury severities of truck drivers in single-and multi-vehicle accidents on rural highways. *Accident Analysis & Prevention*. 2011; 43(05):1677–1688. <https://doi.org/10.1016/j.aap.2011.03.026> PMID: 21658494
23. Asgharzadeh A, Nadarajah S, Sharafi F. Weibull Lindley distribution. *Revstat*. 2018; 16:87–113.
24. Sadani Sahar, Abdollahnezhad Kamel, Teimouri Mahdi, Ranjbar Vahid. A new estimator for Weibull distribution parameters: Comprehensive comparative study for Weibull Distribution. *arXiv preprint arXiv:1902.05658*. 2019.
25. Gul Ahtasham, Mohsin Muhammad. Recurrence relations for moments of order statistics from half logistic-truncated exponential distribution. *Communications in Statistics-Theory and Methods*. 2020:1–14. <https://doi.org/10.1080/03610926.2019.1710753>

26. Alzaatreh Ayman, Lee Carl, Famoye Felix. A new method for generating families of continuous distributions. *Metron*.2013: 71(01);63–79. <https://doi.org/10.1007/s40300-013-0007-y>
27. Eugene Nicholas, Lee Carl, Famoye Felix. Beta-normal distribution and its applications. *Communications in Statistics-Theory and Methods*.2002: 31(04);497–512. <https://doi.org/10.1081/STA-120003130>
28. Provost Serge B, Saboor Abdus, Ahmad Munir. THE GAMMA-WEIBULL DISTRIBUTION. *Pak. J. Statistics*.2011: 27(02);111–131.
29. Smith Richard L, Naylor JC. A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution. *Applied Statistics*.1987:358–369. <https://doi.org/10.2307/2347795>
30. Jeffrey Alan, Zwillinger Daniel. *Table of integrals, series, and products*. Elsevier.2007.
31. Johnson R, Kliche D, Smith P. Maximum likelihood estimation from a left-truncated distribution. 34th Conference on Radar Meteorology, AM Soc., Ed. P.2009:02.
32. Linhart Heinz, Zucchini Walter. *Model selection*. John Wiley & Sons.1986.
33. Gupta Rameshwar D, Kundu Debasis. Exponentiated exponential family: an alternative to gamma and Weibull distributions. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*.2001: 43(01);117–130. [https://doi.org/10.1002/1521-4036\(200102\)43:1%3C117::AID-BIMJ117%3E3.0.CO;2-R](https://doi.org/10.1002/1521-4036(200102)43:1%3C117::AID-BIMJ117%3E3.0.CO;2-R)
34. Merovci Faton, Elbatal Ibrahim. The McDonald modified Weibull distribution: Properties and applications. *arXiv preprint arXiv:1309.2961*.2013.
35. Lawless Jerald F. *Statistical models and methods for lifetime data*. John Wiley & Sons.2011:362.
36. Meeker William Q, Escobar Luis A. *Statistical methods for reliability data*. John Wiley & Sons.2014.
37. Bhaumik Dulal K, Kapur Kush, Gibbons Robert D. Testing parameters of a gamma distribution for small samples. *Technometrics*.2009: 51(03);326–334. <https://doi.org/10.1198/tech.2009.07038>
38. Ghosh Indranil, Bourguignon Marcelo. A New Extended Burr XII Distribution. *Austrian Journal of Statistics*.2017: 46(01);33–39. <https://doi.org/10.17713/ajs.v46i1.139>
39. Hogg Robert V, Klugman Stuart A. *Loss distributions*. John Wiley & Sons.2009.
40. Ihtisham Shumaila, Khalil Alamgir, Manzoor Sadaf, Khan Sajjad Ahmad, Ali Amjad. Alpha-Power Pareto distribution: Its properties and applications. *PloS one*.2019. <https://doi.org/10.1371/journal.pone.0218027> PMID: 31188897
41. Ijaz Muhammad, Syed Muhammad Asim, Alamgir. Lomax exponential distribution with an application to real-life data. *PloS one*.2019. <https://doi.org/10.1371/journal.pone.0225827> PMID: 31826022