Filomat 33:9 (2019), 2635–2652
https://doi.org/10.2298/FIL1909635E

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

The Odd Flexible Weibull-H Family of Distributions: Properties and Estimation with Applications to Complete and Upper Record Data

M. El-Morshedy[a], M. S. Eliwa[a]

[a]*Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt.*

Abstract. In this paper, a new generator of continuous distributions called the odd flexible Weibull-H family is proposed and studied. Some of its statistical properties including quantile, skewness, kurtosis, hazard rate function, moments, incomplete moments, mean deviations, coefficient of variation, Bonferroni and Lorenz curves, moments of the residual (past) lifetimes and entropies are studied. Two special models are introduced and discussed in-detail. The maximum likelihood method is used to estimate the model parameters based on complete and upper record data. A detailed simulation study is carried out to examine the bias and mean square error of maximum likelihood estimators. Finally, three applications to real data sets show the flexibility of the new family.

1. Introduction

Statistical distributions are commonly applied to describe real world phenomena in various areas such as environmental, demography, biological studies, medical sciences, engineering, life testing problems, actuarial and economics. Due to the usefulness of statistical distributions, their theory is widely discussed and new distributions are proposed. The interest in developing more flexible statistical distributions remains strong in statistics profession. However, in many areas such as medical sciences, survival and reliability theory, there is a clear need for extended forms of these distributions, because in many practical situations, classical distributions do not provide adequate fits to real data. Therefore, there has been an increased interest in developing more flexible distributions. Weibull distribution is considered one of these distributions. This model is one of the most popular and widely used models for failure time in reliability theory and life testing. So, some modifications of the Weibull distribution are proposed in the statistical literature to provide more flexibility in lifetime modeling. See for example, Silva et al. (2010), Cordeiro et al. (2012), Saad and Jingsong (2013), Bourguignon et al. (2014), El-Bassiouny et al. (2016, 2017), Jehhan et al. (2018), Eliwa et al. (2019), El-Morshedy et al. (2019), Alizadeh et al. (2019), Eliwa and El-Morshedy (2019) and references cited therein.

In the last decade, Gurvich et al. (1997) introduced a class of distribution characterized by the CDF

$$\Psi(t;\eta) = 1 - e^{-\eta L(t)}; \eta, t > 0, \tag{1}$$

2010 *Mathematics Subject Classification*. 60E05; 62E10; 62N05
Keywords. Flexible Weibull distribution; Hazard rate function; Maximum likelihood method; Upper record values
Received: 03 September 2018; Revised: 12 December 2018; Accepted: 02 February 2019
Communicated by Aleksandar Nastić
ORCID: M. Elmorshedy: https://orcid.org/0000-0002-7585-5519, M. S. Eliwa: https://orcid.org/0000-0001-5619-210X
Email addresses: mah_elmorshedy@mans.edu.eg (M. El-Morshedy), mseliwa@mans.edu.eg (M. S. Eliwa)

where $L(t)$ is a monotonically increasing function of t. Recently, Bebbington et al. (2007) introduced the flexible Weibull (FW) distribution by putting $\eta = 1$ and $L(t) = e^{\alpha t - \frac{\beta}{t}}$ for $\alpha, \beta > 0$ in Equation (1). Thus, the non-negative random variable T is said to have the FW distribution if its CDF can be expressed as follows

$$\Pi(t; \alpha, \beta) = 1 - e^{-e^{\alpha t - \frac{\beta}{t}}}; t > 0. \tag{2}$$

The exponential distribution can be obtained as a special case from the FW distribution when $\alpha = \log(\lambda)$ and $\beta = 0$. Several modifications of the FW distribution are introduced and discussed in the literature. See for example, El-Gohary et al. (2015a, 2015b), El-Morshedy et al. (2017), Sangun and Jiwhan (2018), among others.

Let $\pi(t)$ be the PDF of a random variable $T \in [\xi, \xi^*]$ for $-\infty < \xi < \xi^* < \infty$ and let $U^*[H(x;\kappa)]$ be a function of the CDF of baseline model such that $U^*[H(x;\kappa)]$ satisfies the following conditions:

1. $U^*[H(x;\kappa)] \in [\xi, \xi^*]$ where κ is a parameter vector $1 \times k$.
2. $U^*[H(x;\kappa)]$ is differentiable and monotonically non-decreasing.
3. $U^*[H(x;\kappa)] \to \xi$ as $x \to -\infty$ but $U^*[H(x;\kappa)] \to \xi^*$ as $x \to \infty$.

Recently, Alzaatreh et al. (2013) defined the T-X family CDF by

$$F(x) = \int_{\xi}^{U^*[H(x;\kappa)]} \pi(t)dt. \tag{3}$$

The PDF corresponding to Equation (3) can be expressed as follows

$$f(x) = \frac{d}{dx}\left(U^*[H(x;\kappa)]\right)\pi\left(U^*[H(x;\kappa)]\right). \tag{4}$$

In this paper, we propose and study a new class of continuous distributions called the odd flexible Weibull-H (OFW-H) family by taking $U^*[H(x;\kappa)] = \frac{H(x;\kappa)}{1-H(x;\kappa)}$ and

$$\pi(t; \alpha, \beta) = \left(\alpha + \frac{\beta}{t^2}\right) e^{\alpha t - \frac{\beta}{t}} e^{-e^{\alpha t - \frac{\beta}{t}}}. \tag{5}$$

Note: If X be a lifetime random variable having a certain continuous distribution $H(x;\kappa)$, then the odds ratio that an individual following the lifetime X will die at time x is $\frac{H(x;\kappa)}{1-H(x;\kappa)}$ (see Cooray, 2006).

Our motivations for using the OFW-H family are the following:

1. To define special models with all types of the hazard rate function (HRF).
2. To make the kurtosis and the skewness more flexible compared to the baseline model.
3. To provide consistently better fits than other generated models under the same baseline distribution.
4. To generate distributions with right-skewed, left-skewed and symmetric shaped.

2. The OFW-H Family of Distributions

Let $h(x;\kappa)$, $H(x;\kappa)$, and $\overline{H}(x;\kappa) = 1 - H(x;\kappa)$, respectively, denote the PDF, CDF and the reliability function (RF) of a baseline model with parameter vector κ. Then, the CDF of the OFW-H family can be expressed as follows

$$F(x; \alpha, \beta, \kappa) = \int_0^{\frac{H(x;\kappa)}{\overline{H}(x;\kappa)}} \pi(t)dt = 1 - \exp\left\{-\exp\left[\alpha\frac{H(x;\kappa)}{\overline{H}(x;\kappa)} - \beta\frac{\overline{H}(x;\kappa)}{H(x;\kappa)}\right]\right\}; \; x, \alpha, \beta > 0. \tag{6}$$

The corresponding PDF of Equation (6) can be written as follows

$$f(x;\alpha,\beta,\kappa)=\frac{\alpha H(x;\kappa)^2+\beta\overline{H}(x;\kappa)^2}{H(x;\kappa)^2\overline{H}(x;\kappa)^2}h(x;\kappa)\exp\left\{\alpha\frac{H(x;\kappa)}{\overline{H}(x;\kappa)}-\beta\frac{\overline{H}(x;\kappa)}{H(x;\kappa)}\right\}$$
$$\times\exp\left\{-\exp\left[\alpha\frac{H(x;\kappa)}{\overline{H}(x;\kappa)}-\beta\frac{\overline{H}(x;\kappa)}{H(x;\kappa)}\right]\right\}\ ;\ x,\alpha,\beta>0. \tag{7}$$

Using the power series for the exponential function and the generalized binomial expansion, Equations (6) and (7) can be represented as an infinite mixture of exponential-H ($Exp-H$) density functions as follows

$$F(x;\alpha,\beta,\kappa)=\sum_{i=1,j=0}^{\infty}\frac{(-1)^i\,i^j}{i!\,j!}\left\{\alpha\frac{H(x;\kappa)}{\overline{H}(x;\kappa)}-\beta\frac{\overline{H}(x;\kappa)}{H(x;\kappa)}\right\}^j$$
$$=\sum_{i=1,j=0}^{\infty}\sum_{n=0}^{j}\sum_{m=0}^{\infty}\Upsilon_{i,j}^{(n,m)}G_{2n+m-j}(x), \tag{8}$$

and

$$f(x;\alpha,\beta,\kappa)=\sum_{i=1,j=0}^{\infty}\sum_{n=0}^{j}\sum_{m=0}^{\infty}\Upsilon_{i,j}^{(n,m)}g_{2n+m-j}(x), \tag{9}$$

respectively, where $G_{2n+m-j}(x)=H(x;\kappa)^{2n+m-j}$ is the $Exp-H$ family with power parameter $(2n+m-j)$, $g_{2n+m-j}(x)=(2n+m-j)\,h(x;\kappa)H(x;\kappa)^{2n+m-j-1}$, and

$$\Upsilon_{i,j}^{(n,m)}=\frac{(-1)^{i+n}\,i^j\,\alpha^n\beta^{j-n}}{i!\ j!\ m!}\binom{j}{n}\frac{\Gamma(2n+m-j)}{\Gamma(2n-j)},$$

where $(2n+m-j)\neq-1,-2,-3,...$ and $(2n-j)\neq-1,-2,-3,....$ Moreover, the RF of the OFW-H family can be expressed as follows

$$R(x;\alpha,\beta,\kappa)=\exp\left\{-\exp\left[\alpha\frac{H(x;\kappa)}{\overline{H}(x;\kappa)}-\beta\frac{\overline{H}(x;\kappa)}{H(x;\kappa)}\right]\right\}\ ;\ x,\alpha,\beta>0. \tag{10}$$

The hazard rate function (HRF) can be written as follows

$$h(x;\alpha,\beta,\kappa)=\frac{\alpha H(x;\kappa)^2+\beta\overline{H}(x;\kappa)^2}{H(x;\kappa)^2\overline{H}(x;\kappa)^2}h(x;\kappa)\exp\left\{\alpha\frac{H(x;\kappa)}{\overline{H}(x;\kappa)}-\beta\frac{\overline{H}(x;\kappa)}{H(x;\kappa)}\right\}\ ;\ x,\alpha,\beta>0. \tag{11}$$

On the other hand, record values (RC-V) are of importance in many fields like weather, Olympic records, economic and life testing. For example, in industry many products fail under stress and a battery dies under the stress of time. Also, one of the most important applications of RC-V in reliability studies, the structure $k-out-of-n$:Good and $k-out-of-n$:Fail. Let $X_1,X_2,...,X_n$ be a sequence of independent and identically distributed (IID) random variables. Let a set $Z=\max(X_1,X_2,...,X_n);n\geq 1$, then X_j is said to be an upper record (UP-RC) and is denoted by $X_{U(j)}$ if $X_j>X_{j-1}, j>1$. Due to its importance in applied statistics, especially in prediction, many authors studied different types of RC-V. See for example, Ahsanullah (2004), Soliman et al. (2010), El-Bassiouny et al. (2015), and references cited therein. Let $X_{U(1)},X_{U(2)},...,X_{U(n)}$ be the first UP-RC values of size n. Then, the PDF $f_n(x;\alpha,\beta,\kappa)$ of UP-RC values $x_{U(n)}$ can be expressed as follows

$$f_n(x;\alpha,\beta,\kappa)=\frac{1}{\Gamma(n)}\frac{\alpha H(x;\kappa)^2+\beta\overline{H}(x;\kappa)^2}{H(x;\kappa)^2\overline{H}(x;\kappa)^2}h(x;\kappa)\exp\left\{n\alpha\frac{H(x;\kappa)}{\overline{H}(x;\kappa)}-n\beta\frac{\overline{H}(x;\kappa)}{H(x;\kappa)}\right\}$$
$$\times\exp\left\{-\exp\left[\alpha\frac{H(x;\kappa)}{\overline{H}(x;\kappa)}-\beta\frac{\overline{H}(x;\kappa)}{H(x;\kappa)}\right]\right\};\ x>0, \tag{12}$$

where $f_n(x;\alpha,\beta,\kappa) = \frac{1}{\Gamma(n)}\left[-\ln R(x;\alpha,\beta,\kappa)\right]^{n-1} f(x;\alpha,\beta,\kappa)$.

3. Main Properties

3.1. Quantile function

For any $q \in (0,1)$, the qth quantile function, say $Q^*(q)$, of the OFW-H family can be expressed as follows

$$Q^*(q) = H^{-1}\left(\frac{-2\beta + \ln\left[-\ln(1-q)\right] + \sqrt{4\alpha\beta + \left\{\ln\left[-\ln(1-q)\right]\right\}^2}}{2\left\{\alpha-\beta-\ln\left[-\ln(1-q)\right]\right\}}\right), \tag{13}$$

where H^{-1} represents the baseline quantile function. Setting $q = 0.5$, we get the median of the OFW-H family. The effects of the shape parameters on the skewness (see Kenney and Keeping, 1962), say Sk, and kurtosis (see Moors, 1998), say Ku, can be based on quantile measures, where $Sk = \frac{Q^*(\frac{3}{4})+Q^*(\frac{1}{4})-2Q^*(\frac{1}{2})}{Q^*(\frac{3}{4})-Q^*(\frac{1}{4})}$ and $Ku = \frac{Q^*(\frac{3}{8})-Q^*(\frac{1}{8})+Q^*(\frac{7}{8})-Q^*(\frac{5}{8})}{Q^*(\frac{6}{8})-Q^*(\frac{2}{8})}$, respectively.

3.2. Moments and incomplete moments

Assume non-negative random variable $X \sim OFW-H\ (\alpha,\beta,\kappa)$, then the rth moment ($\mu_r^{'}$) of X, is given as follows

$$\begin{aligned}
\mu_r^{'} &= \mathbf{E}(X^r) = \int_0^\infty x^r f(x;\alpha,\beta,\kappa)\, dx \\
&= \sum_{i=1,j=0}^{\infty}\sum_{n=0}^{j}\sum_{m=0}^{\infty}\Upsilon_{i,j}^{(n,m)}\int_0^\infty x^r g_{2n+m-j}(x)dx \\
&= \sum_{i=1,j=0}^{\infty}\sum_{n=0}^{j}\sum_{m=0}^{\infty}\Upsilon_{i,j}^{(n,m)}\mathbf{E}(Z_{2n+m-j}^r),
\end{aligned} \tag{14}$$

where $Z_{2n+m-j}^r \sim Exp - H$ with power parameter $(2n+m-j)$. By setting $r = 1$ in Equation (14), we get the mean of X. For lifetime models, it is also of interest to obtain the incomplete moments. The incomplete moments play an important role for measuring inequality. The m^*th incomplete moment of X can be expressed as follows

$$\Phi_{(m^*)}(t) = \sum_{i=1,j=0}^{\infty}\sum_{n=0}^{j}\sum_{m=0}^{\infty}\Upsilon_{i,j}^{(n,m)}\Phi_{(m^*)}^*(t), \tag{15}$$

where $\Phi_{(m^*)}^*(t) = \int_0^t x^{m^*} g_{2n+m-j}(x)dx$. Furthermore, the main application of the first incomplete moment refers to the Bonferroni and Lorenz curves.

3.3. Mean deviations and coefficient of variation

In statistics, the mean deviations about the mean and median measure the amount of scatter in a population. For a random variable $X \sim OFW-H\ (\alpha,\beta,\kappa)$, the mean deviations about the mean and median can be expressed as follows

$$\begin{aligned}
\lambda_1 &= \int_0^\infty \left|x-\mu_1^{'}\right| f(x;\alpha,\beta,\kappa)dx \\
&= 2\mu_1^{'}F(\mu_1^{'};\alpha,\beta,\kappa) - 2\mu_1^{'} + 2\int_{\mu_1^{'}}^\infty x f(x;\alpha,\beta,\kappa)dx \\
&= 2\mu_1^{'}F(\mu_1^{'}) - 2\Phi_{(1)}(\mu_1^{'})
\end{aligned} \tag{16}$$

and

$$\lambda_2 = \int_0^\infty \left| x - Q^*(\frac{1}{2}) \right| f(x;\alpha,\beta,\kappa)dx = -\mu_1^{'} + 2\int_{Q^*(\frac{1}{2})}^\infty x f(x;\alpha,\beta,\kappa)dx$$
$$= \mu_1^{'} - 2\Phi_{(1)}(Q^*(0.5)), \tag{17}$$

respectively. Further, the coefficient of variation ($\mathbb{CV}$) is a measure of variability in the data. If $X \sim OFW-H\ (\alpha,\beta,\kappa)$, then the $\mathbb{CV}$ can be represented as $\mathbb{CV} = \sqrt{\mu_2^{'} - \mu_1^{'2}} / \left|\mu_1^{'}\right|$. When the value of $\mathbb{CV}$ is high (low), it means that the data has high (low) variability and low (high) stability.

3.4. Bonferroni and Lorenz curves

Bonferroni and Lorenz curves have applications in many fields like reliability, economics, demography, medicine and insurance. If $X \sim OFW-H\ (\alpha,\beta,\kappa)$, then the Bonferroni curve, say $B^*(s)$, is given by

$$B^*(s) = \sum_{i=1,j=0}^{\infty} \sum_{n=0}^{j} \sum_{m=0}^{\infty} \Upsilon_{i,j}^{(n,m)} B_{2n+m-j}^{**}(s), \tag{18}$$

where $B_{2n+m-j}^{**}(s) = \frac{1}{\mu^* s}\int_0^{H^{-1}(s)} x g_{2n+m-j}(x)dx$ is the Bonferroni curve of $Exp-H$ family with power parameter $(2n+m-j)$, and μ^* denotes the average. Also, the Lorenz curve, say $L^*(s)$, is given by

$$L^*(s) = \sum_{i=1,j=0}^{\infty} \sum_{n=0}^{j} \sum_{m=0}^{\infty} \Upsilon_{i,j}^{(n,m)} L_{2n+m-j}^{**}(s), \tag{19}$$

where $L_{2n+m-j}^{**}(s) = \frac{1}{\mu^*}\int_0^{H^{-1}(s)} x g_{2n+m-j}(x)dx$ is the Lorenz curve of $Exp-H$ family with power parameter $(2n+m-j)$.

3.5. Moments of residual and past lifetimes

To describe different maintenance strategies, it is an important to calculate the mean residual lifetime (MRL) and mean past lifetime (MPL), which can be derived from their moments. So, the n^*th moment of the residual lifetime, say $M_{RL}^{(n^*)}(t)$, is given as follows

$$M_{RL}^{(n^*)}(t) = E((T-t)^{n^*} \mid T > t)\ ; n^* = 1,2,3,\ldots. \tag{20}$$

Therefore, if the non-negative random variable $T \sim OFW-H\ (\alpha,\beta,\kappa)$, then

$$M_{RL}^{(n^*)}(t) = \frac{1}{R(t)}\int_t^\infty (x-t)^{n^*} f(x)\, dx$$
$$= \frac{1}{R(t)} \sum_{i=1,j=0}^{\infty} \sum_{n=0}^{j} \sum_{m=0}^{\infty} \sum_{l=0}^{n^*} (-1)^l \binom{n^*}{l} t^l\, \Upsilon_{i,j}^{(n,m)} \int_t^\infty x^{n^*-l} g_{2n+m-j}(x)\, dx$$
$$= \frac{1}{R(t)} \sum_{i=1,j=0}^{\infty} \sum_{n=0}^{j} \sum_{m=0}^{\infty} \sum_{l=0}^{n^*} (-t)^l \binom{n^*}{l} \Upsilon_{i,j}^{(n,m)} M_*^{(n^*)}(t), \tag{21}$$

where $M_*^{(n^*)}(t) = \int_t^\infty x^{n^*-l} g_{2n+m-j}(x)\, dx$. Setting $n=1$ in Equation (21), we get the MRL. The importance of the MRL function is due to it is uniquely determination of the lifetime model as well as the HRF. Also, the n^*th moment of the past lifetime or the n^*th moment of the waiting time, say $M_{IT}^{(n^*)}(t)$, is given as follow

$$M_{IT}^{(n^*)}(t) = E((t-T)^{n*} \mid T \le t)\ ; n^* = 1,2,3,\ldots. \tag{22}$$

So, if $T \sim OFW-H\,(\alpha,\beta,\kappa)$, then

$$\begin{aligned} M_{IT}^{(n^*)}(t) &= \frac{1}{F(t)}\int_0^t (t-x)^{n^*}\, f(x)\, dx \\ &= \frac{1}{F(t)}\sum_{i=1,j=0}^{\infty}\sum_{n=0}^{j}\sum_{m=0}^{\infty}\sum_{l=0}^{n^*}(-1)^l \binom{n^*}{l}\, t^{n^*-l}\,\Upsilon_{i,j}^{(n,m)} M_{**}^{(n^*)}(t), \end{aligned} \tag{23}$$

where $M_{**}^{(n^*)}(t) = \int_0^t x^l g_{2n+m-j}(x)\, dx$. Setting $n = 1$ in Equation (23), we get the MPL.

3.6. Entropies

Entropy is a measure of variation or uncertainty of a random variable X. It has many applications not only in survival analysis, but also in computer science and econometrics. Two popular entropy measures are the Rényi and Shannon entropies. The Rényi entropy of the random variable X can be expressed as follows

$$I_\delta(X) = \frac{1}{1-\delta}\log\int_0^\infty f^\delta(x;\alpha,\beta,\kappa)\, dx\,, \tag{24}$$

where $\delta > 0$ and $\delta \neq 1$. Using Equation (7), we can obtain after some algebra

$$f^\delta(x;\alpha,\beta,\kappa) = \sum_{i=0}^{\infty}\sum_{j=0}^{i}\sum_{m=0}^{\infty}\sum_{n=0}^{i-j+m}\mho_{i,j}^{n,m}\, h(x;\kappa)^\delta H(x;\kappa)^\varkappa, \tag{25}$$

where $\varkappa = i - j + m + l - 2n - 2s$ and

$$\mho_{i,j}^{n,m} = \sum_{l=0}^{\infty}\sum_{s=0}^{\delta}\frac{(-1)^{j+n}\,\delta^i j^m}{i!\; m!\; l!}\alpha^{i-j+m-n+\delta-s}\beta^{n+s}\binom{i}{j}\binom{\delta}{s}\binom{i-j+m}{n}\frac{\Gamma(\varkappa+2\delta)}{\Gamma(\varkappa+2\delta-l)},$$

for $(\varkappa + 2\delta) \neq -1, -2, -3, ...$ and $(\varkappa + 2\delta - l) \neq -1, -2, -3, ...$ Thus,

$$I_\delta(X) = \frac{1}{1-\delta}\log\left\{\sum_{i=0}^{\infty}\sum_{j=0}^{i}\sum_{m=0}^{\infty}\sum_{n=0}^{i-j+m}\mho_{i,j}^{n,m}\int_0^\infty h(x;\kappa)^\delta H(x;\kappa)^\varkappa dx\right\}. \tag{26}$$

Note that the integral term only depends on the $h(x;\kappa)$ and $H(x;\kappa)$ of baseline distribution. This integral is so difficult to obtain, but it can be computed numerically for most α, β and κ parameter values. Shannon Entropy (SEn) can be obtained as a particular case of Rényi entropy when δ tends to 1 where SEn $= \mathbf{E}\left[-\log f(x;\alpha,\beta,\kappa)\right]$.

4. Special Models

4.1. The OFW-Fréchet (OFWFr) distribution

Consider the CDF of the Fréchet (Fr) distribution is given by $H(x;a,b) = e^{-\left(\frac{a}{x}\right)^b}$. Then, the CDF of the OFWFr(α,β,a,b) distribution can be represented as follows

$$F(x;\alpha,\beta,a,b) \;=\; 1-\exp\left\{-\exp\left[\alpha\left(e^{\left(\frac{a}{x}\right)^b}-1\right)^{-1}-\beta\left(e^{\left(\frac{a}{x}\right)^b}-1\right)\right]\right\}; \alpha,\beta,a,b,x>0. \tag{27}$$

Figure 1 shows the PDF and HRF of the OFWFr distribution for various values of the parameters.

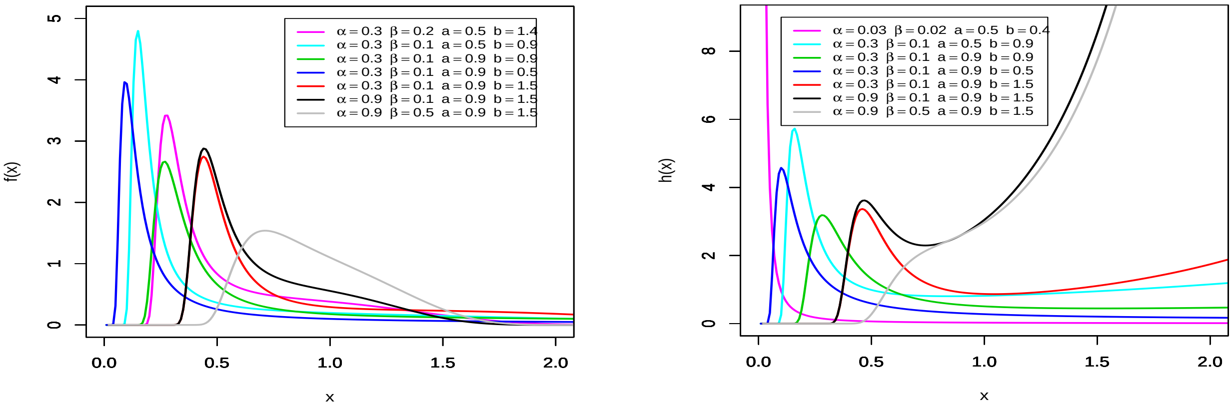

Figure 1. The plots of the PDF (left panel) and HRF (right panel) of the OFWFr distribution.

Figure 1 shows the HRF of the OFWFr distribution can be decreasing, increasing, unimodal or of unimodal-bathtub shape, which makes the distribution more flexible to fit different lifetime data sets. On the other hand, the skewness and kurtosis of the OFWFr distribution for some choices of a, b and β as function of α (alpha) are displayed in Figure 2. We take $a = 0.3$ and $b = 0.9$ for the plots of the skewness and kurtosis.

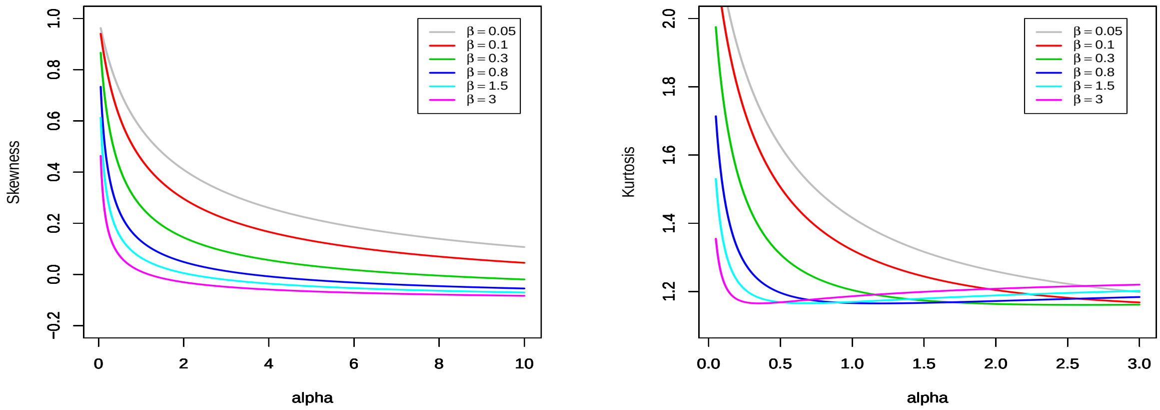

Figure 2. The plots of the skewness (left panel) and kurtosis (right panel) of the OFWFr distribution.

The plots in Figure 2 show that the shapes of the OFWFr distribution have strong dependence on the values of α and β. Furthermore, it can be used in modelling symmetric, positive and negative skewness data sets.

4.2. The OFW-exponential (OFWE) distribution

Consider the CDF of the exponential (E) distribution is given by $H(x; a) = 1 - e^{-ax}$. Then, the CDF of the OFWE(α, β, a) distribution can be expressed as follows

$$F(x; \alpha, \beta, a) = 1 - \exp\left\{-\exp\left[\alpha\left(e^{ax} - 1\right) - \beta\left(e^{ax} - 1\right)^{-1}\right]\right\}; \alpha, \beta, a, x > 0.$$

Figure 3 shows the PDFs and the HRFs of the OFWE distribution for various values of the parameters.

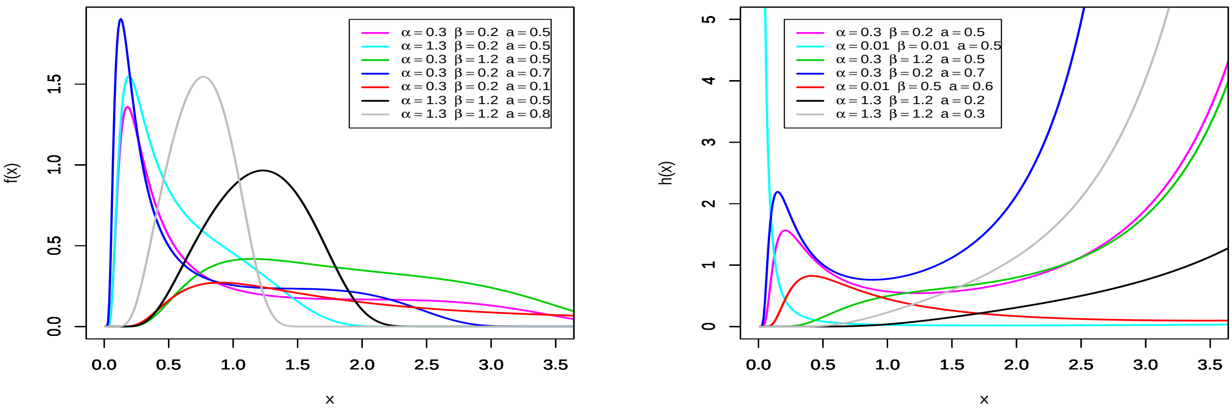

Figure 3. The plots of the PDF (left panel) and HRF (right panel) of the OFWE distribution.

From Figure 3, it is immediate that the HRF can be either unimodal, decreasing, increasing or of unimodal-bathtub shape. Moreover, the skewness and kurtosis of the OFWE distribution for some choices of a and β as function of α are displayed in Figure 4. We take $a = 0.3$ for the plots of the skewness and kurtosis.

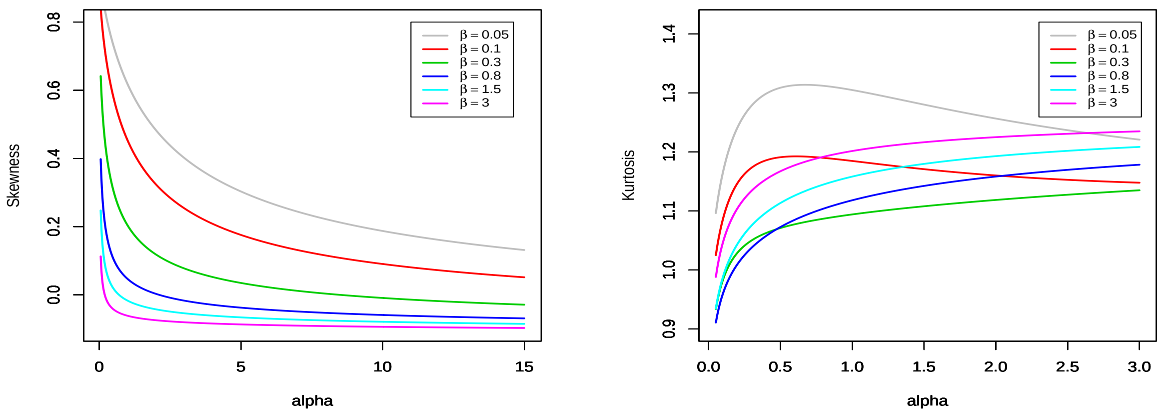

Figure 4. The plots of the skewness (left panel) and kurtosis (right panel) of the OFWE distribution.

The plots of skewness and kurtosis show that the shapes of the OFWE distribution have strong dependence on the values of α and β.

5. Estimation Based On Complete Data

There exist many different methods for parameter estimation but the most commonly used method is the maximum likelihood (ML) method. For this reason, we provide the estimation of the parameters α, β and κ for the OFW-H family using this method. Let $X_1, X_2, ..., X_n$ be a random sample of size n from the

OFW-H family, then the log-likelihood function (LL) can be expressed as

$$LL(\alpha,\beta,\kappa|x) = \sum_{i=1}^{n} \ln\left[\alpha H(x_i;\kappa)^2 + \beta\overline{H}(x_i;\kappa)^2\right] - 2\sum_{i=1}^{n} \ln\left[H(x_i;\kappa)\overline{H}(x_i;\kappa)\right] + \sum_{i=1}^{n} \ln\left[h(x_i;\kappa)\right]$$
$$+ \alpha\sum_{i=1}^{n}\frac{H(x_i;\kappa)}{\overline{H}(x_i;\kappa)} - \beta\sum_{i=1}^{n}\frac{\overline{H}(x_i;\kappa)}{H(x_i;\kappa)} - \sum_{i=1}^{n}\exp\left(\alpha\frac{H(x_i;\kappa)}{\overline{H}(x_i;\kappa)} - \beta\frac{\overline{H}(x_i;\kappa)}{H(x_i;\kappa)}\right). \quad (28)$$

We derive the MLEs of the parameters α,β and κ by solving the nonlinear likelihood equations obtained by differentiating Equation (28) with respect to α,β and κ, respectively. The components of the score vector $\mathbf{V}(\alpha,\beta,\kappa) = (\frac{\partial LL}{\partial\alpha}, \frac{\partial LL}{\partial\beta}, \frac{\partial LL}{\partial\kappa})^T$ are given by

$$V_\alpha = \sum_{i=1}^{n}\frac{H(x_i;\kappa)^2}{\alpha H(x_i;\kappa)^2 + \beta\overline{H}(x_i;\kappa)^2} + \sum_{i=1}^{n}\frac{H(x_i;\kappa)}{\overline{H}(x_i;\kappa)} - \sum_{i=1}^{n}\frac{H(x_i;\kappa)}{\overline{H}(x_i;\kappa)}\exp\left(\alpha\frac{H(x_i;\kappa)}{\overline{H}(x_i;\kappa)} - \beta\frac{\overline{H}(x_i;\kappa)}{H(x_i;\kappa)}\right), \quad (29)$$

$$V_\beta = \sum_{i=1}^{n}\frac{\overline{H}(x_i;\kappa)^2}{\alpha H(x_i;\kappa)^2 + \beta\overline{H}(x_i;\kappa)^2} - \sum_{i=1}^{n}\frac{\overline{H}(x_i;\kappa)}{H(x_i;\kappa)} + \sum_{i=1}^{n}\frac{\overline{H}(x_i;\kappa)}{H(x_i;\kappa)}\exp\left(\alpha\frac{H(x_i;\kappa)}{\overline{H}(x_i;\kappa)} - \beta\frac{\overline{H}(x_i;\kappa)}{H(x_i;\kappa)}\right) \quad (30)$$

and

$$V_{\kappa_j} = 2\sum_{i=1}^{n}\frac{\left(\alpha H(x_i;\kappa) - \beta\overline{H}(x_i;\kappa)\right)\left[H'(x_i;\kappa)\right]_{\kappa_j}}{\alpha H(x_i;\kappa)^2 + \beta\overline{H}(x_i;\kappa)^2} + \sum_{i=1}^{n}\frac{\left[h'(x_i;\kappa)\right]_{\kappa_j}}{h(x_i;\kappa)}$$
$$+ \sum_{i=1}^{n}\left[1 - \exp\left(\alpha\frac{H(x_i;\kappa)}{\overline{H}(x_i;\kappa)} - \beta\frac{\overline{H}(x_i;\kappa)}{H(x_i;\kappa)}\right)\right]\left[\alpha\frac{\left[H'(x_i;\kappa)\right]_{\kappa_j}}{\overline{H}(x_i;\kappa)^2} + \beta\frac{\left[H'(x_i;\kappa)\right]_{\kappa_j}}{H(x_i;\kappa)^2}\right]$$
$$+ 2\sum_{i=1}^{n}\frac{(2H(x_i;\kappa) - 1)\left[H'(x_i;\kappa)\right]_{\kappa_j}}{H(x_i;\kappa)\overline{H}(x_i;\kappa)}, \quad (31)$$

where $\left[h'(x_i;\kappa)\right]_{\kappa_j} = \partial h(x_i;\kappa)/\partial\kappa_j$ and $\left[H'(x_i;\kappa)\right]_{\kappa_j} = \partial H(x_i;\kappa)/\partial\kappa_j$ for $j = 1, 2, .., k$.

Setting the Equations (29-31) to zero and solving them, immediately yields the MLEs for the OFW-H family parameters. These equations cannot be solved analytically; therefore analytical software is required to solve them numerically.

6. Estimation Based On UP-RC Values

Consider n UP-RC values $X = \left\{X_{U(1)}, X_{U(2)}, ..., X_{U(n)}\right\}$ from a sequence of IID random variables following the OFW-H family are observed. Then, the likelihood function can be expressed as follows

$$L^*(\alpha,\beta,\kappa|x) = f(x_{U(n)};\alpha,\beta,\kappa)\prod_{i=1}^{n-1}\frac{f(x_{U(i)};\alpha,\beta,\kappa)}{R(x_{U(i)};\alpha,\beta,\kappa)};\ 0 \le x_{U(1)} < x_{U(2)} < ... < x_{U(n)} < \infty, \quad (32)$$

where $f(x_{U(i)};\alpha,\beta,\kappa)$ and $R(x_{U(i)};\alpha,\beta,\kappa)$ can be obtained respectively from Equations (7) and (10). Differentiating Equation (32) with respect to α,β and κ, we get the nonlinear likelihood equations. These equations need to an iterative procedure like Newton Raphson to solve them numerically.

7. Simulation Study

In this section, we assess the performance of the MLE based on complete and UP-RC data sets with respect to sample size n using R software. The assessment is based on a simulation study:

1. Generate 500 samples of size $n = 20, 25, 30, 35, ...,100$ from OFWFr(1.6, 1.8, 2.1, 1.5) and OFWE(1.3, 1.1, 0.3), respectively.
2. Compute the MLEs for the 500 samples, say $\widehat{\alpha}_s, \widehat{\beta}_s, \widehat{a}_s$ and $\widehat{b}_s$ for $s = 1, 2, ..., 500$.
3. Compute the biases and mean-squared errors (MSEs), where

$$\text{bias} = \frac{1}{500}\sum_{s=1}^{500}\left(\widehat{\varsigma}_s - \varsigma\right) \text{ and MSE} = \frac{1}{500}\sum_{s=1}^{500}\left(\widehat{\varsigma}_s - \varsigma\right)^2.$$

4. The empirical results based on complete data are given in Figures 5, 6, 7 and 8.

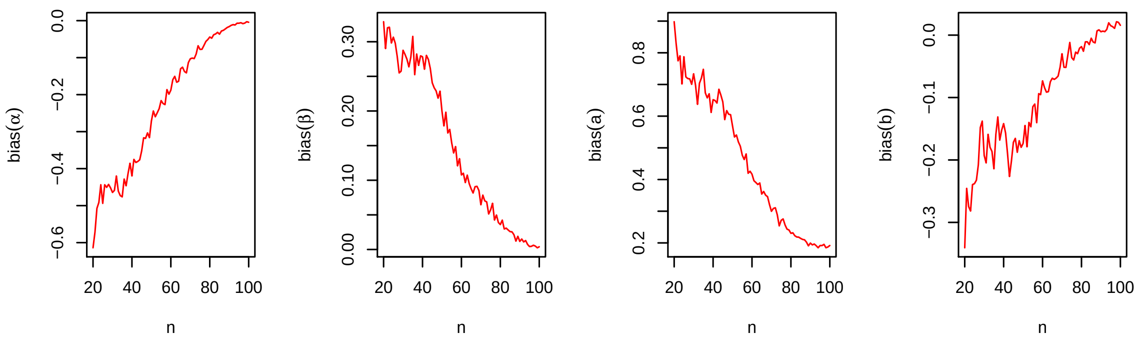

Figure 5. The bias of $\widehat{\alpha}, \widehat{\beta}, \widehat{a}$ and $\widehat{b}$ versus n for the OFWFr model when $(\alpha, \beta, a, b) = (1.6, 1.8, 2.1, 1.5)$.

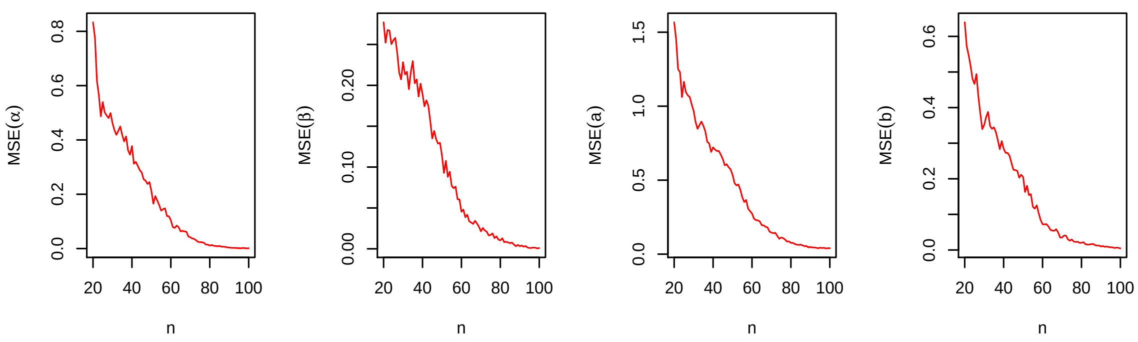

Figure 6. The MSE of $\widehat{\alpha}, \widehat{\beta}, \widehat{a}$ and $\widehat{b}$ versus n for the OFWFr model when $(\alpha, \beta, a, b) = (1.6, 1.8, 2.1, 1.5)$.

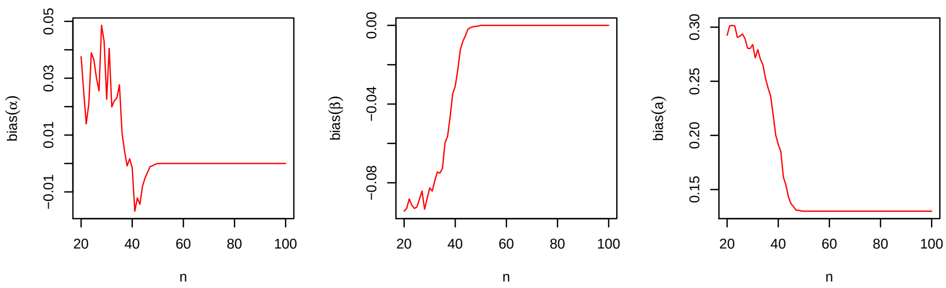

Figure 7. The bias of $\widehat{\alpha}, \widehat{\beta}$ and $\widehat{a}$ versus n for the OFWE model when $(\alpha, \beta, a) = (1.3, 1.1, 0.3)$.

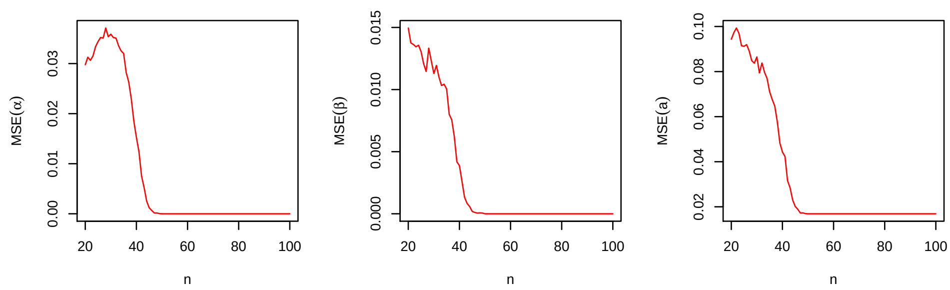

Figure 8. The MSE of $\widehat{\alpha}, \widehat{\beta}$ and $\widehat{a}$ versus n for the OFWE model when $(\alpha, \beta, a) = (1.3, 1.1, 0.3)$.

From Figures 5, 6, 7 and 8 the following observations can be made:

(a) The magnitude of bias always decreases to zero as $n \to \infty$.
(b) The MSEs always decrease to zero as $n \to \infty$. This shows the consistency of the estimators.
(c) The performance of the MLE method is very good. Thus, it is a proper method to estimate the model parameters.

Regarding the simulation of the UP-RC values, it is found that the results of this simulation (Figures and discussion) are very similar as compared to complete data. Thus, we avoided to show the same results.

8. Data Analysis

In this section, we illustrate the empirical importance of the OFWFr and OFWE distributions using three applications to real data. The data set I and II are analyzed based on complete data; but the data set III is analyzed based on UP-RC values. The fitted distributions are compared using some criteria namely, the maximized log-likelihood (LL), Akaike information criterion (AIC), corrected AIC (CAIC), Bayesian IC (BIC), Hannan-Quinn IC (HQIC), Cramér-Von Mises ($\mathbf{W}^*$), Anderson-Darling ($\mathbf{A}^*$) statistics; in addition to the Kolmogorov-Smirnov (KS) statistic and its p-value.

8.1. Data set I (Aluminum Coupons)

The data represents the fatigue time of 101 6061-T6 aluminum coupons cut parallel to the direction of rolling and oscillated at 18 cycles per second (cps) (see Birnbaum and Saunders, 1969). This data set is used to compare the fits of the OFWFr model with some competitive models like Topp-Leone Fréchet (ToLFr), transmuted Fréchet (TrFr), exponentiated transmuted Fréchet (ETrFr), Gumble Fréchet (GuFr), type I general exponential Fréchet (TIGEFr), exponentiated Fréchet (EFr) and Fréchet (Fr) by using R software.

Tables 1 and 2 provide the MLEs, KS and p-values; in addition to the value of LL, AIC, CAIC, BIC, HQIC, A^* and W^*, respectively.

Table 1. The MLEs, KS and p-values for data set I.

Model	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{a}$	$\widehat{b}$	KS	p-value
OFWFr	9.310	0.380	312.686	0.736	0.090	0.383
ToLFr	35.077	0.783	59.691	4.088	0.121	0.102
TrFr	1.00	–	136.952	3.980	0.120	0.105
ETrFr	−0.997	0.259	136.956	5.582	0.123	0.092
GuFr	1.968	0.029	3.457	0.107	0.135	0.050
TIGEFr	16648.993	74.473	7.531	5.056	0.133	0.057
EFr	–	73.221	51.679	5.057	0.133	0.056
Fr	–	–	120.782	5.057	0.133	0.056

Table 2. The LL, AIC, CAIC, BIC, HQIC, $\mathbf{A}^*$ and $\mathbf{W}^*$ values for data set I.

Model	$-LL$	AIC	CAIC	BIC	HQIC	A^*	W^*
OFWFr	459.69	927.38	927.79	937.84	931.62	0.672	0.099
ToLFr	466.35	940.71	941.13	951.17	944.94	1.543	0.270
TrFr	466.41	438.81	939.06	946.66	941.99	1.565	0.275
ETrFr	468.94	945.88	946.29	956.34	950.11	1.923	0.347
GuFr	475.73	959.46	959.88	969.92	963.69	2.558	0.443
TIGEFr	475.19	958.38	958.79	968.84	962.62	2.496	0.432
EFr	475.18	956.37	956.62	964.22	959.55	2.497	0.433
Fr	475.18	954.37	954.49	959.60	956.49	2.497	0.433

From Tables 1 and 2, it is clear that the OFWFr model provides the best fit among all tested models because it has the smallest value among all goodness-of-fit measures as well as it has the highest p-value. The empirical PDFs, CDFs and P-P plots are displayed in Figures 9 and 10, which support the results of Tables 1 and 2.

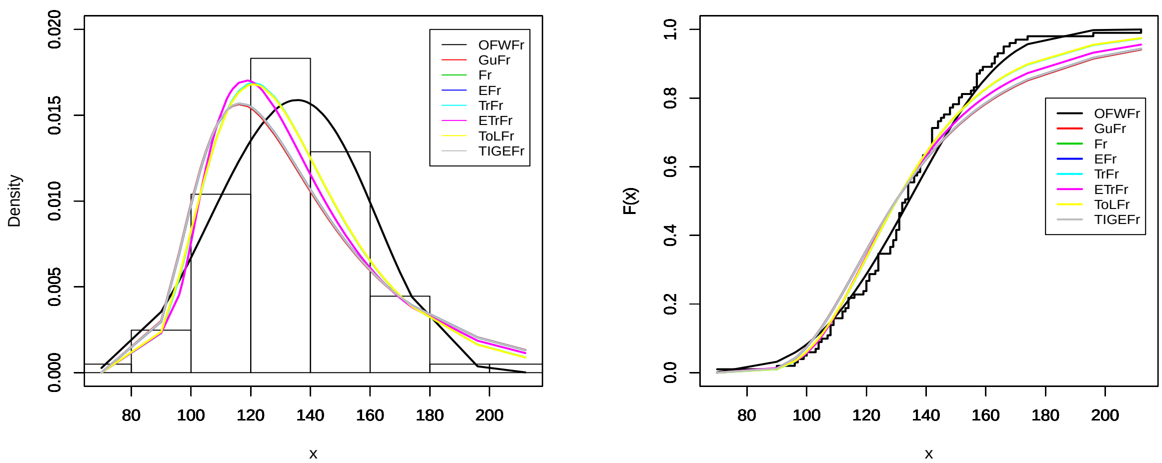

Figure 9. Estimated PDFs (left panel) and CDFs (right panel) for data set I.

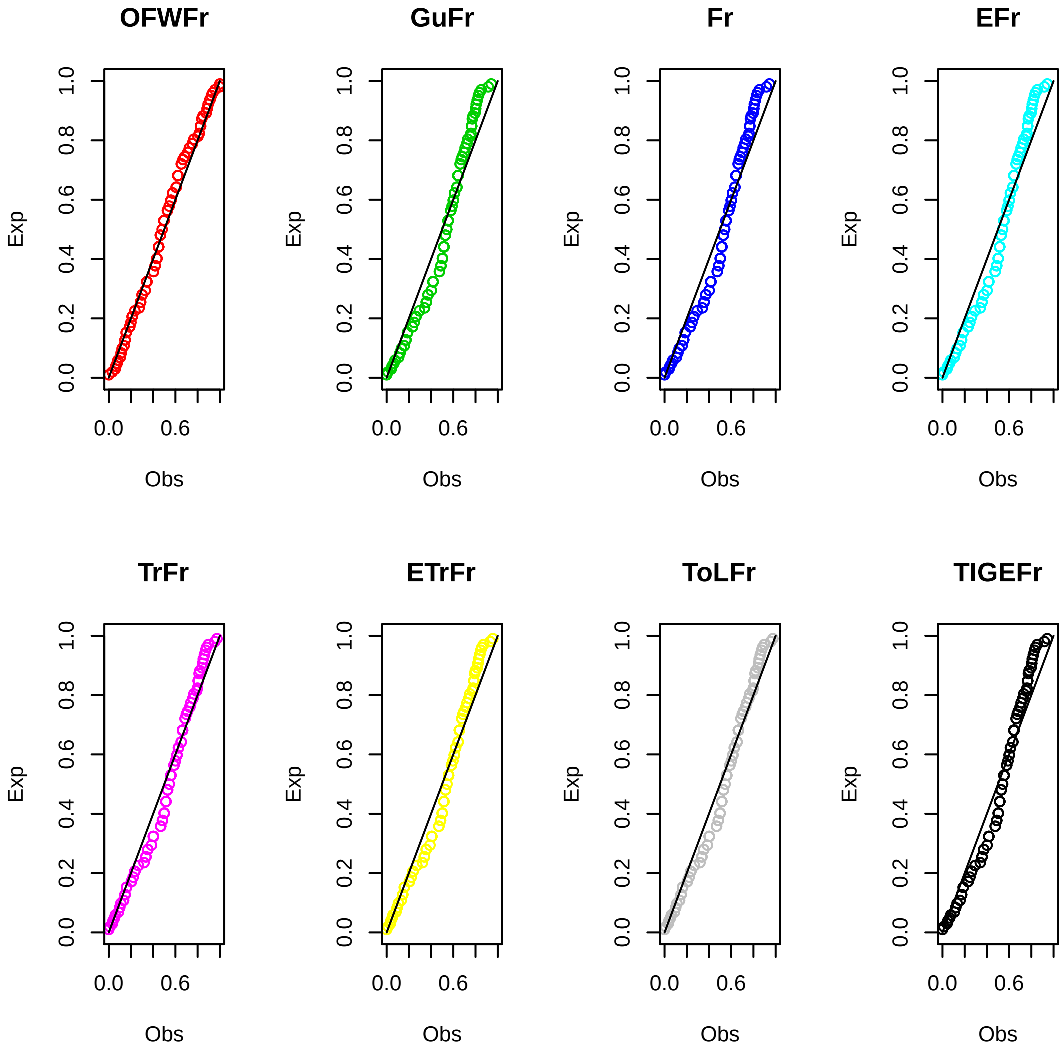

Figure 10. The P-P plots for data set I.

From Figure 10, it is clear that the aluminum coupons data set plausibly came from all tested distributions. But, the OFWFr model is the best. Figure 11 shows the profiles of the *LL* function for OFWFr model.

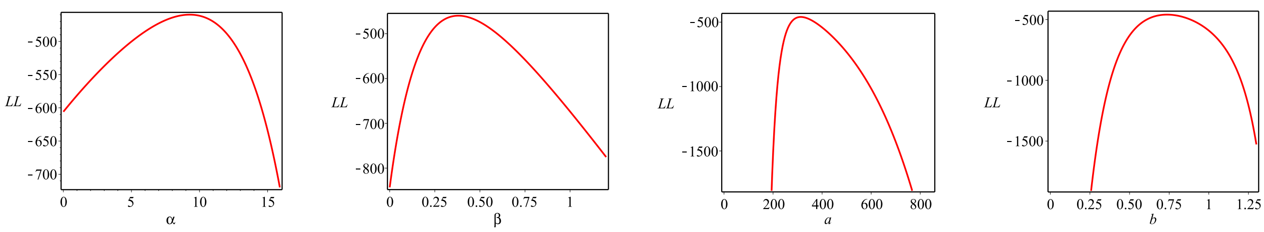

Figure 11. The profiles of the *LL* function for data set I.

It is clear that the parameters are unimodal functions. Some statistics for data set I are listed in Table 3 using the OFWFr model.

Table 3. Some statistics for data set I.

Method ↓ Measure →	**Mean**	**Variance**	**Skewness**	**Kurtosis**
MLE	128.397	491.369	0.102	1.123

It is observed that this data is suffering from over-dispersed where the index of dispersion ($\frac{\text{Variance}}{\text{Mean}} > 1$). Moreover, this data is moderately skewed right: its right tail is longer and most of the distribution is at the left with platykurtic.

8.2. *Data set II (Glass Fibers)*

The data is reported in Smith and Naylor (1987), which consists of 63 observations of the strengths of 1.5 cm glass fibers. This data set is used to compare the fits of the OFWE model with some competitive models like Topp-Leone exponential (ToLE), Gumbel exponential (GuE), exponentiated transmuted exponential (ETrE), genaralized exponential (GE), transmuted exponential (TrE) and exponential (E) distributions by using *R* software. The goodness-of-fit measures are reported in Tables 5 and 6.

Table 5. The MLEs, KS and p-values for data set II.

Model	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{a}$	KS	p-value
OFWE	3.293	1.419	0.311	0.166	0.062
ToLE	55.088	0.435	2.196	0.225	0.003
GuE	34.50	4.309	11.447	0.222	0.004
ETrE	−0.434	22.23	2.619	0.228	0.003
GE	−	31.37	2.611	0.229	0.003
TrE	−1.00	−	0.963	0.343	7.3×10^{-7}
E	−	−	0.664	0.418	5.5×10^{-10}

Table 6. The *LL*, AIC, CAIC, BIC, HQIC, $\mathbf{A}^*$ and $\mathbf{W}^*$ values for data set II.

Model	$-LL$	AIC	CAIC	BIC	HQIC	A^*	W^*
OFWE	16.09	38.19	38.60	44.62	40.72	1.497	0.273
ToLE	26.09	58.19	58.59	64.61	60.71	3.469	0.634
GuE	30.53	67.06	67.46	73.49	69.59	4.129	0.755
ETrE	31.23	68.47	68.88	74.89	70.99	4.267	0.783
GE	31.39	66.79	66.99	71.07	68.47	4.298	0.789
TrE	68.55	141.09	141.29	145.38	142.78	3.219	0.588
E	88.84	179.67	179.74	181.82	180.52	3.138	0.573

From Tables 5 and 6, it is observed that the OFWE model provides the best fit. The empirical PDFs, CDFs and P-P plots are displayed in Figures 12 and 13, which support the results of Tables 5 and 6.

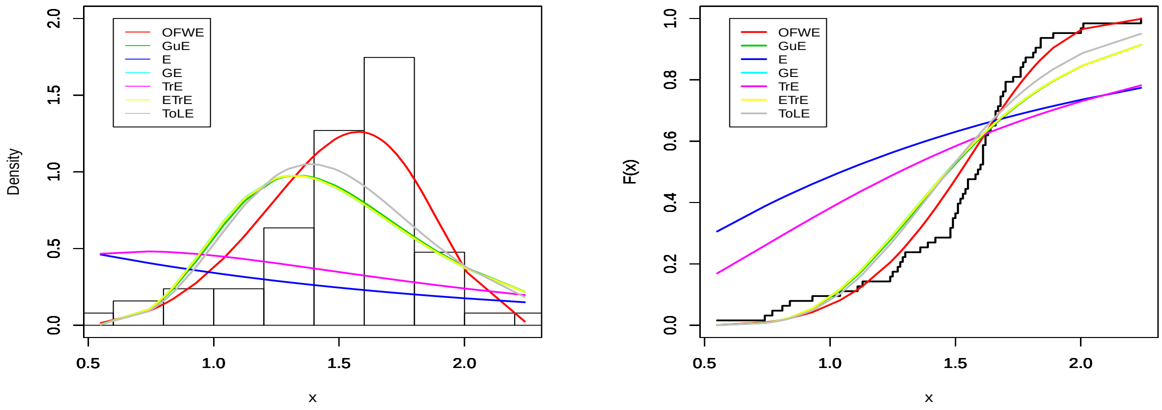

Figure 12. Estimated PDFs (left panel) and CDFs (right panel) for data set II.

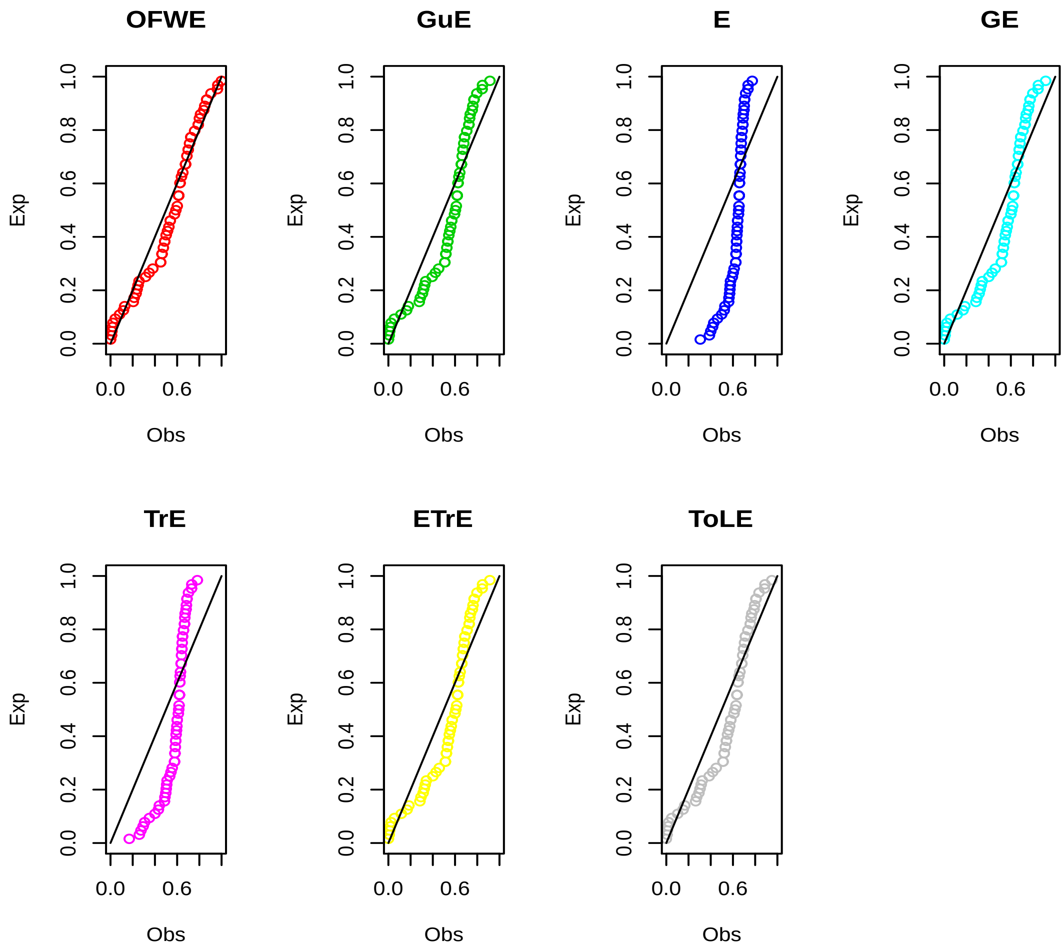

Figure 13. The P-P plots for data set II.

From Figure 13, it is clear that the glass fibers data set plausibly came from some tested distributions, especially, the OFWE model. Figure 14 shows the profiles of the LL function for OFWE model, which represents that the estimators are unique.

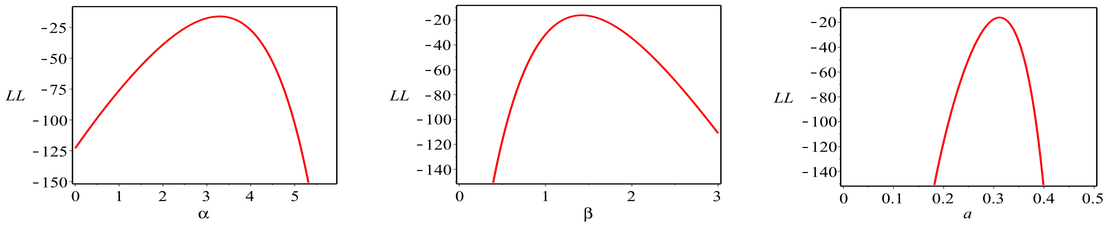

Figure 14. The profiles of the LL function for data set II.

Some statistics for data set II are listed in Table 7 using the OFWE model.

Table 7. Some statistics for data set II.

Method ↓ Measure →	**Mean**	**Variance**	**Skewness**	**Kurtosis**
MLE	1.103	0.087	−0.127	1.243

It is clear that this data is suffering from under-dispersed where the index of dispersion ($\frac{\text{Variance}}{\text{Mean}} < 1$). Further, this data is moderately skewed right with platykurtic.

8.3. *Data set III (Electrical Insulating Fluid)*

The data is reported in Lawless (1982) which consists of $n = 11$ times to breakdown of electrical insulating fluid subjected to 30 kilovolts. The data under a logarithm transformation is 2.836, 3.120, 3.045, 5.169, 4.934, 4.970, 3.018, 3.770, 5.272, 3.856 and 2.046. From this data, it is found that the UP-RC values are 2.836, 3.120, 5.169 and 5.272. The OFWFr distribution is used to analyze this data by using Mathcad software based on initial values $\alpha = 0.12$, $\beta = 0.3$, $a = 1.3$ and $b = 0.12$. The MLE(s), L^*, KS and its p-value are listed in Table 8.

Table 8. The MLEs, L^*, KS and p-values using UP-RC values from data set III.

Model	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{a}$	$\widehat{b}$	$-L^*$	KS	p-value
OFWFr	6.754	0.0184	35.981	0.578	5.89	0.377	0.619

From Table 8, it is clear that the OFWFr model provides a good fit for this data. The empirical PDF, CDF and P-P plots are displayed in Figure 15 which supports the results of Table 8.

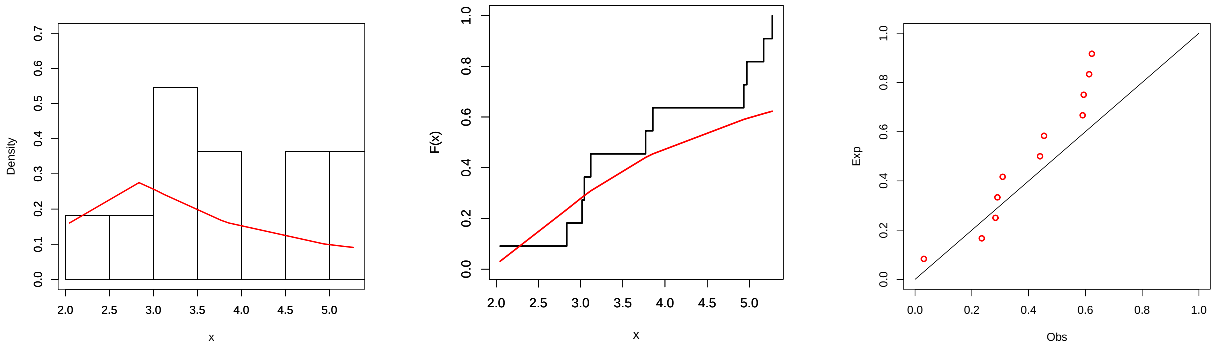

Figure 15. The fitted PDF (left panel), the estimated CDF (middle panel) and P-P plot (right panel) of OFWFr using UP-RC values from data set III.

Figure 16 shows the TTT plots for data sets I, II and III, and it is observed that the TTT plots indicate to these data sets have an increasing HRF.

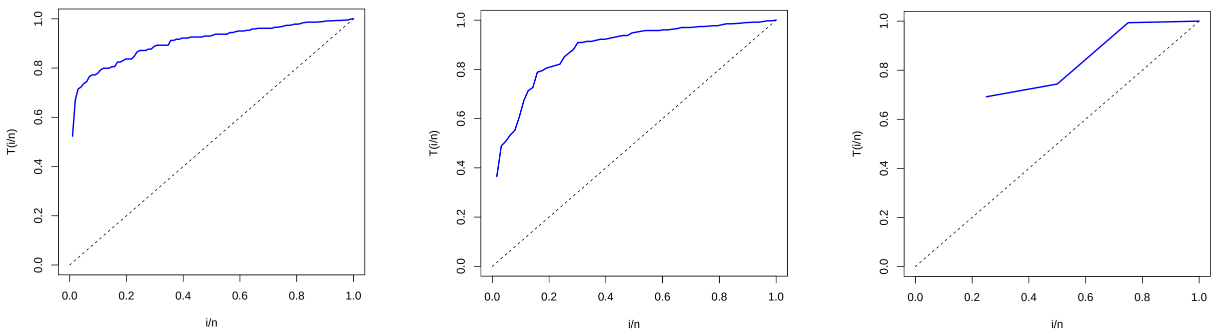

Figure 16. The TTT plots of data set I (left panel), data set II (middle panel) and UP-RC values of data set III.

9. Concluding Remarks

In this article, we have proposed a new family of distributions called the odd flexible Weibull-H (OFW-H) family. Some of its mathematical and statistical properties have been derived. It is found that the HRF of the OFW-H family can be increasing, decreasing, unimodal and bathtub-shaped. Furthermore, the OFW-H family is capable of modelling symmetric and negative as well as positive skewness data sets; in addition the generation of random samples from the OFW-H family is very simple. Therefore, Monte Carlo simulation can be performed very easily for different statistical inference purpose. The model parameters have been estimated by the maximum likelihood method based on complete and upper record data. Finally, the flexibility of the OFW-H family has been illustrated by means of three different real data sets.

References

[1] Ahsanullah, M., (2004). Record values: theory and application. University Press of America, New York.
[2] Alizadeh, M., Afify, A. Z., Eliwa, M. S., and Sajid Ali, (2019). The odd log-logistic Lindley-G family of distributions: properties, Bayesian and non-Bayesian estimation with applications. Computational statistics, DOI: 10.1007/s00180-019-00932-9.

[3] Alzaatreh, A., Lee, C., and Famoye, F., (2013). A new method for generating families of continuous distributions. Metron, 71(1), 63-79.

[4] Bebbington, M. S., Lai, C. D., and Zitikis, R., (2007). A flexible Weibull extension. Reliability engineering and system safety, 92(6), 719-726.

[5] Birnbaum, Z. W., and Saunders, S. C., (1969). Estimation for a family of life distributions with applications to fatigue. Journal of applied probability, 6, 328-347.

[6] Bourguignon, M., Silva, R. B., and Cordeiro, G. M., (2014). The Weibull-G family of probability distributions. Journal of data science, 12, 53-68.

[7] Cooray, K., (2006). Generalization of the Weibull distribution: The odd Weibull family. Statistical modelling, 6, 265-277.

[8] Cordeiro, G. M., Silva, G. O., and Ortega, E. M., (2012). The beta extended Weibull distribution. Journal of probability and statistical science, 10, 15-40.

[9] El-Bassiouny, A. H., Medhat EL-Damcese, Abdelfattah Mustafa, and Eliwa, M. S., (2017). Exponentiated generalized Weibull-Gompertz distribution with application in survival analysis. Journal of statistics applications and probability, 6(1), 7-16.

[10] El-Bassiouny, A. H., Medhat EL-Damcese, Abdelfattah Mustafa, and Eliwa, M. S., (2016). Mixture of exponentiated generalized Weibull-Gompertz distribution and its applications in reliability. Journal of statistics applications and probability, 5(3), 455-468.

[11] El-Bassiouny, A. H., Medhat EL-Damcese, Abdelfattah Mustafa, and Eliwa, M. S., (2015). Characterization of the generalized Weibull-Gompertz distribution based on the upper record values. International journal of mathematics and its applications, 3(3), 13-22.

[12] El-Gohary, A., EL-Bassiouny, A. H., and El-Morshedy, M., (2015a). Exponentiated flexible Weibull extension distribution. International journal of mathematics and its applications, 3(A), 1-12.

[13] El-Gohary, A., El-Bassiouny, A. H., and El-Morshedy, M., (2015b). Inverse flexible Weibull extension distribution. International journal of computer applications, 115, 46-51.

[14] Eliwa, M. S., and El-Morshedy, M., (2019). Bivariate odd Weibull-V family of distributions: properties, Bayesian and non-Bayesian estimation with bootstrap confidence intervals and application. Journal of Taibah University for science. To appear.

[15] Eliwa, M. S., El-Morshedy, M. and Afify, A. Z. (2019). The odd Chen generator of distributions: properties and estimation methods with applications in medicine and engineering. Journal of the national science foundation of Sri Lanka. To appear.

[16] El-Morshedy, M., El-Bassiouny, A. H., and El-Gohary, A., (2017). Exponentiated inverse flexible Weibull extension distribution. Journal of statistics applications and probability, 6(1), 169-183.

[17] El-Morshedy, M., Eliwa, M. S., El-Gohary, A., and Khalil, A. A., (2019). Bivariate exponentiated discrete Weibull distribution: Statistical properties, estimation, simulation and applications. Mathematical sciences, DOI 10.1007s40096-019-00313-9.

[18] Gurvich, M. R., Dibenedetto, A. T., and Rande, S. V., (1997). A new statistical distribution for characterizing the random length of brittle materials. Journal of materials science, 32, 2559-2564.

[19] Jehhan, A., Mohamed, I. , Eliwa, M. S., Al-mualim, S., and Yousof, H. M., (2018). The two-parameter odd lindley Weibull lifetime model with properties and applications. International journal of statistics and probability, 7(2), 57-68.

[20] Kenney, J. F., and Keeping, E. S., (1962). Mathematics of statistics. Princeton, New Jersey, Part 1, 3rd edition, 101-102.

[21] Lawless, J. F., (1982). Statistical models and methods for lifetime data, 2nd Edition, Wiley, New York.

[22] Moors, J. J., (1998). A quantile alternative for kurtosis. Journal of the Royal statistical society series D: The Statistician, 37, 25-32.

[23] Saad, J. A., and Jingsong,Y., (2013). A new modified Weibull distribution. Reliability engineering and system safety, 111, 164-170.

[24] Sangun, P., and Jiwhan, P., (2018). A general class of flexible Weibull distributions. Communications in statistics: theory and methods, 47(4), 767-778.

[25] Silva, G. O., Ortega, E. M., and Cordeiro, G. M., (2010). The beta modified Weibull distribution. Lifetime data analysis, 16, 409-430.

[26] Soliman, A., Amin, E. A., and Abd-El Aziz, A. A., (2010). Estimation and prediction from inverse Rayleigh distribution based on lower record values. Applied Mathematical Sciences, 62, 3057-3066.

[27] Smith, R. L., and Naylor, J. C., (1987). A comparison of maximum likelihood and Bayesian estimatorsfor the three-parameter Weibull distribution. Applied statistics, 36, 358-369.