Available online at www.worldscientificnews.com

World Scientific News

An International Scientific Journal

WSN 191 (2024) 81-101 EISSN 2392-2192

Extended odd Frechet half logistic distribution with it's properties and applications

Aisha Sani Bawa*, **Sani Ibrahim Doguwa**, **Tasi'u Musa** and **Musaddiq Sirajo**

Department of Statistics, Ahmadu Bello University, Zaria, Nigeria

*E-mail address: sanibawaaisha30@gmail.com

ABSTRACT

Lifetime distributions describe the behavior of the length of life of individuals or components in survival or reliability analyses. They are important tools in modeling the different characteristics of lifetime data sets emanating from various fields of human endeavor. Many lifetime distributions exist in the statistical literature but are commonly characterized with having many parameters which may cause estimation related problems. To trade-off between simplicity and flexibility in modeling lifetime data sets with half logistic distribution, a new extension is proposed in this paper by using the extended odd Frechet-G family of distributions. The new distribution has only two parameters and simple mathematical form that can be interpreted in terms of odds ratio. The statistical properties of the distribution, including moments, quantile function and order statistics are studied. The unknown parameters were estimated by using two different estimation methods, namely, maximum likelihood and maximum product of spacing. Monte Carlo simulation study is undertaken to compare the finite sample performance of these parameter estimation methods based on generated samples using the quantile function of the new distribution. To demonstrate suitability in favor of the proposed distribution, three real data sets were analyzed and compared with four competitive models, two from the extended odd Frechet-G family and the remaining two having the same baseline distribution as the proposed. Empirical findings show that the new two-parameter distribution compared well to the four-parameter distributions of the same family and produced better results than the other extensions of half logistic distribution.

***Keywords*:** Half logistic, Extended odd Frechet-G family, Maximum likelihood, Maximum product of spacing, Quantile function, Order statistics

(Received 11 February 2024; Accepted 04 March 2024; Date of Publication 06 March 2024)

1. INTRODUCTION

The Half logistic (HL) distribution is equivalent to a logistic distribution with only positive values and a zero mean. This popular classical distribution is generally utilized for modeling failure times of components. It has an increasing failure rate function, and a decreasing probability density function (pdf) given by

$$g(x)=\frac{2e^{-x}}{\left(1+e^{-x}\right)^{2}},\quad x>0. \tag{1}$$

The equivalent cumulative distribution function (cdf) is given by

$$G(x)=\frac{1-e^{-x}}{1+e^{-x}},\quad x>0. \tag{2}$$

Despite its popularity in many fields such as in survival analysis and reliability studies, the HL distribution have one major disadvantage, that is, it cannot model unimodal data with non-monotone failure rates. To address this, many extensions have been proposed. A good example is the type I HL family by Cordeiro *et al*. (2016). However, none of the existing extensions of HL distribution is a member of extended odd Frechet-G (EOFG) family. It is shown in (Nasiru, 2018) by means of two successful applications that members of the EOFG family are easily applicable for modeling purposes. However, only two special members of the EOFG family were comprehensively studied.

These are the four-parameter extended odd Frechet Nadarajah-Haghighi (EOFNH) and four-parameter extended odd Frechet Weibull (EOFW) distributions which were each generated based on two-parameter baseline distributions, resulting to distributions with increased number of parameters. Having too many parameters may cause a problem in estimation results.

Results in statistical modeling suggest that decreasing the number of parameters to be estimated can help detect misspecification errors (Jackson, 2007). Thus, a two-parameter extended odd Frechet HL (EOFHL) distribution is hereby proposed. To the best of our knowledge, the direction of this work remains new and promising in view of the respective qualities of HL distribution and EOFG family.

For any given baseline cdf, $G(x)$, and pdf, $g(x)$, Nasiru (2018) defined the EOFG family of distributions with respective cdf, pdf and failure rate function given by

$$F(x)=e^{-\left[\left(1-G(x)^{\alpha}\right)/G(x)^{\alpha}\right]^{\theta}}, \tag{3}$$

$$f(x)=\frac{\alpha\theta g(x)\left(1-G(x)^{\alpha}\right)^{\theta-1}}{G(x)^{\alpha\theta+1}}e^{-\left[\left(1-G(x)^{\alpha}\right)/G(x)^{\alpha}\right]^{\theta}}, \tag{4}$$

$$h(x)=\frac{\alpha\theta g(x)\left(1-G(x)^{\alpha}\right)^{\theta-1}}{G(x)^{\alpha\theta+1}\left(1-e^{-\left[\left(1-G(x)^{\alpha}\right)/G(x)^{\alpha}\right]^{\theta}}\right)}e^{-\left[\left(1-G(x)^{\alpha}\right)/G(x)^{\alpha}\right]^{\theta}}. \quad (5)$$

where $x\in\mathrm{R}$, and $\alpha,\theta>0$ are shape parameters.

In this paper, HL distribution is used as the baseline. The rest of this paper is organized as follows. In Section 2, the EOFHL is defined and some of its statistical properties established, including parameter estimation using the methods of maximum likelihood and maximum product of spacing. In section 3, a Monte Carlo simulation study was carried out to assess the maximum product of spacing as well as maximum likelihood estimated parameters. An illustration on the basis of real datasets is provided in Section 4. Finally, Section 5 concludes the paper.

2. THE EXTENDED ODD FRECHET HALF LOGISTIC DISTRIBUTION

By substituting equations (1) and (2) into equations (3) and (4), the EOFHL distribution is defined with pdf and cdf respectively obtained as

$$f(x)=\frac{2\alpha\theta e^{-x}\left[1-\left(\frac{1-e^{-x}}{1+e^{-x}}\right)^{\alpha}\right]^{\theta-1}\Big/\left(1+e^{-x}\right)^{2}}{\left(\frac{1-e^{-x}}{1+e^{-x}}\right)^{\alpha\theta+1}}\times\exp\left\{-\left[\left(1-\left(\frac{1-e^{-x}}{1+e^{-x}}\right)^{\alpha}\right)\Big/\left(\frac{1-e^{-x}}{1+e^{-x}}\right)^{\alpha}\right]^{\theta}\right\}, \quad (6)$$

$$F(x)=\exp\left\{-\left[\left(1-\left(\frac{1-e^{-x}}{1+e^{-x}}\right)^{\alpha}\right)\Big/\left(\frac{1-e^{-x}}{1+e^{-x}}\right)^{\alpha}\right]^{\theta}\right\},\ x,\alpha,\theta>0. \quad (7)$$

The physical interpretation of the EOFHL distribution can be given as follows. Let the random variable *X* represents lifetimes of individuals following the HL distribution, given in equation (1). Consider that we are interested to model the odds that an individual dies before a given time *X* which is given by $F(x)/(1-F(x))$ where $F(x)$ is the cdf of the HL distribution, given in equation (2).

Let's consider that we are required to model the randomness of the odds by the random variable *Y*, which follows the Frechet distribution. Then, we can use the EOFHL distribution. The pdf of EOFHL model is unimodal and exhibits different shapes such as near symmetric, right-skewed, and decreasing. Figures 1 and 2 demonstrate the different shapes which are novel and adaptable.

The failure rate function is defined from equation (5) as

$$h(x)=\frac{2\alpha\theta e^{-x}\left[1-\left(\frac{1-e^{-x}}{1+e^{-x}}\right)^{\alpha}\right]^{\theta-1}\Big/\left(1+e^{-x}\right)^{2}}{\left(\frac{1-e^{-x}}{1+e^{-x}}\right)^{\alpha\theta+1}\left\{1-\exp\left(-\left[\left(1-\left(\frac{1-e^{-x}}{1+e^{-x}}\right)^{\alpha}\right)\Big/\left(\frac{1-e^{-x}}{1+e^{-x}}\right)^{\alpha}\right]^{\theta}\right)\right\}}\times\exp\left\{-\left[\left(1-\left(\frac{1-e^{-x}}{1+e^{-x}}\right)^{\alpha}\right)\Big/\left(\frac{1-e^{-x}}{1+e^{-x}}\right)^{\alpha}\right]^{\theta}\right\}. \quad (8)$$

where α, $\theta > 0$ are shape parameters and $x > 0$.

The failure rate function can assume decreasing, constant, and other non-monotonic failure rate forms. This is demonstrated in Figures 3 and 4. Therefore, if the empirical study suggests a non-monotone failure rate function which has a unimodal shape, then the EOFHL distribution may be adopted for the analysis of such data sets.

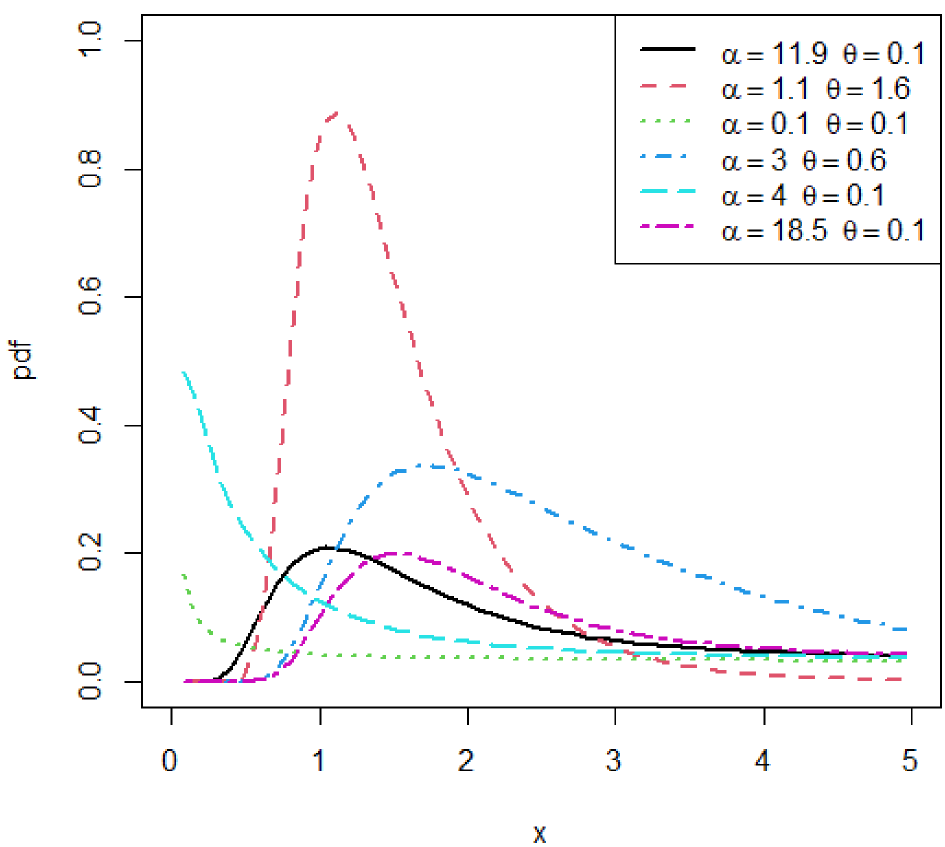

Fig. 1. Pdf of EOFHL distribution at some selected parameter values

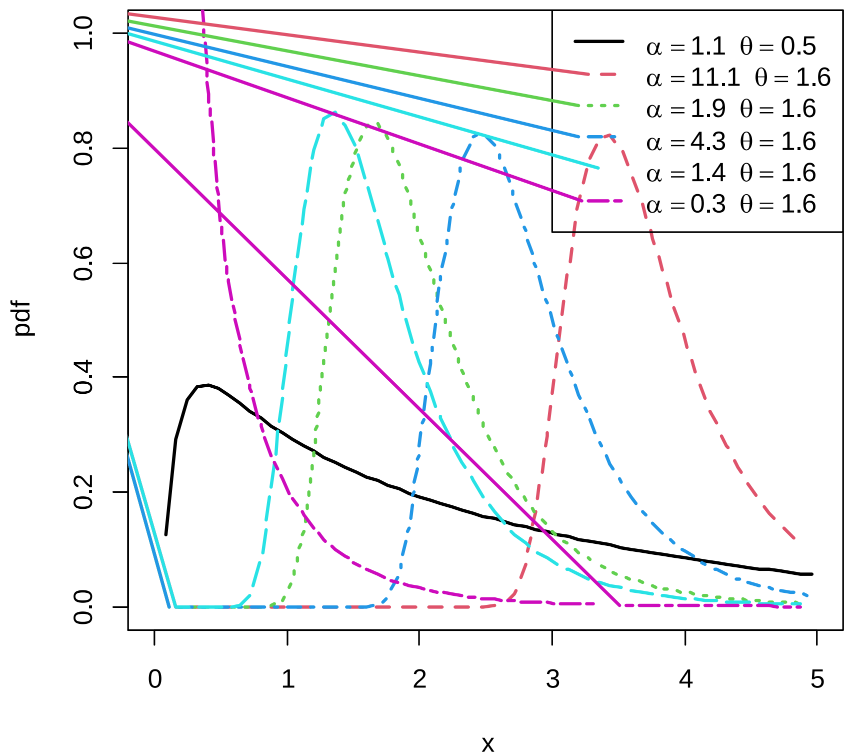

Fig. 2. Pdf of EOFHL distribution at some selected parameter values

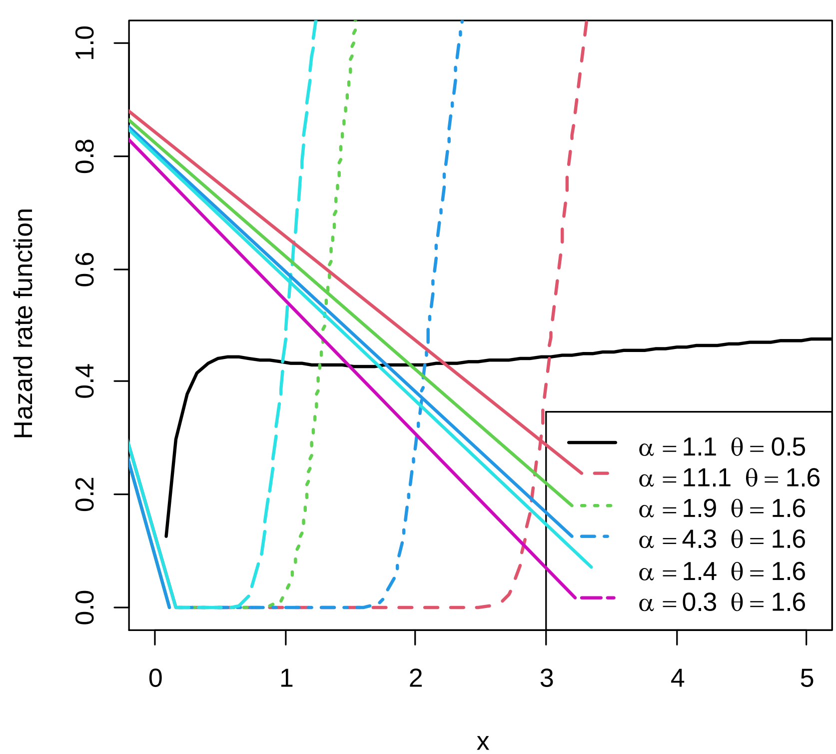

Fig. 3. Failure rate of EOFHL distribution at some selected parameter values

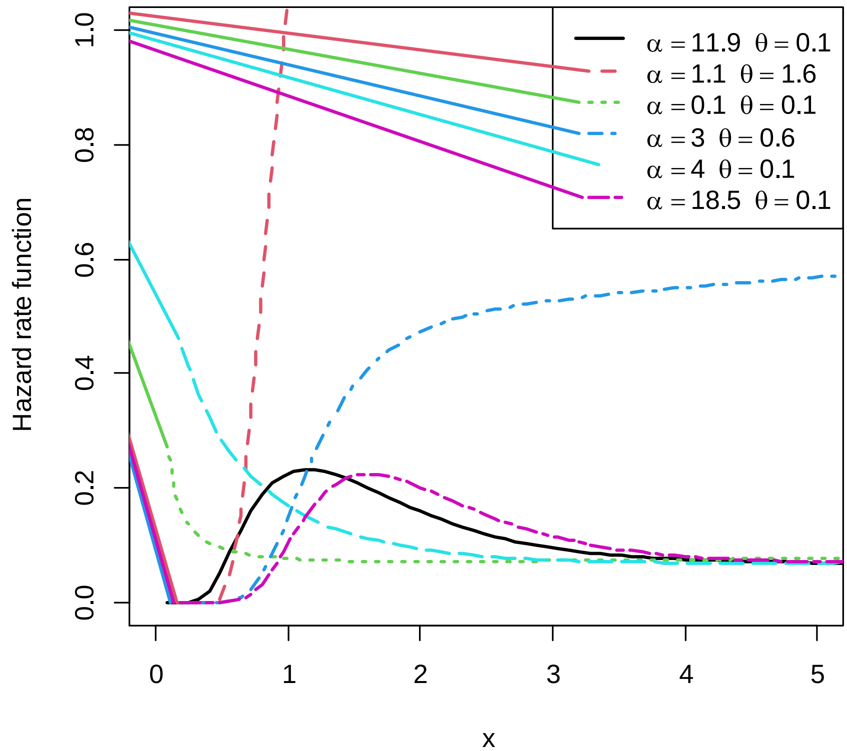

Fig. 4. Failure rate of EOFHL distribution at selected parameter values

2. 1. Mixture Representation

The pdf of EOFHL distribution can be represented in mixture form in terms of exponentiated-G density which will be useful when deriving the statistical properties of the new distribution. For details, readers are referred to Nasiru (2018).

Using the generalized binomial series

$$(1-z)^{\eta-1}=\sum_{j=0}^{\infty}(-1)^j\binom{\eta-1}{j}z^j, \quad |z|<1,$$

The pdf of EOFG family is defined as

$$f(x)=\alpha\theta\sum_{i,j=0}^{\infty}\sum_{k,m=0}^{\infty}\sum_{q=0}^{m+j}\omega_{ijkmq}(q+1)g(x)G(x)^q, \tag{9}$$

where $\omega_{ijkmq}=\frac{(-1)^{i+k+m+q}}{i!(q+1)}\binom{\alpha\theta(i+1)+j}{j}\binom{\theta(i+1)-1}{k}\binom{\alpha k}{m}\binom{m+j}{q}$, and $(q+1)^{-1}g(x)G(x)^q$ is the pdf of exponentiated-G with parameters α,θ and $(q+1)$.

The mixture representation of the cdf of EOFG is given by

$$F(x)=\alpha\theta\sum_{i,j=0}^{\infty}\sum_{k,m=0}^{\infty}\sum_{q=0}^{m+j}\omega_{ijkmq}G(x)^{q+1}, \tag{10}$$

where $G(x)^{q+1}$ is the cdf of exponentiated-G with power parameter $q+1$.

Consequently, the density of EOFHL distribution represented as a combination of exponentiated half logistic (EHL) distribution, is defined, using equation (9), by

$$f(x)=\alpha\theta\sum_{i,j=0}^{\infty}\sum_{k,m=0}^{\infty}\sum_{q=0}^{m+j}\omega_{ijkmq}\pi_{q+1}(x), \tag{11}$$

where $\pi_{q+1}(x)$ is the pdf of the EHL distribution with parameter $q+1$. This means that the EOFHL density can be expressed as a linear combination of EHL densities. Consequently, the properties of EOFHL distribution can be derived as linear combinations of those of the EHL distribution. For details on EHL distribution, readers are referred to Seo and Kang (2015).

2. 2. Moments

If sample X follows an EOFHL distribution with parameters α,θ and q, then the corresponding rth moment can be obtained from equation (11) as

$$E(X^r)=\alpha\theta\sum_{i,j=0}^{\infty}\sum_{k,m=0}^{\infty}\sum_{q=0}^{m+j}q\omega_{ijkmq}\int_0^{\infty}\left(\frac{1-e^{-x}}{1+e^{-x}}\right)^q\frac{2e^{-x}}{1-e^{-2x}}x^r dx.$$

Letting $t=\left(\frac{1-e^{-x}}{1+e^{-x}}\right)^q$,

$$E(X^r)=\alpha\theta\sum_{i,j=0}^{\infty}\sum_{k,m=0}^{\infty}\sum_{q=0}^{m+j}q\omega_{ijkmq}\int_0^{1}\left[\log\left(\frac{1+t^{1/q}}{1-t^{1/q}}\right)\right]^r dt.$$

Define $\log\left(\frac{1+w}{1-w}\right)=2\sum_{p=1}^{\infty}\frac{w^{2p-1}}{2p-1},$ for $w^2<1$.

The first moment is obtained as

$$E(X)=\alpha\theta\sum_{i,j=0}^{\infty}\sum_{k,m=0}^{\infty}\sum_{q=0}^{m+j}q\omega_{ijkmq}\int_0^{1}\left[\log\left(\frac{1+t^{1/q}}{1-t^{1/q}}\right)\right]dt,$$

$$E(X)=\alpha\theta\sum_{i,j=0}^{\infty}\sum_{k,m=0}^{\infty}\sum_{q=0}^{m+j}q\omega_{ijkmq}\,2q\,y_1(q),$$

$$E(X)=2\alpha\theta\sum_{i,j=0}^{\infty}\sum_{k,m=0}^{\infty}\sum_{q=0}^{m+j}q^{2}\omega_{ijkmq}\,y_{1}(q),\qquad(12)$$

where $y_{1}(q)=\sum_{p=1}^{\infty}\frac{1}{(2p-1)(2p-1+q)}$.

Similarly,

$$E\left(X^{2}\right)=\alpha\theta\sum_{i,j=0}^{\infty}\sum_{k,m=0}^{\infty}\sum_{q=0}^{m+j}q\omega_{ijkmq}\int_{0}^{1}\left[\log\left(\frac{1+t^{1/q}}{1-t^{1/q}}\right)\right]^{2}dt.$$

The integrand may be decomposed as follows

$$\left[\log\left(\frac{1+t^{1/q}}{1-t^{1/q}}\right)\right]^{2}=\left[\log\left(1+t^{1/q}\right)\right]^{2}+\left[\log\left(1-t^{1/q}\right)\right]^{2}-2\log\left(1+t^{1/q}\right)\log\left(1-t^{1/q}\right).$$

Then, with the use of the two series expansions

$$\left[\log(1\pm w)\right]^{2}=2\sum_{p=1}^{\infty}(\pm1)^{p+1}\frac{w^{p+1}}{p+1}\sum_{z=1}^{p}\frac{1}{z}\qquad\text{for}\qquad w^{2}<1,$$

$$\log(1+w)\log(1-w)=-\sum_{p=1}^{\infty}\frac{w^{2p}}{p}\sum_{z=1}^{2p-1}\frac{(-1)^{z+1}}{z}\qquad\text{for}\qquad w^{2}<1.$$

The second moment is thus obtained as

$$E\left(X^{2}\right)=\alpha\theta\sum_{i,j=0}^{\infty}\sum_{k,m=0}^{\infty}\sum_{q=0}^{m+j}q\omega_{ijkmq}\,2q\left[y_{2}(q)+y_{3}(q)\right],$$

$$E\left(X^{2}\right)=2\alpha\theta\sum_{i,j=0}^{\infty}\sum_{k,m=0}^{\infty}\sum_{q=0}^{m+j}q^{2}\omega_{ijkmq}\left[y_{2}(q)+y_{3}(q)\right],\qquad(13)$$

where

$$y_{2}(q)=\sum_{p=1}^{\infty}\frac{1+(-1)^{p+1}}{(p+1)(p+1+q)}\sum_{z=1}^{p}\frac{1}{z},\quad\text{and}\quad y_{3}(q)=\sum_{p=1}^{\infty}\frac{1}{p(2p+q)}\sum_{z=1}^{2p-1}\frac{(-1)^{z+1}}{z}.$$

More detailed accounts on the derivation can be found in Seo & Kang (2015).

2. 3. Quantile Function

The quantile function of EOFHL distribution can be obtained in a closed form and it is given as defined in Theorem 1.

Theorem 1. For any fixed $\alpha, \theta > 0$, the quantile function of EOFHL random variable is

$$x = \log\left\{1+\left[1+\left(-\log u\right)^{1/\theta}\right]^{-1/\alpha}\right\} - \log\left\{1-\left[1+\left(-\log u\right)^{1/\theta}\right]^{-1/\alpha}\right\}, \quad u \in [0,1]. \tag{14}$$

Proof. Using equation (7) to solve for x in the equation $F(x) = u$, the following is obtained

$$F(x) = \exp\left\{-\left[\frac{1-\left(\frac{1-e^{-x}}{1+e^{-x}}\right)^{\alpha}}{\left(\frac{1-e^{-x}}{1+e^{-x}}\right)^{\alpha}}\right]^{\theta}\right\} = u,$$

$$\left[\frac{1-\left(\frac{1-e^{-x}}{1+e^{-x}}\right)^{\alpha}}{\left(\frac{1-e^{-x}}{1+e^{-x}}\right)^{\alpha}}\right]^{\theta} = -\log u,$$

$$\left(\frac{1-e^{-x}}{1+e^{-x}}\right)^{-\alpha} - 1 = \left(-\log u\right)^{1/\theta},$$

$$\frac{1-e^{-x}}{1+e^{-x}} = \left[1+\left(-\log u\right)^{1/\theta}\right]^{-\frac{1}{\alpha}}.$$

On letting $k = \left[1+\left(-\log u\right)^{1/\theta}\right]^{-\frac{1}{\alpha}}$,

$$1-e^{-x} = k\left(1+e^{-x}\right) \Rightarrow e^{-x} = \frac{1-k}{1+k},$$

So that, $x = -\log\left(\frac{1-k}{1+k}\right)$, or equivalently $x = \log(1+k) - \log(1-k)$, which leads to the result.

The random variables from EOFHL distribution can be generated by using the quantile function in equation (14). To do this, the following algorithm steps can be implemented.

1). Set the parameters $\alpha, \theta > 0$.

2). Generate $u \sim U(0,1)$.

3). Using the generated u, calculate

$x = \log\left\{1+\left[1+(-\log u)^{1/\theta}\right]^{-1/\alpha}\right\} - \log\left\{1-\left[1+(-\log u)^{1/\theta}\right]^{-1/\alpha}\right\}.$

4). Repeat the steps 2 and 3 N times.

2. 4. Order statistics

If X is EOFG random variable, then the pdf of the r*th* order statistics can be expressed in terms of exponentiated-G density as

$$f_{r:n}(x) = \frac{\alpha\theta n!}{(r-1)!(n-r)!}\sum_{j,k=0}^{\infty}\sum_{q,s=0}^{\infty}\sum_{\omega=0}^{k+s}\sum_{i=0}^{n-r}\varphi_{ijkqs\omega}\Delta_{\omega+1}(x),$$

where $\varphi_{ijkqs\omega} = \frac{(-1)^{i+j+q+s+\omega}(r+i)^{j}}{(\omega+1)\,j!}\binom{n-r}{i}\binom{\alpha\theta(j+1)+k}{k}\binom{\theta(j+1)-1}{q}\binom{\alpha q}{s}\binom{k+s}{\omega}$ and $\Delta_{\omega+1}(x) = (\omega+1)g(x)G(x)^{\omega}$ is an exponentiated-G density with power parameter $\omega+1$.

Following this definition and using the EHL density, the density of the rth order statistics of EOFHL distribution is obtained as,

$$f_{r:n}(x) = \frac{\alpha\theta n!}{(r-1)!(n-r)!}\sum_{j,k=0}^{\infty}\sum_{q,s=0}^{\infty}\sum_{\omega=0}^{k+s}\sum_{i=0}^{n-r}\varphi_{ijkqs\omega}\pi_{q+1}(x). \tag{15}$$

Consequently, the moments of the order statistics can be computed in similar fashion as the ordinary moments defined in (12) and (13).

2. 5. Maximum likelihood estimation

Let $x_1, x_2, ..., x_n$ be a random sample from sample X that follows EOFG, then the log-likelihood function is given by

$$ll = n\log(\alpha\theta) + \sum_{i=1}^{n}\log g(x) + (\theta-1)\sum_{i=1}^{n}\log\left[1-G(x)^{\alpha}\right] - (\alpha\theta+1)\sum_{i=1}^{n}\log G(x) - \sum_{i=1}^{n}\left[\frac{1-G(x)^{\alpha}}{G(x)}\right]. \tag{16}$$

The maximum likelihood (ML) estimates of the parameters can be obtained by maximizing equation (16). Let the vector of parameters be $(\alpha, \theta, \vartheta)$, then the partial derivatives of (16) are as follows

$$\frac{\partial ll}{\partial \alpha}=\frac{n}{\alpha}+(\theta-1)\sum_{i=1}^{n}\frac{G(x)^{\alpha}\log G(x)}{1-G(x)}-\theta\sum_{i=1}^{n}\log G(x)+\theta\sum_{i=1}^{n}\frac{\left[1-G(x)^{\alpha}\right]^{\theta-1}\log G(x)}{G(x)^{\alpha\theta}}.$$

$$\frac{\partial ll}{\partial \theta}=\frac{n}{\theta}+\sum_{i=1}^{n}\log\left[1-G(x)^{\alpha}\right]-\alpha\sum_{i=1}^{n}\log G(x)-\sum_{i=1}^{n}\left[\frac{1-G(x)^{\alpha}}{G(x)^{\alpha}}\right]^{\theta}\log\left[\frac{1-G(x)^{\alpha}}{G(x)^{\alpha}}\right].$$

Generally,

$$\frac{\partial ll}{\partial \vartheta}=\sum_{i=1}^{n}\frac{g^{'}(x;\vartheta)}{g(x;\vartheta)}+\alpha(\theta-1)\sum_{i=1}^{n}\frac{G^{'}(x;\vartheta)G(x;\vartheta)^{\alpha-1}}{1-G(x;\vartheta)}$$
$$-(\alpha\theta+1)\sum_{i=1}^{n}\frac{G^{'}(x;\vartheta)}{G(x;\vartheta)}+\alpha\theta\sum_{i=1}^{n}\frac{G^{'}(x;\vartheta)\left[1-G(x;\vartheta)^{\alpha}\right]^{\theta-1}}{G(x;\vartheta)^{\alpha\theta+1}},$$

where $g^{'}(x;\vartheta)=\partial g(x;\vartheta)/\partial\vartheta$ and $G^{'}(x;\vartheta)=\partial G(x;\vartheta)/\partial\vartheta$.

The log-likelihood function and its partial derivatives for the EOFHL distribution can easily be obtained by setting $g(x)$ and $G(x)$ equals the pdf and cdf of HL distribution defined in equations (1) and (2) respectively.

2. 6. Maximum product of spacing estimation

Let $F(X_{j:n})$ denote the distribution function of the ordered random variables $X_{1:n}<X_{2:n}<...<X_{n:n}$ where $\{X_1,X_2,...,X_n\}$ is a random sample of size n from the EOFHL distribution function $F(x|\alpha,\theta)$. Using this definition, the uniform spacings of a random sample from the EOFHL distribution is defined as:

$$D_j(\alpha,\theta)=F(x_{j:n}|\alpha,\theta)-F(x_{j-1:n}|\alpha,\theta),\ i=1,2,...,n. \tag{17}$$

where $F(x_{0:n}|\alpha,\theta)=0,\ F(x_{n+1:n}|\alpha,\theta)=1,$ and $\sum_{j=1}^{n+1}D_j(\alpha,\theta)=1.$

Following Sirajo *et al*. (2021) and the references therein, the maximum product of spacing (MPS) estimates of α and θ are obtained by maximizing the geometric mean of the spacings with respect to α and θ.

The geometric mean is defined as

$$G(\alpha,\theta)=\left[\prod_{j=1}^{n+1}D_j(\alpha,\theta)\right]^{\frac{1}{n+1}}. \tag{18}$$

3. SIMULATION

In this section, simulation study is carried out to assess the accuracy and consistency of the MPS and ML estimators in estimating the parameters of EOFHL distribution. For the purpose of illustration, the simulation is performed using the maximum likelihood estimators of the parameters of the EOFHL distribution as well as the maximum product of spacing estimators. The quantile function defined in equation (14) is used to generate data consisting of random observations from the EOFHL distribution. The algorithm below is used for the Monte Carlo simulation.

1 000 iteration number were determined.
Different sample sizes n are considered to be 30, 75, 100, 300.
Different cases of actual values were used as follows:
Case I: $\alpha = 3,\ \theta = 0.6$.
Case II: $\alpha = 1.1,\ \theta = 0.5$.
Case III: $\alpha = 1.9,\ \theta = 1.6$.
Random sample X generated from the EOFHL distribution.
ML and MPS estimation methods used to estimate the unknown parameters of EOFHL distribution.
Steps 4 and 5 were repeated to obtain the estimates for 1 000 iteration.
Different measures were used as follows: $Bias = \left(\hat{\Theta} - \Theta\right)$, $MSE = Mean\left(\hat{\Theta} - \Theta\right)^2$ and $RMSE = \sqrt{MSE}$.

Table 1 presents the average estimated parameter values (Mean) with associated bias and root mean square error (RMSE) for the estimators of the parameters using the methods of ML and MPS. The results indicated that the RMSEs decrease as the sample size increases. These results clearly show the accuracy and the consistency of the maximum product of spacing as well as maximum likelihood estimators. Also, the average values are quite close to the nominal values, especially for the maximum product of spacing. Thus, these methods work very well to estimate the parameters of the EOFHL distribution.

Table 1. ML and MPS estimates of the parameters of EOFHL distribution

Case	n	Parameter	MLE			MPS		
			Mean	Bias	RMSE	Mean	Bias	RMSE
I	30	$\alpha = 3$	2.9172	-0.0828	0.1434	3.0180	0.0180	0.5932
		$\theta = 0.6$	0.6443	0.0443	0.0767	0.5760	-0.0240	0.1036
	75	$\alpha = 3$	3.2268	0.2268	0.3928	2.9988	-0.0012	0.3613
		$\theta = 0.6$	0.5883	-0.0117	0.0202	0.5860	-0.0140	0.0657
	100	$\alpha = 3$	3.1486	0.1486	0.2573	3.0054	0.0054	0.3127
		$\theta = 0.6$	0.6054	0.0054	0.0094	0.5876	-0.0124	0.0552
	300	$\alpha = 3$	3.0331	0.0331	0.0572	2.9970	-0.0030	0.1739
		$\theta = 0.6$	0.6001	0.0001	0.0002	0.5926	-0.0074	0.0318

	30	$\alpha=1.1$	1.0735	-0.0265	0.0459	1.1095	0.0095	0.2466
		$\theta=0.5$	0.5351	0.0351	0.0609	0.4809	-0.0191	0.0882
	75	$\alpha=1.1$	1.1898	0.0898	0.1555	1.0997	-0.0003	0.1493
		$\theta=0.5$	0.4896	-0.0104	0.0179	0.4889	-0.0111	0.0561
II	100	$\alpha=1.1$	1.1575	0.0575	0.0995	1.1024	0.0024	0.1286
		$\theta=0.5$	0.5044	0.0044	0.0075	0.4900	-0.0100	0.0471
	300	$\alpha=1.1$	1.1150	0.0150	0.0260	1.0985	-0.0015	0.0713
		$\theta=0.5$	0.4996	-0.0004	0.0007	0.4940	-0.0060	0.0272
	30	$\alpha=1.9$	1.8665	-0.0335	0.0581	1.8970	-0.0030	0.1563
		$\theta=1.6$	1.7204	0.1204	0.2085	1.5256	-0.0744	0.2575
	75	$\alpha=1.9$	1.9613	0.0613	0.1062	1.8968	-0.0032	0.1002
		$\theta=1.6$	1.5866	-0.0134	0.0232	1.5574	-0.0426	0.1605
III	100	$\alpha=1.9$	1.9450	0.0450	0.0780	1.9000	0.0000	0.0869
		$\theta=1.6$	1.6235	0.0235	0.0408	1.5632	-0.0368	0.1348
	300	$\alpha=1.9$	1.9058	0.0058	0.0100	1.8996	-0.0004	0.0484
		$\theta=1.6$	1.6056	0.0056	0.0097	1.5779	-0.0221	0.0775

4. APPLICATION

The fitting performance of the EOFHL distribution is compared with the following competitive models.

Exponentiated Half Logistic (EHL):

Seo & Kang (2015) defined a random variable X said to have EHL distribution with shape parameter $\alpha>0$ if the cdf is given by

$$F_{EHL}(x;\alpha)=\left(\frac{1-e^{-x}}{1+e^{-x}}\right)^{\alpha}, \qquad x>0.$$

Generalized Half Logistic (GHL):

Olapade (2014) defined the type 1 GHL distribution with a shape parameter $\alpha>0$ and cdf given by

$$F_{GHL}(x;\alpha)=1-\left(\frac{2e^{-x}}{1+e^{-x}}\right)^{\alpha}, \qquad x>0.$$

Exponentiated Odd Frechet Nadarajah-Haghighi (EOFNH):

According to Nasiru (2018), a random variable X is said to have EOFNH distribution with positive scale parameter $\lambda>0$ and shape parameters $\alpha,\beta,\theta>0$ if the cdf is given by

$$F_{\mathrm{EOFNH}}\left(x;\alpha,\beta,\lambda,\theta\right)=\exp\left\{-\left[\left(1-\exp\left(1-\left(1+\lambda x\right)^{\beta}\right)\right)^{-\alpha}-1\right]^{\theta}\right\},\quad x>0.$$

Exponentiated Odd Frechet Weibull (EOFW):

Nasiru (2018) also defined the EOFW with scale parameter $\lambda>0$ and shape parameters $\alpha,\beta,\theta>0$ given the cdf

$$F_{\mathrm{EOFW}}\left(x;\alpha,\beta,\lambda,\theta\right)=\exp\left\{-\left[\left(1-\exp\left(-\lambda x^{\beta}\right)\right)^{-\alpha}-1\right]^{\theta}\right\},\quad x>0.$$

4. 1. Application 1: Failure Time Data

The data consists of the pooled times of successive failures of the air conditioning system of each member of a fleet of Boeing 720 jet airplanes studied by Sirajo *et al*. (2022) and earlier reported by Proschan (1963).

The performance of the EOFHL distribution is compared with that of the competitive distributions using the Akaike information criterion (AIC) and Bayesian information criterion (BIC). The maximum likelihood estimates of the parameters of the fitted distributions are computed by maximizing the log-likelihood function using the *Fitdistrplus* package in the R software.

Table 2. Maximum likelihood estimates and goodness of fit criteria for failure time data

Model	MLEs (S.E)	$-\ell$	AIC	BIC
EOFHL	$\hat{\alpha}=1.06162$ (0.0500)	-77.1350	158.2701	163.5003
	$\hat{\theta}=1.5587$ (0.1254)			
EOFNH	$\hat{\alpha}=1.2650$ (1.3413)	-76.7716	161.5434	172.0039
	$\hat{\beta}=2.4777$ (1.5790)			
	$\hat{\lambda}=0.2458$ (0.3110)			
	$\hat{\theta}=1.3919$ (1.0074)			
EOFW	$\hat{\alpha}=0.4123$ (0.4854)	-76.8341	161.6683	172.1288
	$\hat{\beta}=1.8598$ (0.8081)			
	$\hat{\lambda}=0.1564$ (0.3566)			
	$\hat{\theta}=1.9368$ (1.4005)			
EHL	$\hat{\alpha}=1.7955$ (0.1786)	-109.4201	220.8401	223.4553
GHL	$\hat{\alpha}=1.0060$ (0.1001)	-123.7850	249.5701	252.1852

The maximum likelihood estimated parameters of the competitive distributions with their respective standard errors in bracket as well as $-\ell$ are reported in Table 2 along with their associated goodness of fit criteria, AIC and BIC. Lower values of these criteria indicate a better fit. The EOFHL demonstrates a better performance.

The density plots are shown in Figure 5.

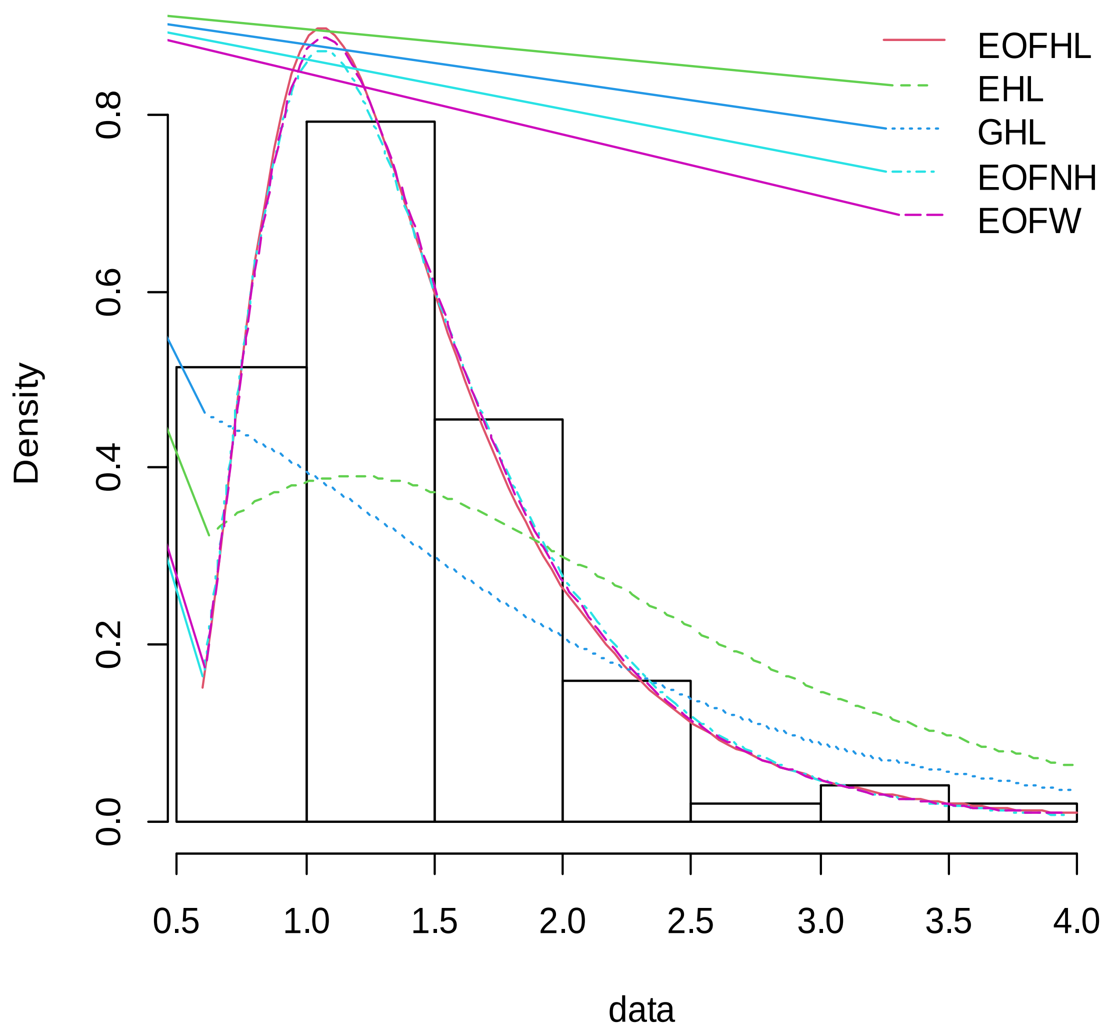

Figure 5. Histogram and pdfs of the fitted models for failure time data

Figure 5 displays the fitted pdfs of the competitive models on the histogram of the data. As can be seen from the plots, all members of the EOFG performed relatively and closely well to describe the characteristics of the modeled data set. Figure 6 displays the empirical quantiles against the theoretical quantiles.

This figure also reveals that the EOFG family distributions provide superior fit to the modelled data set.

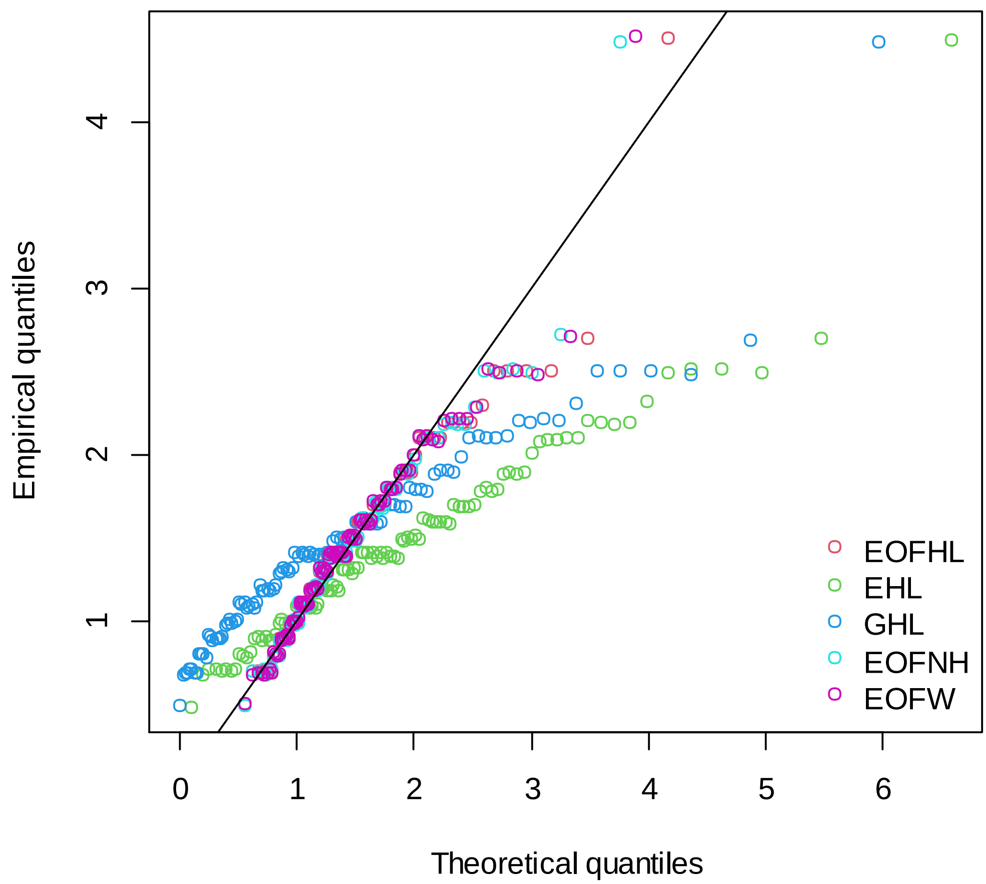

Figure 6. The Q-Q plots of the fitted models for failure time data

4. 2. Application 2: Italy COVID-19 Data

The second application represents COVID-19 data belonging to Italy of 172 days, from 1st March to 21st August 2020. The data can be obtained electronically on the following website: https://github.com/CSSEGISandData/COVID-19/. Figure 7 displays the histogram of the data with the fitted densities, while the empirical quantiles with the theoretical quantiles are displayed in Figure 8.

Table 3. Maximum likelihood estimates and goodness of fit criteria for Covid-19 data

Model	MLEs (S.E)	ℓ	AIC	BIC
EOFHL	$\hat{\alpha}=0.1987$ (0.0050)	157.7811	-311.5623	-305.2673
	$\hat{\theta}=2.2846$ (0.1298)			
EOFNH	$\hat{\alpha}=2.1436$ (0.2274)	197.8378	-387.6756	-375.0856
	$\hat{\beta}=3.3149$ (0.00282)			
	$\hat{\lambda}=3.9826$ (0.0028)			
	$\hat{\theta}=0.2752$ (0.0203)			

EOFW	$\hat{\alpha} = 1.1111$ (0.1202) $\hat{\beta} = 2.0114$ (0.0027) $\hat{\lambda} = 148.1495$ (0.0027) $\hat{\theta} = 0.2536$ (0.0189)	199.4703	-390.9406	-378.3507
EHL	$\hat{\alpha} = 0.3264$ (0.0248)	42.1092	-82.2184	-79.0709
GHL	$\hat{\alpha} = 14.6086$ (1.1138)	180.7370	-359.4740	-356.3265

From Table 3, it is not surprising that EOFHL distribution out performed EHL distribution since it has earlier been shown that EOHL density is a linear combination of EHL densities. Furthermore, even though GHL distribution ranked over EOFHL with lower AIC value, the standard errors of estimated parameters of EOFHL were smaller compared to that of the GHL parameter, indicating that a more precise estimation is achieved using the EOFHL distribution and the other EOFG family members. This means that the EOFG family members would lead to inference with estimates that are closer to the true population parameter.

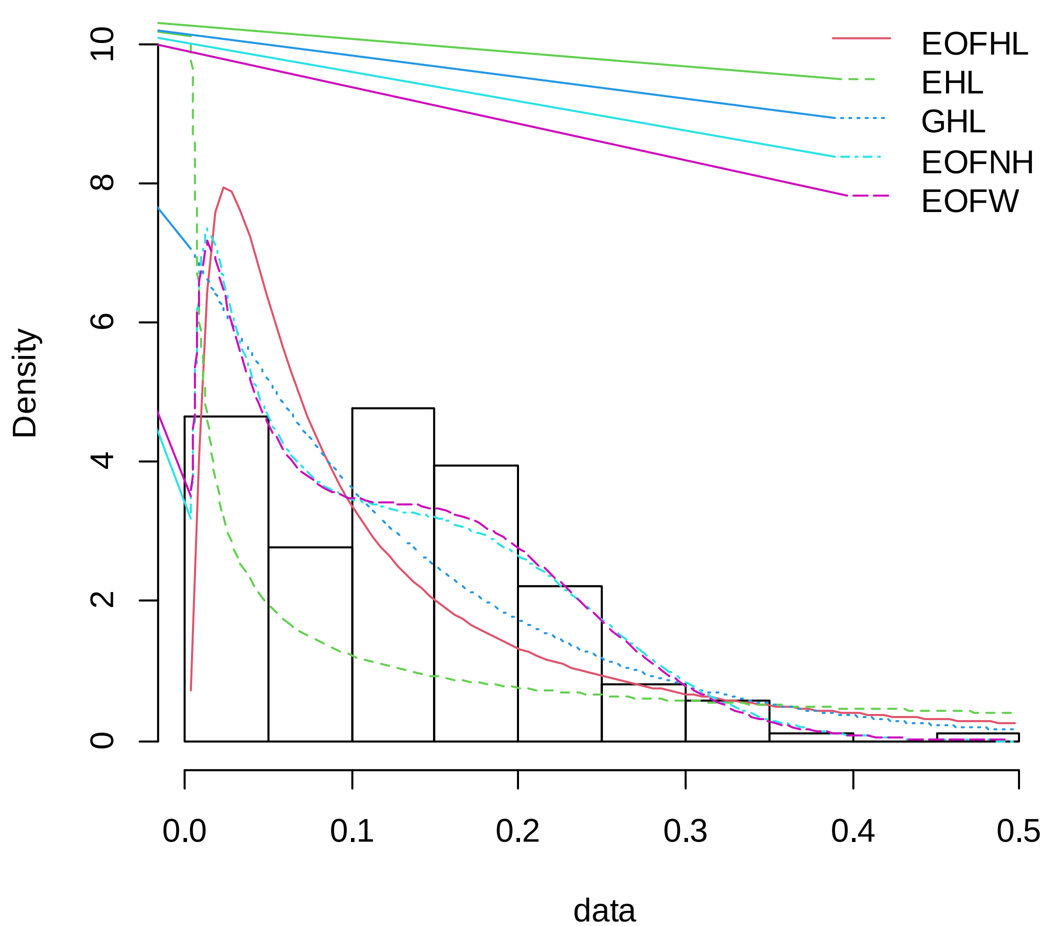

Figure 7. Histogram and pdfs of the fitted models for Italy COVID-19 data

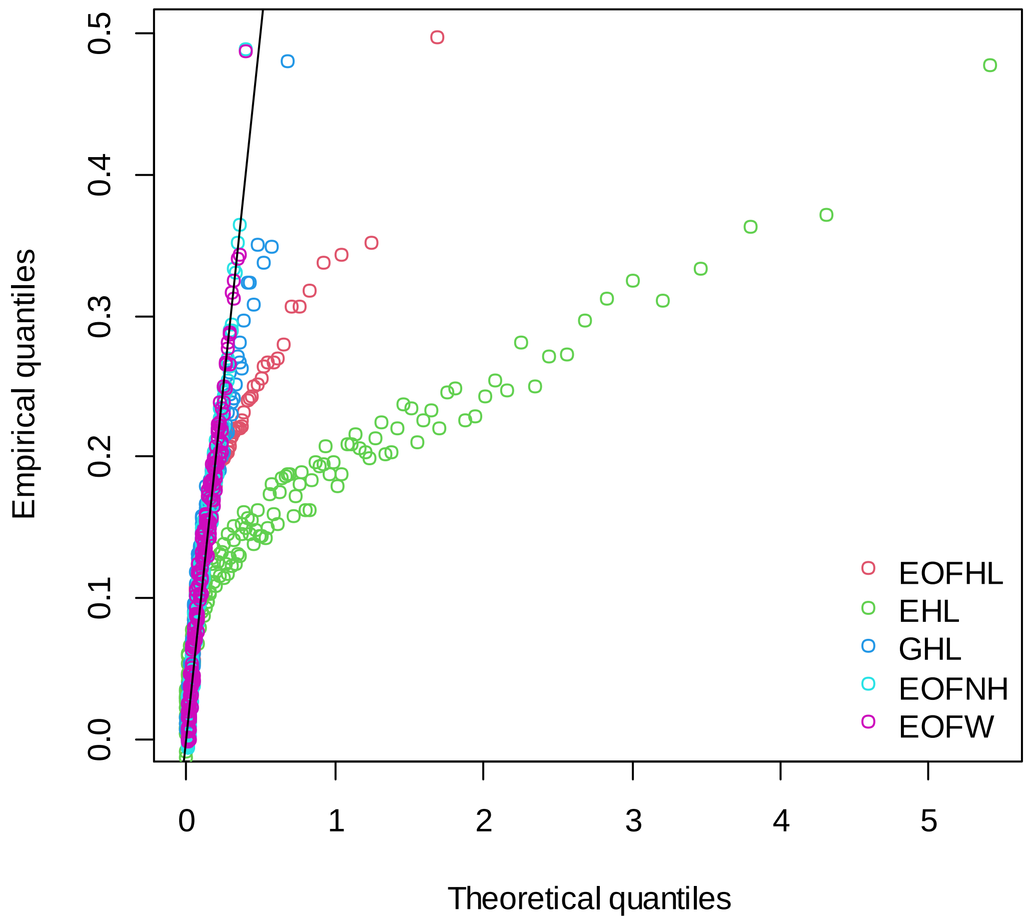

Figure 8. The Q-Q plots of the fitted models for Italy COVID-19 data

4. 3. Application 3: Rainfall Data

The third application employed the real data from Seo & Kang (2015), which represent 30 successive values for precipitation (in inches) in March for the Minneapolis/St. Paul area over a 30-year period. Table 4 and Fig. 9 report maximum likelihood estimates with associated criteria for selection and a histogram for the data, respectively. It is observed that the distribution of this data is skewed to the right. The standard error of estimates is smaller for the parameters of EOFHL distribution than the other models. Therefore, it can be seen that the uncertainty about parameter estimates is greater when models other than EOFHL are used. To further verify the fit of the models, the Q-Q plot is presented in Fig. 10. These results indicate that the EOFHL distribution fits the rainfall data very well.

Table 4. Maximum likelihood estimates and goodness of fit criteria for rainfall data

Model	MLEs (S.E)	$-\ell$	AIC	BIC
EOFHL	$\hat{\alpha}=1.0053$ (0.1381)	-38.3487	80.6974	83.499
	$\hat{\theta}=0.9244$ (0.1440)			

EOFNH	$\hat{\alpha}=1.3460$ (3.2781)	-37.9154	83.8309	89.4357
	$\hat{\beta}=3.0088$ (7.2344)			
	$\hat{\lambda}=0.2093$ (0.8986)			
	$\hat{\theta}=0.7186$ (1.1187)			
EOFW	$\hat{\alpha}=0.5154$ (3.7766)	-37.9085	83.8170	89.4218
	$\hat{\beta}=2.0031$ (4.3885)			
	$\hat{\lambda}=0.2262$ (2.7623)			
	$\hat{\theta}=0.8297$ (2.9286)			
EHL	$\hat{\alpha}=1.7280$ (0.3154)	-39.2342	80.4684	81.8696
GHL	$\hat{\alpha}=0.8280$ (0.1511)	-42.4349	86.8699	88.2711

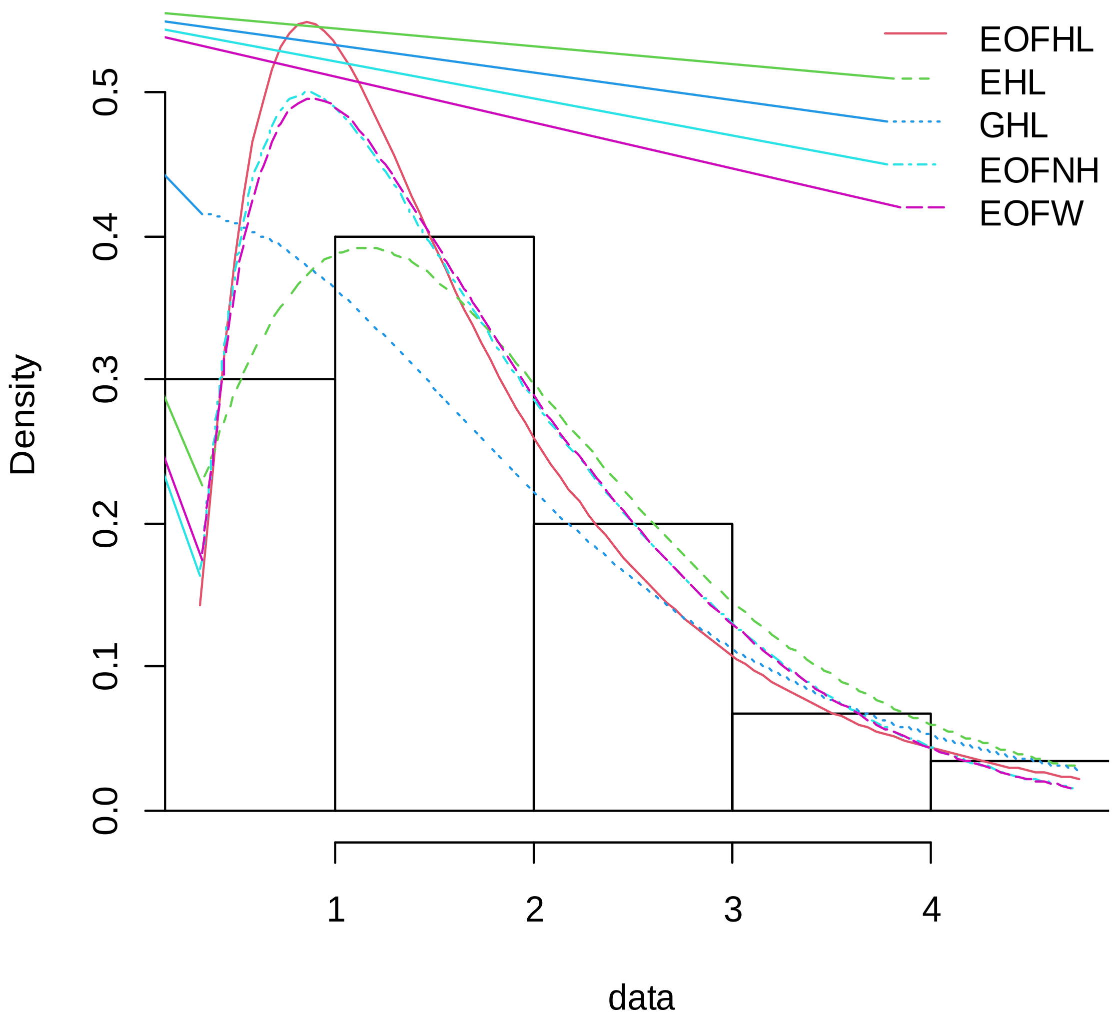

Figure 9. Histogram and pdfs of the fitted models for rainfall data

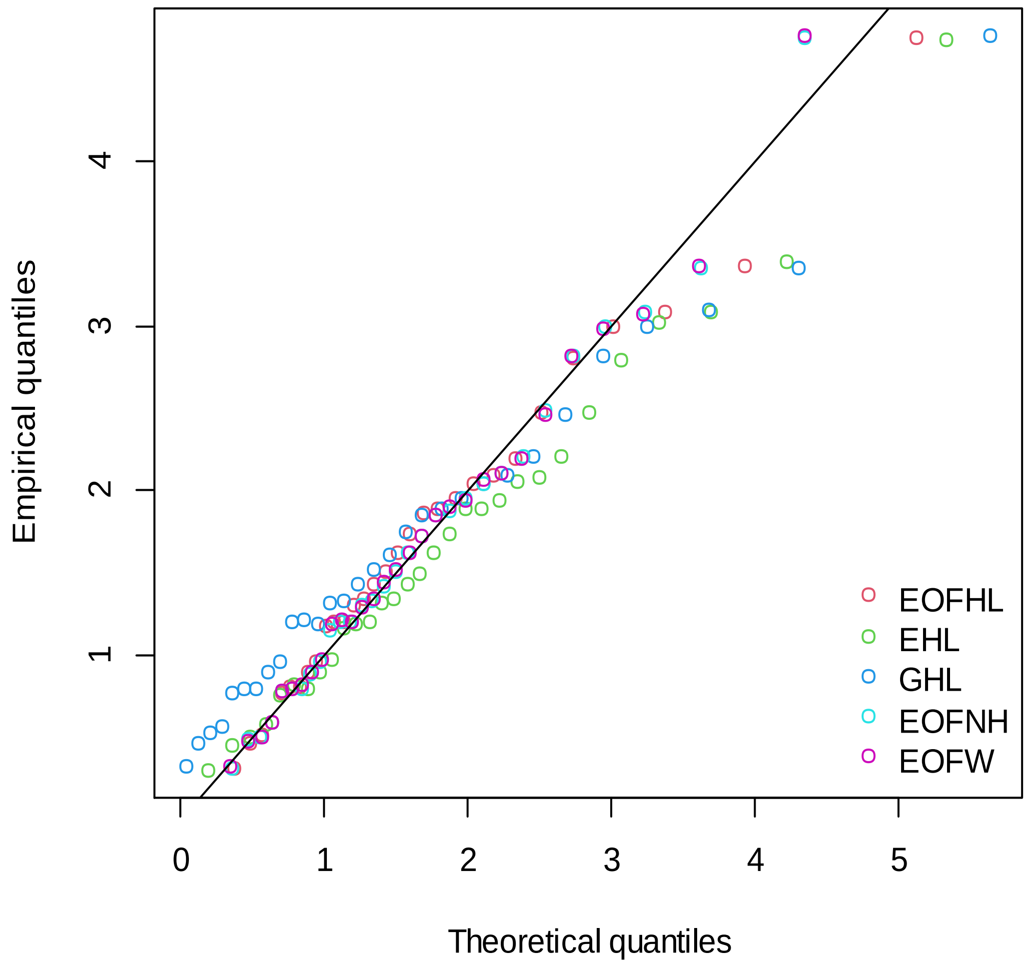

Figure 10. The Q-Q plots of the fitted models for rainfall data

5. CONCLUSION

The lifetime distribution, EOFHL, proposed in this dissertation has the advantage of having only two parameters that control the shape of the distribution which can be right-skewed, decreasing, or nearly symmetric and unimodal. This flexibility of the EOFHL distribution provides an opportunity to data scientists in modeling the different types of data sets. The general finding of the research is that member distributions developed from the EOFG family can provide reasonable parametric fit to lifetime data sets exhibiting both monotone and non-monotone failure rates. Specifically, EOFHL with fewer number of parameters, performed relatively and closely well with EOFNH and EOFW distributions to describe the characteristics of the modeled data sets.

The work presented in this paper provided a flexible extension of half logistic distribution by addressing the long standing issue of monotone failure rate function associated with the classical half logistic distribution. Furthermore, the estimated parameters of EOFHL had consistently produced smaller standard errors over three different applications to real-life data sets. It was therefore noted that uncertainty about parameter estimates is greater when models other than EOFHL are used, indicating that the reduced number of parameters in EOFHL among other EOFG family members had actually helped reduced misspecification errors.

As a future work related to this study, researchers may develop an EOFHL regression model with its influence diagnostics and residuals analysis, which will be a useful choice for practitioners studying in the field of reliability and survival analyses. This is due to the flexibility of the failure rate function and the ease with which parameters of the EOFHL distribution can be estimated. It is hoped that the EOFHL distribution will find a wider application area in the near future.

References

[1] Cordeiro, G. M., Alizadeh, M., & Diniz Marinho, P. R. (2016). The type I half-logistic family of distributions. *Journal of Statistical Computation and Simulation,* 86(4), 707-728.

[2] Jackson, D. L. (2007). The effect of the number of observations per parameter in misspecified confirmatory factor analytic models. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(1), 48-76

[3] Nasiru, S. (2018). Extended odd Fréchet-G family of distributions. *Journal of Probability and Statistics* 2018, Article ID 2931326

[4] Olapade, A. K. (2014). The type I generalized half logistic distribution. *Journal of the Iranian Statistical Society*, 13(1), 69-82.

[5] Proschan, F. (1963). Theoretical explanation of observed decreasing failure rate. *Technometrics*, 5(3), 375-383

[6] Seo, J. I., & Kang, S. B. (2015). Notes on the exponentiated half logistic distribution. *Applied Mathematical Modelling* 39 (1), 6491–6500

[7] Sirajo, M., Falgore, J. Y., Najmuddeen, M. S., & Umar, A. A. (2021). A New Reflected Minimax Distribution on a Bounded Domain: Theory and Application. *Proceeding of the 5th International Conference of PSSN*, 5(1)

[8] Sirajo, M., Shakil, M., & Ahsanullah, M. (2022). A Muth-Pareto Distribution: Properties, Estimation, Characterizations and Applications. *Statistica*, 82(3), 243-274