Pak. J. Statist.
2023 Vol. 39(3), 213-236

EXPLORING THE CAPABILITIES OF A NEW BATHTUB MODEL WITH APPLICATION TO LIFETIME DATA

Muhammad Zeshan Arshad[1§], **Muhammad Zafar Iqbal**[2]
Oluwafemi Samson Balogun[2], **Muhammad Kashif**[3]
and **Zahoor Ahmad**[3]

[1,2,3] Department of Mathematics and Statistics
University of Agriculture, Faisalabad, Pakistan.

[2] School of Computing, University of Eastern Finland
Kuopio, Northern Europe, 70211, Finland.

[3] Department of Mathematics and Statistics
University of Lahore, Sargodha Campus, Pakistan.

[§] Corresponding author Email: profarshad@yahoo.com

ABSTRACT

This research introduces a shifted Kumaraswamy distribution (SKD). The SKD is capable of modelling real-world phenomena with asymmetrical and bathtub-shaped characteristics with exceptional proficiency. The mathematical and reliability characteristics are thoroughly examined. To better comprehend its behaviour, density and hazard rate functions are plotted. The parameters of the model are estimated through the use of the maximum likelihood estimation method. The effectiveness of the SKD is demonstrated through its application to four real-world lifetime datasets, resulting in significantly improved outcomes compared to other well-known models and providing the closest fit to the data. This superiority confirms the SKD as a preferable choice over the baseline model.

KEYWORDS

Kumaraswamy distribution; location parameter; moments; entropy; maximum likelihood estimation.

1. INTRODUCTION

The Kumaraswamy distribution (KD) is a two-parameter continuous probability distribution that was first introduced by K. Kumaraswamy in 1980. It is a generalization of the Beta distribution and is commonly used to model the distribution of continuous random variables that are bounded between 0 and 1. The shape parameters of the KD determine the shape of the distribution, and they control the skewness, kurtosis, and mode of the distribution. The distribution is commonly used in various fields such as engineering, finance, and reliability analysis. The cumulative distribution function (denoted as CDF) and the probability density function (denoted as PDF) of the KD can be easily computed using its mathematical formulation.

$$P_t|_{c,d} = 1 - [1 - t^c]^d\,, t \in (0,1), \tag{1}$$

and

$$p_t|_{c,d} = cdt^{c-1}[1 - t^c]^{d-1}\,, \tag{2}$$

where $c, d > 0$ are two shape parameters.

The KD is a versatile and flexible distribution that has gained popularity due to its ability to model various types of data with different shapes. Unlike other commonly used distributions, such as the normal or exponential distribution, the KD is able to model distributions with different shapes, including distributions that are multimodal or skewed. This makes it well-suited for modelling real-world data, where the underlying distribution may not always be easily characterized by a single standard distribution. Additionally, the KD has a simple mathematical formulation, which makes it easy to implement in statistical software and to calculate various summary statistics, such as the mean, variance, and skewness. It also has an intuitive interpretation as the distribution of a ratio of random variables, making it easily understood by both practitioners and researchers. Overall, the KD is a valuable tool for modelling and analyzing data in various fields. Its flexibility, ease of implementation, and ability to handle a wide range of data shapes make it a useful tool for data scientists and practitioners alike.

In order to provide the reader with further information, the authors recommend reputable sources connected to Kumaraswamy generated (K-G) class such as Jones (2009), McDonald (1984), and Cordeiro and de Castro (2011). Bourguignon et al. (2013), Lemonte et al. (2013), Alizadeh et al. (2015), Afify et al. (2016), Ibrahim (2017), Bursa and Ozel (2017), Mahmoud et al. (2018), Nawaz et al. (2018), Silva et al. (2019), Neto and Santos (2019), Tahir et al. (2014), Hemeda et al. (2020), Hassan et al. (2021), Arshad et al. (2021), and Al-Babtain et al. (2021), Tahir et al. (2020),and Ramzan et al. (2022),.

Despite its usefulness, there is still a significant gap in the research when it comes to exploring shifted distributions, which refers to adding a constant value to each data point, resulting in a new distribution with the same shape but with its mean and median shifted. Shifting a distribution can be useful in various contexts, such as when comparing different distributions that are not centered on the same value or when the units of measurement or the origin of a variable changes. To address this gap, this study proposes a modified version of the KD called the shifted Kumaraswamy distribution (SKD), which includes an additional shifting parameter (m). The SKD can better capture the complex behavior of real-world systems and their failure patterns over time, including the bathtub-shaped failure rate pattern, which is commonly observed in various fields. The SKD provides a more accurate representation of the underlying data and can lead to better decision making in various applications. The purpose of this study is to introduce the SKD and provide a comprehensive analysis of its properties, including its mean, variance, quantile functions, maximum likelihood estimation, and many others. The study also aims to bridge the gap in the research related to shifted distributions and provide practitioners and researchers with a useful tool for modeling and analyzing data in various fields. It is standard practice to simulate shifted distributions by making use of data pertaining to dependability, breakage, percentage, and proportion. More research is being done on it in relation to medical and veterinary issues.

For further information, authors refer the readers to a number of recently published investigations, such as those conducted by Yang et al. (2019), Madi and Leonard (1996), Cousineau (2009), Ikechukwu et al. (2020), Jodrá (2020), Gómez-Déniz et al. (2020), Arshad et al. (2022), Eledum and Alaa (2023), Eldessouky et al. (2023), amongst others.

1.1 Definition

A random variable T~SKD $(t;c,d)$ with $c,d>0$, two shape parameters for $t\in(m,1)$, having CDF corresponding PDF that begin at m with $1/[1-m]$ normalizing constant, are, respectively, given by

$$F_t|_{c,d}=1-\left[1-\left(\frac{t-m}{1-m}\right)^c\right]^d, \tag{3}$$

and

$$f_t|_{c,d}=cd(t-m)^{c-1}(1-m)^{-c}\left[1-\left(\frac{t-m}{1-m}\right)^c\right]^{d-1}. \tag{4}$$

Table 1
Sub-Models of SKD

Parameter	Model	Reference
$m=0$	Kumaraswamy Distribution	Kumaraswamy (1980)
$m=0$, and $c=1$	Lehmann Type – II distribution	Lehmann-II (1953)

The rest of the paper is structured in the following manner: In the 1st Section, a new model is presented for consideration. In the 2nd Section of this paper, the general characteristics of the model are investigated. In the 3rd Section, we discuss various types of reliability measurements. In 4th Section, we discussed various additional characteristics. In 5th Section, an estimating approach based on the maximum likelihood, as well as a simulation study, are carried out in Section 6th. In 7th Section, the application of the model is discussed, and the findings and future directions are analyzed and summarized in 8th Section.

2. MATHEMATICAL PROPERTIES OF SHIFTED KUMARASWAMY DISTRIBUTION

2.1 Infinite Linear Combinations (ILCs)

The ILCs of CDF and PDF for SKD may provide a more straightforward approach than the traditional integral computation. The binomial expansion for T is taken into account as:

$$(1-t)^b=\sum_{l=0}^{\infty}\binom{b}{l}(-1)^l t^l,\ |t|<1.$$

As a result of (3), ILCs of the CDF is given as follows:

$$F_t|_{c,d}=1-\sum_{i_1,i_2=0}^{\infty}\binom{d}{i_1}\binom{ci_1}{i_2}(-1)^{i_1+i_2}\frac{m^{i_2}}{(1-m)^{ai_2}}t^{ci_1-i_2}, \tag{5}$$

$$F_t|_{c,d} = 1 - \sum_{i_1,i_2=0}^{\infty} \zeta_{i_1,i_2} t^{\varepsilon_{i_1,i_2}} . \tag{6}$$

As a result of (4), ILCs of PDF is given as follows:

$$f_t|_{c,d} = \frac{cd}{(1-m)^a}\left[\sum_{i_1,i_2,l=0}^{\infty} \begin{matrix}\binom{c-1}{i_1}\binom{d-1}{i_2}\binom{ci_2}{l}(-1)^{i_1+i_2+l} m^{i_1-ci_2+l} \times \\ (1-m)^{-ci_2} t^{c-1-i_1+ci_2-l}\end{matrix}\right]. \tag{7}$$

$$f_t|_{c,d} = \frac{cd}{(1-m)^c} \sum_{i_1,i_2,l=0}^{\infty} \eta_{i_1,i_2,l} t^{\phi_{i_1,i_2,l}}, \tag{8}$$

where

$$\varepsilon_{i_1,i_2} = ci_1 - i_2, \phi_{i_1,i_2,l} = c - 1 - i_1 + ci_2 - l, \zeta_{i_1,i_2} = \binom{d}{i_1}\binom{ci_1}{i_2}(-1)^{i_1+i_2}\frac{m^{i_2}}{(1-k)^{ci_2}},$$
$$\eta_{i,j,l} = \binom{c-1}{i_1}\binom{d-1}{i_2}\binom{ci_2}{l}(-1)^{i_1+i_2+l} m^{i_1-ci_2+l}(1-m)^{-ci_2}.$$

Theorem 1

A random variable *T~SKD (t;c,d)* with $c, d > 0$, two shape parameters, having CDF corresponding PDF that begin at m with $1/[1-m]$ normalizing constant, then ordinary moment ($\mu'_t|_{r,c,d}$) of T is given by

$$\mu'_t|_{r,c,d} = \frac{cd}{(1-m)^c} \sum_{i_1,i_2,l=0}^{\infty} \eta_{i_1,i_2,l} \frac{1}{\nabla_{i_1,i_2,l,r}}\left[1 - m^{\nabla_{i_1,i_2,l,r}}\right],$$

where

$$\eta_{i_1,i_2,l} = \binom{c-1}{i_1}\binom{d-1}{i_2}\binom{ci_2}{l}(-1)^{i_1+i_2+l}\frac{m^{i_1-ci_2+l}}{(1-m)^{ci_2}}, \nabla_{i_1,i_2,l,r}$$
$$= c - i_1 + ci_2 - l + r.$$

Proof:

$\mu'_t|_{r,c,d}$ is can be written as $\mu'_t|_{r,c,d} = \int_m^1 t^r f_t|_{c,d} dt$ and it can be written directly followed by (4)

$$\mu'_t|_{r,c,d} = \int_m^1 t^r\, cd(t-m)^{c-1}(1-m)^{-c}\left[1-(t-m)^c(1-m)^{-c}\right]^{d-1} dt.$$

Through the use of the binomial expansion, the last expression is simplified and presented in its most straightforward form as follows:

$$\mu'_t|_{r,c,d} = \frac{cd}{(1-m)^c}$$
$$\sum_{i_1,i_2,l=0}^{\infty}\binom{c-1}{i_1}\binom{d-1}{i_2}\binom{ci_2}{l}(-1)^{i_1+i_2+l}\frac{m^{i_1-ci_2+l}}{(1-m)^{ci_2}}\int_m^1 t^{c-i_1+ci_2-l+r-1}\, dt.$$

Hence, through the process of simple integration, a final expression of $\mu'_t|_{r,c,d}$ can be obtained and is presented as follows:

$$\mu'_t|_{r,c,d} = \frac{cd}{(1-m)^c} \sum_{i_1,i_2,l=0}^{\infty} \frac{\eta_{i_1,i_2,l}}{\nabla_{i_1,i_2,l,r}} \left[1 - m^{\nabla_{i_1,i_2,l,r}}\right], \tag{9}$$

where

$$\eta_{i_1,i_2,l} = \binom{c-1}{i_1}\binom{d-1}{i_2}\binom{ci_2}{l}(-1)^{i_1+i_2+l}\frac{m^{i_1-ci_2+l}}{(1-m)^{ci_2}}, \nabla_{i_1,i_2,l,r} = c - i_1 + ci_2 - l + r.$$

The equation represented by (9) has the potential to play a valuable impact in formation of various statistical metrics. As an illustration: to calculate mean (denoted as $\mu'_t|_{1,c,d}$), three moments (denoted as $\mu'_t|_{2,c,d}$, $\mu'_t|_{3,c,d}$, $\mu'_t|_{4,c,d}$) and negative moments (denoted as $\mu'_t|_{-v,c,d}$) of *T*, replace *r* with 1,2,3,4 and -*v* in (9), respectively. It is presented as follows:

$$\mu'_t|_{1,c,d} = \frac{cd}{(1-m)^c} \sum_{i_1,i_2,l=0}^{\infty} \frac{\eta_{i_1,i_2,l}}{\nabla_{i_1,i_2,l,1}} \left[1 - m^{\nabla_{i_1,i_2,l,1}}\right], \tag{10}$$

$$\mu'_t|_{2,c,d} = \frac{cd}{(1-m)^c} \sum_{i_1,i_2,l=0}^{\infty} \frac{\eta_{i_1,i_2,l}}{\nabla_{i_1,i_2,l,2}} \left[1 - m^{\nabla_{i_1,i_2,l,2}}\right],$$

$$\mu'_t|_{3,c,d} = \frac{cd}{(1-m)^c} \sum_{i_1,i_2,l=0}^{\infty} \frac{\eta_{i_1,i_2,l}}{\nabla_{i_1,i_2,l,3}} \left[1 - m^{\nabla_{i_1,i_2,l,3}}\right],$$

$$\mu'_t|_{4,c,d} = \frac{cd}{(1-m)^c} \sum_{i_1,i_2,l=0}^{\infty} \frac{\eta_{i_1,i_2,l}}{\nabla_{i_1,i_2,l,4}} \left[1 - m^{\nabla_{i_1,i_2,l,4}}\right],$$

$$\mu'_t|_{-v,c,d} = \frac{cd}{(1-m)^c} \sum_{i_1,i_2,l=0}^{\infty} \frac{\eta_{i_1,i_2,l}}{\nabla_{i_1,i_2,l,(-v)}} \left[1 - m^{\nabla_{i_1,i_2,l,(-v)}}\right].$$

Tables 2 and 3 exhibits the values of first four ordinary moments, variance (denoted as $\sigma^2{}_t|_{c,d}$), skewness (indicated as $\pi_{sk}|_{c,d}$), and kurtosis (designated as $\pi_{kr}|_{c,d}$) for selected model parameters of *m* = 0.001.

Table 2
Some Numerical Results of Statistics

Statistics	$\boldsymbol{c = 0.1}$					**Comments**
$\mu'_t\|_{r,c,d}$	$d = 1.2$	$d = 1.3$	$d = 1.4$	$d = 1.5$	$d = 1.6$	
$\mu'_t\|_{1,c,d}$	0.0623	0.0518	0.0431	0.0362	0.0306	Decreasing
$\mu'_t\|_{2,c,d}$	0.0282	0.0221	0.0173	0.0136	0.0107	
$\mu'_t\|_{3,c,d}$	0.0179	0.0133	0.1001	0.0076	0.0058	
$\mu'_t\|_{4,c,d}$	0.0128	0.0093	0.0068	0.0050	0.0096	
$\sigma^2{}_t\|_{c,d}$	0.0245	0.0194	0.0153	0.0121	0.0096	
$\pi_{sk}\|_{c,d}$	7.7304	10.2106	13.0813	16.3151	20.0917	Increasing
$\pi_{kr}\|_{c,d}$	11.5508	14.8118	18.5862	22.9534	28.0026	Increasing

Table 3
Some Numerical Results of Statistics

Statistics	$\boldsymbol{d} = \mathbf{0.9}$					**Comments**
$\mu'_t\|_{r,c,d}$	$c = 0.1$	$c = 0.2$	$c = 0.3$	$c = 0.4$	$c = 0.5$	
$\mu'_t\|_{1,c,d}$	0.1125	0.1941	0.2604	0.3161	0.3637	Decreasing
$\mu'_t\|_{2,c,d}$	0.0623	0.1118	0.1549	0.1935	0.2283	
$\mu'_t\|_{3,c,d}$	0.0439	0.0796	0.1117	0.1411	0.1682	
$\mu'_t\|_{4,c,d}$	0.0341	0.0623	0.0879	0.1117	0.1340	
$\sigma^2_t\|_{c,d}$	0.0492	0.0713	0.0793	0.0778	0.0691	
$\pi_{sk}\|_{c,d}$	2.1879	0.2687	0.0125	0.0066	0.0252	Decreasing
$\pi_{kr}\|_{c,d}$	4.1553	1.2238	0.5411	0.3441	0.2748	Decreasing

Moment generating function (mgf) of T is defined as $M_T(w)_t|_{c,d} = \sum_{r=0}^{\infty} \frac{w^r}{r!} \mu'_t|_{r,c,d}$ and mathematical expression is presented as follows:

$$M_T(w)_t|_{c,d} = \frac{cd}{(1-m)^c} \sum_{r=0}^{\infty} \frac{w^r}{r!} \sum_{i_1,i_2,l=0}^{\infty} \frac{\eta_{i_1,i_2,l}}{\nabla_{i_1,i_2,l,r}} \left[1 - m^{\nabla_{i_1,i_2,l,r}}\right].$$

The characteristic function of T is defined as $\emptyset_T(w)_t|_{c,d} = \sum_{r=0}^{\infty} \frac{(iw)^r}{r!} \mu'_t|_{r,c,d}$. It is presented as follows:

$$\emptyset_T(w)_t|_{c,d} = \frac{cd}{(1-m)^c} \sum_{r=0}^{\infty} \frac{(iw)^r}{r!} \sum_{i_1,i_2,l=0}^{\infty} \frac{\eta_{i_1,i_2,l}}{\nabla_{i_1,i_2,l,r}} \left[1 - m^{\nabla_{i_1,i_2,l,r}}\right].$$

The first incomplete moment of U is defined as $\Phi_u|_{1,c,d} = \int_0^u t f_t|_{c,d} dt$ and it is obtained by substituting $r = 1$ in (11), as

$$\Phi_u|_{1,c,d} = \frac{cd}{(1-m)^c} \sum_{i_1,i_2,l=0}^{\infty} \frac{\eta_{i_1,i_2,l}}{\nabla_{i_1,i_2,l,r}} \left[u^{\nabla_{i_1,i_2,l,1}} - m^{\nabla_{i_1,i_2,l,r}}\right]. \quad (12)$$

As is seen in detail in Section 2.4, the idea of the $\Phi_u|_{1,c,d}$ may be put to use in a variety of contexts when analyzing Bonferroni and Lorenz curves.

2.4 Inequality Curves

The Bonferroni and Lorenz curves, which are defined as $B_t|_{c,d}$ and $L_t|_{c,d}$, respectively, play a crucial role in a wide array of fields including economics, where they are used to analyze the distribution of income, poverty, and wealth. These curves are also of great significance in fields such as insurance, demography, medicine, reliability and others. Their extensive applications can be attributed to their ability to provide an accurate representation of the cumulative distribution of a random variable, which is crucial in decision making processes. These curves are formally defined as follows:

$$L_t|_{c,d} = \frac{\Phi_u|_{1,c,d}}{\mu'_t|_{1,c,d}}, \qquad B_t|_{c,d} = \frac{L_t|_{c,d}}{F_t|_{c,d}}.$$

The Lorenz curve of T

$$L_t|_{c,d} = \frac{\sum_{i_1,i_2,l=0}^{\infty} \rho_{i_1,i_2,l,r}\left[u^{\nabla_{i_1,i_2,l,1}} - m^{\nabla_{i_1,i_2,l,r}}\right]}{\sum_{i_1,i_2,l=0}^{\infty} \rho_{i_1,i_2,l,1}\left[1 - m^{\nabla_{i_1,i_2,l,1}}\right]}, \tag{13}$$

and Bonferroni curve along with their plots are represented, respectively

$$B_t|_{c,d} = \frac{1}{F|_{t,c,d}}\left[\frac{\sum_{i_1,i_2,l=0}^{\infty} \rho_{i_1,i_2,l,r}\left[u^{\nabla_{i_1,i_2,l,1}} - m^{\nabla_{i_1,i_2,l,r}}\right]}{\sum_{i_1,i_2,l=0}^{\infty} \rho_{i_1,i_2,l,1}\left[1 - m^{\nabla_{i_1,i_2,l,1}}\right]}\right],$$

where

$$\rho_{i_1,i_2,l,r} = \frac{\eta_{i_1,i_2,l}}{\nabla_{i_1,i_2,l,r}}, \quad \rho_{i_1,i_2,l,1} = \frac{\eta_{i_1,i_2,l}}{\nabla_{i_1,i_2,l,1}}.$$

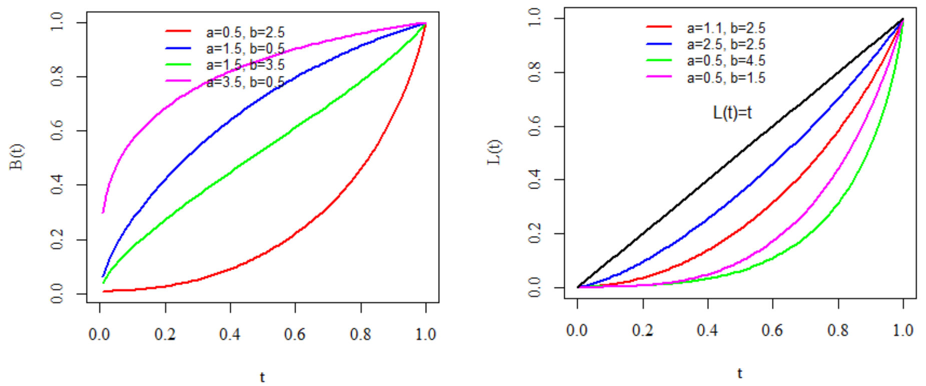


Figure 1: Bonferroni and Lorenz Cures for SKD

The residual life and reverse residual life functions for T are formally defined as $R|_{v/t} = S|_{(t+v)}/S|_v$ and $\bar{R}|_{v/t} = S|_{(t-v)}/S|_v$ and mathematical expressions are, respectively, presented as follows:

$$R|_{v/t,c,d} = y|_v\left[1 - \left(\frac{t+v-m}{1-m}\right)^c\right]^d,$$

and

$$\bar{R}|_{v/t,c,d} = y|_v\left[1 - \left(\frac{t-v+m}{1-m}\right)^c\right]^d,$$

where

$$y|_v = 1 - \left[1 - \left(\frac{v-m}{1-m}\right)^c\right]^d.$$

3. RELIABILITY MEASURES

The survival function which is denoted as $S_t|_{c,d}$, that represents the probability that a component will last until time t, can be expressed as $S_t|_{c,d} = 1 - F_t|_{c,d}$. The $S_t|_{c,d}$ of T is described as follows:

$$S_t|_{c,d} = \left[1 - \left(\frac{t-m}{1-m}\right)^c\right]^d.$$

The hazard rate function which is denoted as $h_t|_{c,d}$, that quantifies the rate at which a component fails at time *t*, is expressed mathematically as $h_t|_{c,d} = f_t|_{c,d}/S_t|_{c,d}$. The $h_t|_{c,d}$ of *T* is defined as:

$$h_t|_{c,d} = \frac{cd(t-m)^{c-1}}{(1-m)^c - (t-m)^c}.$$

In addition to the survival and hazard rate functions, other noteworthy reliability metrics can be derived for *T*. One such metric is the reversed hazard rate function, which is defined as $rh_t|_{c,d} = f_t|_{c,d}/F_t|_{c,d}$. This function provides a different perspective on the failure rate of a component, taking into consideration not only the remaining time until failure, but also the amount of time that has already elapsed. The $rh_t|_{c,d}$ can be a valuable tool for evaluating the reliability of a component over its lifetime and making informed decisions about maintenance and replacement strategies. The expression for the $rh_t|_{c,d}$ of *T* is:

$$rh_t|_{c,d} = \frac{cd(t-m)^{c-1}\left[1 - \left(\frac{t-m}{1-m}\right)^c\right]^{d-1}}{(1-m)^c\left[1 - \left\{1 - \left(\frac{t-m}{1-m}\right)^c\right\}^d\right]}.$$

The Mills ratio which is denoted as $M_t|_{c,d}$, is another important reliability measure that provides insight into the relationship between the survival and failure rate functions of a component. It is defined as $M_t|_{c,d} = S_t|_{c,d}/f_t|_{c,d}$, and provides a measure of the excess waiting time between failures. The $M_t|_{c,d}$ of *T* can be expressed as:

$$M_t|_{c,d} = \frac{(1-m)^c - (t-m)^c}{cd(t-m)^{c-1}}.$$

The Odd function is another reliability metric that provides a measure of the likelihood of a component surviving or failing. It is expressed as $O_t|_{c,d} = F_t|_{c,d}/S_t|_{c,d}$, and provides a measure of the odds of a component failing at time *t*, given that it has survived until that point. The $O_t|_{c,d}$ of *T* can be expressed as:

$$O_t|_{c,d} = \left[1 - \left(\frac{t-m}{1-m}\right)^c\right]^{-d} - 1.$$

4. ADDITIONAL CHARACTERISTICS

4.1 Limiting Behaviour (LB)

The LB of CDF ($F_t|_{c,d}$), PDF ($f_t|_{c,d}$) and HRF ($h_t|_{c,d}$) of *T* for limit $t \to m$ and limit $t \to 1$ is explored in propositions I and II, respectively.

Proposition I

For limit $t \to m$, $F_t|_{c,d}$, $f_t|_{c,d}$ and $h_t|_{c,d}$ are, respectively, illustrated as

$$F_m|_{c,d} \sim 0,$$
$$f_m|_{c,d} \sim 0,$$
$$h_m|_{c,d} \sim 0.$$

Proposition II.

For limit $t \to 1$, $F_t|_{c,d}$, $f_t|_{c,d}$ and $h_t|_{c,d}$ are, respectively, illustrated as

$$F_1|_{c,d} \sim 1,$$
$$f_1|_{c,d} \sim 0,$$
$$h_1|_{c,d} \sim \text{Undefined}.$$

4.2 Shapes

In Figure 2, a variety of shapes for the PDF (I and II) and HRF (III and IV) of SKD are presented, based on different model parameters. It should be noted that the shapes of both the PDF and HRF can greatly impact the reliability of the component being analyzed.

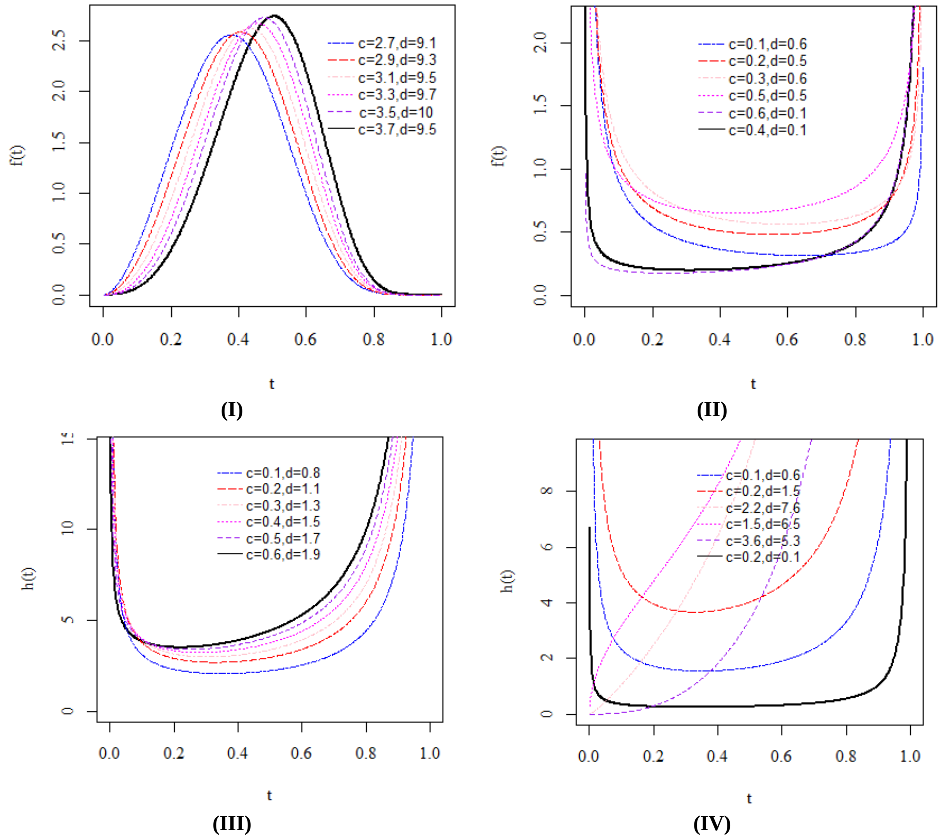


Figure 2: PDF and HRF curves for SKD

4.3 Quantile Function

The q^{th} quantile function (defined as $qf_q\big|_{c,d}$) of the SKD, which is obtained by reversing the CDF as presented in (3). Mathematically, it is defined as $qf_q\big|_{c,d} = P\left(T \leq t_q\right), q \in (0,1)$. The $qf_q\big|_{c,d}$ of T is then given by this expression.

$$qf_q\big|_{c,d} = m + (1-m)\left(1-(1-q)^{1/d}\right)^{1/c}. \tag{14}$$

For q = 0.25, 0.5, and 0.75 respectively one may find out $Q_1|_{c,d}$, $Q_2|_{c,d}$, and $Q_3|_{c,d}$.To generate random numbers, if CDF of proposed SKD follows to $u = U$ (0, 1) (uniform distribution).

4.4 Entropy Measures

Entropy measures, including Rényi and Tsallis entropy, are used to quantify the amount of disorder or randomness in a system. These measures are commonly used in statistical mechanics and information theory to describe complex systems. Rényi entropy generalizes Shannon entropy, while Tsallis entropy provides an alternative to Boltzmann entropy. Each measure has its own unique formulation and interpretation, allowing for analysis of a broad range of physical and informational systems. The Rényi (1961) entropy of T is defined by

$$H_R|_{t,c,d,\vartheta} = \frac{1}{1-\vartheta} log \int_m^1 \left[f_t|_{c,d}\right]^{\vartheta} dt \ , \vartheta > 0 \ and \vartheta \neq 1. \tag{15}$$

Initially, simplify $\left[f_t|_{a,b}\right]^{\vartheta}$ by utilizing (4)

$$\left[f_t|_{c,d}\right]^{\vartheta} = \frac{(cd)^{\vartheta}}{(1-m)^{c\vartheta}}(t-m)^{\vartheta(c-1)}\left[1-\left(\frac{t-m}{1-m}\right)^{c}\right]^{\vartheta(d-1)},$$

and substitute in (15), we obtain $H_T|_{c,d,\vartheta}$ of T

$$H_R|_{t,c,d,\vartheta} = \left[\begin{array}{c} \dfrac{(cd)^{\vartheta}}{(1-\vartheta)(1-m)^{c\vartheta}} log \displaystyle\sum_{i_1,i_2,l=0}^{\infty} \binom{\vartheta(c-1)}{i_1}\binom{\vartheta(d-1)}{i_2}\binom{ci_2}{l} \times \\ (-1)^{i_1+i_2+l} m^{i_1-ci_2+l} \displaystyle\int_m^1 t^{\vartheta(c-1)-i_1+ci_2-l} dt \end{array}\right].$$

Hence, by mathematical operations to last expression provide $H_T|_{c,d,\vartheta}$ which is presented as follows:

$$H_R|_{t,c,d,\vartheta} = \frac{(cd)^{\vartheta}}{(1-\vartheta)(1-m)^{c\vartheta}} log \sum_{i_1,i_2,l=0}^{\infty} A_{i_1,i_2,l,\vartheta} \frac{1}{\tau_{i_1,i_2,l,\vartheta}}\left[1-m^{\tau_{i_1,i_2,l,\vartheta}}\right], \tag{16}$$

where $\tau_{i_1,i_2,l,\vartheta} = \vartheta(c-1) - i_1 + ci_2 - l$,

$$A_{i_1,i_2,l,\vartheta} = \binom{\vartheta(c-1)}{i_1}\binom{\vartheta(d-1)}{i_2}\binom{ci_2}{l}(-1)^{i_1+i_2+l} m^{i_1-ci_2+l}.$$

The Rényi entropy is a useful tool for finding the Tsallis entropy (short $H_T|_{t,c,d,\vartheta}$). $H_T|_{t,c,d,\vartheta} = \frac{1}{\vartheta-1}\int_m^1 \left[f_t|_{c,d}\right]^{\vartheta-1} dt \ ,\vartheta > 0 \ and \vartheta \neq 1$. Hence, the final version of the $H_T|_{t,c,d,\vartheta}$ is defined as follows.

$$H_T|_{t,c,d,\vartheta} = \frac{(cd)^{(\vartheta-1)}}{(\vartheta-1)(1-m)^{c(\vartheta-1)}} \sum_{i_1,i_2,l=0}^{\infty} A^*_{i_1,i_2,l,(\vartheta-1)} \frac{1}{\tau^*_{i_1,i_2,l,(\vartheta-1)}} \left[1 - m^{\tau^*_{i_1,i_2,l,(\vartheta-1)}}\right],$$

where $\tau^*_{i_1,i_2,l,(\vartheta-1)} = (\vartheta-1)(c-1) - i_1 + ci_2 - l$,

$$A^*_{i_1,i_2,l,(\vartheta-1)} = \binom{(\vartheta-1)(c-1)}{i_1}\binom{(\vartheta-1)(d-1)}{i_2}\binom{ci_2}{l}(-1)^{i_1+i_2+l} m^{i_1-ci_2+l}.$$

The lack of closed-form expressions for the Rényi and Tsallis entropies means that Table 4 provides numerical results based on possible parameter combinations, which can help in comprehending the behavior of these entropies.

Table 4
Numerical Results for Rényi Entropy and Tsallis Entropy.

Parameter Combinations			Rényi Entropy	Comments	Tsallis Entropy	Comments
$c = 1.1$	$d = 0.5$	$\vartheta = 0.5$	-0.2270	Decreasing	0.3188	Increasing
	$d = 1.5$	$\vartheta = 0.6$	-0.1016		0.1144	
	$d = 2.0$	$\vartheta = 0.7$	-0.5821		0.4384	
	$d = 2.5$	$\vartheta = 0.8$	-1.9827		0.8075	
	$d = 3.0$	$\vartheta = 0.9$	-7.0437		1.0432	
$d = 1.5$	$c = 1.1$	$\vartheta = 0.5$	-0.0677	Decreasing	0.1295	Decreasing than Increase
	$c = 1.2$	$\vartheta = 0.6$	-0.0813		0.0914	
	$c = 1.3$	$\vartheta = 0.7$	-0.1182		0.0748	
	$c = 1.4$	$\vartheta = 0.8$	-0.2222		-0.0219	
	$c = 1.5$	$\vartheta = 0.9$	-0.6077		0.0775	

4.5 Order Statistics (OS)

Let $T_1, T_2, T_3, \ldots, T_n$ be a random sample of size n follows to T and $\{T_{(1)} < T_{(2)} < T_{(3)} < \ldots < T_{(n)}\}$ be the corresponding OS. Then PDF of i-th ($i=1, 2, 3, \ldots, n$) OS is defined as

$$f_t|_{c,d,(i:n)} = C|_{B(.)}\left[F_t|_{c,d}\right]^{i-1}\left[1 - F_t|_{c,d}\right]^{n-i} f_t|_{c,d}.$$

The i-th OS $f_t|_{c,d,(i:n)}$ is defined as

$$f_t|_{c,d,(i:n)} = \begin{bmatrix} C|_{B(.)}[1-[1-p^c]^d]^{i-1}[\{1-p^c\}^d]^{n-i} \\ \times\left[\frac{cd(t-m)^{c-1}}{(1-m)^c}[1-p^c]^{d-1}\right] \end{bmatrix}. \quad (17)$$

For minimum $f_t|_{c,d,(i:n)}$, replace (i = 1) in (17), we get

$$f_t|_{c,d,(1:n)} = C|_{B(.)}[\{1-p^c\}^d]^{n-1} \frac{cd(t-m)^{c-1}}{(1-m)^c}[1-p^c]^{d-1}.$$

For maximum $f_t|_{c,d,(i:n)}$, replace (i = n) in (17), we get

$$f_t|_{c,d,(n:n)} = C|_{B(.)}[1-\{1-p^c\}^d]^{n-1}\left[\frac{cd(t-m)^{c-1}}{(1-m)^c}[1-p^c]^{d-1}\right].$$

where

$$p=\left(\frac{t-m}{1-m}\right), C|_{B(.)} = \frac{1}{B(i,n-i+1)!}.$$

5. INFERENCE

Let $T_1, T_2, \dots, T_n$ be a random sample of size *n* from *T*, then the log-likelihood function $l_t|_{c,d}$ of *T* is obtained as

$$l_t|_{c,d} = \begin{bmatrix} n\{logc + logd - clog(1-m)\} + (c-1)\sum_{i=1}^{n} log(t_i - m) \\ +(d-1)\sum_{i=1}^{n} log(1-p_i{}^c) \end{bmatrix}. \tag{18}$$

Let's suppose $p_i = (t_i - m)/(1-m)$ then the partial derivatives *w.r.t* c and d of $l_t|_{c,d,\tau}$ provides

$$\frac{\partial l_t|_{c,d}}{\partial c} = \frac{n}{c} - nlog(1-m) + \sum_{i=1}^{n} log(t_i - m) - (d-1)\sum_{i=1}^{n} \frac{p_i{}^c logp_i}{(1-p_i{}^c)(1-m)},$$

$$\frac{\partial l_t|_{c,d}}{\partial d} = \frac{n}{d} - \sum_{i=1}^{n} log(1-p_i{}^c).$$

The Maximum Likelihood Estimates (MLEs) for the SKD, represented as ($\hat{\tau}_i = \hat{c}, \hat{d}$), can be obtained through two methods: maximizing (18) or solving a set of non-linear equations. However, these equations lack an analytical solution for MLEs of parameters c and d. To find their optimal values, numerical methods like the Newton-Raphson algorithm are often used as they provide an efficient solution for MLEs.

6. SIMULATION EXPERIMENT

In this section, we evaluate the asymptotic performance of the MLEs $\hat{\tau}_i = \hat{c}, \hat{d}$ using the following algorithm:

i. Obtain random samples of sizes $n =$ 25, 50, 100, 200, 300, 400, 500, and 1000 from (14).

ii. Obtain results based on different combinations of model parameters from $S|_I$ = (c = 2.2, d = 2.5), $S|_{II}$ = (c = 1.5, d = 0.9), and $S|_{III}$= (c = 3.1, d = 0.5).

iii. Calculate the various known statistics including mean (denoted as $E|_{\hat{\tau}}$), variance (denoted as $Var|_{\hat{\tau}}$), bias (denoted as $Bias|_{\hat{\tau}}$), mean square error (denoted as$MSE|_{\hat{\tau}}$), coverage probability (denoted as$CP|_{\hat{\tau}}$), and average width (denoted as$AW|_{\hat{\tau}}$) using R, illustrated in Tables 5-10.

iv. Replicate each sample 1000 times.

v. Observe a gradual decrease in mean, biases, MSEs, and Var with an increase in sample sizes.

vi. Observe that CPs of all parameters $\tau = (c, d)$ approaching to 0.95 and AW decreases as sample sizes increase.

vii. The following are the statistics.

$$E|_{\hat{\tau}} = \frac{1}{1000}\sum_{i=1}^{1000}\hat{\tau}_i\,, Var|_{\hat{\tau}} = \frac{1}{1000}\sum_{i=1}^{1000}(\tau - \bar{\tau}_i)^2\,, Bias|_{\hat{\tau}} = \frac{1}{1000}\sum_{i=1}^{1000}(\hat{\tau}_i - \tau),$$

$$MSE|_{\hat{\tau}} = \frac{1}{1000}\sum_{i=1}^{1000}(\hat{\tau}_i - \tau)^2\,, CP|_{\hat{\tau}} = \sum_{i=1}^{1000} I\big(\hat{\tau}_i - 0.95SE|_{\hat{\tau}_i}, \hat{\tau}_i + 0.95SE|_{\hat{\tau}_i}\big),$$

$$AW|_{\hat{\tau}} = \frac{1}{1000}\sum_{i=1}^{1000}|(\hat{\tau}_i + 0.95) - (\hat{\tau}_i - 0.95)|,$$

where, Indicator function I(.), $SE|_{\hat{\tau}_i} = \sqrt{var\hat{\tau}_i}$.

7. REAL-LIFE EXAMPLES

This section discusses the applicability of the SKD and its relevance to different fields. To demonstrate the versatility of the SKD, we explore four distinct data sets related to textile engineering, reliability engineering, petroleum engineering, and agriculture sciences. By examining the SKD's performance on these data sets, we aim to highlight its ability to effectively analyze data from a variety of domains and provide meaningful insights and results. This also showcases the versatility and broad applicability of the SKD, making it a valuable tool for data analysis and decision-making in many fields. Additionally, this exploration of different data sets helps to further validate the robustness and reliability of the SKD, ensuring that it can be trusted to deliver accurate results in diverse situations. In order to conserve space and improve the readability of the main text, the authors have chosen to provide the data sets in the appendix section for easy access and reference. For the most up-to-date information, we highly recommend that readers refer to the comprehensive research conducted by Alsadat et al. (2023), Tashkandy et al. (2023), and Alghamdi et al. (2023).

The first data set under consideration in this study comprises of 30 measurements of tensile strength of polyester fibers. This data set was originally presented by Quesenberry and Hales (1980) and provides valuable insights into the tensile strength of polyester fibers. A set of descriptive statistics for a particular data set includes mean = 0.3659, skewness = 0.5193, kurtosis = 2.2293, minimum = 0.0230, maximum = 0.9260, first quartile = 0.1323, third quartile = 0.5265, and 95% confidence interval = (0.2655, 0.4661). The second data set includes 30 measurements of electronic device lifetimes. A set of descriptive statistics for a particular data set includes mean = 0.4940, skewness = 0.0616, kurtosis = 1.3128, minimum = 0.0200, maximum = 0.9900, first quartile = 0.1432, third quartile = 0.8920, and 95% confidence interval = (0.3534, 0.6345). This data set was originally presented by Rahman et al. (2019). Third data set consists of measurements of the shape perimeter to area ratio in petroleum rock samples. A set of descriptive statistics for a particular data set includes mean = 0.2181, skewness = 1.1329, kurtosis = 4.1098, minimum = 0.0903, maximum = 0.4641, first quartile = 0.1622, third quartile = 0.2626, and 95% confidence interval = (0.1938, 0.2423). This data set was originally presented by Cordeiro and Brito (2012). The fourth data set involves the analysis of soil fertility and the biological fixation of nitrogen in relation to the growth of 128 Dimorphandra wilsonii Rizz plants, specifically looking at phosphorus concentration in the leaves. The results of this study, as discussed by Oliveira et al. (2013), will provide insights into the relationship between soil fertility, nitrogen fixation, and plant growth, and will be of interest to those working in agriculture sciences and plant biology. A set of descriptive statistics for a particular data set includes mean = 0.1408, skewness = 0.4490, kurtosis = 2.3552, minimum = 0.0500, maximum = 0.2800, first quartile = 0.1000, third quartile = 0.1800, and 95% confidence interval = (0.1312, 0.1502).

Table 5
Statistical Measures for $S|_{III}$

Sample	$Mean|_c$	$Var|_c$	$Bias|_c$	$MSE|_c$	$CP|_c$	$AW|_c$
25	2.3874	0.2874	0.1874	0.5679	0.884	1.6378
50	2.2839	0.1290	0.0839	0.3688	0.889	1.1196
100	2.2452	0.0575	0.0452	0.2441	0.899	0.7828
200	2.2202	0.0306	0.0202	0.1763	0.890	0.5490
300	2.2170	0.0201	0.0170	0.1429	0.888	0.4479
400	2.2125	0.0149	0.0125	0.1228	0.892	0.3870
500	2.2089	0.0115	0.0089	0.1078	0.905	0.3460
1000	2.2056	0.0055	0.0056	0.0747	0.904	0.2444

Table 6
Statistical Measures for $S|_{III}$

Sample	$Mean|_d$	$Var|_d$	$Bias|_d$	$MSE|_d$	$CP|_d$	$AW|_d$
25	3.0028	1.3495	0.5028	1.2658	0.941	3.3009
50	2.7226	0.4566	0.2226	0.7115	0.916	2.0429
100	2.6115	0.1905	0.1115	0.4505	0.910	1.3668
200	2.5516	0.0902	0.0517	0.3047	0.895	0.9375
300	2.5376	0.0558	0.0377	0.2392	0899	0.7597
400	2.5270	0.0420	0.0270	0.2067	0.894	0.6543
500	2.5201	0.0323	0.0201	0.1809	0.900	0.5831
1000	2.5102	0.0157	0.0102	0.1257	0.900	0.4102

Table 7
Statistical Measures for $S|_{IV}$

Sample	$Mean|_c$	$Var|_c$	$Bias|_c$	$MSE|_c$	$CP|_c$	$AW|_c$
25	1.6781	0.2308	0.1781	0.5124	0.883	1.4418
50	1.5815	0.1004	0.0815	0.3272	0.899	0.9754
100	1.5421	0.0436	0.0421	0.2130	0.904	0.6778
200	1.5189	0.0228	0.0189	0.1524	0.877	0.4740
300	1.5154	0.0149	0.0154	0.1234	0.886	0.3866
400	1.5113	0.0111	0.0113	0.1057	0.894	0.3341
500	1.5082	0.0085	0.0082	0.0931	0.901	0.2984
1000	1.5048	0.0042	0.0048	0.0643	0.905	0.2107

Table 8
Statistical Measures for $S|_{IV}$

Sample	$Mean|_d$	$Var|_d$	$Bias|_d$	$MSE|_d$	$CP|_d$	$AW|_d$
25	1.0230	0.0946	0.1230	0.3313	0.917	0.8938
50	0.9575	0.0357	0.0575	0.1976	0.908	0.5809
100	0.9287	0.0157	0.0287	0.1288	0.904	0.3955
200	0.9132	0.0073	0.0132	0.0868	0.904	0.2739
300	0.9094	0.0046	0.0094	0.0685	0.906	0.2225
400	0.9066	0.0035	0.0066	0.0595	0.898	0.1920
500	0.9050	0.0027	0.0051	0.0526	0.908	0.1714
1000	0.9023	0.0013	0.0023	0.0369	0.901	0.1207

Table 9
Statistical Measures for $S|_{V}$

Sample	$Mean|_c$	$Var|_c$	$Bias|_c$	$MSE|_c$	$CP|_c$	$AW|_c$
25	3.6049	1.610	0.5049	1.3657	0.893	3.706
50	3.3378	0.6695	0.2378	0.8521	0.897	2.4665
100	3.2167	0.2770	0.1167	0.5391	0.909	1.6953
200	3.1524	0.1416	0.0524	0.3799	0.883	1.1802
300	3.1423	0.0934	0.0423	0.3086	0.894	0.9619
400	3.1328	0.0685	0.0328	0.2638	0.893	0.8311
500	3.1234	0.0534	0.0234	0.2324	0.897	0.7416
1000	3.1122	0.0253	0.0122	0.1596	0.914	0.5230

Table 10
Statistical Measures for $S|_V$

Sample	$Mean|_d$	$Var|_d$	$Bias|_d$	$MSE|_d$	$CP|_d$	$AW|_d$
25	0.5571	0.0226	0.0571	0.1608	0.917	0.4411
50	0.5277	0.0088	0.0277	0.0981	0.907	0.2922
100	0.5137	0.0040	0.0137	0.0648	0.897	0.2002
200	0.5062	0.0018	0.0062	0.0433	0.908	0.1391
300	0.5044	0.0011	0.0044	0.0344	0.903	0.1131
400	0.5032	0.0009	0.0033	0.0300	0.901	0.0977
500	0.5023	0.0007	0.0023	0.0265	0.907	0.0872
1000	0.5010	0.0003	0.0010	0.0187	0.889	0.0614

The proposed SKD is compared to its rival models (CDFs listed in Table 11) based on evaluation criteria (denoted as EC) such as the Akaike information criterion (denoted as$AIC|_{EC}$) and Bayesian information criterion (denoted as$BIC|_{EC}$), as well as fit statistics Cramer-Von Mises (denoted as$CVM|_{EC}$), Anderson-Darling (denoted as$AD|_{EC}$), and Kolmogorov Smirnov (denoted as$KS|_{EC}$) with its p-value. The parameters' estimates and fit statistics are shown in Tables 12 - 15. The model with the lowest fit statistics is deemed to be the best-fit model and in this case, the SKD outperforms its competitors in all four data sets.

Additionally, to provide further evidence for the superiority of the proposed SKD model, the fitted density and distribution functions, as well as the Probability-Probability (P-P) and Kaplan-Meier survival plots, are depicted in Figures 3-6. The total time on test transform (denoted as TTT) and box plots are also defined to give a comprehensive overview of the model's performance. These visual representations demonstrate the close agreement between the fitted model and the actual datasets. The calculations are carried out using the R software and its dedicated Adequacy Model package.

Table 11
List of Some Competitive Models CDFs

Model	Competitive Models CDF	Support	Author(s)
K	$1-[1-t^c]^d$	$c,d>0, 0<t<1$	Kumaraswamy (1980)
WPF	$1-\ e^{-e[t^c/(M^c-t^c)]^d}$	$c,d,e>0, 0<t\le M$	Tahir et al. (2014)
KPF	$1-\ [1-(t/M)^{cd}]^e$	$c,d,e>0, 0<t\le M$	Abdul-Moniem (2017)
GPF	$1-[M-t]^c[M-m]^{-c}$	$c>0, m\le t\le M$	Saran and Pandey (2004)
Beta	$I_t(c,d)$	$c,d>0, 0<t<1$	Mood et al. (1974)

Table 12
Estimates and Fit Statistics for First Data

Model	$\hat{c}$	$\hat{d}$	$\hat{e}$	$AIC\|_{EC}$	$BIC\|_{EC}$	$CVM\|_{EC}$	$AD\|_{EC}$	$KS\|_{EC}$	p-value
SKD	0.8182	1.4522	-	-5.3404	-2.5380	0.0097	0.0786	0.0480	1.0000
WPF	3.0299	1.3464	0.7957	0.2444	4.4480	0.0174	0.1382	0.0611	0.9995
K	0.9627	1.6081	-	-2.6221	0.1803	0.0183	0.1550	0.0650	0.9987
Beta	0.9666	1.6204	-	-2.6101	0.1923	0.0184	0.1559	0.0669	0.9979
KPF	1.7343	0.4737	1.1223	-0.8739	3.3297	0.0401	0.3217	0.0958	0.9214
GPF	1.3187	-	-	-5.5597	-4.1585	0.0248	0.1953	0.1287	0.6557

Table 13
Estimates and Fit Statistics for Second Data

Model	$\hat{c}$	$\hat{d}$	$\hat{e}$	$AIC\|_{EC}$	$BIC\|_{EC}$	$CVM\|_{EC}$	$AD\|_{EC}$	$KS\|_{EC}$	p-value
SKD	0.5082	0.5826	-	-7.1348	-4.3324	0.0845	0.5146	0.1429	0.5263
Beta	0.6062	0.5911	-	-3.2498	-0.4474	0.1038	0.6483	0.1550	0.4242
K	0.5875	0.6115	-	-3.0050	-0.2026	0.1059	0.6625	0.1600	0.3850
KPF	7.9804	0.0713	0.5807	-2.7157	1.4879	0.1002	0.6207	0.1616	0.3731
WPF	0.7992	0.6687	0.9600	12.8687	17.0723	0.1666	1.0490	0.1671	0.3340
GPF	0.7525	-	-	-1.8849	-0.4837	0.0784	0.4681	0.2729	0.0183

Table 14
Estimates and Fit Statistics for Third Data

Model	$\hat{c}$	$\hat{d}$	$\hat{e}$	$AIC\|_{EC}$	$BIC\|_{EC}$	$CVM\|_{EC}$	$AD\|_{EC}$	$KS\|_{EC}$	p-value
SKD	1.6281	18.375	-	-110.6602	-106.9178	0.0789	0.5019	0.1128	0.5750
Beta	5.9417	21.205	-	-107.2004	-103.4580	0.1281	0.7782	0.1428	0.2820
WPF	42.995	8.7742	0.3131	-99.4829	-93.8693	0.2000	1.2259	0.1499	0.2308
K	2.7187	44.6671	-	-100.9831	-97.2407	0.2084	1.2803	0.1533	0.2095
GPF	1.7877	-	-	-103.4055	-101.5343	0.2316	1.4418	0.1558	0.1944
KPF	1.4411	1.4050	2.6326	-86.0842	-80.4706	0.4173	2.5450	0.1863	0.0716

Table 15
Estimates and Fit Statistics for Fourth Data

Model	$\hat{c}$	$\hat{d}$	$\hat{e}$	$AIC\|_{EC}$	$BIC\|_{EC}$	$CVM\|_{EC}$	$AD\|_{EC}$	$KS\|_{EC}$	p-value
SKD	1.7889	48.976	-	-395.1622	-389.4582	0.0884	0.5177	0.0791	0.3990
Beta	5.7256	34.9435	-	-390.1785	-384.4745	0.1384	0.7698	0.0971	0.1792
WPF	62.087	1.9504	1.3536	-382.8504	-374.2943	0.2147	1.2038	0.1160	0.0638
K	2.8104	176.343	-	-385.6015	-379.8974	0.2078	1.1612	0.1181	0.0562
GPF	1.6077	-	-	-388.1287	-385.2766	0.0983	0.5482	0.1301	0.0262
KPF	1.2431	1.8299	2.9017	-375.2356	-366.6795	0.2823	1.6755	0.1304	0.0257

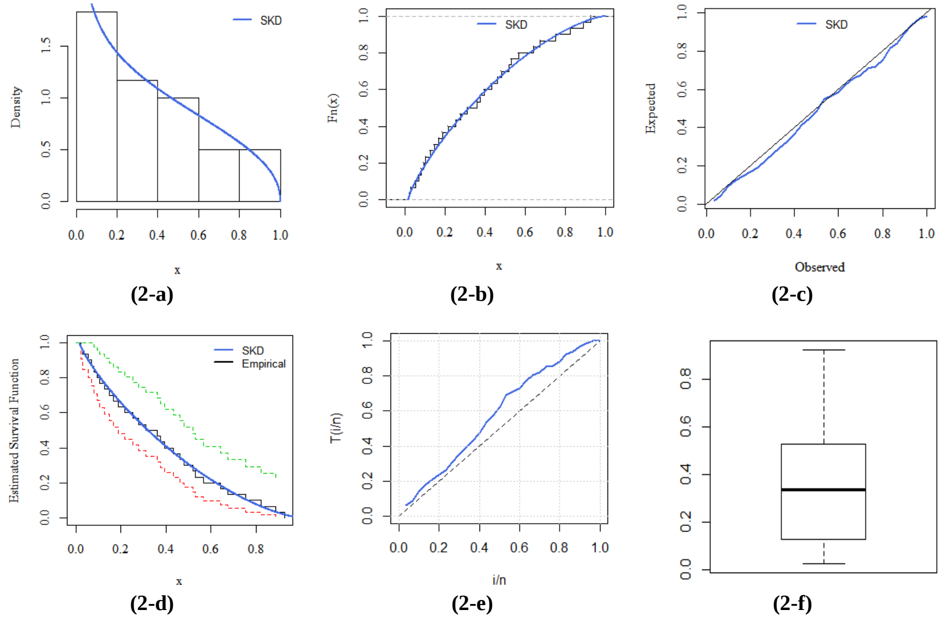


Figure 3: Fitted Plots for First Data

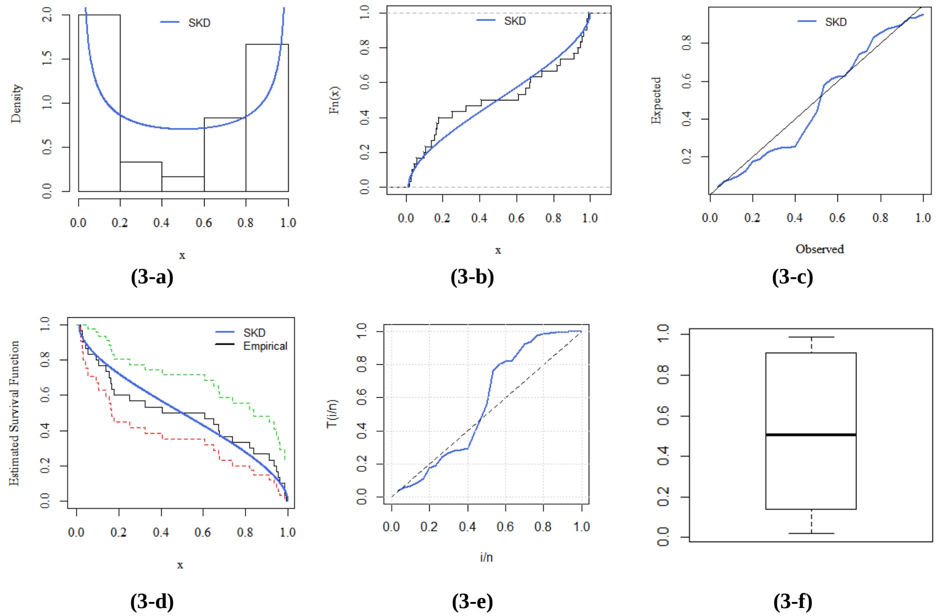
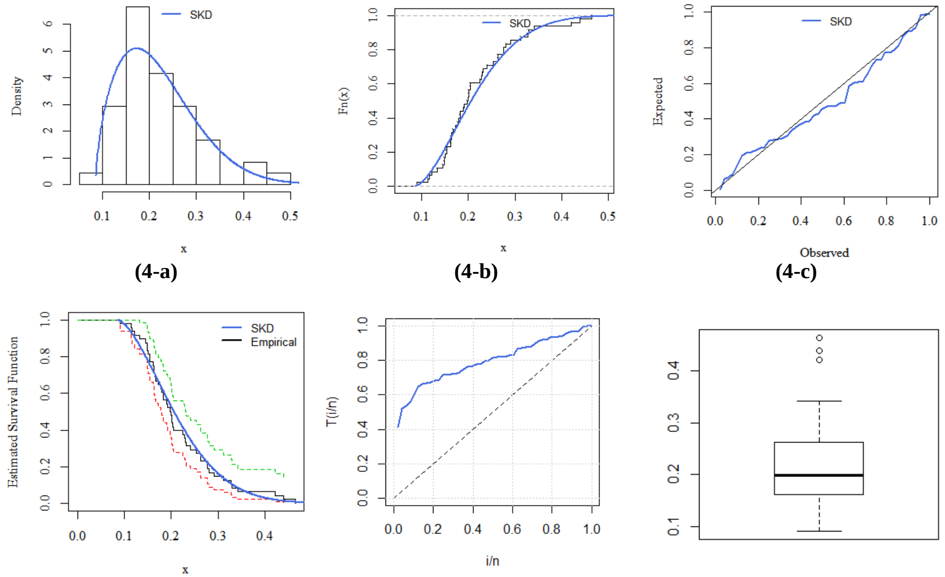


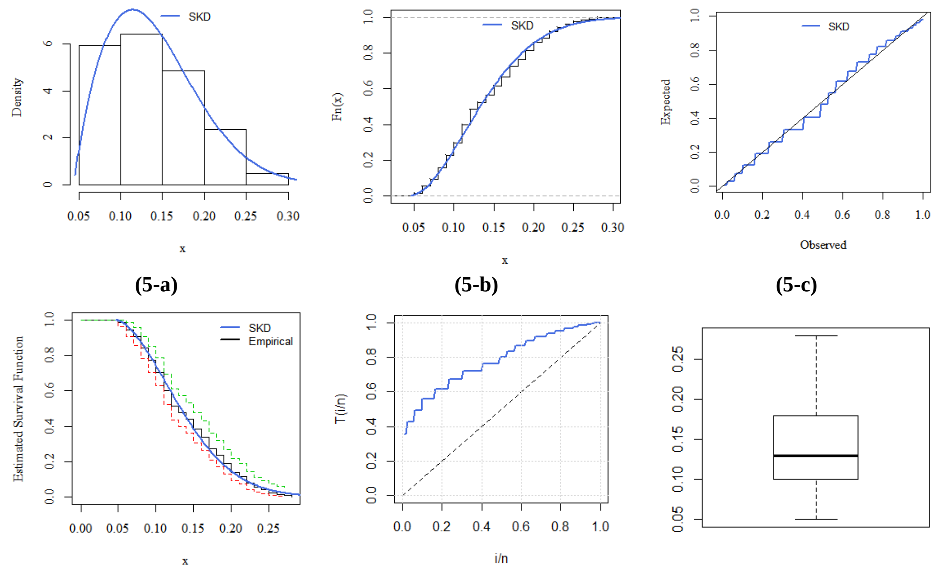
Figure 4: Fitted Plots for Second Data



(4-a) (4-b) (4-c)

(4-d) (4-e) (4-f)

Figure 5: Fitted Plots for Third Data



(5-a) (5-b) (5-c)

(5-d) (5-e) (5-f)

Figure 6: Fitted Plots for Fourth Data

8. CONCLUSION & FUTURE DIRECTIONS

This study introduced the shifted Kumaraswamy distribution (SKD), which demonstrated exceptional performance in modeling asymmetrical and bathtub-shaped real-world phenomena. The study derived and analyzed the SKD's moment-generating function, first incomplete moment, Bonferroni and Lorenz functions, residual life and reverse residual life functions, reliability measures, probability density function, hazard rate function, Rényi and Tsallis entropy, and order statistics. To demonstrate the versatility of the SKD, four distinct datasets related to textile engineering, reliability engineering, petroleum engineering, and agriculture sciences were explored. The SKD's parameters were estimated using maximum likelihood estimation, and the model was tested on these datasets, outperforming baseline models. This study provides robust evidence for the effectiveness and usefulness of the SKD in modeling complex real-world phenomena.

It is believed that the SKD will be a valuable option in real-world applications in the future. In future research, several directions can be taken to further explore the capabilities of the new bathtub model for predicting failure rates in lifetime data: (a) Extension to censored data: The current model can be extended to handle censored data, which is a common scenario in lifetime data analysis. (b) Incorporation of covariates: The effect of incorporating relevant covariates on the prediction accuracy can be investigated.

COMPETING INTERESTS

No competing interests were disclosed

ACKNOWLEDGEMENTS

We would like to extend our sincere gratitude to the anonymous referees and the editor for their valuable feedback and constructive comments on earlier versions of this manuscript.

REFERENCES

1. Abdul-Moniem, I.B. (2017). The Kumaraswamy power function distribution. *Journal of Statistics and Applications & Probability*, 6(1), 81-90.
2. Afify, A.Z., Yousof, H.M., Cordeiro, G.M., Nofal, Z.M. and Ahmad, M. (2016). The Kumaraswamy Marshall-Olkin Fréchet distribution with applications. *Journal of ISOSS*, 2(2), 151-168.
3. Al-Babtain, A.A., Elbatal, I., Chesneau, C. and Elgarhy, M. (2021). Estimation of different types of entropies for the Kumaraswamy distribution. *PLoS One,* 16(3), 1-21.
4. Alghamdi, S.M., Shrahili, M., Hassan, A.S., Gemeay, A.M., Elbatal, I. and Elgarhy, M. (2023). Statistical Inference of the Half Logistic Modified Kies Exponential Model with Modeling to Engineering Data. *Symmetry*, 15(3), 586.

5. Alizadeh, M., Tahir, H.M., Cordeiro, G.M., Mansoor, M., Zubair, M. and Hamedani, G.G. (2015). The Kumaraswamy Marshall-Olkin family of distributions. *Journal of the Egyptian Mathematical Society*, 23(3), 546-557.
6. Alsadat, Najwan, Aijaz, A., Muzamil, J., Ahmed M.G., Mohammed A.M., Eslam H., Ehab M.E. and Md Moyazzem H. (2023). The novel Kumaraswamy power Fréchet distribution with data analysis related to diverse scientific areas. *Alex. Eng. J.*, 70, 651-664.
7. Arshad, M.Z., Balogun, O.S., Iqbal, M.Z. and Oguntunde, P.E. (2022). On Some Properties of a New Truncated Model With Applications to Lifetime Data. *International Journal of Analysis and Applications*, 20, 23-23.
8. Arshad, M.Z., Iqbal, M.Z. and Al Mutairi, A. (2021). Statistical Properties of a New Bathtub Shaped Failure Rate Model with Applications in Survival and Failure Rate Data. *International Journal of Statistics and Probability*, 10(3), 49-68.
9. Bourguignon, M., Silva, R.B., Zea, L.M. and Cordeiro, G.M. (2013). The Kumaraswamy Pareto distribution. *Journal of Statistical Theory and Practice*, 12(2), 129-144.
10. Bursa, N. and Ozel, G. (2017). The exponentiated Kumaraswamy power function distribution. *Hacettepe Journal of Mathematics and Statistics*, 46(2), 277-292.
11. Cordeiro, G.M. and Brito, R.D.S. (2012). The beta power distributions. *Brazilian Journal of Probability and Statistics*, 26(1), 88-112.
12. Cordeiro, G.M. and De Castro, M. (2011). A new family of generalized distributions. *Journal of Statistical Computation and Simulation*, 81(7), 883-893.
13. Cousineau, D. (2009). Fitting the Three-Parameter Weibull Distribution: Review and Evaluation of Existing and New Methods. *IEEE Transactions on Dielectrics and Electrical Insulation*, 16(1), 281-288.
14. Eledum, H. and Alaa, R. (2023). Discrete Kumaraswamy Erlang-Truncated Exponential Distribution with Applications to Count Data. *J. Stat. Appl. Pro*. 12(2), 725-739.
15. Eldessouky, E.A., Hassan, O.H.M., Elgarhy, M., Hassan, E.A., Elbatal, I. and Almetwally, E.M. (2023). A New Extension of the Kumaraswamy Exponential Model with Modeling of Food Chain Data. *Axioms*, 12(4), 379.
16. Gómez-Déniz, E., Pérez-Rodríguez, J.V., Reyes, J. and Gómez, H.W. (2020). A Bimodal Discrete Shifted Poisson Distribution. A Case Study of Tourists' Length of Stay. *Symmetry*, 12(3), 2-15.
17. Hassan, A.S., Almetwally, E.M. and Ibrahim, G.M. (2021). Kumaraswamy Inverted Topp-Leone Distribution with Applications to COVID-19 Data. *Computers, Materials and Continua*, 68(1), 337-358.
18. Hemeda, S.E., Pramanik, S. and Maitic, S.S. (2020). Kumaraswamy Inverse Lindley Distribution with Stress-Strength Reliability. *Gazi Journal of Engineering Sciences*, 6(3), 255-264.
19. Ibrahim, M. (2017). The Kumaraswamy power function distribution. *Journal of Applied Probability and Statistics*, 6(1), 81-90.

20. Ikechukwu, A.F., Jaspa, O.E., Emmanuel, R.F. and Adeyinka, O. (2020). A Three Parameter Shifted Exponential Distribution: Properties and Applications. *Thailand Statistician*, 18(4), 470-480.
21. Jodrá, P. (2020). A bounded distribution derived from the shifted Gompertz law. *Journal of King Saud University - Science*, 32(1), 523-536.
22. Jones, M.C. (2009). Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. *Statistical Methodology*, 6(1), 70-81.
23. Kumaraswamy, P. (1980). Generalized probability density-function for double-bounded random-processes. *Journal of Hydrology*, 46(1-2), 79-88.
24. Lehmann, E.L. (1953). The power of rank tests. *Annals of Mathematical Statistics,* 24(1), 23-43.
25. Lemonte, A.J., Souza, W.B. and Cordeiro, G.M. (2013). The exponentiated Kumaraswamy distribution and its log-transform. *Brazilian Journal of Probability and Statistics*, 27(1), 31-53.
26. Madi, M.T. and Leonard, T. (1996). Bayesian estimation for shifted exponential distributions. *Journal of Statistical Planning and Inference*, 55(3), 345-351.
27. Mahmoud, M.M., Aryal, G., Afify, A.Z. and Ahmad, M. (2018). The Kumaraswamy Exponentiated Fréchet Distribution. *Pakistan Journal of Statistics,* 34(3), 177-193.
28. McDonald, J.B. (1984). Some generalized functions for the size distribution of income. *Econometrica,* 52(3), 647-663.
29. Mood, A.M., Graybill, F.A. and Boes, D.C. (1974). *Introduction to the Theory of Statistics*. McGraw-Hill: New York, USA.
30. Nawaz, T., Hussain, S., Ahmad, T., Naz, F. and Abid, M. (2018). Kumaraswamy generalized Kappa distribution with application to stream flow data. *Journal of King Saud University - Science*, 32(1), 1-11.
31. Neto, F.C. and Santos, J. (2019). Inflated Kumaraswamy distributions. *Anais da Academia Brasileira de Ciências,* 91(2), 1-18.
32. Oliveira, J., Santos, J., Xavier, C., Trindade, D. and Cordeiro, G.M. (2013). The McDonald half-logistic distribution. *Communications in Statistics - Theory and Methods*, 45(7), 2005-2022.
33. Quesenberry, C.P. and Hales, C. (1980). Concentration bands for uniformity plots. *Journal of Statistical Computation and Simulation*, 11(1), 41-53.
34. Rahman, Md. M., Al-Zahrani, B., Shahbaz, S.H. and Shahbaz, M.Q. (2019). Cubic Transmuted Uniform Distribution: An Alternative to Beta and Kumaraswamy Distributions. *European Journal of Pure and Applied Mathematics*, 12(3), 1106-1121.
35. Ramzan, Q., Qamar, S., Amin, M., Alshanbari, H.M., Nazeer, A. and Elhassanein, A. (2022). On the extended generalized inverted Kumaraswamy distribution. *Computational Intelligence and Neuroscience*, https://doi.org/10.1155/2022/1612959
36. Rényi, A. (1961). On measures of entropy and information, In: *Proceedings of the 4th Fourth Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, 547- 561.
37. Saran, J. and Pandey, A. (2004). Estimation of parameters of power function distribution and its characterization by the k-th record values. *Statistica*, 64(3), 523-536.

38. Silva, R., Silva, F.G., Ramos, M., Cordeiro, G.M., Marinho, P. and Andrade, T. (2019). The exponentiated Kumaraswamy-G class: general properties and application. *Revista Colombiana de Estadistica*, 42(1), 1-33.
39. Tahir, H.M., Alizadeh, M., Mansoor, M., Cordeiro, G.M. and Zubair, M. (2014). The Weibull power function distribution with applications. *Hacettepe Journal of Mathematics and Statistics*, 45(1), 245-265.
40. Tahir, M.H., Hussain, M.A., Cordeiro, G.M., El-Morshedy, M. and Eliwa, M.S. (2020). A new Kumaraswamy generalized family of distributions with properties, applications, and bivariate extension. *Mathematics*, 8(11), 1989.
41. Tashkandy, Y.A., Nagy, M., Akbar, M., Mahmood, Z., Gemeay, A.M., Hossain, M.M. and Muse, A.H. (2023). The Exponentiated Cotangent Generalized Distributions: Characteristics and Applications Patients of Chemotherapy Treatments Data. *IEEE Access*.
42. Yang, F., Ren, H. and Hu, Z. (2019). Maximum Likelihood Estimation for Three-Parameter Weibull Distribution Using Evolutionary Strategy. *Mathematical Problems in Engineering*, https://doi.org/10.1155/2019/6281781

APPENDIX

The following data set has been used and analyzed in real life examples

Data Set 1: 30 Measurements of Tensile Strength of Polyester Fibers.
0.023, 0.032, 0.054, 0.069, 0.081, 0.094, 0.105, 0.127, 0.148, 0.169, 0.188, 0.216, 0.255, 0.277, 0.311, 0.361, 0.376, 0.395, 0.432, 0.463, 0.481, 0.519, 0.529, 0.567, 0.642, 0.674, 0.752, 0.823, 0.887, 0.926.

Data Set 2: 30 Measurements of Electronic Device Lifetimes.
0.020, 0.029, 0.034, 0.044, 0.057, 0.096, 0.106, 0.139, 0.156, 0.164, 0.167, 0.177, 0.250, 0.326, 0.406, 0.607, 0.650, 0.672, 0.676, 0.736, 0.817, 0.838, 0.910, 0.931, 0.946, 0.953, 0.961, 0.981, 0.982 , 0.990.

Data Set 3: Shape Perimeter to Area Ratio in Petroleum Rock Samples.
0.0903296, 0.2036540, 0.2043140, 0.2808870, 0.1976530, 0.3286410, 0.1486220, 0.1623940, 0.2627270, 0.1794550, 0.3266350, 0.2300810, 0.1833120, 0.1509440, 0.2000710, 0.1918020, 0.1541920, 0.4641250, 0.1170630, 0.1481410, 0.1448100, 0.1330830, 0.2760160, 0.4204770, 0.1224170, 0.2285950, 0.1138520, 0.2252140, 0.1769690, 0.2007440, 0.1670450, 0.2316230, 0.2910290, 0.3412730, 0.4387120, 0.2626510, 0.1896510, 0.1725670, 0.2400770, 0.3116460, 0.1635860, 0.1824530, 0.1641270, 0.1534810, 0.1618650, 0.2760160, 0.2538320, 0.2004470.

Data Set 4: Phosphorus Concentration in the Leaves.
0.22, 0.17, 0.11, 0.10, 0.15, 0.06, 0.05, 0.07, 0.12, 0.09, 0.23, 0.25, 0.23, 0.24, 0.20, 0.08, 0.11, 0.12, 0.10, 0.06, 0.20, 0.17, 0.20, 0.11, 0.16, 0.09, 0.10, 0.12, 0.12, 0.10, 0.09, 0.17, 0.19, 0.21, 0.18, 0.26, 0.19, 0.17, 0.18, 0.20, 0.24, 0.19, 0.21, 0.22, 0.17, 0.08, 0.08, 0.06, 0.09, 0.22, 0.23, 0.22, 0.19, 0.27, 0.16, 0.28, 0.11, 0.10, 0.20, 0.12, 0.15, 0.08, 0.12, 0.09, 0.14, 0.07, 0.09, 0.05, 0.06, 0.11, 0.16, 0.20, 0.25, 0.16, 0.13, 0.11, 0.11, 0.11, 0.08, 0.22, 0.11, 0.13, 0.12, 0.15, 0.12, 0.11, 0.11, 0.15, 0.10, 0.15, 0.17, 0.14, 0.12, 0.18, 0.14, 0.18, 0.13, 0.12, 0.14, 0.09, 0.10, 0.13, 0.09, 0.11, 0.11, 0.14, 0.07, 0.07, 0.19, 0.17, 0.18, 0.16, 0.19, 0.15, 0.07, 0.09, 0.17, 0.10, 0.08, 0.15, 0.21, 0.16, 0.08, 0.10, 0.06, 0.08, 0.12, 0.13.

List of Acronyms and Abbreviations
Kumaraswamy = K,
Weibull Power Function = WPF,
Kumaraswamy Power Function = KPF,
Generalized Power Function = GPF.