

RESEARCH

Open Access



The Kumaraswamy transmuted Pareto distribution

Sher B. Chhetri¹, Alfred A. Akinsete^{2*} , Gokarna Aryal³ and Hongwei Long¹

*Correspondence:

akinsete@marshall.edu

²Department of Mathematics,
Marshall University, Huntington,
WV, USA

Full list of author information is
available at the end of the article

Abstract

In this work, a new five-parameter Kumaraswamy transmuted Pareto (*KwTP*) distribution is introduced and studied. We discuss various mathematical and statistical properties of the distribution including obtaining expressions for the moments, quantiles, mean deviations, skewness, kurtosis, reliability and order statistics. The estimation of the model parameters is performed by the method of maximum likelihood. We compare the distribution with few other distributions to show its versatility in modeling data with heavy tail.

Keywords: Kumaraswamy distribution, Quadratic rank transmutation map (QRTM), Pareto distribution, Hazard function, Maximum likelihood estimation

AMS Subject Classification: 60E05, 62E15, 62H12

1 Introduction

Furtherance to the work by Eugene et al. (2002), who proposed and defined the beta-generated class of distributions for a continuous random variable, derived from the logit of the beta random variable, many statistical distributions have been proposed and studied by numerous authors. According to Eugene et al. (2002), suppose X is a random variable with cumulative distribution function (CDF) $F(x)$, the CDF for the beta-generated family is obtained by applying the inverse probability transformation to the beta density function. The CDF for the beta-generated family of distributions is given by

$$G(x) = \frac{1}{B(u, v)} \int_0^{F(x)} t^{u-1} (1-t)^{v-1} dt, \quad 0 < u, v < \infty, \quad (1)$$

where $B(u, v) = \Gamma(u)\Gamma(v)/\Gamma(u+v)$. The corresponding probability density function (PDF) is given by

$$g(x) = \frac{1}{B(u, v)} [F(x)]^{u-1} [1-F(x)]^{v-1} \left[\frac{d}{dx} F(x) \right]. \quad (2)$$

A closely related generalized distribution similar to using the beta random variable as the baseline distribution as defined in Eqs. (1) and (2), is the work due to Cordeiro and de Castro (2011), where the authors combined the works of Eugene et al. (2002) and Jones (2009) by replacing the baseline distribution with the Kumaraswamy (*Kw*) distribution to

construct a new class of *Kw-G* distributions. See also Bourguignon et al. (2013), Cordeiro et al. (2010), and Elbatal (2013). The work by Shams (2013) also introduced and studied another type of generalization of the Kumaraswamy distribution.

A new approach of constructing and generalizing statistical distributions was proposed by Shaw and Buckley (2007,2009). Their concept is defined by transmutation maps that are functional composition of the cumulative distribution function of one distribution with the quantile function of another distribution. One of the maps is referred to as the Quadratic Rank Transmutation Map (QRTM). Specifically, given two distributions with a common sample space with CDFs F_1 and F_2 , we can define a pair of general rank transmutation map

$$G_{R_{12}}(u) = F_2 \left(F_1^{-1}(u) \right), \quad G_{R_{21}}(u) = F_1 \left(F_2^{-1}(u) \right),$$

where the pair of maps $\{G_{R_{12}}(u), G_{R_{21}}(u)\}$ takes the unit interval $I_{[0,1]}$ into itself, and under suitable assumptions are mutual inverses, satisfying the conditions $G_{ij}(0) = 0$ and $G_{ij}(1) = 1$. A useful additional condition requires that the rank transmutation maps are monotone and continuously differentiable, otherwise transmuted density may be discontinuous.

In this paper, we use the transmutation map approach outlined in Shaw and Buckley (2007), and also defined in Shaw and Buckley (2009). This approach considered for $|\lambda| \leq 1$,

$$G_{R_{12}}(u) = u + \lambda u(1 - u).$$

The consequence of the above results in the relationship

$$F_2(x) = (1 + \lambda)F_1(x) - \lambda F_1(x)^2,$$

which on differentiation yields,

$$f_2(x) = f_1(x)[(1 + \lambda) - 2\lambda F_1(x)],$$

where $f_1(x)$ and $f_2(x)$ are the corresponding PDFs associated with $F_1(x)$ and $F_2(x)$ respectively.

We will use the above formulation for a pair of distributions $F(x)$ and $G(x)$, where $G(x)$ is a submodel or the baseline distribution of $F(x)$. By a formal definition, a random variable X is said to have a transmuted probability distribution with CDF $F(x)$ if

$$F(x) = (1 + \lambda)G(x) - \lambda G(x)^2, \quad |\lambda| \leq 1 \tag{3}$$

The corresponding PDF of the transmuted probability distribution is

$$f(x) = g(x)(1 + \lambda - 2\lambda G(x)).$$

Note that when $\lambda = 0$, the above relation reduces to the distribution ($G(x)$) of the baseline random variable.

Following the work by Aryal and Tsokos (2009) on the transmuted extreme value distribution, a number of transmuted family of distributions have been proposed and discussed by many authors in the literature. A few of them are the work by Aryal and Tsokos (2011): transmuted Weibull distribution; Khan and King (2013): transmuted modified Weibull distribution; Ashour and Eltehiwy (2013): transmuted Lomax distribution; Hady and Ebraheim (2014): exponentiated transmuted Weibull distribution; Hussian (2014): transmuted exponentiated gamma distribution; Merovci and Puka (2014): transmuted Pareto distribution; Owokolo et al. (2015): transmuted exponential distribution, and Khan et al. (2016): transmuted Kumaraswamy distribution; Afify et al. (2014): transmuted complementary Weibull geometric distribution; Yousof et al. (2015): the transmuted exponentiated generalized-G family of distributions, and a host of many others. Tahir and Cordeiro (2016) provided a comprehensive list of contributed works on transmuted distributions. This paper is outlined as follows. In Section 2, we introduce the Kumaraswamy transmuted Pareto distribution. Some of its mathematical properties are discussed in Sections 3, 4, 5, 6, 7, 8 and 9. Section 10 discusses the estimation of the parameters of the distribution by the method of maximum likelihood. We provide the application of the distribution to two real life data in Section 11. In Section 12 we have simulation, and concluding remarks in Section 13.

2 The Kumaraswamy transmuted Pareto (KwTP) distribution

Following the transmutation map by Shaw and Buckley (2007) as defined in Eq. (3), we combine the Kumaraswamy distribution and the transmuted Pareto to form what we refer to as the Kumaraswamy transmuted Pareto (*KwTP*) distribution. Without loss of generality, we give a brief description of these *component* distributions in what follows.

2.1 Component distributions

2.1.1 The Kumaraswamy-Kw distribution

Kumaraswamy (1980) proposed and discussed a probability distribution for handling double-bounded random processes with varied hydrological applications. Let T be a random variable with the Kumaraswamy's distribution. The PDF and CDF are defined, respectively, as

$$f_T(t) = abt^{a-1}(1-t^a)^{b-1}, \quad 0 < t < 1,$$

and

$$F_T(t) = 1 - (1-t^a)^b, \quad 0 < t < 1, \quad (4)$$

where both $a > 0$ and $b > 0$ are shape parameters. A generalized form of this distribution is obtained when t is replaced by the CDF $G(t)$ of another random variable, to have what is referred to as the Kumaraswamy-G (*Kw-G*) distribution. The beta and Kumaraswamy distributions share similar properties. For example, the Kumaraswamy distribution, also referred to as the minimax distribution, is unimodal, uniantimodal, increasing, decreasing or constant depending on the values of its parameters. A more detailed description, background, genesis, and properties of Kumaraswamy distribution are outlined in Jones (2009), where the author highlighted

several advantages of the Kumaraswamy distribution over the beta distribution, namely; its simple normalizing constant, simple explicit formulas for the distribution and quantile functions, and simple random variate generation procedure. There have been many contributions to the theory and applications of the Kumaraswamy distribution in the literature. See for example the works by Cordeiro et al. (2010): Kumaraswamy Weibull; Cordeiro et al. (2012): Kumaraswamy Gumbel; Pascoa et al. (2011): Kumaraswamy generalized gamma; Saulo (2012): Kumaraswamy Birnbaum-Saunders; Akinsete et al. (2014): Kumaraswamy geometric, and a host of many others.

2.1.2 The Pareto and the transmuted Pareto (TP) distributions

The Pareto distribution was named after a sociologist Vilfredo Pareto. It is used in modelling the distribution of incomes and other financial variables, and in the description of social and other phenomena. Many distributions have been derived using the Pareto distribution. A few examples are the gamma-Pareto distribution by Alzaatreh et al. (2012), the beta exponentiated Pareto distribution by Zea et al. (2012), and the beta-Pareto distribution by Akinsete et al. (2008).

The PDF and CDF of the Pareto distribution are given respectively by

$$r(x) = \frac{k\theta^k}{x^{k+1}},$$

and

$$R(x) = 1 - (\theta/x)^k, \quad k > 0, \quad \theta > 0, \quad x \geq \theta.$$

Using the quadratic rank transmutation map, Merovci and Puka (2014) generalized the Pareto distribution to obtain what the authors called the transmuted Pareto distribution. They provided a comprehensive description of the mathematical properties of the distribution and its application in modeling real life data. Formally, a random variable X is said to have the transmuted Pareto distribution with scale parameter $\theta > 0$, shape parameter $\alpha > 0$, and the transmuted parameter λ ($|\lambda| \leq 1$), if its PDF is given by

$$g(x; \alpha, \theta, \lambda) = \frac{\alpha\theta^\alpha}{x^{\alpha+1}} \left[1 - \lambda + 2\lambda \left(\frac{\theta}{x}\right)^\alpha \right],$$

where θ is the (necessarily positive) minimum possible value of X . The corresponding CDF of the transmuted Pareto distribution is given by

$$G(x; \alpha, \theta, \lambda) = \left[1 - \left(\frac{\theta}{x}\right)^\alpha \right] \left[1 + \lambda \left(\frac{\theta}{x}\right)^\alpha \right]. \tag{5}$$

The distribution in Eq. (5) becomes the Pareto when $\lambda = 0$.

2.2 The KwTP distribution

By substituting Eq. (5) into Eq. (4), the cumulative distribution function of the five-parameter Kumaraswamy transmuted Pareto (KwTP) distribution is given by

$$\begin{aligned} F(x; \alpha, \theta, \lambda, a, b) &= 1 - [1 - G(x; \alpha, \theta, \lambda)]^a]^b \\ &= 1 - \left\{ 1 - \left[1 - \left(\frac{\theta}{x}\right)^\alpha \right]^a \left[1 + \lambda \left(\frac{\theta}{x}\right)^\alpha \right]^a \right\}^b, \end{aligned} \tag{6}$$

where $x > \theta$, $\alpha > 0$, $a > 0$, and $b > 0$ are shape parameters, $\theta > 0$ is the scale parameter, while $|\lambda| \leq 1$ is the transmuted parameter. The corresponding probability density function of the $KwTP$ distribution may be expressed as

$$\begin{aligned}
 f(x; \alpha, \theta, \lambda, a, b) &= abg(x)G(x)^{a-1}[1 - G(x)^a]^{b-1} \\
 &= \frac{ab\alpha\theta^\alpha}{x^{\alpha+1}} \left[1 - \lambda + 2\lambda \left(\frac{\theta}{x}\right)^\alpha \right] \left\{ \left[1 - \left(\frac{\theta}{x}\right)^\alpha \right] \left[1 + \lambda \left(\frac{\theta}{x}\right)^\alpha \right] \right\}^{a-1} \\
 &\quad \times \left\{ 1 - \left[1 - \left(\frac{\theta}{x}\right)^\alpha \right]^a \left[1 + \lambda \left(\frac{\theta}{x}\right)^\alpha \right]^a \right\}^{b-1}. \tag{7}
 \end{aligned}$$

Figure 1 shows the graphs of the PDF of $KwTP$ distribution for selected values of the parameters. The graphs indicate that $KwTP$ is unimodal, and may be used in modeling different phenomena exhibited by real world data. Figure 2 also shows the graphs of the CDF for various values of the parameters.

3 Special cases of $KwTP$

Some special submodels may be obtained from $KwTP$ by varying the values of the parameters. Here is a list of few these submodels.

- When $\lambda = 0$, $KwTP$ reduces to KwP distribution by Bourguignon et al. (2013) with PDF

$$f(x; \alpha, \theta, a, b) = \frac{ab\alpha\theta^\alpha}{x^{\alpha+1}} \times \left\{ \left[1 - \left(\frac{\theta}{x}\right)^\alpha \right] \right\}^{a-1} \left\{ 1 - \left[1 - \left(\frac{\theta}{x}\right)^\alpha \right]^a \right\}^{b-1}.$$

- When $a = b = 1$, $KwTP$ reduces to the transmuted Pareto distribution by Merovci and Puka (2014) with PDF

$$f(x; \alpha, \theta, \lambda) = \frac{\alpha\theta^\alpha}{x^{\alpha+1}} \left[1 - \lambda + 2\lambda \left(\frac{\theta}{x}\right)^\alpha \right].$$

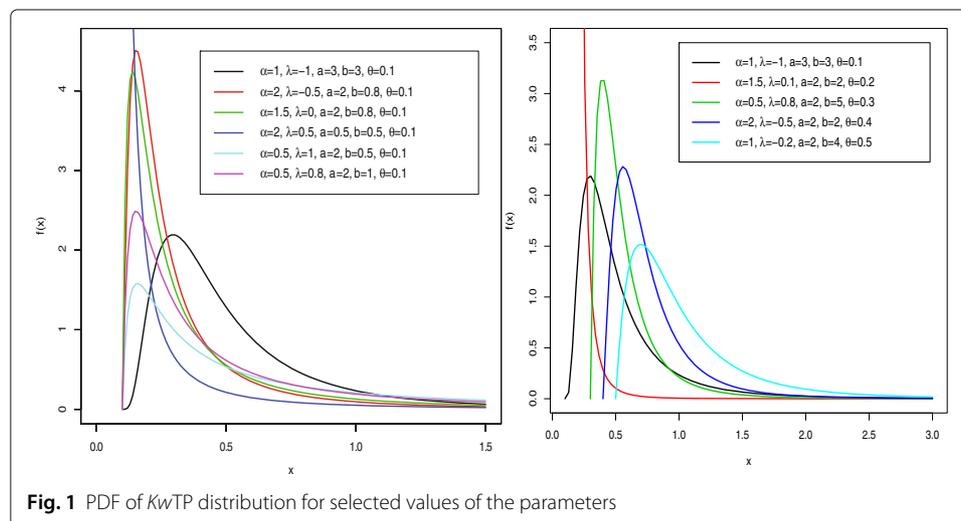


Fig. 1 PDF of $KwTP$ distribution for selected values of the parameters

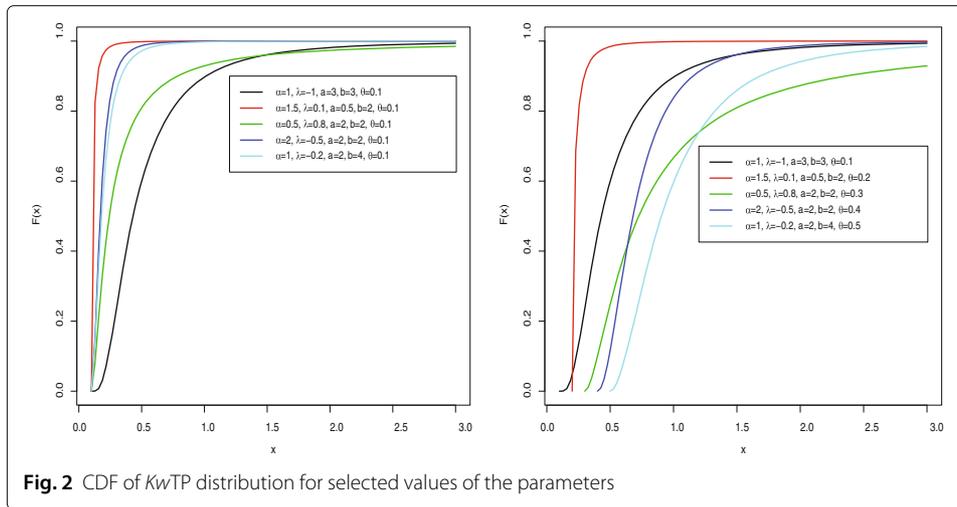


Fig. 2 CDF of *KwTP* distribution for selected values of the parameters

- By setting $b = 1$ and $\lambda = 0$, the *KwTP* reduces to the exponentiated Pareto distribution, defined by Nadarajah (2005), with PDF

$$f(x; \alpha, \theta, a) = \frac{a\alpha\theta^\alpha}{x^{\alpha+1}} \times \left\{ \left[1 - \left(\frac{\theta}{x} \right)^\alpha \right] \right\}^{a-1}.$$

- When $a = 1 = b$ and $\lambda = 0$, *KwTP* reduces to the classical Pareto distribution with PDF

$$f(x; \alpha, \theta) = \frac{\alpha\theta^\alpha}{x^{\alpha+1}}.$$

4 Mixture representation

In this section, we provide alternative and useful expressions for the PDF and the CDF of *KwTP* distribution. Let X be a random variable having the *KwTP* density (7).

For any $b > 0$ real non-integer and $|z| < 1$, consider the power series

$$(1 - z)^{b-1} = \sum_{k=0}^{\infty} \frac{(-1)^k \Gamma(b)}{k! \Gamma(b - k)} z^k. \tag{8}$$

As mentioned by Afify et al. (2016), applying the power series (8) to Eq. (7), we can write the PDF as a mixture of exponentiated-transmuted Pareto with the power parameter $(k + 1)a$ as

$$f(x) = \sum_{k=0}^{\infty} v_k g(x) G(x)^{(k+1)a-1}, \tag{9}$$

where

$$v_k = \frac{(-1)^k ab \Gamma(b)}{k! \Gamma(b - k)} = ab \binom{b-1}{k} (-1)^k,$$

and $g(x)$ and $G(x)$ are the PDF and CDF of the transmuted Pareto distribution respectively.

The PDF (9) can be expressed as a mixture of exp-G densities

$$f(x) = \sum_{k=0}^{\infty} w_k \gamma_k g(x) G(x)^{\gamma_k - 1}, \tag{10}$$

where $w_k = \frac{\gamma_k}{(k+1)a}$, and $\gamma_k = (k+1)a$. By integrating (10), we get a similar mixture representation for the CDF

$$F(x) = \sum_{k=0}^{\infty} w_k G(x)^{(k+1)a}, \tag{11}$$

where $G(x)$ is the CDF of the transmuted Pareto distribution with power parameter $(k+1)a$. Using the binomial expansions of the expressions

$$\left[1 - (1 - (\theta/x)^\alpha) (1 + \lambda(\theta/x)^\alpha)\right]^{b-1}, \quad (1 - (\theta/x)^\alpha)^{(k+1)a-1} \quad \text{and} \quad (1 + \lambda(\theta/x)^\alpha)^{(k+1)a-1},$$

the PDF of the KwTP distribution can be expressed as

$$\begin{aligned} f(x) &= \frac{ab\alpha\theta^\alpha}{x^{\alpha+1}} \left[1 - \lambda + 2\lambda \left(\frac{\theta}{x}\right)^\alpha\right] \times \left\{ \left[1 - \left(\frac{\theta}{x}\right)^\alpha\right] \left[1 + \lambda \left(\frac{\theta}{x}\right)^\alpha\right] \right\}^{a-1} \\ &\quad \times \left\{ 1 - \left[1 - \left(\frac{\theta}{x}\right)^\alpha\right]^a \left[1 + \lambda \left(\frac{\theta}{x}\right)^\alpha\right]^a \right\}^{b-1} \\ &= \sum_{k=0}^{\infty} \frac{ab\alpha\theta^\alpha}{x^{\alpha+1}} \binom{b-1}{k} (-1)^k \left[1 - \lambda + 2\lambda \left(\frac{\theta}{x}\right)^\alpha\right] \\ &\quad \times \left[1 - \left(\frac{\theta}{x}\right)^\alpha\right]^{(k+1)a-1} \left[1 + \lambda \left(\frac{\theta}{x}\right)^\alpha\right]^{(k+1)a-1} \\ &= \sum_{k=0}^{\infty} \frac{ab\alpha\theta^\alpha}{x^{\alpha+1}} \binom{b-1}{k} (-1)^k \left[1 - \lambda + 2\lambda \left(\frac{\theta}{x}\right)^\alpha\right] \\ &\quad \times \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \binom{(k+1)a-1}{i} \binom{(k+1)a-1}{j} (-1)^{i+j} \lambda^j \left(\frac{\theta}{x}\right)^{\alpha(i+j)}, \end{aligned}$$

where $\alpha > 0, a > 0, b > 0, \theta > 0, |\lambda| \leq 1$ and $x > \theta$. Let

$$w_{ij}^k = \binom{(k+1)a-1}{i} \binom{(k+1)a-1}{j} (-1)^{i+j} \lambda^j.$$

By setting $i + j = m$, and $w_{k,m}^* = \sum_{i,j:i+j=m} w_{ij}^k = \sum_{i=0}^m \sum_{j=0}^{m-i} w_{ij}^k$, we can further express the PDF of the KwTP in the form

$$f(x) = \frac{ab\alpha\theta^\alpha}{x^{\alpha+1}} \left[1 - \lambda + 2\lambda \left(\frac{\theta}{x}\right)^\alpha\right] \sum_{m=0}^{\infty} u_m \left(\frac{\theta}{x}\right)^{\alpha m}, \tag{12}$$

where

$$u_m = \sum_{k=0}^{\infty} w_{k,m}^* \binom{b-1}{k} (-1)^k. \tag{13}$$

In particular,

$$u_0 = \sum_{k=0}^{\infty} w_{k,0}^* \binom{b-1}{k} (-1)^k, \quad \text{where} \quad w_{k,0}^* = \sum_{i=0}^0 \sum_{j=0}^{0-i} w_{ij}^k.$$

That is, $m = 0 \implies i = 0, j = 0$. From $w_{ij}^k = \binom{(k+1)a-1}{i} \binom{(k+1)a-1}{j} (-1)^i \lambda^j$ and $i = 0, j = 0$, we have $w_{0,0} = 1$. Thus, $w_{k,0}^* = 1$, and hence, $u_0 = \sum_{k=0}^{\infty} \binom{b-1}{k} (-1)^k$.

Similarly, using (11) we can express $F(x)$ as

$$\begin{aligned} F(x) &= \sum_{k=0}^{\infty} w_k G(x)^{(k+1)a} \\ &= \sum_{k=0}^{\infty} w_k \left[1 - \left(\frac{\theta}{x}\right)^\alpha \right]^{(k+1)a} \times \left[1 + \lambda \left(\frac{\theta}{x}\right)^\alpha \right]^{(k+1)a} \\ &= \sum_{k=0}^{\infty} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} w_k (-1)^i \lambda^j \binom{(k+1)a}{i} \binom{(k+1)a}{j} \left(\frac{\theta}{x}\right)^{\alpha(i+j)}. \end{aligned}$$

Let $v_{ij}^k = \binom{(k+1)a}{i} \binom{(k+1)a}{j} (-1)^i \lambda^j$, where in particular $v_{00}^k = 1$.

Setting $i + j = m$, $v_{k,m}^* = \sum_{i,j:i+j=m} v_{ij}^k = \sum_{i=0}^m \sum_{j=0}^{m-i} v_{ij}^k$, $v = \left(\frac{\theta}{x}\right)^\alpha$, and

$$\rho_m = \sum_{k=0}^{\infty} v_{k,m}^* w_k, \tag{14}$$

we can express the CDF of the $KwTP$ as,

$$\begin{aligned} F(x) &= \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} v_{k,m}^* w_k \left(\frac{\theta}{x}\right)^{\alpha m} \\ &= \sum_{m=0}^{\infty} \left(\frac{\theta}{x}\right)^{\alpha m} \rho_m \\ &= \sum_{m=0}^{\infty} \rho_m v^m, \end{aligned} \tag{15}$$

where ρ_0 may be expressed as

$$\begin{aligned} \rho_0 &= \sum_{k=0}^{\infty} v_{k,0}^* w_k = \sum_{k=0}^{\infty} v_{00}^k w_k = \sum_{k=0}^{\infty} w_k \\ &= \sum_{k=0}^{\infty} \frac{v_k}{a(k+1)} = \sum_{k=0}^{\infty} (-1)^k \frac{ab \binom{b-1}{k}}{a(k+1)} \\ &= b \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} \binom{b-1}{k}. \end{aligned} \tag{16}$$

5 Quantile function

The quantile function of a distribution is the real solution of $F(x_q) = q$ for $0 \leq q \leq 1$. It is defined as the inverse function of the CDF of a random variable X . Using Eq. (6), the quantiles (x_q) of $KwTP$ distribution may be expressed in the form

$$x_q = F^{-1}(q) = \theta \left(\frac{\lambda - 1 + \sqrt{(1 + \lambda)^2 - 4\lambda \left(\left\{ 1 - (1 - q)^{\frac{1}{b}} \right\}^{\frac{1}{a}} \right)}}{2\lambda} \right)^{-\frac{1}{\alpha}}.$$

For specified values of parameters $\alpha, \theta, \lambda, a$ and b , and a set of random number $u \in (0, 1)$, we can generate a random variable X having the $KwTP$ distribution (6), and satisfying

$$X = Q_u = F^{-1}(u) = \theta \left(\frac{\lambda - 1 + \sqrt{(1 + \lambda)^2 - 4\lambda \left(\left\{ 1 - (1 - u)^{\frac{1}{b}} \right\}^{\frac{1}{a}} \right)}}{2\lambda} \right)^{-\frac{1}{\alpha}}. \tag{17}$$

In general, the quantile function enables one to find the relationship between one random variable and another random variable. For example, if Y is a random variable with CDF $G(y)$, we can show that $X \sim KwTP(x|\alpha, \theta, \lambda, a, b)$ if

$$X = \theta \left(\frac{2\lambda}{\lambda - 1 + \sqrt{(1 + \lambda)^2 - 4\lambda \left(\left\{ 1 - (G(y))^{\frac{1}{b}} \right\}^{\frac{1}{a}} \right)}} \right)^{\frac{1}{\alpha}}.$$

6 Moment functions

We define in this section, various moment expressions for the $KwTP$ distribution. These consist of the raw moments and their related functions, e.g., skewness and kurtosis, and the moment generating function. By definition, the r^{th} -moment of the $KwTP$ random variable is defined by

$$\begin{aligned} E(X^r) &= \mu'_r(\alpha, \theta, \lambda, a, b) = \int_{\theta}^{\infty} x^r f(x) dx \\ &= ab\alpha\theta^{\alpha} \left[(1 - \lambda) \sum_{m=0}^{\infty} u_m \theta^{\alpha m} \int_{\theta}^{\infty} x^{r-1-\alpha-\alpha m} dx + 2\lambda \sum_{m=0}^{\infty} u_m \theta^{\alpha+\alpha m} \right. \\ &\quad \left. \times \int_{\theta}^{\infty} x^{r-1-2\alpha-\alpha m} dx \right] \\ &= ab\alpha\theta^{\alpha} \left\{ (1 - \lambda) \sum_{m=0}^{\infty} u_m \theta^{\alpha m} \left[\frac{x^{r-\alpha-\alpha m}}{r - \alpha - \alpha m} \right] \Big|_{\theta}^{\infty} \right. \\ &\quad \left. + 2\lambda \sum_{m=0}^{\infty} u_m \theta^{\alpha+\alpha m} \left[\frac{x^{r-2\alpha-\alpha m}}{r - 2\alpha - \alpha m} \right] \Big|_{\theta}^{\infty} \right\} \\ &= (1 - \lambda) ab\alpha\theta^r \sum_{m=0}^{\infty} \frac{u_m}{\alpha m + \alpha - r} + 2\lambda ab\alpha\theta^r \sum_{m=0}^{\infty} \frac{u_m}{\alpha m + 2\alpha - r} \\ &= ab\alpha\theta^r \sum_{m=0}^{\infty} u_m \left[\frac{1 - \lambda}{\alpha + \alpha m - r} + \frac{2\lambda}{2\alpha + \alpha m - r} \right], \quad r < \alpha. \tag{18} \end{aligned}$$

In particular if $\lambda = 0, a = b = 1$, we have the Pareto distribution. In this case, we must have $m = 0, u_0 = 1$, so that the moment expression becomes

$$E(X^r) = \mu'_r(\alpha, \theta, 0, 1, 1) = \frac{\alpha\theta^r}{\alpha - r}, \quad r < \alpha.$$

It is interesting to note that the mean of X is

$$E(X) = \mu'_1(\alpha, \theta, 0, 1, 1) = \frac{\alpha\theta}{\alpha - 1}, \quad \alpha > 1.$$

By similar procedure, the moment generating function is defined as follows:

$$\begin{aligned}
 M_X(t) &= \int_{\theta}^{\infty} e^{tx} f(x) dx \\
 &= ab\alpha\theta^{\alpha} \left[(1-\lambda) \sum_{m=0}^{\infty} u_m \theta^{\alpha m} \int_{\theta}^{\infty} e^{tx} x^{-\alpha-\alpha m-1} dx + 2\lambda \sum_{m=0}^{\infty} u_m \theta^{\alpha+\alpha m} \right. \\
 &\quad \left. \times \int_{\theta}^{\infty} e^{tx} x^{-\alpha m-2\alpha-1} dx \right] \\
 &= ab\alpha \left[(1-\lambda) \sum_{m=0}^{\infty} u_m (-t\theta)^{\alpha+\alpha m} \Gamma(-\alpha-\alpha m, -t\theta) \right. \\
 &\quad \left. + 2\lambda \sum_{m=0}^{\infty} u_m (-t\theta)^{2\alpha+\alpha m} \Gamma(-2\alpha-\alpha m, -t\theta) \right], \\
 &= ab\alpha \sum_{m=0}^{\infty} u_m (-t\theta)^{\alpha+\alpha m} \left[(1-\lambda) \Gamma(-\alpha-\alpha m, -t\theta) \right. \\
 &\quad \left. + 2\lambda (-t\theta)^{\alpha} \Gamma(-2\alpha-\alpha m, -t\theta) \right], \quad t < 0,
 \end{aligned}$$

where $\Gamma(.,.)$ denotes the upper incomplete gamma function. That is, $\Gamma(s, \theta) = \int_{\theta}^{\infty} e^{-t} t^{s-1} dt$.

The variance, skewness and kurtosis of the *KwTP* distribution can be calculated from (18) using the relations given below.

$$\begin{aligned}
 \text{Variance}(X) &= E(X^2) - [E(X)]^2, \quad \alpha > 2 \\
 \text{Skewness}(X) &= \frac{E(X^3) - 3E(X)E(X^2) + 2E^3(X)}{\text{Var}^{3/2}(X)}, \quad \alpha > 3 \\
 \text{Kurtosis}(X) &= \frac{E(X^4) - 4E(X)E(X^3) + 6E(X^2)E^2(X) - 3E^4(X)}{\text{Var}^2(X)}, \quad \alpha > 4.
 \end{aligned}$$

We provide in Table 1, the numerical measures of the median, mean, variance, skewness and kurtosis of the *KwTP* distribution for selected values of the parameters, to illustrate the effects of the parameters on these measures. We had arbitrarily fixed the value of parameter θ at 0.5.

The values of the statistics; mean, median, variance, skewness and kurtosis increase or decrease according to the changes in parameter values. For example, these statistics decrease as α or b increases, while other parameters are held constant. We also observe that each statistic in the set consisting of the mean, median and variance increases by increasing the value of parameter a , while each statistic in the set decreases as λ increases. On the other hand, the skewness and kurtosis both decrease for increasing value of a , while the values in this set both increase, and then decrease, for increasing value of λ . The effects of parameters a and b on the mean, variance, skewness and kurtosis for given values of λ, θ and α are displayed in Figs. 3 and 4 respectively.

In addition to the traditional measures of skewness and kurtosis, we also obtain the Bowley's measure of skewness, introduced by Bowley (1920), and the Moors' measure of kurtosis, introduced by Moors (1988). These measures are quartile alternatives to the traditional skewness and kurtosis, and are more robust estimation of these measures, see Kenney and Keeping (1962).

Table 1 Median, mean, variance, skewness, kurtosis for selected values of the parameters

α	λ	a	b	Median	Mean	Variance	Skewness	Kurtosis
5	0.2	1	1	0.5636	0.6111	0.0224	4.9196	81.8250
6	0.2	1	1	0.5525	0.5891	0.0130	4.0458	42.9488
7	0.2	1	1	0.5446	0.5744	0.0084	3.5980	30.9806
8	0.2	1	1	0.5388	0.5638	0.0059	3.3243	25.2955
10	0.2	1	1	0.5309	0.5497	0.0034	3.0056	19.8668
5	-1	1	1	0.6392	0.6944	0.0386	4.1113	60.5271
5	-0.5	1	1	0.6061	0.6997	0.0335	4.2366	63.3030
5	0.1	1	1	0.5688	0.6181	0.0243	4.7743	77.4267
5	0.5	1	1	0.5505	0.5903	0.0162	5.4825	101.8249
5	0.8	1	1	0.5408	0.5694	0.0091	6.0430	135.7943
5	1	1	1	0.5359	0.5556	0.0039	2.8111	17.8286
5	0.5	1	1	0.5505	0.5903	0.0162	5.4825	101.8249
5	0.5	2	1	0.5964	0.6422	0.0249	4.7467	78.6519
5	0.5	3	1	0.6304	0.6802	0.0313	4.4534	70.3107
5	0.5	4	1	0.6578	0.7106	0.0366	4.2881	65.8432
5	0.5	5	1	0.6810	0.7362	0.0411	4.1793	63.0006
5	0.2	1	1	0.5636	0.6111	0.0224	4.9196	81.8250
5	0.2	1	2	0.5303	0.5480	0.0031	3.0122	20.2613
5	0.2	1	3	0.5199	0.5305	0.0011	2.6351	15.2360
5	0.2	1	4	0.5150	0.5224	0.0006	2.4635	13.3142
5	0.2	1	5	0.5118	0.5176	0.0003	2.3646	12.2905

- Bowley skewness (B_{sk}): By definition, the Bowley’s measure of skewness is expressed as

$$B_{sk} = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1} = \frac{Q_{0.75} - 2Q_{0.5} + Q_{0.25}}{Q_{0.75} - Q_{0.25}}$$

- Moors kurtosis M_{kur} : This is defined as

$$M_{kur} = \frac{(E_7 - E_5) + (E_3 - E_1)}{E_6 - E_2} = \frac{Q_{0.875} - Q_{0.625} + Q_{0.375} - Q_{0.125}}{Q_{0.75} - Q_{0.25}}$$

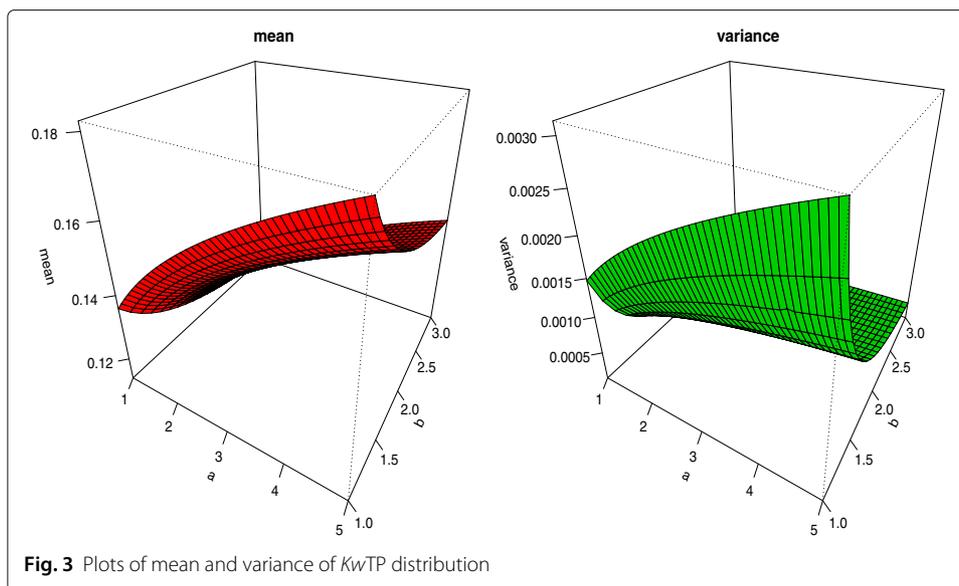
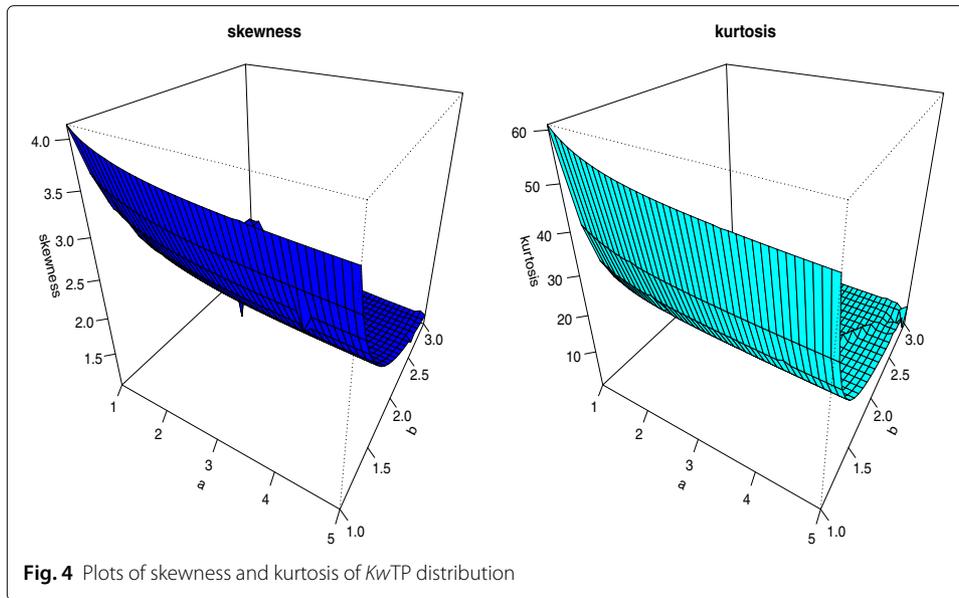


Fig. 3 Plots of mean and variance of KwTP distribution

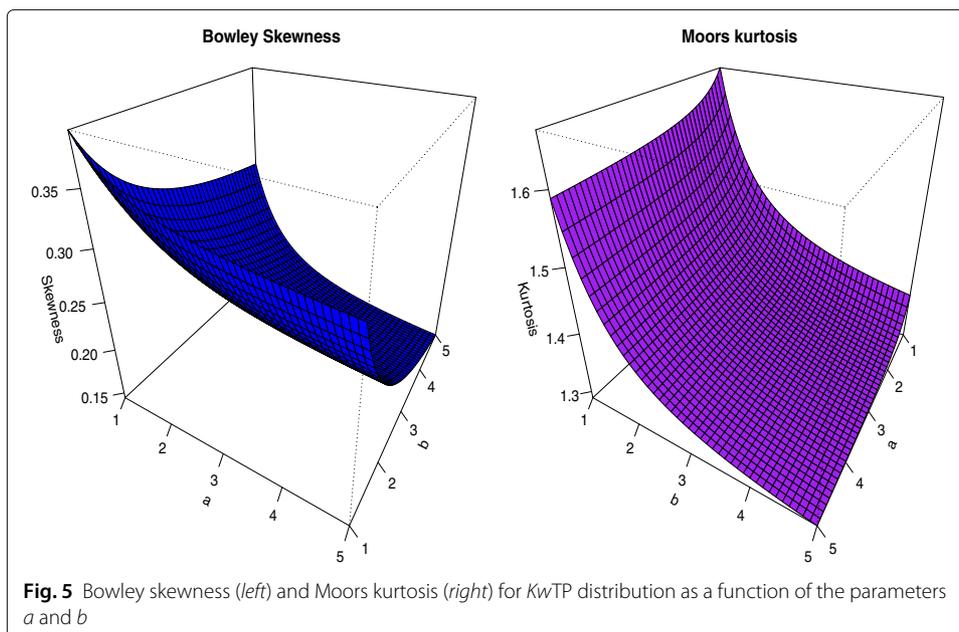


where Q_i is the i^{th} quartile for $i = 1, 2, 3$, and $E_i = F^{-1}(i/8)$, $i = 1, 2, \dots, 7$ represents the i^{th} octile.

The graphs of the Bowley skewness and Moors kurtosis for different values of a and b , with $\lambda = 0.2, \alpha = 2$ and $\theta = 0.1$, are shown in Fig. 5. The graphs show that both Bowley skewness and Moors kurtosis depend on the choice of the parameters.

7 Mean deviation

The deviation from the mean, in the case of symmetric distributions, or deviation from the median, in the case of skewed distributions, can be used as a measure of



spread in a population. Given that the random variable X has the $KwTP$ distribution. Let $\mu = E(X)$ be the mean and M be the median of X . The mean deviation about the mean $D(\mu)$ and the mean deviation about the median $D(M)$ can be expressed respectively as

$$\begin{aligned}
 D(\mu) &= \int_{\theta}^{\infty} |x - \mu|f(x)dx = \int_{\theta}^{\mu} (\mu - x)f(x)dx + \int_{\mu}^{\infty} (x - \mu)f(x)dx \\
 &= 2\mu F(\mu) - 2J(\mu),
 \end{aligned}
 \tag{19}$$

and

$$\begin{aligned}
 D(M) &= \int_{\theta}^{\infty} |x - M|f(x)dx = \int_{\theta}^M (M - x)f(x)dx + \int_M^{\infty} (x - M)f(x)dx \\
 &= \mu - 2J(M),
 \end{aligned}
 \tag{20}$$

where $F(\cdot)$ is the CDF of the $KwTP$ distribution, and $J(t) = \int_{\theta}^t xf(x)dx$.

We compute $J(t)$ as follows:

$$\begin{aligned}
 J(t) &= \int_{\theta}^t xf(x)dx \\
 &= ab\alpha\theta^{\alpha} \left[(1 - \lambda) \sum_{m=0}^{\infty} u_m\theta^{\alpha m} \int_{\theta}^t x^{-\alpha-\alpha m}dx + 2\lambda \sum_{m=0}^{\infty} u_m\theta^{\alpha+\alpha m} \right. \\
 &\quad \left. \times \int_{\theta}^t x^{-2\alpha-\alpha m}dx \right] \\
 &= ab\alpha \left[(1 - \lambda) \sum_{m=0}^{\infty} u_m\theta^{\alpha+\alpha m} \left(\frac{t^{-\alpha-\alpha m+1}}{1 - \alpha - \alpha m} + \frac{\theta^{-\alpha-\alpha m+1}}{\alpha + \alpha m - 1} \right) \right. \\
 &\quad \left. + 2\lambda \sum_{m=0}^{\infty} u_m\theta^{2\alpha+\alpha m} \left(\frac{t^{-2\alpha-\alpha m+1}}{1 - 2\alpha - \alpha m} + \frac{\theta^{-2\alpha-\alpha m+1}}{2\alpha + \alpha m - 1} \right) \right] \\
 &= ab\alpha \sum_{m=0}^{\infty} u_m \left[\frac{(1 - \lambda)t}{1 - \alpha - \alpha m} \left(\frac{\theta}{t} \right)^{\alpha} + \frac{(1 - \lambda)\theta}{\alpha + \alpha m - 1} + \frac{2\lambda t}{1 - 2\alpha - \alpha m} \left(\frac{\theta}{t} \right)^{2\alpha+\alpha m} \right. \\
 &\quad \left. + \frac{2\lambda\theta}{2\alpha + \alpha m - 1} \right].
 \end{aligned}
 \tag{21}$$

Using Eq. (21), we can write appropriate expressions for $J(\mu)$ and $J(M)$. Combining these with the expression for $F(\mu)$, it is easy to obtain the expressions for $D(\mu)$ and $D(M)$ in Eqs. (19) and (20) respectively.

8 Reliability

According to Ashour and Eltehiwy (2013), the reliability function, also called the survival function, is the characteristic of an explanatory variable that maps a set of events, usually associated with mortality or failure of some system onto time. In other words, the

reliability function $R(t)$ is the probability of an item not failing prior to some time t , and is defined by $R(t) = 1 - F(t)$. The reliability function of the $KwTP$ distribution is

$$R(t) = 1 - F(t) = \left\{ 1 - \left[1 - \left(\frac{\theta}{t} \right)^\alpha \right]^a \left[1 + \lambda \left(\frac{\theta}{t} \right)^\alpha \right]^b \right\}.$$

The other characteristic of interest of a random variable is the hazard rate function $h(t)$ which is also known as instantaneous failure rate. The hazard function of a random variable X with PDF $f(x)$ and associated CDF $F(x)$ is defined as

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

Substituting the PDF and CDF for the $KwTP$ distribution into the above expression, we have

$$h(t) = \frac{f(t)}{1 - F(t)} = \frac{ab\alpha\theta^\alpha \left[1 - \lambda + 2\lambda \left(\frac{\theta}{t} \right)^\alpha \right] \left[1 - \left(\frac{\theta}{t} \right)^\alpha \right]^{a-1} \left[1 + \lambda \left(\frac{\theta}{t} \right)^\alpha \right]^{a-1}}{t^{\alpha+1} \left\{ 1 - \left[1 - \left(\frac{\theta}{t} \right)^\alpha \right]^a \left[1 + \lambda \left(\frac{\theta}{t} \right)^\alpha \right]^a \right\}}.$$

The flexibility of $KwTP$ distribution to model reliability data is illustrated by varying shapes of the hazard rate function in Fig. 6.

The following lemma shows the limiting behavior of the hazard function:

Lemma 1 *If $h(t)$ is the hazard function of the Kumaraswamy transmuted Pareto distribution, then*

$$\lim_{t \rightarrow \theta} h(t) = \begin{cases} 0, & a > 1 \\ \frac{ab\alpha(1+\lambda)}{\theta}, & a = 1 \\ \infty, & a < 1, \end{cases}$$

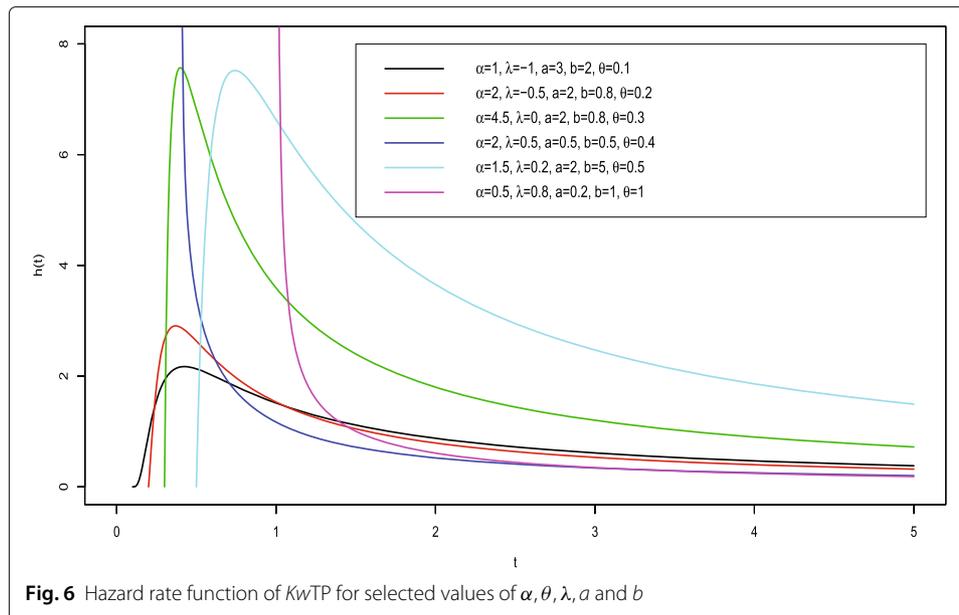


Fig. 6 Hazard rate function of $KwTP$ for selected values of $\alpha, \theta, \lambda, a$ and b

and

$$\lim_{t \rightarrow \infty} h(t) = 0.$$

Proof First note that

$$\lim_{t \rightarrow \theta} G(t) = \lim_{t \rightarrow \theta} \left[1 - \left(\frac{\theta}{t} \right)^\alpha \right] \left[1 + \lambda \left(\frac{\theta}{t} \right)^\alpha \right] = 0,$$

and

$$\lim_{t \rightarrow \theta} g(t) = \lim_{t \rightarrow \theta} \frac{\alpha \theta^\alpha}{t^{\alpha+1}} \left[1 - \lambda + 2\lambda \left(\frac{\theta}{t} \right)^\alpha \right] = \frac{\alpha(1 + \lambda)}{\theta}.$$

Then we have

$$\begin{aligned} \lim_{t \rightarrow \theta} h(t) &= \lim_{t \rightarrow \theta} \frac{f(t)}{1 - F(t)} = \lim_{t \rightarrow \theta} \frac{abg(t)G(t)^{a-1}}{1 - G(t)^a} \\ &= \begin{cases} 0, & \text{if } a > 1 \\ \frac{ab\alpha(1+\lambda)}{\theta}, & \text{if } a = 1 \\ \infty, & \text{if } a < 1. \end{cases} \end{aligned}$$

For the second part, we already know that $\lim_{t \rightarrow \infty} G(t) = 1$, $\lim_{t \rightarrow \infty} g(t) = 0$, $\lim_{t \rightarrow \infty} F(t) = 1$, and $\lim_{t \rightarrow \infty} f(t) = 0$. Then using the L'Hospital's rule, we have

$$\begin{aligned} \lim_{t \rightarrow \infty} h(t) &= \lim_{t \rightarrow \infty} \frac{abg(t)G(t)^{a-1}}{1 - G(t)^a} \\ &= \lim_{t \rightarrow \infty} \frac{abg(t)}{G(t)^{1-a} - G(t)} \\ &= \lim_{t \rightarrow \infty} \frac{abg'(t)}{(1 - a)G(t)^{-a}g(t) - g(t)} \\ &= \lim_{t \rightarrow \infty} \frac{ab}{[(1 - a)G(t)^{-a} - 1]} \times \lim_{t \rightarrow \infty} \frac{g'(t)}{g(t)} \\ &= -b \lim_{t \rightarrow \infty} \frac{g'(t)}{g(t)} = 0. \end{aligned}$$

□

9 Order statistics

Let X_1, X_2, \dots, X_n be a simple random sample from $KwTP(x; \alpha, \theta, \lambda, a, b)$ with CDF (6) and PDF (7). Also let $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ denote the order statistics from this sample. The PDF $f_{(i:n)}(x)$ of the i -th order statistics is given by

$$f_{(i:n)}(x) = \frac{1}{B(i, n - i + 1)} [F(x)]^{i-1} [1 - F(x)]^{n-i} f(x), \tag{22}$$

and the CDF is given by

$$F_{(i:n)}(x) = \sum_{k=i}^n \binom{n}{k} [F(x)]^k [1 - F(x)]^{n-k} = \int_0^{F(x)} \frac{1}{B(i, n - i + 1)} t^{i-1} (1 - t)^{n-i} dt.$$

Using Eqs. (12), (13), (14), (15) and binomial expansions, Eq. (22) becomes

$$\begin{aligned}
 f_{i:n}(x) &= \frac{1}{B(i, n-i+1)} f(x) \sum_{s=0}^{n-i} (-1)^s \binom{n-i}{s} F(x)^{i+s-1} \\
 &= \frac{1}{B(i, n-i+1)} \left(\frac{ab\alpha\theta^\alpha}{x^{\alpha+1}} \left[1 - \lambda + 2\lambda \left(\frac{\theta}{x}\right)^\alpha \right] \sum_{l=0}^{\infty} u_l \left(\frac{\theta}{x}\right)^{\alpha l} \right) \\
 &\quad \times \sum_{s=0}^{n-i} (-1)^s \binom{n-i}{s} \times \left(\sum_{m=0}^{\infty} \rho_m v^m \right)^{i+s-1}. \tag{23}
 \end{aligned}$$

From Gradshteyn and Ryzhik (2000, sec. 0.314), for any positive integer r ,

$$\left(\sum_{m=0}^{\infty} \rho_m v^m \right)^r = \sum_{m=0}^{\infty} d_{r,m} v^m,$$

where the coefficients $d_{r,m}$, for $m = 1, 2, 3, \dots$, can be determined from the recurrence equation

$$d_{r,m} = (m\rho_0)^{-1} \sum_{q=1}^m \{q(r+1) - m\} \rho_q d_{r,m-q}$$

and $d_{r,0} = (\rho_0^*)^r$.

We can obtain $d_{r,m}$ from $d_{r,0}, \dots, d_{r,m-1}$ and, therefore, from $\rho_0, \rho_1, \dots, \rho_m$.

Hence (23) becomes

$$\begin{aligned}
 f_{i:n}(x) &= \frac{ab\alpha\theta^\alpha (1 - \lambda + 2\lambda v)}{x^{\alpha+1} B(i, n-i+1)} \sum_{l=0}^{\infty} u_l v^l \\
 &\quad \times \sum_{s=0}^{n-i} (-1)^s \binom{n-i}{s} \times \left(\sum_{m=0}^{\infty} d_{i+s-1,m} v^m \right) \\
 &= \frac{ab\alpha\theta^\alpha (1 - \lambda + 2\lambda v)}{x^{\alpha+1} B(i, n-i+1)} \sum_{l=0}^{\infty} u_l v^l \sum_{m=0}^{\infty} c_i(m) v^m \\
 &= \frac{ab\alpha\theta^\alpha (1 - \lambda + 2\lambda v)}{x^{\alpha+1} B(i, n-i+1)} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} u_l c_i(m) v^{l+m},
 \end{aligned}$$

where

$$v = \left(\frac{\theta}{x}\right)^\alpha, \quad c_i(m) = \sum_{s=0}^{n-i} (-1)^s \binom{n-i}{s} d_{i+s-1,m},$$

$$d_{i+s-1,m} = (m\rho_0)^{-1} \sum_{q=1}^m [q(i+s) - m] \rho_q d_{i+s-1,m-q}$$

and $d_{i+s-1,0} = (\rho_0)^{i+s-1}$, where ρ_0 is defined in Eq. (16).

10 Parameter estimation

In this section we estimate the parameters of the Kumaraswamy transmuted Pareto distribution by the method of maximum likelihood estimation.

Let X_1, X_2, \dots, X_n be a random sample from the Kumaraswamy transmuted Pareto distribution with observed values x_1, x_2, \dots, x_n , and $\boldsymbol{\gamma} = (\alpha, \theta, \lambda, a, b)^T$ be the vector of the model parameters. The log-likelihood function for $\boldsymbol{\gamma}$ can be written as

$$\begin{aligned}
 l(\boldsymbol{\gamma}) &= \log L(x; \alpha, \theta, \lambda, a, b) \\
 &= n \log a + n \log b + n \log \alpha + n\alpha \log \theta - (\alpha + 1) \sum_{i=1}^n \log(x_i) \\
 &\quad + \sum_{i=1}^n \log \left(1 - \lambda + 2\lambda \left(\frac{\theta}{x_i} \right)^\alpha \right) \\
 &\quad + (a - 1) \left[\sum_{i=1}^n \log \left(1 - \left(\frac{\theta}{x_i} \right)^\alpha \right) + \sum_{i=1}^n \log \left(1 + \lambda \left(\frac{\theta}{x_i} \right)^\alpha \right) \right] \\
 &\quad + (b - 1) \sum_{i=1}^n \log \left\{ 1 - \left[1 - \left(\frac{\theta}{x_i} \right)^\alpha \right]^a \left[1 + \lambda \left(\frac{\theta}{x_i} \right)^\alpha \right]^a \right\}. \tag{24}
 \end{aligned}$$

Since $x \in (\theta, \infty)$, the maximum likelihood estimator of θ is the first order statistic $X_{(1)}$. Following the normal routine of parameter estimation for the maximum likelihood estimation of α, λ, a and b , we differentiate Eq. (24) with respect to α, λ, a and b to obtain the score vector $\left(\frac{\partial l}{\partial \alpha}, \frac{\partial l}{\partial \lambda}, \frac{\partial l}{\partial a}, \frac{\partial l}{\partial b} \right)^T$. The elements of score vectors are

$$\begin{aligned}
 \frac{\partial l}{\partial \alpha} &= \frac{n}{\alpha} + n \log \theta - \sum_{i=1}^n \log(x_i) + \sum_{i=1}^n \frac{2\lambda \left(\frac{\theta}{x_i} \right)^\alpha \cdot \ln \left(\frac{\theta}{x_i} \right)}{\left(1 - \lambda + 2\lambda \left(\frac{\theta}{x_i} \right)^\alpha \right)} \\
 &\quad - (a - 1) \left[\sum_{i=1}^n \frac{\left(\frac{\theta}{x_i} \right)^\alpha \cdot \ln \left(\frac{\theta}{x_i} \right)}{\left(1 - \left(\frac{\theta}{x_i} \right)^\alpha \right)} - \sum_{i=1}^n \frac{\lambda \left(\frac{\theta}{x_i} \right)^\alpha \ln \left(\frac{\theta}{x_i} \right)}{\left(1 + \lambda \left(\frac{\theta}{x_i} \right)^\alpha \right)} \right] \tag{25}
 \end{aligned}$$

$$\begin{aligned}
 &\quad - a(b - 1) \sum_{i=1}^n \frac{G^{a-1}(x_i; \alpha, \theta, \lambda) \left(\frac{\theta}{x_i} \right)^\alpha \left[1 - \lambda + 2\lambda \left(\frac{\theta}{x_i} \right)^\alpha \right] \ln \left(\frac{\theta}{x_i} \right)}{1 - G^a(x_i; \alpha, \theta, \lambda)}, \\
 \frac{\partial l}{\partial \lambda} &= \sum_{i=1}^n \frac{2 \left(\frac{\theta}{x_i} \right)^\alpha - 1}{1 - \lambda + 2\lambda \left(\frac{\theta}{x_i} \right)^\alpha} + (a - 1) \sum_{i=1}^n \frac{\left(\frac{\theta}{x_i} \right)^\alpha}{1 + \lambda \left(\frac{\theta}{x_i} \right)^\alpha} \\
 &\quad - a(b - 1) \sum_{i=1}^n \frac{\left(\frac{\theta}{x_i} \right)^\alpha \left[1 - \left(\frac{\theta}{x_i} \right)^\alpha \right]^a \left[1 + \lambda \left(\frac{\theta}{x_i} \right)^\alpha \right]^{a-1}}{1 - G^a(x_i; \alpha, \theta, \lambda)}, \tag{26}
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial l}{\partial a} &= \frac{n}{a} + \left[\sum_{i=1}^n \log \left(1 - \left(\frac{\theta}{x_i} \right)^\alpha \right) + \sum_{i=1}^n \log \left(1 + \lambda \left(\frac{\theta}{x_i} \right)^\alpha \right) \right] \\
 &\quad - (b - 1) \sum_{i=1}^n \frac{G^a(x_i; \alpha, \theta, \lambda) \times \log\{G^a(x_i; \alpha, \theta, \lambda)\}}{1 - G^a(x_i; \alpha, \theta, \lambda)}, \tag{27}
 \end{aligned}$$

and

$$\frac{\partial l}{\partial b} = \frac{n}{b} + \sum_{i=1}^n \log \{ 1 - G^a(x_i; \alpha, \theta, \lambda) \}, \tag{28}$$

where $G(x_i; \alpha, \theta, \lambda)$ is the PDF of the transmuted Pareto distribution.

The maximum likelihood estimates $\hat{\alpha}, \hat{\lambda}, \hat{a}, \hat{b}$ of the unknown parameters α, λ, a, b respectively, can be obtained by setting Eqs. (25) - (28) equal zero and solving for the parameters. We can use numerical methods such as the quasi-Newton algorithm to numerically optimize the log-likelihood function given in Eq. (24), to get the maximum likelihood estimates of the parameters α, λ, a, b . To compute the standard error and the asymptotic confidence interval, we use the usual large sample approximation in which the maximum likelihood estimators for $\boldsymbol{\gamma}$ can be treated as being approximately normal. This will require the computations of the second order derivatives of Eq. (24) with respect to the vector of parameters.

This procedure will result in

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\lambda} \\ \hat{a} \\ \hat{b} \end{pmatrix} \sim Normal \left[\begin{pmatrix} \alpha \\ \lambda \\ a \\ b \end{pmatrix}, \begin{pmatrix} \hat{V}_{\alpha\alpha} & \hat{V}_{\alpha\lambda} & \hat{V}_{\alpha a} & \hat{V}_{\alpha b} \\ \hat{V}_{\lambda\alpha} & \hat{V}_{\lambda\lambda} & \hat{V}_{\lambda a} & \hat{V}_{\lambda b} \\ \hat{V}_{a\alpha} & \hat{V}_{a\lambda} & \hat{V}_{aa} & \hat{V}_{ab} \\ \hat{V}_{b\alpha} & \hat{V}_{b\lambda} & \hat{V}_{ba} & \hat{V}_{bb} \end{pmatrix} \right]$$

with $\hat{V}_{u,v} = V_{(u,v)}|_{(u,v)=(\hat{u},\hat{v})}$, and the asymptotic variance-covariance matrix of the MLEs is,

$$\begin{pmatrix} \hat{V}_{\alpha\alpha} & \hat{V}_{\alpha\lambda} & \hat{V}_{\alpha a} & \hat{V}_{\alpha b} \\ \hat{V}_{\lambda\alpha} & \hat{V}_{\lambda\lambda} & \hat{V}_{\lambda a} & \hat{V}_{\lambda b} \\ \hat{V}_{a\alpha} & \hat{V}_{a\lambda} & \hat{V}_{aa} & \hat{V}_{ab} \\ \hat{V}_{b\alpha} & \hat{V}_{b\lambda} & \hat{V}_{ba} & \hat{V}_{bb} \end{pmatrix} = -E \begin{pmatrix} V_{\alpha\alpha} & V_{\alpha\lambda} & V_{\alpha a} & V_{\alpha b} \\ V_{\lambda\alpha} & V_{\lambda\lambda} & V_{\lambda a} & V_{\lambda b} \\ V_{a\alpha} & V_{a\lambda} & V_{aa} & V_{ab} \\ V_{b\alpha} & V_{b\lambda} & V_{ba} & V_{bb} \end{pmatrix}^{-1}$$

where entries are obtained from

$$V_{\alpha\alpha} = \frac{\partial^2 l}{\partial \alpha^2}, \quad V_{\alpha\lambda} = \frac{\partial^2 l}{\partial \alpha \partial \lambda}, \quad V_{\alpha a} = \frac{\partial^2 l}{\partial \alpha \partial a}, \quad V_{\alpha b} = \frac{\partial^2 l}{\partial \alpha \partial b}$$

$$V_{\lambda\alpha} = \frac{\partial^2 l}{\partial \lambda \partial \alpha}, \quad V_{\lambda\lambda} = \frac{\partial^2 l}{\partial \lambda^2}, \quad V_{\lambda a} = \frac{\partial^2 l}{\partial \lambda \partial a}, \quad V_{\lambda b} = \frac{\partial^2 l}{\partial \lambda \partial b}$$

$$V_{a\alpha} = \frac{\partial^2 l}{\partial a \partial \alpha}, \quad V_{a\lambda} = \frac{\partial^2 l}{\partial a \partial \lambda}, \quad V_{aa} = \frac{\partial^2 l}{\partial a^2}, \quad V_{ab} = \frac{\partial^2 l}{\partial a \partial b}$$

$$V_{b\alpha} = \frac{\partial^2 l}{\partial b \partial \alpha}, \quad V_{b\lambda} = \frac{\partial^2 l}{\partial b \partial \lambda}, \quad V_{ba} = \frac{\partial^2 l}{\partial b \partial a}, \quad V_{bb} = \frac{\partial^2 l}{\partial b^2}$$

and l is the log-likelihood function given in (24). Approximate $100(1 - \phi)\%$ two sided confidence intervals for α, λ, a and b are, respectively, given by

$$\hat{\alpha} \pm z_{\frac{\phi}{2}} \sqrt{\hat{V}_{\alpha\alpha}}, \quad \hat{\lambda} \pm z_{\frac{\phi}{2}} \sqrt{\hat{V}_{\lambda\lambda}}, \quad \hat{a} \pm z_{\frac{\phi}{2}} \sqrt{\hat{V}_{aa}}, \quad \text{and} \quad \hat{b} \pm z_{\frac{\phi}{2}} \sqrt{\hat{V}_{bb}},$$

where z_{ϕ} is the upper ϕ -th percentile of the standard normal distribution.

Table 2 Descriptive statistics for the Wheaton river data

Min.	Q1	Median	Mean	Q3	Max.	Skewness	Kurtosis
0.100	2.125	9.500	12.200	20.12	64.00	1.4725	5.8895

11 Applications of KwTP

In this section we apply the KwTP to two data sets. These are exceedances of flood peaks (in m^3/s) of the Wheaton River near Carcross in Yukon Territory, Canada, and the Norwegian fire insurance claims data.

11.1 The exceedances of flood peaks (in m^3/s) of the Wheaton River near Carcross in Yukon Territory, Canada

The data are the exceedances of flood peaks (in m^3/s) of the Wheaton River near Carcross in Yukon Territory, Canada. The data consists of 72 exceedances for the years 1958-1984. These data have been analyzed by many authors including Choulakian and Stephens (2001), Nadarajah (2005), Akinsete et al. (2008), Bourguignon et al. (2013), Merovci and Puka (2014), and Chhetri et al. (2017), among others. A summary of the descriptive statistics of the data set is given in Table 2. The analysis of these data with the KwTP is compared with the results from few commonly used distributions in the literature. All required computations are carried out using the AdequacyModel script of R-package (Marinho et al. (2016)). In particular we compare the KwTP with the Kumaraswamy Pareto distribution (KwP), the transmuted Pareto distribution (TP), exponentiated Pareto distribution (EP), beta transmuted Pareto distribution (BTP), beta Pareto distribution (BP), and the Pareto distribution (P). Using the method of maximum likelihood estimation, we estimate the distribution parameters. The goodness of fit measures, including the log-likelihood function ($-\ell$), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Consistent Akaike Information Criterion (CAIC), and Hannan-Quinn Information Criteria (HQIC) are obtained for the fitted models.

Table 3 lists the values the MLEs and their standard errors, whereas the values of $-\hat{\ell}$, AIC, CAIC, HQIC, BIC and Kolmogorov (K-S) statistic are given in Table 4. From Table 4, it is noted that the KwTP distribution has the lowest values for

Table 3 Estimated parameters and their standard errors for the Wheaton river data

Model	\hat{a}	\hat{b}	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\theta}$
KwTP	4.2684 (1.5669)	17.0139 (12.6727)	-0.3687 (0.5308)	0.2003 (0.0609)	0.1 -
KwP	2.8553 (0.3371)	85.8468 (60.4213)	-	0.0528 (0.0185)	0.1 -
BTP	3.9118 (1.8159)	17.3874 (11.7365)	-0.8518 (0.2588)	0.1159 (0.0509)	0.1 -
BP	3.1473 (0.4993)	85.7508 (0.0001)	-	0.0088 (0.0015)	0.1 -
TP	1 -	1 -	-0.952 (0.089)	0.3490 (0.072)	0.1 -
EP	2.8797 (0.4911)	1 -	-	0.4241 (0.0463)	0.1 -
P	1 -	1 -	-	0.2438 (0.0287)	0.1 -

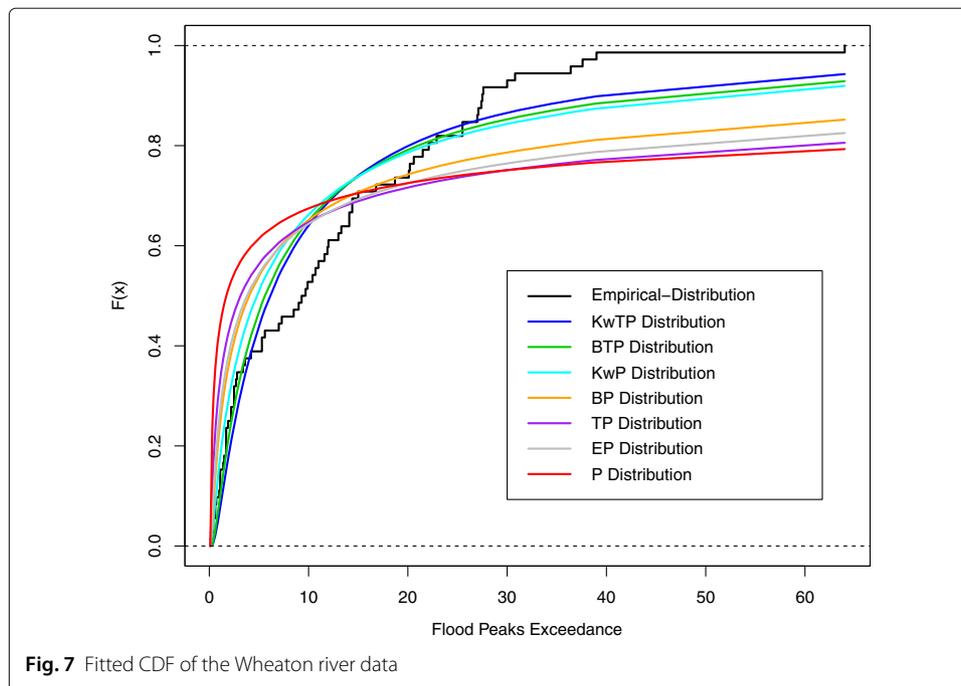
Table 4 The AIC, CAIC, BIC, HQIC and K-S test statistic of the Wheaton river data

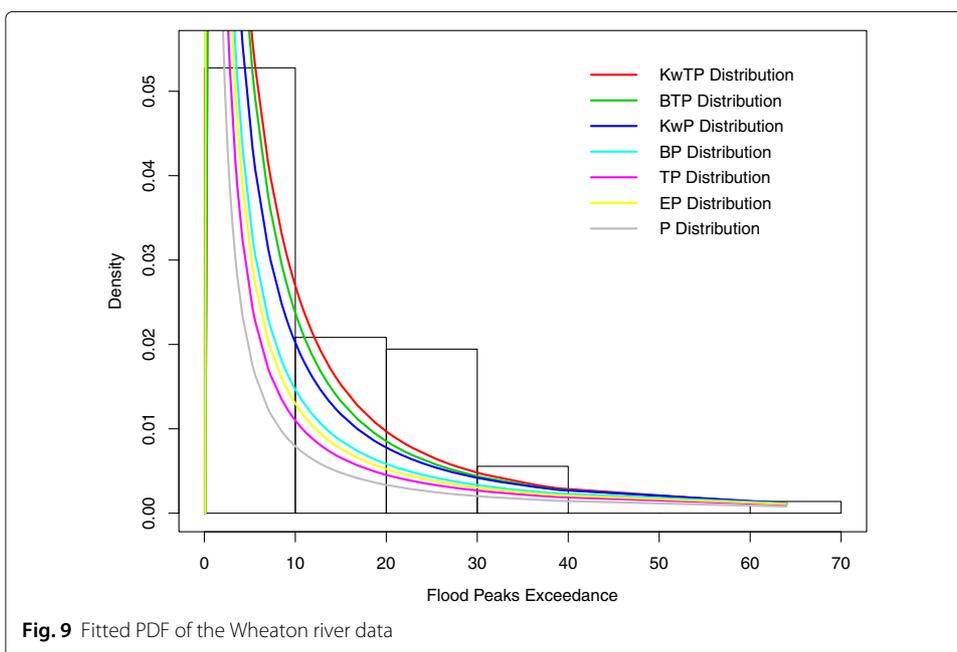
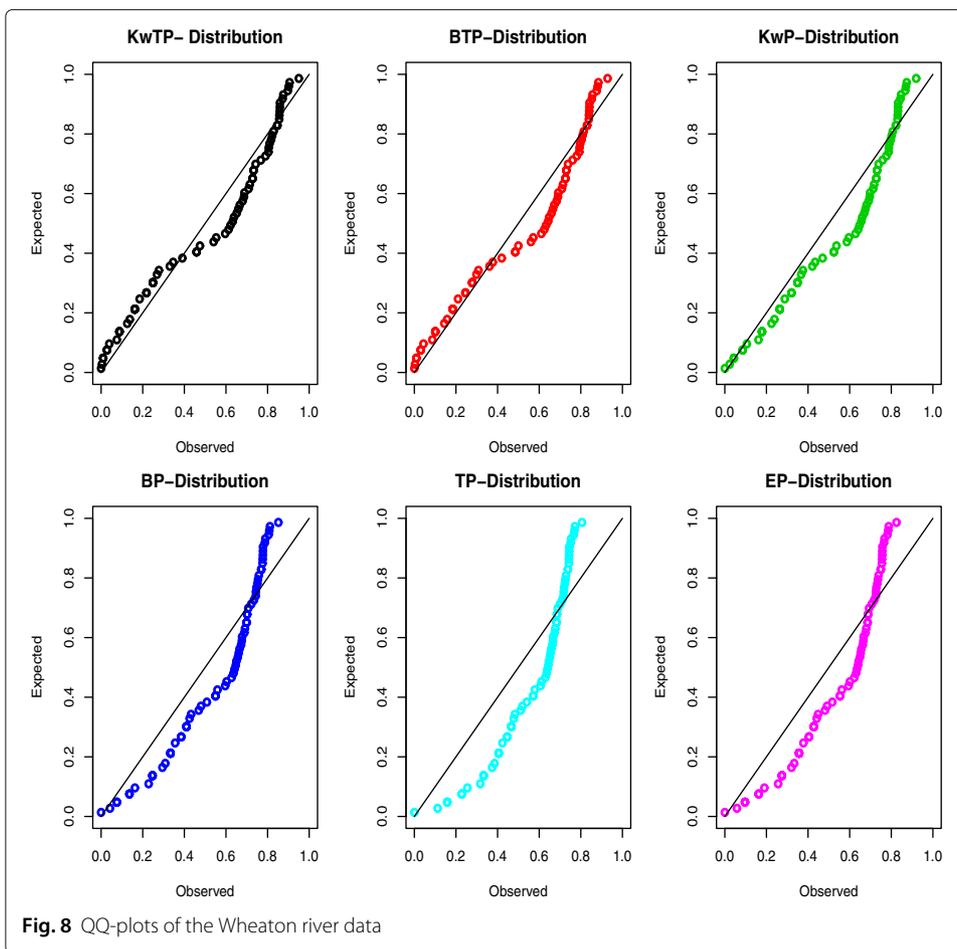
Model	Statistics					
	$-\hat{\ell}(.,x)$	AIC	CAIC	BIC	HQIC	KS
<i>KwTP</i>	254.017	516.034	516.641	525.085	519.634	0.147
BTP	256.577	521.154	521.760	530.204	524.753	0.160
<i>KwP</i>	271.200	548.400	548.753	555.230	551.119	0.170
BP	283.700	573.400	573.753	580.230	576.119	0.175
TP	286.201	576.402	576.575	580.954	578.214	0.287
EP	287.300	578.600	578.774	583.153	580.413	0.199
P	303.100	608.200	608.257	610.477	609.106	0.332

the $-\hat{\ell}$, AIC, CAIC, HQIC, BIC, and K-S statistic. One can employ the likelihood ratio statistic to test the superiority of the *KwTP* distribution over the other distributions. The plots comparing the *KwTP* distribution with other distributions are given in Figs. 7, 8 and 9. The values in Table 4, and the plots, all indicate that the *KwTP* distribution is more superior and fits the data more adequately than the other comparative distributions.

11.2 The Norwegian fire insurance data (in kr)

We apply the *KwTP* to the Norwegian fire insurance claims data for the years 1988 and 1990. Several authors have analyzed these data. See for example, Mdziniso and Cooray (2017) and related references in the paper. The 1988 Norwegian fire claims data consist of 827 fire insurance losses in thousand Norwegian kroner, ranging from 500 to 465,365 thousand Norwegian kroner. The 1990 Norwegian fire data consist of 628 fire insurance losses in thousand Norwegian kroner, ranging from 500 to 78,537 thousand Norwegian kroner. Data for both years are highly positively skewed. Using these data, Mdziniso and





Cooray (2017) compared the performances of the Pareto (P), the 3-parameter generalized Pareto (GP), the odd-Pareto (OP) and its extension (OP*), the 3-parameter Burr, the exponentiated Pareto (EP), the exponentiated odd Pareto (EOP) and the odd generalized Pareto (OGP) distributions.

We provide in Table 5, the parameter estimation of the insurance data for both 1988 and 1990. Comparing the performance of *KwTP* with the distributions listed in Table 4 of Mdziniso and Cooray (2017) for the 1988 data, we see from the values of the AIC, A-D, and K-S, that *KwTP* displays superiority over the P, GP, OP, OP*, Burr, and EP distributions. It also performs better than the OGP judging from their K-S values. The same pattern of performances of *KwTP* and the distributions listed in Table 5 of the paper are observed in the case for the 1990 data. With the exception of the exponentiated odd Pareto (EOP) distribution, the *KwTP* is a better fit for these sets of data.

12 Simulation study

In this section we perform some simulation study to assess the reliability of the MLEs. An ideal technique for simulating random numbers from the *KwTP* distribution is the inversion method. For fixed selected combinations of α, λ, a and b in Eq. (17), we generate samples of sizes $n = 5, 10, 20, 50, 100, 200, 500$ and 1000 from the *KwTP* distribution. We repeated the simulations $N = 1,000$ times and calculated the mean estimates, the root mean square errors (RMSEs) and the mean absolute errors (MAEs). The empirical results obtained using R are given in Table 6. Observe that RMSEs decay as the sample size increases, while the maximum likelihood estimates get closer to their true values.

13 Conclusion

We have proposed in this article, a new distribution that is being referred to as the Kumaraswamy transmuted Pareto (*KwTP*). The transmuted Pareto distribution is used as a baseline distribution in the Kumaraswamy distribution to construct the *KwTP* distribution. Many mathematical and statistical properties and special cases of the *KwTP* are obtained. The estimation of the model parameters is performed by the maximum likelihood method. We compare the distribution with few other distributions in modeling two real datasets. Various statistics indicate that *KwTP* better fits the Wheaton river data set than other comparative distributions, and majority of its competitive distributions for the Norwegian insurance claims data. We conducted a simulation

Table 5 Norwegian fire insurance claims: Estimated values for *KwTP*

Parameters				$-\ell$	AIC	BIC	A-D	K-S	p -value
(Year = 1988, $n = 827$)									
$\hat{\alpha}$	$\hat{\lambda}$	\hat{a}	\hat{b}						
1.0197	-0.7034	0.9698	1.2808	6751.24	13510.49	13529.29	0.46	0.03	0.62
(0.4104)	(0.2148)	(0.1609)	(0.6186)						
(Year = 1990, $n = 628$)									
$\hat{\alpha}$	$\hat{\lambda}$	\hat{a}	\hat{b}						
0.6310	-0.7285	1.0873	3.1417	5039.91	10087.82	10105.57	0.96	0.04	0.40
(0.2739)	(0.1873)	(0.1669)	(1.8428)						

Table 6 Simulation from *KwTP* of the MLE for $\alpha = 2, \lambda = 0.1, a = 3,$ and $b = 2.5$

Sample size (<i>n</i>)	Parameter	Estimate	RMSE	MAE
5	$\hat{\alpha}$	2.0490	0.7727	0.6466
	$\hat{\lambda}$	-0.1342	0.6446	0.5640
	\hat{a}	2.7096	1.0869	0.8522
	\hat{b}	2.4806	1.1296	0.9987
10	$\hat{\alpha}$	2.0722	0.6983	0.5810
	$\hat{\lambda}$	-0.1046	0.6089	0.5274
	\hat{a}	2.7421	0.9994	0.7871
	\hat{b}	2.5427	1.0715	0.9389
20	$\hat{\alpha}$	2.0429	0.6157	0.5030
	$\hat{\lambda}$	-0.0228	0.5786	0.5022
	\hat{a}	2.8104	0.9036	0.7274
	\hat{b}	2.5201	0.9400	0.8048
50	$\hat{\alpha}$	2.0078	0.5674	0.4486
	$\hat{\lambda}$	0.0330	0.5507	0.4692
	\hat{a}	2.8363	0.8251	0.6755
	\hat{b}	2.4863	0.8682	0.7382
100	$\hat{\alpha}$	1.9644	0.5361	0.4302
	$\hat{\lambda}$	0.0590	0.5288	0.4445
	\hat{a}	2.8337	0.7392	0.5914
	\hat{b}	2.4907	0.7856	0.6617
200	$\hat{\alpha}$	1.9783	0.5214	0.4246
	$\hat{\lambda}$	0.0628	0.5082	0.4210
	\hat{a}	2.8550	0.6567	0.5191
	\hat{b}	2.5186	0.7352	0.6141
500	$\hat{\alpha}$	1.9691	0.5515	0.4607
	$\hat{\lambda}$	0.0846	0.4972	0.4057
	\hat{a}	2.8452	0.6056	0.4562
	\hat{b}	2.5001	0.7121	0.5980
1000	$\hat{\alpha}$	1.9802	0.5694	0.4841
	$\hat{\lambda}$	0.1073	0.4818	0.3830
	\hat{a}	2.8629	0.5301	0.3895
	\hat{b}	2.4704	0.6748	0.5693

study to assess the performance of the maximum likelihood estimation procedure for estimating the parameters of the *KwTP* distribution. It is expected that the *KwTP* distribution will serve as a better alternative in modeling data sets exhibiting the extreme value properties.

Acknowledgements

The authors would like to thank the Editor and the anonymous reviewers for their valuable comments and suggestions which have greatly improved the quality of the paper.

Authors' contributions

The authors, SBC, AAA, GA and HL with the consultation of each other carried out this work and drafted the manuscript together. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Mathematical Sciences, Florida Atlantic University, Bacon Raton, FL, USA. ²Department of Mathematics, Marshall University, Huntington, WV, USA. ³Department of Mathematics, Statistics and Computer Science, Purdue University Northwest, Hammond, IN, USA.

Received: 22 March 2017 Accepted: 6 July 2017

Published online: 15 August 2017

References

- Afify, AZ, Cordeiro, GM, Yousof, HM, Alzaatreh, A, Nofal, ZM: The Kumaraswamy transmuted-g family of distributions: Properties and applications. *J. Data Sci.* **14**, 245–270 (2016)
- Afify, AZ, Nofal, ZM, Butt, NS: Transmuted complementary Weibull geometric distribution. *Pak. J. Stat. Oper. Res.* **4**, 435–454 (2014)
- Akinsete, A, Famoye, F, Lee, C: The beta-Pareto distributions. *Statistics.* **42**(6), 547–563 (2008)
- Akinsete, AA, Famoye, F, Lee, C: The Kumaraswamy - geometric distribution. *J. Stat. Distrib. Appl.* **1**, 1–21 (2014)
- Aryal, GR, Tsokos, CP: On the transmuted extreme value distribution with application. *Nonlinear Anal. Theory Methods and Appl.* **7**, 1401–1407 (2009)
- Aryal, GR, Tsokos, CP: Transmuted Weibull distribution: A generalization of the Weibull probability distribution. *Eur. J. Pur. Appl. Math.* **2**(2), 89–102 (2011)
- Ashour, SK, Eltehiwy, MA: Transmuted Lomax distribution. *Am. J. Appl. Math. Stat.* **1**(6), 121–127 (2013)
- Bourguignon, MB, Silva, RB, Zea, LM, Cordeiro, GM: The Kumaraswamy Pareto distribution. *J. Stat. Theory Appl.* **12**(2), 129–144 (2013)
- Bowley, AL: *Elements of Statistics*. 4th edn. Charles Scribner, New York, NY (1920)
- Chhetri, S, Long, H, Aryal, G: The beta transmuted Pareto distribution: Theory and applications. *J. Stat. Appl. Prob.* **6**(2), 243–258 (2017)
- Choulakian, V, Stephens, MA: Goodness-of-fit for the generalized Pareto distribution. *Technometrics.* **43**, 478–484 (2001)
- Cordeiro, GM, de Castro, M: A new family of generalized distributions. *J. Stat. Comput. Simul.* **81**, 883–898 (2011)
- Cordeiro, GM, Ortega, EMM, Nadarajah, S: The Kumaraswamy Weibull distribution with application to failure data. *J. Frankl. Inst.* **347**, 1399–1429 (2010)
- de Pascoa, MAR, Ortega, EMM, Cordeiro, GM: The Kumaraswamy generalized gamma distribution with application in survival analysis. *Stat Methodol.* **8**, 411–433 (2011)
- Elbatal, I: The Kumaraswamy exponentiated Pareto distribution. *Econ. Qual. Control.* **28**(1), 1–8 (2013)
- Eugene, N, Lee, C, Famoye, F: The beta-normal distribution and its applications. *Commun. Stat. Theory Methods.* **31**, 497–512 (2002)
- Gradshteyn, IS, Ryzhik, IM: *Table of Integrals, Series, and Products*. 3rd edn. Academic Press, New York (2000)
- Hady, AE, Ebraheim, N: Exponentiated transmuted Weibull distribution: A generalization of the Weibull distribution. *Int. Sch. Sci. Res. Innov.* **8**(6), 903–911 (2014)
- Hussian, MA: Transmuted exponentiated gamma distribution: A generalization of the exponentiated gamma probability distribution. *Appl. Math. Sci.* **8**(27), 1297–1310 (2014)
- Jones, MC: Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. *Stat Methodol.* **6**(1), 70–81 (2009)
- Kenny, JF, Keeping, ES: *Mathematics of Statistics*. 3rd edn. D. Van Nostrand Co., New Jersey (1962)
- Khan, MS, King, R: Transmuted modified Weibull distribution: A generalization of the modified Weibull probability distribution. *Eur. J. Pur. Appl. Math.* **6**(1), 66–88 (2013)
- Khan, MS, King, R, Hudson, IL: Transmuted Kumaraswamy distribution. *Stat. Transit.* **17**(2), 183–210 (2016)
- Kumaraswamy, P: A generalized probability density function for double-bounded random processes. *Hydrology.* **46**, 79–88 (1980)
- Marinho, PRD, Bourguignon, M, Dias, CRB: R Package - AdequacyModel (2016). <https://cran.r-project.org/web/packages/AdequacyModel/AdequacyModel.pdf>
- Mdziniso, NC, Cooray, K: Odd Pareto families of distributions for modelling loss payment data. *Scand. Actuar. J.* 1–22 (2017). <http://dx.doi.org/10.1080/03461238.2017.1280527>
- Merovci, F, Puka, L: Transmuted Pareto distribution. *ProbStat Forum.* **07**, 1–11 (2014)
- Moors, JJA: A quantile alternative for kurtosis. *J. R. Stat. Series D Stat.* **37**(1), 25–32 (1988)
- Nadarajah, S: Exponentiated Pareto distribution. *Stat. J. Theor. Appl. Stat.* **39**(3), 255–260 (2005)
- Owokolo, EA, Oguntade, PE, Adejumo, AO: Performance rating of the transmuted exponential distribution: an analytical approach. *SpringerPlus.* **4**, 818 (2015). doi:10.1186/s40064-015-1590-6
- Saulo, H: The Kumaraswamy Birnbaum-Saunders distribution. *J. Stat. Theory Pract.* **6**(4), 745–759 (2012)
- Shams, TM: The Kumaraswamy-generalized Lomax distribution. *Middle East J. Sci. Res.* **17**(5), 641–646 (2013)
- Shaw, WT, Buckley, IRC: The alchemy of probability distributions: beyond Gram-Charlier & Cornish-Fisher expansions, and skew-normal or kurtotic-normal distributions. Conference on Computational Finance, IMA, 6670 (2007). Research Report
- Shaw, WT, Buckley, IRC: The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. Conference on Computational Finance, IMA, 0901–0434 (2009). Research Report
- Tahir, MH, Cordeiro, GM: Compounding of distributions: a survey and new generalized classes. *J. Stat. Distrib. Appl.* **3**, 183–210 (2016)
- Yousof, HM, Afify, AZ, Alizadeh, M, Butt, NS, Hamedani, GG, Ali, MM: The transmuted exponentiated generalized-g family of distributions. *Pak. J. Stat. Oper. Res.* **4**, 441–464 (2015)