Information Sciences Letters
An International Journal

http://dx.doi.org/10.18576/isl/110128

A New Generalization of the Gumbel Distribution With Climate Application

Hussein Eledum[1,*] *and Mohammed Ahmmed Mosilhy*[1,2]

[1]Department of Statistics, Faculty of Science, University of Tabuk, KSA
[2]Department of Mathematics, Faculty of Science, Cairo University, Egypt

Received: 28 Jan. 2021, Revised: 20 Oct. 2021, Accepted: 24 Oct. 2021
Published online: 1 Jan. 2022

Abstract: In this paper, we propose a new generalization of the Gumbel distribution named the Cubic Transmuted Gumbel distribution (CTGD) based on a cubic ranking transmutation map. Statistical properties of CTGD such as reliability function, hazard function, moments, moment generating function, quantile function, and simulation of the random sample are studied. The parameters of CTGD are estimated using the Maximum Likelihood method. Finally, an application of CTGD using two real data sets on climate change is conducted to illustrate and compare with the base Gumbel distribution (GD) and transmuted Gumbel distribution (TGD). CTGD was found to be a better fit than GD and TGD.

Keywords: Cubic transmutation, Maximum likelihood estimation, Extreme value distribution , Reliability function

1 Introduction

The Gumbel distribution (GD), named after Gumbel [1] is also referred to as the Smallest Extreme Value (SEV) distribution or Type I Extreme Value distribution. The GD is a very popular distribution due to its extensive applicability in several areas, and its wide applications have been reported by Kotz and Nadarajah [2]. Koutsoyiannis [3] studied GD to observe the appropriateness of this distribution in modeling extreme rainfall. Aryal and Tsokos [4] have given the necessary formulation of the GD to study the airline spill data. The applicability of GD in the field of flood frequency analysis, network, space, software reliability, structural, and wind engineering are reported by Cardeiro et al. [5]. Generally, GD is used to analyze and model the behavior of random phenomena that occur in engineering, biology, environment among others. The cumulative distribution function (cdf) and the probability density function (pdf) of Gumbel random variable X are defined, respectively, as

$$G(x) = e^{-z} \quad ; \quad x \in R \tag{1}$$

and

$$g(x) = \frac{1}{\beta} z e^{-z} \tag{2}$$

where $z = e^{-(x-\mu)/\beta}$, while $\beta \in [0, \infty)$ and $\mu \in R$ are a scale and location parameters, respectively.
Shaw and Buckley [6] used the rank transmutation map to propose a new method for generating a family of distribution. According to [6], the cumulative distribution function of the ranking quadratic transformation map (QRTM) is:

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

where $G(x)$ is the cumulative distribution function (cdf) of the base distribution. Observe that, when $\lambda = 0$, the new distribution becomes the original one. Using ORTM of Eq.(3), Aryal and Tsokos [7] developed transmuted Gumbel distribution (TGD).
Abed Al-Kadim [8] proposed a generalized formula for transmuted distribution presented by [6], the cumulative distribution function of the Cubic Ranking transformation map (CRTM) is:

$$F(x) = (1+\lambda)G(x) - \lambda G^2(x) + \lambda G^3(x); \qquad |\lambda| \leq 1 \tag{4}$$

This method used by Abed Al-Kadim and Mohammed [9] to develop a cubic transmuted Weibull distribution. Another two classes of Cubic Transmuted distributions with two transmuted parameters have been developed, one by Granzatto et al. [10], the other by Rahman et al.

* Corresponding author e-mail: heledum@yahoo.com

[11]. Rahman formula used to propose some new distributions, for examples, Pareto distribution [12], Weibull distribution[13] and Frechet distribution [14].
Based on the CRTM proposed by [11], the cdf and pdf respectively given as

$$F(x) = (1+\lambda_1)G(x) + (\lambda_2 - \lambda_1)G^2(x) - \lambda_2 G^3(x) \quad (5)$$

and

$$f(x) = g(x)[(1+\lambda_1) + 2(\lambda_2 - \lambda_1)G(x) - 3\lambda_2 G^2(x)] \quad (6)$$

where $\lambda_1, \lambda_2 \in [-1,1], -2 \leq \lambda_1 + \lambda_2 \leq 1$ and $G(x), g(x)$ are the cdf and pdf of the base distribution respectively.
In this article, we use CRTM suggested by [11] to propose a new distribution which generalizes the Gumbel distribution. This new version of the Gumbel distribution is called Cubic Transmuted Gumbel (CTGD). Some statistical properties are studied, and the model parameters are estimated using the maximum likelihood method. Moreover, an application to two real data sets on climate change is illustrated and compared with the base GD and TGD.
The rest of this paper is structured as follows: The new proposed distribution Cubic Transmuted Gumbel (CTGD) is presented in Section 2 . We have investigated some statistical properties such as reliability function, hazard function, moments and moment generating function for CTGD in Section 3. Section 4 provides the parameter estimation of CTGD. An application of the CTGD to two real data sets for the purpose of illustration is conducted in section 5. Finally, Section 6 gives some concluding remarks.
The new Gumbel model is motivated because it exhibits a maximum or minimum of a number of samples with data close to normal distribution or approximately positively skewed, as illustrated in Fig. 1. The justification for the new generalized extreme value (GEV) model's practicality is based on its ability to model the wind velocity and the snow accumulation data sets, as illustrated in Section 5.

2 Cubic Transmuted Gumbel Distribution (CTGD)

In this section, the new proposed distribution CTGD is demonstrated. Including the cumulative distribution function (cdf), probability density function (pdf), survival and hazard function.

2.1 Cumulative and density functions for CTGD

Theorem 1. *Let X be a random variable with the CTGD. The cdf and pdf are defined, respectively, as*

$$F(x) = e^{-z}[(1+\lambda_1) + (\lambda_2 - \lambda_1)e^{-z} - \lambda_2 e^{-2z}] \quad (7)$$

and

$$f(x) = \frac{1}{\beta} z e^{-z}[(1+\lambda_1) + 2(\lambda_2 - \lambda_1)e^{-z} - 3\lambda_2 e^{-2z}]; \quad x \in R \quad (8)$$

where $z = e^{-(x-\mu)/\beta}$, $\beta \geq 0$ and $\mu \in R$ are a scale and location parameters respectively, λ_1, $\lambda_2 \in [-1,1]$ and $-2 \leq \lambda_1 + \lambda_2 \leq 1$.
Proof. The proof is straightforward. Eq.(7) is obtained by substituting Eq.(1) into Eq.(5) and Eq.(8) is gotten from substituting Eq.(2) into Eq.(6).
Proposition 1. The limit of CTGD density as $x \to \infty$ and $x \to -\infty$ is 0
Proof.

$$\lim_{x\to\infty} f(x) = \frac{1}{\beta} \lim_{z\to 0} z e^{-z}[(1+\lambda_1) + 2(\lambda_2 - \lambda_1)e^{-z} - 3\lambda_2 e^{-2z}] = 0$$

whilst from L'Hopital's rule we get

$$\lim_{x\to-\infty} f(x) = \frac{1}{\beta} \lim_{z\to\infty} z e^{-z}[(1+\lambda_1) + 2(\lambda_2 - \lambda_1)e^{-z} - 3\lambda_2 e^{-2z}] = 0$$

Proposition 2. $f(x)$ of Eq. (8) is a pdf.
Proof. To show this proposition, we must prove that $f(x) \geq 0$ and $\int_{-\infty}^{\infty} f(x)dx = 1$
From Proposition 1 $\lim_{x\to-\infty} f(x) = 0$ and $\lim_{x\to\infty} f(x) = 0$. It follows that $f(x) \geq 0$.
Proof of $\int_{-\infty}^{\infty} f(x)dx = 1$

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{\infty} e^{e^{-(x-\mu)/\beta}} [(1+\lambda_1) + 2(\lambda_2 - \lambda_1) \times e^{-e^{-(x-\mu)/\beta}} - 3\lambda_2 e^{-2e^{-(x-\mu)/\beta}}] dx$$

let $z = e^{-(x-\mu)/\beta} \Rightarrow \ln z = -(x-\mu)/\beta$ and $x = \mu - \beta \ln z$ also if $x \to -\infty \Rightarrow z \to \infty$ and if $x \to \infty \Rightarrow z \to 0$ then

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)dx &= \int_{\infty}^{0} f(z) \frac{-\beta}{z} dz \\ &= \int_{0}^{\infty} \left(\frac{1}{z}\right) z e^{-z}[(1+\lambda_1) + 2(\lambda_2 - \lambda_1)e^{-z} - 3\lambda_2 e^{-2z}] dz \\ &= (1+\lambda_1) + (\lambda_2 - \lambda_1) - \lambda_2 \\ &= 1\end{aligned}$$

Therefore, the proposition is proved.
Fig. 1 and 2 illustrate respectively, some of possible shapes of the pdf and cdf of CTGD for selected values of parameters λ_1 and λ_2 where $\mu = 2$ and $\beta = 2$.

From plot of pdf of Fig. 1, we can observe that for the positive values of both transmuted parameters λ_1 and λ_2, the distribution is approximately mesokurtic symmetrical shape, while for the negative values of λ_1 and λ_2 the distribution is shifting right with platykurtic shape.

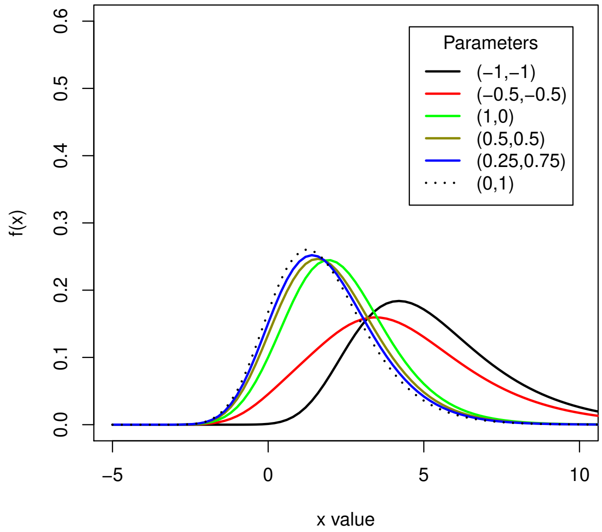

Fig. 1: The pdf of CTGD for different value of λ_1 and λ_2 where $\mu = 2$ and $\beta = 2$.

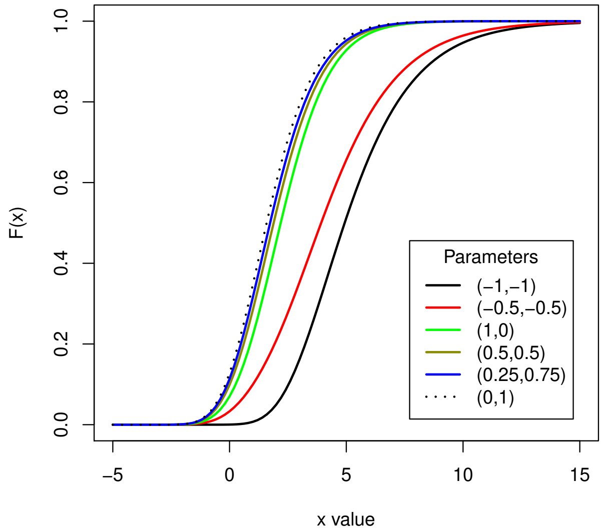

Fig. 2: The cdf of CTGD for different value of λ_1 and λ_2 where $\mu = 2$ and $\beta = 2$.

2.2 Survival and Hazard function

The survival function is defined as $S(x) = 1 - F(x)$ and for the CTGD is given as

$$s(x) = 1 - e^{-z}[(1+\lambda_1) + (\lambda_2 - \lambda_1)e^{-z} - \lambda_2 e^{-2z}]$$

The hazard function is defined as $h(x) = \frac{f(x)}{S(x)}$ and for the CTGD is given as

$$h(x) = \frac{z[(1+\lambda_1) + 2(\lambda_2 - \lambda_1)e^{-z} - 3\lambda_2 e^{-2z}]}{\beta[(1+\lambda_1) + (\lambda_2 - \lambda_1)e^{-z} - \lambda_2 e^{-2z}]}$$

Fig. 3 and 4 show respectively some possible shapes of the survival and hazard functions for the CTGD using different combination of model parameters λ_1 and λ_2 where $\mu = 2$ and $\beta = 2$

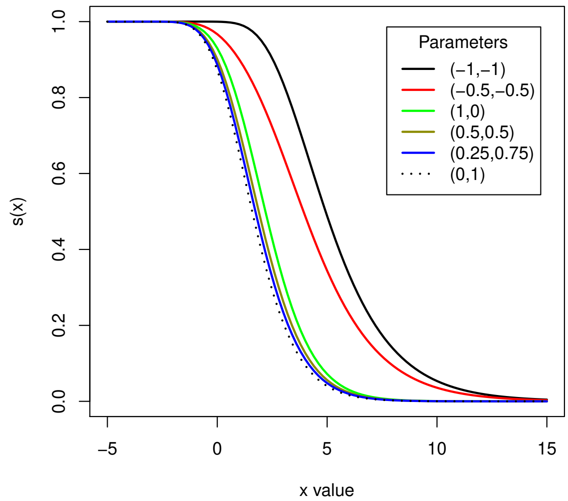

Fig. 3: The $s(x)$ of CTGD for different value of λ_1 and λ_2 where $\mu = 2$ and $\beta = 2$.

3 Statistical Properties

In this section, some statistical properties for the proposed distribution, CTGD is discussed. These properties involve moments, moment generating function, quantile function and simulation of the random sample.

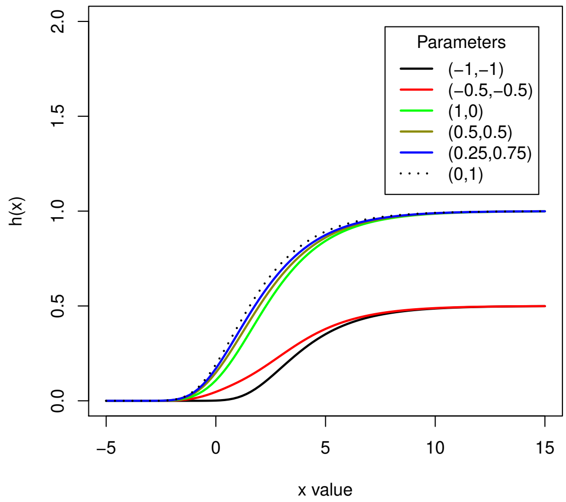

Fig. 4: The $h(x)$ of CTGD for different value of λ_1 and λ_2 where $\mu = 2$ and $\beta = 2$.

3.1 The Moments

Theorem 2. *Let X be a random variable having the CTGD , then the r^{th} moment of X about the origin is*

$$E(X^r) = \sum_{i=0}^{r}\binom{r}{i}\mu^{(r-i)}(-\beta)^i\Bigg[(1+\lambda_1)\Gamma^{(i)}(1) + \sum_{k=0}^{i}\binom{i}{k}\Gamma^{(k)}(1)T_{i,k}\Bigg];\ r = 0,1,2,\ldots \tag{9}$$

where $\Gamma^{(i)}(1) = \int_0^\infty [\ln(z)]^i e^{-z}dz$ is the i^{th} derivative of gamma function and

$$T_{i,k} = \left[(\lambda_2-\lambda_1)(\ln(\tfrac{1}{2}))^{(i-k)} - \lambda_2(\ln(\tfrac{1}{3}))^{(i-k)}\right].$$

Proof. we know that

$$\begin{aligned}E(X^r) &= \int_{-\infty}^{\infty} x^r f(x)dx \\ &= \int_{\infty}^{0} [\mu - \beta\ln(z)]^r f(z)\left(\frac{-\beta}{z}\right)dz \ \ \text{(see Prop. 2)} \\ &= \int_{0}^{\infty} [\mu - \beta\ln(z)]^r e^{-z}\left[(1+\lambda_1) + 2(\lambda_2-\lambda_1)e^{-z} - 3\lambda_2 e^{-2z}\right]dz\end{aligned}$$

therefore,

$$E(X^r) = (1+\lambda_1)I_1 + 2(\lambda_2-\lambda_1)I_2 - 3\lambda_2 I_3 \tag{10}$$

where $I_1 = \int_0^\infty [\mu - \beta\ln(z)]^r e^{-z}dz$,
$I_2 = \int_0^\infty [\mu - \beta\ln(z)]^r e^{-2z}dz$ and
$I_3 = \int_0^\infty [\mu - \beta\ln(z)]^r e^{-3z}dz$
Now, we calculate the value of I_1 using binomial Expansion as follows

$$\begin{aligned}I_1 &= \int_0^\infty [\mu - \beta\ln(z)]^r e^{-z}dz \\ &= \int_0^\infty \sum_{i=0}^{r}\binom{r}{i}\mu^{(r-i)}[-\beta\ln(z)]^i e^{-z}dz \\ &= \sum_{i=0}^{r}\binom{r}{i}\mu^{(r-i)}(-\beta)^i\int_0^\infty [\ln(z)]^i e^{-z}dz \\ &= \sum_{i=0}^{r}\binom{r}{i}\mu^{(r-i)}(-\beta)^i\Gamma^{(i)}(1)\end{aligned} \tag{11}$$

Similarly, we can obtain the results

$$\begin{aligned}I_2 = &\frac{1}{2}\Bigg\{\sum_{i=0}^{r}\binom{r}{i}\mu^{(r-i)}(-\beta)^i\Bigg[\sum_{k=0}^{i}\binom{i}{k} \\ &\times[-\ln(2)]^{(i-k)}\Gamma^{(k)}(1)\Bigg]\Bigg\}\end{aligned} \tag{12}$$

and

$$\begin{aligned}I_3 = &\frac{1}{3}\Bigg\{\sum_{i=0}^{r}\binom{r}{i}\mu^{(r-i)}(-\beta)^i\Bigg[\sum_{k=0}^{i}\binom{i}{k} \\ &\times[-\ln(3)]^{(i-k)}\Gamma^{(k)}(1)\Bigg]\Bigg\}\end{aligned} \tag{13}$$

Substitute equations (11), (12), and (13) in Eq. (10) we get

$$E(X^r) = \sum_{i=0}^{r}\binom{r}{i}\mu^{(r-i)}(-\beta)^i\Bigg[(1+\lambda_1)\Gamma^{(i)}(1) + \sum_{k=0}^{i}\binom{i}{k}\Gamma^{(k)}(1)T_{i,k}\Bigg]$$

Therefore, the theorem is proved.
The mean and variance can be easily obtained by using $r = 1,2$ in Eq(9) such that $\Gamma^{(0)}(1) = 1, \Gamma^{(1)}(1) = -\gamma$ and $\Gamma^{(2)}(1) = \gamma^2 + \frac{\pi^2}{6}$ where $\gamma \approx 0.5772$ is the Euler Mascheroni constant. We get

$$E(X) = \theta(1+\lambda_1) + (\lambda_2-\lambda_1)\left[\theta + \beta ln(2)\right] - \lambda_2\left[\theta + \beta ln(3)\right]$$

$$\begin{aligned}E(X^2) = &(1+\lambda_1)\left[\theta^2 + \frac{(\beta\pi)^2}{6}\right] + (\lambda_2-\lambda_1)\Big\{\theta^2 + \theta\beta\ln(4) \\ &+ \beta^2\Big[\frac{\pi^2}{6} + \left(\ln(2)\right)^2\Big]\Big\} - \lambda_2\Big\{\theta^2 + \theta\beta\ln(9) \\ &+ \beta^2\Big[\frac{\pi^2}{6} + \left(\ln(3)\right)^2\Big]\Big\}\end{aligned}$$

where $\theta = \mu + \beta\gamma$

$$Var[X] = E(X^2) - [E(X)]^2$$

The mean and variance of CTGD for various combinations of model parameters are given in Table 1 and Table 2 respectively. From Tables 1 and 2, it is observed that, holding the location and scale parameters μ and β constants, as the transmuted parameters λ_1 and λ_2 increase the mean and variance of CTGD decrease. Whilst, holding λ_1 and λ_2 constants, as the scale parameter β increases the mean and variance also increase.

Table 1: Mean of the CTGD for various combinations of the parameters

μ	β	λ_2 / λ_1	$\lambda_1 = -1$	$\lambda_1 = -0.5$	$\lambda_1 = 1$	$\lambda_1 = 0.5$	$\lambda_1 = 0$
μ=-2	β=1	$\lambda_2 = -1$	-0.3242	-0.6707	-1.7105	-1.3639	-1.0173
		$\lambda_2 = -0.5$	-0.5269	-0.8735	-1.9132	-1.5666	-1.2201
		$\lambda_2 = 0$	-0.7296	-1.0762	-2.1159	-1.7694	-1.4228
		$\lambda_2 = 0.5$	-0.9324	-1.2789	–	-1.9721	-1.6255
		$\lambda_2 = 1$	-1.1351	-1.4817	–	–	-1.8282
	β=3	$\lambda_2 = -1$	3.0275	1.9878	-1.1314	-0.0917	0.948
		$\lambda_2 = -0.5$	2.4193	1.3796	-1.7396	-0.6999	0.3398
		$\lambda_2 = 0$	1.8111	0.7714	-2.3478	-1.3081	-0.2684
		$\lambda_2 = 0.5$	1.2029	0.1632	–	-1.9163	-0.8766
		$\lambda_2 = 1$	0.5947	-0.445	–	–	-1.4847
μ=0	β=1	$\lambda_2 = -1$	1.6758	1.3293	0.2895	0.6361	0.9827
		$\lambda_2 = -0.5$	1.4731	1.1265	0.0868	0.4334	0.7799
		$\lambda_2 = 0$	1.2704	0.9238	-0.1159	0.2306	0.5772
		$\lambda_2 = 0.5$	1.0676	0.7211	–	0.0279	0.3745
		$\lambda_2 = 1$	0.8649	0.5183	–	–	0.1718
	β=3	-1	5.0275	3.9878	0.8686	1.9083	2.948
		$\lambda_2 = -0.5$	4.4193	3.3796	0.2604	1.3001	2.3398
		$\lambda_2 = 0$	3.8111	2.7714	-0.3478	0.6919	1.7316
		$\lambda_2 = 0.5$	3.2029	2.1632	–	0.0837	1.1234
		$\lambda_2 = 1$	2.5947	1.5550	–	–	0.5153
μ =2	β=1	$\lambda_2 = -1$	3.6758	3.3293	2.2895	2.6361	2.9827
		$\lambda_2 = -0.5$	3.4731	3.1265	2.0868	2.4334	2.7799
		$\lambda_2 = 0$	3.2704	2.9238	1.8841	2.2306	2.5772
		$\lambda_2 = 0.5$	3.0676	2.7211	–	2.0279	2.3745
		$\lambda_2 = 1$	2.8649	2.5183	–	–	2.1718
	β=3	$\lambda_2 = -1$	7.0275	5.9878	2.8686	3.9083	4.948
		$\lambda_2 = -0.5$	6.4193	5.3796	2.2604	3.3001	4.3398
		$\lambda_2 = 0$	5.8111	4.7714	1.6522	2.6919	3.7316
		$\lambda_2 = 0.5$	5.2029	4.1632	–	2.0837	3.1234
		$\lambda_2 = 1$	4.5947	3.555	–	–	2.5153

Table 2: Variance of the CTGD for various combinations of the parameters

μ	β	λ_2 / λ_1	$\lambda_1 = -1$	$\lambda_1 = -0.5$	$\lambda_1 = 1$	$\lambda_1 = 0.5$	$\lambda_1 = 0$
$\mu = -2$	β=1	$\lambda_2 - 1$	1.6449	2.0461	1.8082	2.1277	2.207
		$\lambda_2 = -0.5$	1.686	1.9467	1.2872	1.7473	1.9671
		$\lambda_2 = 0$	1.6449	1.765	0.684	1.2846	1.6451
		$\lambda_2 = 0.5$	1.5216	1.5012	–	0.7397	1.2406
		$\lambda_2 = 1$	1.3161	1.1552	–	–	0.754
	β=3	$\lambda_2 - 1$	14.8044	18.4148	16.2739	19.1496	19.8633
		$\lambda_2 = -0.5$	15.1743	17.52	11.585	15.7254	17.7037
		$\lambda_2 = 0$	14.8044	15.8854	6.1563	11.5613	14.8044
		$\lambda_2 = 0.5$	13.6947	13.511	–	6.6575	11.1653
		$\lambda_2 = 1$	11.8452	10.3968	–	–	6.7863
$\mu = 0$	β=1	$\lambda_2 - 1$	1.6449	2.0461	1.8082	2.1277	2.207
		$\lambda_2 = -0.5$	1.686	1.9467	1.2872	1.7473	1.9671
		$\lambda_2 = 0$	1.6449	1.765	0.684	1.2846	1.6449
		$\lambda_2 = 0.5$	1.5216	1.5012	–	0.7397	1.2406
		$\lambda_2 = 1$	1.3161	1.1552	–	–	0.754
	β=3	$\lambda_2 = -1$	14.8044	18.4148	16.2739	19.1496	19.8633
		$\lambda_2 = -0.5$	15.1743	17.52	11.585	15.7254	17.7037
		$\lambda_2 = 0$	14.8044	15.8854	6.1563	11.5613	14.8044
		$\lambda_2 = 0.5$	13.6947	13.511	–	6.6575	11.1653
		$\lambda_2 = 1$	11.8452	10.3968	–	–	6.7863
$\mu = 2$	β=1	$\lambda_2 = -1$	1.6449	2.0461	1.8082	2.1277	2.207
		$\lambda_2 = -0.5$	1.686	1.9467	1.2872	1.7473	1.9671
		$\lambda_2 = 0$	1.64488	1.765	0.684	1.2846	1.64488
		$\lambda_2 = 0.5$	1.5216	1.5012	–	0.7397	1.2406
		$\lambda_2 = 1$	1.3161	1.1552	–	–	0.754
	β=3	$\lambda_2 = -1$	14.8044	18.4148	16.2739	19.1496	19.8633
		$\lambda_2 = -0.5$	15.1743	17.52	11.585	15.7254	17.7037
		$\lambda_2 = 0$	14.8044	15.8854	6.1563	11.5613	14.8044
		$\lambda_2 = 0.5$	13.6947	13.511	–	6.6575	11.1653
		$\lambda_2 = 1$	11.8452	10.3968	–	–	6.7863

3.2 The Moment Generating Function of CTGD

The moment generating function of CTGD is present in this theorem

Theorem 3. *Let X be a random variable having the CTGD, then the moment generating function (mgf) of X is*

$$M_x(t) = e^{\mu t}\Gamma(1-\beta t)\big[(1+\lambda_1) + 2^{\beta t}(\lambda_2-\lambda_1) - 3^{\beta t}\lambda_2\big] \tag{14}$$

where $\Gamma(1-\beta t)$ is the Gamma function.

Proof. We know that

$$\begin{aligned} M_X(t) &= \int_{-\infty}^{\infty} e^{tx} f(x)dx \\ &= \int_{\infty}^{0} e^{t[\mu-\beta\ln(z)]} f(z)(\frac{-\beta}{z})dz \\ &= \int_{0}^{\infty} e^{t[\mu-\beta\ln(z)]} e^{-z}\big[(1+\lambda_1) \\ &\qquad +2(\lambda_2-\lambda_1)e^{-z} - 3\lambda_2 e^{-2z}\big]dz \end{aligned}$$

$$M_x(t) = e^{t\mu}\left[(1+\lambda_1)I_1 + 2(\lambda_2-\lambda_1)I_2 - 3\lambda_2 I_3\right] \tag{15}$$

where $I_1 = \int_0^\infty z^{-\beta t}e^{-z}dz$, $I_2 = \int_0^\infty z^{-\beta t}e^{-2z}dz$ and $I_3 = \int_0^\infty z^{-\beta t}e^{-3z}dz$
Now, we calculate the value of I_1, I_2 and I_3 using the relation $\int_0^\infty t^b e^{-at}dt = \frac{\Gamma(1+b)}{a^{(1+b)}}$, we obtain
$I_1 = \int_0^\infty z^{-\beta t}e^{-z}dz = \Gamma(1-\beta t)$,
$I_2 = \int_0^\infty z^{-\beta t}e^{-2z}dz = 2^{(\beta t-1)}\Gamma(1-\beta t)$ and $I_3 = \int_0^\infty z^{-\beta t}e^{-3z}dz = 3^{(\beta t-1)}\Gamma(1-\beta t)$.
By Substituting in Eq. (15), we get
$M_x(t) = e^{\mu t}\Gamma(1-\beta t)\big[(1+\lambda_1) + 2^{\beta t}(\lambda_2-\lambda_1) - 3^{\beta t}\lambda_2\big]$
So the theorem is proved.

3.3 Quantile function of CTGD

The quantile function for CTGD is derived by finding the value of Q for which $F(x) = p$:

$$Q(p,\mu,\beta) = \mu - \beta\ln[-(\ln y)] \tag{16}$$

where

$$y = \left[\frac{\sqrt[3]{(\theta_2+\sqrt{h})}}{a\sqrt[3]{54}}\right] - \left[\frac{\theta_1\sqrt[3]{2}}{3a\sqrt[3]{(\theta_2+\sqrt{h})}}\right] - \left(\frac{b}{3a}\right) \tag{17}$$

such that
$\theta_1 = 3ac - b^2$,$\theta_2 = -2b^3 + 9abc - 27da^2$ and $h = 4\theta_1^3 + \theta_2^2$, where $a = \lambda_2$, $b = (\lambda_1 - \lambda_2)$, $c = -(1+\lambda_1)$ and $d = p$ are the coefficients of the cubic equation of y
The three quartiles Q_1, Q_2 and Q_3 can be obtained by using $p = 0.25, 0.50$ and 0.75 in Eq.(16), respectively.

3.4 Simulating the Random Sample of CTGD

Random numbers from the CTGD can be obtained by equating cdf of the distribution in Eq.(7) with a uniform random number and inverting the expression, that is the random number from CTGD is obtained by solving $F(x) = u$ for x. The random sample from CTGD can be further expressed as

$$x = \mu - \beta\ln[-(\ln y)] \tag{18}$$

where y is given in Eq.(17) with $d = u$ and u is an arbitrary continuous uniform point over $(0,1)$.

4 Parameters Estimation

This section pertains to discuss the maximum likelihood estimation (MLE) for parameters of CTGD.
Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from CTGD, then the likelihood function is given by

$$\begin{aligned} L &= \prod_{i=1}^{n} f(x_i,\mu,\beta,\lambda_1,\lambda_2) \\ &= \prod_{i=1}^{n}\Big\{\left(\frac{1}{\beta}\right) z_i\, e^{-z_i}\Big[(1+\lambda_1) + 2(\lambda_2-\lambda_1)e^{-z_i} - 3\lambda_2 e^{-2z_i}\Big]\Big\} \end{aligned}$$

where $z_i = e^{-(\frac{x_i-\mu}{\beta})}$; $i = 1,2,\ldots,n$.
so, the log-likelihood function is

$$\begin{aligned} l = \ln L = -n\ln\beta + \sum_{i=1}^{n}\Big\{\ln z_i - z_i + \ln\Big[(1+\lambda_1) \\ + 2(\lambda_2-\lambda_1)e^{-z_i} - 3\lambda_2 e^{-2z_i}\Big]\Big\} \end{aligned} \tag{19}$$

Therefore, the maximum likelihood estimates of μ, β, λ_1 and λ_2 which maximize Eq. (19), must satisfy the four equations (20), (21), (22) and (23).

$$\frac{\partial l}{\partial \mu} = \frac{1}{\beta}\left[n - \sum_{i=1}^{n}\left(\frac{9\lambda_2 z_i e^{-2z_i} - (1+\lambda_1)z_i}{(1+\lambda_1) + 2(\lambda_2-\lambda_1)e^{-z_i} - 3\lambda_2 e^{-2z_i}}\right)\right]$$

$$\sum_{i=1}^{n}\left[\frac{9\lambda_2 z_i e^{-2z_i} - (1+\lambda_1)z_i}{(1+\lambda_1) + 2(\lambda_2-\lambda_1)e^{-z_i} - 3\lambda_2 e^{-2z_i}}\right] = n \tag{20}$$

$$\begin{aligned} \frac{\partial l}{\partial \beta} &= \frac{-1}{\beta}\Big\{n - \frac{1}{\beta} \\ &\quad \times \sum_{i=1}^{n}\Big[\Big(\frac{9\lambda_2 z_i e^{-2z_i} - (1+\lambda_1)z_i}{(1+\lambda_1) + 2(\lambda_2-\lambda_1)e^{-z_i} - 3\lambda_2 e^{-2z_i}}\Big) \\ &\qquad - \beta\Big]\ln z_i\Big\} \end{aligned}$$

$$\sum_{i=1}^{n}\left[\left(\frac{9\lambda_2 z_i e^{-2z_i} - (1+\lambda_1)z_i}{(1+\lambda_1) + 2(\lambda_2-\lambda_1)e^{-z_i} - 3\lambda_2 e^{-2z_i}}\right) - \beta\right]\ln z_i = n\beta \tag{21}$$

$$\frac{\partial l}{\partial \lambda_1} = \sum_{i=1}^{n} \left[\frac{1-2e^{-z_i}}{(1+\lambda_1)+2(\lambda_2-\lambda_1)e^{-z_i}-3\lambda_2 e^{-2z_i}} \right]$$

$$\sum_{i=1}^{n} \left[\frac{1-2e^{-z_i}}{(1+\lambda_1)+2(\lambda_2-\lambda_1)e^{-z_i}-3\lambda_2 e^{-2z_i}} \right] = 0 \qquad (22)$$

$$\frac{\partial l}{\partial \lambda_2} = \sum_{i=1}^{n} \left[\frac{2e^{-z_i}-3e^{-2z_i}}{(1+\lambda_1)+2(\lambda_2-\lambda_1)e^{-z_i}-3\lambda_2 e^{-2z_i}} \right]$$

$$\sum_{i=1}^{n} \left[\frac{2e^{-z_i}-3e^{-2z_i}}{(1+\lambda_1)+2(\lambda_2-\lambda_1)e^{-z_i}-3\lambda_2 e^{-2z_i}} \right] = 0 \qquad (23)$$

The maximum likelihood estimates $\hat{\theta} = (\hat{\mu}, \hat{\beta}, \hat{\lambda}_1, \hat{\lambda}_2)$ of $\theta = (\mu, \beta, \lambda_1\ , \lambda_2)$ is obtained by solving the non-linear system of equations (20), (21), (22) and (23).

5 Application of CTGD

In this section, the CTGD is applied to two real data sets. The first data in Table 3 corresponds to a wind velocity (WVD) involving 246 observations of the maximum of monthly wind speed (mph) in Palm Beach, Florida (USA) for the months January 1984 to December, 2005. Data is available for download from NOAA website

The second data in Table 4 is related to the snow accumulation (SAD) in inches in the Raleigh-Durham airport, North Carolina, from 1948 to 2000. The data set contains 63 observations. Summary statistics of the two data sets is demonstrated in Table 5. The maximum likelihood estimates, the log-likelihood value (-Log(L)), the Kolmogorov-Smirnov (k-s) test statistic and the p-value for the k-s statistic for the fitted distributions are demonstrated in Tables 6 and 7 respectively. Latterly, Yolanda et al. [15] used the data in Tables 3 and 4 to investigate model fitting for Gumbel (GD) , Slash (SD) and Slashed Gumbel (SGD) distributions.

Table 3: : the wind velocity (WVD) data set

33	40	46	41	31	37	41	56	45	31
40	35	33	43	36	36	48	45	51	44
38	36	40	32	51	37	43	33	35	44
41	41	33	45	38	43	62	45	51	39
35	58	48	35	43	49	43	39	39	40
39	45	48	43	45	36	40	36	47	35
40	39	44	37	36	38	37	41	38	36
36	48	37	40	38	37	37	38	49	66
39	45	37	35	39	52	66	51	39	64
59	36	36	36	41	41	39	45	40	37
33	66	38	59	38	41	45	35	43	39
74	63	37	45	52	43	44	52	36	43
46	40	43	29	39	53	32	41	52	31
46	48	49	41	32	37	29	43	40	47
45	38	28	30	40	36	37	38	37	33
30	34	38	45	40	31	39	31	31	38
32	34	45	39	31	29	39	36	34	55
38	37	36	34	44	32	54	30	39	30
41	33	36	39	33	33	30	40	44	61
34	26	38	26	34	36	28	36	43	35
43	37	40	35	36	28	41	30	31	48
43	43	49	36	38	30	33	35	36	45
29	43	33	39	38	29	38	41	31	35
40	33	51	33	40	45	32	29	35	37
35	30	32	39	32	39	38	39	83	30
33	39	33	36	39	44	31	43	44	43
41	101	37	33						

Table 4: The snow accumulation (SAD) data set

1.0	2.5	1.2	1.2	4.1	9.0	3.0	1.0	1.4
2.0	3.0	1.7	1.2	1.2	1.1	1.5	5	1.6
2.0	0.1	0.4	0.8	3.7	1.3	3.8	0.1	0.1
0.2	2.0	7.6	0.1	1.8	0.5	0.5	0.5	1.1
1.4	1.0	1.0	0.7	5.7	0.4	0.3	1.8	0.4
1.0	1.2	2.6	1.0	5.0	1.7	2.4	0.1	0.5
7.1	0.2	0.7	0.1	2.7	2.9	0.4	2.0	20.3

Table 5: Summary Statistics for WVD and SAD data sets

Data	n	Mean	Median	Skewness	kurtosis
WVD	264	40.11	39.00	2.3438	13.030
SAD	63	2.125	1.200	4.01015	23.250

Table 6: Parameters estimates, -log (L), k-s test value and p-value for Gumbel, Transmuted Gumbel, and Cubic Transmuted Gumbel for WVD data

Model	Parameters estimates		-log(L)	k-s	P-value
GD	$\beta = 1.3293$	$\mu = 1.7361$	126.036	0.1486	0.1236
TGD	$\beta = 1.5036$ $\lambda = 0.5592$	$\mu = 1.5711$	124.089	0.1335	0.2115
CTGD	$\beta = 1.5419$ $\lambda_1 = 0.7089$	$\mu = 1.6799$ $\lambda_2 = -0.2859$	**123.217**	**0.1275**	**0.2288**

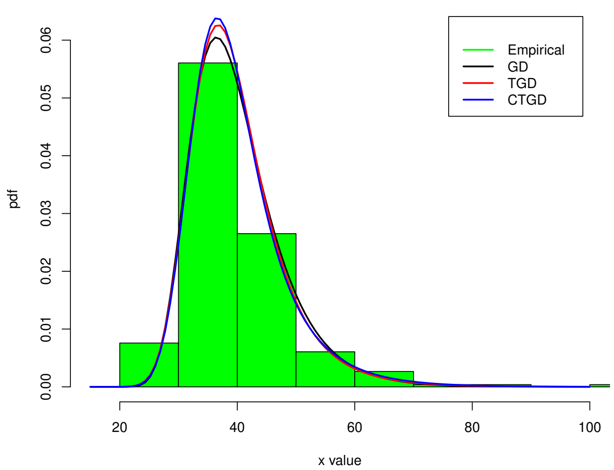

Fig. 5: The pdf of GD, TGD and CTGD for WVD data set

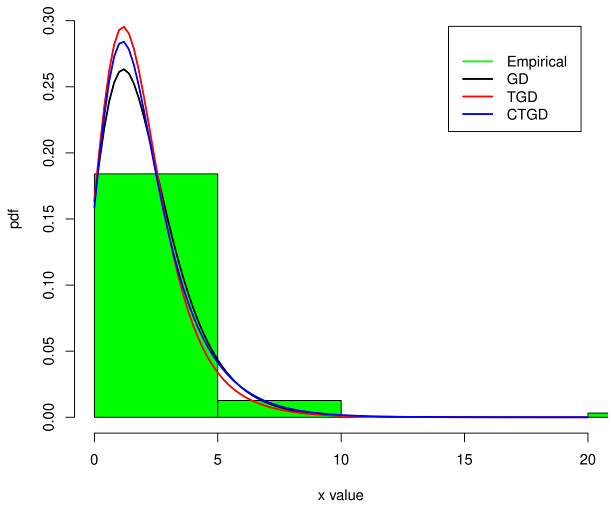

Fig. 6: The pdf of GD, TGD and CTGD for SAD data set

Table 7: Parameters estimates, -log (L), k-s test value and p-value for Gumbel, Transmuted Gumbel, and Cubic Transmuted Gumbel for SAD data

Model	Parameters estimates		-log(L)	k-s	P-value
GD	$\beta = 6.0802$	$\mu = 36.466$	898.880	0.0671	0.1848
TGD	$\beta = 7.260$ $\lambda = 0.6233$	$\mu = 38.837$	896.752	**0.0612**	**0.2759**
CTGD	$\beta = 7.5512$ $\lambda_1 = 0.99$	$\mu = 39.7094$ $\lambda_2 = -0.360$	**895.894**	0.0629	0.2456

By comparing the goodness of fit statistics in Table 6 among the three distributions, it is clear that all distributions are competitors and fit the wind velocity data well but the proposed distribution CTGD leads to a better fit than the other two distributions (see Fig. 5). Moreover, basing on -2log (L) criteria (the smaller the better), CTGD performs better than GD and TGD. Regarding the snow accumulation data, from Table 7 it is clear that the three distributions fit the data well but the TGD is the best (see Fig. 6), whilst, basing on -2log (L) criteria, the CTGD performs better than GD and TGD.

6 Conclusions

In this paper, a new generalized version of the Gumbel distribution which called cubic transmuted Gumbel distribution (CTGD) is introduced by using the generalization formula for transmuted distribution proposed by [11]. Some statistical properties of CTGD are derived. The model parameters are estimated by the maximum likelihood method. Finally, an application of CTGD to two real data sets and compared with some distributions based on Gumbel distribution is explained. We conclude that the applications suggest that the proposed distribution CTGD fits the two data sets very well. Generally, CTGD performs better than the other distributions. We recommend the CTGD for modelling extreme data sets in the field of flood frequency analysis, network, space, software reliability, structural and wind engineering and hope that it would receive significant applications in the future.

Acknowledgement

The authors are grateful to the editor and the anonymous reviewers for their valuable comments and suggestions, which have substantially improved this paper.

Conflicts of Interests

The authors declare that they have no conflicts of interests.

References

[1] E.J. Gumbel, *Statistics of Extremes*. Columbia University Press, New York, (1958).
[2] S. Kotz and S. Nadarajah, *Extreme value distributions: theory and applications*, World Scientific, (2000).
[3] D. Koutsoyiannis, *On the appropriateness of the Gumbel distribution for modelling extreme rainfall*, In Proceedings of the ESF LESC Exploratory Workshop held at Bologna, 24-25, (2003), doi:10.13140/RG.2.1.3811.6080.

[4] G. Aryal and C.P. Tsokos, Airline spill analysis using Gumbel and Moyal distributions, *Neural, Parallel and Scientific Computations*, **16**, 35-43 (2008).
[5] G.M. Cardeiro, S. Nadarajah and E.M.M. Ortega, The Kumaraswamy Gumbel distribution, *Statistical Methods and Applications*, **21(2)**, 139-168 (2012).
[6] W. Shaw and I. Buckley, *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, Research Report, (2009).
[7] G.R. Aryal and C. P. Tsokos, On the transmuted extreme value distribution with application, *Nonlinear Analysis*, **71(12)**, e1401-e1407 (2009), doi:10.1016/j.na.2009.01.168.
[8] K. Abed Al-Kadim, Proposed Generalized Formula for Transmuted Distribution, *Journal of University of Babylon, Pure and Applied Sciences*, **26(4)** (2018).
[9] K. Abed AL Kadim and M. Mohammed, The Cubic Transmuted Weibull Distribution, *Journal of Babylon University, Pure and Applied Sciences*, **25(3)**, 862-876 (2017).
[10] D. Granzatto, f. Louzada and BN. alakrishnan, Cubic rank transmuted distributions: inferential issues and applications, *Journal of Statistical Computation and Simulation*, **87(14)**, 2760-2778 (2017), doi:10.1080/00949655.2017.1344239.
[11] M. Rahman, B. Al-Zahrani and M. Shahbaz, A general transmuted family of distributions, *Pakistan Journal of Statistics and Operation Research*, **14(4)**, 451-469 (2018), doi:10.18187/pjsor.v14i4.2527.
[12] M. Rahman, B. Al-Zahrani and M. Shahbaz, Cubic transmuted Pareto distribution, *Annals of Data Science*, **7(1)**, 91-108 (2020), doi:10.1007/s40745-018-0178-8.
[13] M. Rahman, B. Al-Zahrani and M. Shahbaz, Cubic Transmuted Weibull Distribution: Properties and Applications, *Annals of Data Science*, **6(1)**, 83-102 (2019), doi:10.1007/s40745-018-00188.
[14] M. A. Mosilhy and H. Eledum, Cubic Transmuted Frechet Distribution, *Journal of Statistics Applications and Probability*, **11(1)**, 135-145 (2022), doi:10.18576/jsap/110110.
[15] M. Yolanda, B. Heleno and W. Hector, Gumbel distribution with heavy tails and applications to environmental Data, *Mathematics and Computers in Simulation*, **157**, 115-129 (2019), doi:10.1016/j.matcom.2018.10.003.