Article

A Revised Bimodal Generalized Extreme Value Distribution: Theory and Climate Data Application

Cira E. G. Otiniano [1]*, Mathews N. S. Lisboa [1] and Terezinha K. A. Ribeiro[1]

[1] Statistics Department, University of Brasília, Brasília, DF, 70910-900, Brazil; cira@unb.br; mnsl.meiter@gmail.com; terezinha.ribeiro@unb.br
* Correspondence: cira@unb.br

Abstract: The bimodal generalized extreme value (BGEV) distribution was first introduced by [16]. This distribution offers greater flexibility than the generalized extreme value (GEV) distribution for modeling extreme and heterogeneous (bimodal) events. However, applying this model requires a data-centering technique, as it lacks a location parameter. In this work, we investigate the properties of the BGEV distribution as redefined in [17], which incorporates a location parameter, thereby enhancing its flexibility in practical applications. We derive explicit expressions for the probability density, the hazard rate, and the quantile function. Furthermore, we establish the identifiability property of this new class of BGEV distributions and compute expressions for the moments, the moment-generating function, and entropy. The applicability of the new model is illustrated using climate data.

Keywords: Heterogeneous data; Bimodal GEV distribution; Properties.

Received:
Revised:
Accepted:
Published:

Citation: Otiniano, C. E. G.; Lisboa, M. N. S.; Ribeiro, T. K. A. A Revised Bimodal Generalized Extreme Value Distribution: Theory and Climate Data Application. *Entropy* **2025**, *1*, 0. https://doi.org/

1. Introduction

The Fréchet, Weibull, and Gumbel extreme value distributions ([6], [8]) are genuine probabilistic models for extreme event data, as they correspond to the asymptotic distribution of statistics extreme of independent and identically distributed random variables. The generalized extreme value (GEV) distribution, presented by [10], summarizes the three extreme distributions. For this reason, the GEV distribution is widely used to model extreme events across various fields, including insurance, finance, and hydrology. The theory and applications of the GEV distribution are thoroughly discussed in the books [5], [12], [9], [19], and [21], among others.

A continuous random variable X has a GEV distribution, $X \sim \text{GEV}(\xi, \mu, \sigma)$, if its cumulative distribution function (CDF) and probability density function (PDF) are given, respectively, by

$$F(y;\xi,\mu,\sigma) = \begin{cases} \exp\left\{-\left[1+\xi\left(\frac{y-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}, & \text{if } \xi \neq 0, \\ \exp\left\{-\exp\left[-\left(\frac{y-\mu}{\sigma}\right)\right]\right\}, & \text{if } \xi = 0, \end{cases} \tag{1}$$

and

$$f(y;\xi,\mu,\sigma) = \begin{cases} \frac{1}{\sigma}\left[1+\xi\left(\frac{y-\mu}{\sigma}\right)\right]^{-1/\xi-1} \exp\left\{-\left[1+\xi\left(\frac{y-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}, & \text{if } \xi \neq 0, \\ \frac{1}{\sigma}\exp\left\{-\left(\frac{y-\mu}{\sigma}\right) - \exp\left[-\left(\frac{y-\mu}{\sigma}\right)\right]\right\}, & \text{if } \xi = 0, \end{cases} \tag{2}$$

with shape parameter $\xi \in \mathbb{R}$, scale parameter $\sigma > 0$, and location parameter $\mu \in \mathbb{R}$. The parameter ξ determines the weight of the tail of the distribution. The GEV distribution accommodates heavy-tailed and light-tailed distributions and is characterized by its unimodal shape. Some of the unimodal generalizations of the GEV distribution are: the transmuted GEV distribution ([1], [14],[15]); the dual gamma generalized extreme value distribution, the exponentiated generalized extreme value distribution [14]; the blended generalized extreme value distribution [13].

In various applications, extreme climate data, such as wind speed, humidity, and temperature, exhibit heterogeneous (bimodal) densities with rare events and heavy tails. A very promising model for extreme heterogeneous data is the bimodal GEV distribution, as defined in [16].

Following [16], a random variable X has a bimodal GEV (BGEV) distribution; $X \sim F_{\text{BGEV}}(\cdot;\xi,\mu,\sigma,\delta)$, if its cumulative distribution function is given by

$$F_{\text{BGEV}}(x;\quad \xi,\mu,\sigma,\delta) = F_{\xi,\mu,1}\big(T_{\sigma,\delta}(x)\big), \tag{3}$$

with $\xi \in \mathbb{R}$, $\mu \in \mathbb{R}$, $\sigma > 0$, $\delta > -1$, and the transformation $T_{\sigma,\delta}(\cdot)$ defined by

$$T_{\sigma,\delta}(x) = \sigma x|x|^{\delta}, \quad x \in \mathbb{R}, \tag{4}$$

is invertible and differentiable.

The disadvantage of the model (3) is that its four parameters are shape parameters. In this distribution, there are no location and scale parameters. In other words, σ is not a scale parameter, and μ is not a location parameter, as is the case with the GEV distribution. Furthermore, the local minimum of the PDF is always located at zero. This limitation complicates its applicability, as real bimodal data can have local minimum at any value of the real line.

The chief goal of this paper is to examine several properties of the new BGEV distribution, redefined by [17], which includes a location parameter, and to illustrate its applicability. Specifically, this paper complements the work of [17] in three directions. First, it presents the proof of the identifiability of the new bimodal GEV, which is crucial for the practical application of this model. Second, it presents expressions for the moments, the moment-generating function, and the differential entropy of the new BGEV model. Third, it presents a real data application of the BGEV distribution in a scenario where a bimodal model for extreme data is needed.

The remainder of this paper is structured as follows. Section 2 presents the main results of this work. We begin in Subsection 2.1, with the definition of the main functions related to the new BGEV model. Next, in Subsection 2.2, we provide a graphical illustration of the new BGEV model. Finally, in Subsection 2.3, we show the main properties of the new BGEV distribution. Section 3 contains an application of the new bimodal BGEV model to climate data. Finally, Section 4 closes the paper with some concluding remarks.

2. The new BGEV distribution

Initially, in Subsection 2.1, we show how the model (3) was redefined by [17], presenting the cumulative distribution function, probability density function, failure rate function, and quantile function. The versatility of the BGEV distribution is illustrated through graphical representations in Subsection 2.2. The main results of this work are the properties of the new BGEV distribution, which are in Subsection 2.3.

2.1. The redefined BGEV distribution

Definition 1. *A random variable X has bimodal GEV distribution with location parameter,* $X \sim \mathrm{BGEV}(\xi, \mu, \sigma, \delta)$*, if CDF is given by*

$$F(x;\xi,\mu,\sigma,\delta) = F(T(x);\xi,0,\sigma), \tag{5}$$

where

$$T(x) = (x-\mu)|x-\mu|^{\delta}, \ \ \delta > -1, \mu \in \mathbb{R}. \tag{6}$$

The inverse and derivative functions of $T(\cdot)$ are, respectively, given by

$$T^{-1}(x) = \mathrm{sgn}(x)|x|^{1/(\delta+1)} + \mu \tag{7}$$

and

$$T'(x) = (\delta+1)|x-\mu|^{\delta}. \tag{8}$$

The expressions (7) and (8) allow us to obtain the following PDF of $X \sim \mathrm{BGEV}(\xi,\mu,\sigma,\delta)$ given by

$$f(x;\xi,\mu,\sigma,\delta) = \begin{cases} \dfrac{1}{\sigma}\left[1+\xi\left(\dfrac{T(x)}{\sigma}\right)\right]^{(-1/\xi)-1} \exp\left[-\left[1+\xi\left(\dfrac{T(x)}{\sigma}\right)\right]^{-1/\xi}\right] T'(x), & \xi \neq 0, \\ \dfrac{1}{\sigma}\exp\left(-\dfrac{T(x)}{\sigma}\right)\exp\left[-\exp\left(-\dfrac{T(x)}{\sigma}\right)\right] T'(x), & \xi = 0, \end{cases} \tag{9}$$

whose support is

$$\mathrm{Support}(f(\cdot;\xi,\mu,\sigma,\delta)) = \begin{cases} \left[\mu - \left(\dfrac{\sigma}{\xi}\right)^{1/(\delta+1)}, +\infty\right), & \text{if } \xi > 0, \\ \left(-\infty, \mu + \left|\dfrac{\sigma}{\xi}\right|^{1/(\delta+1)}\right], & \text{if } \xi < 0, \\ (-\infty,+\infty), & \text{if } \xi = 0. \end{cases}$$

In the BGEV model, ξ, δ, and σ are shape parameters, while μ is a location parameter. It is important to note that in the GEV distribution (1), σ is a scale parameter; however, in (5) σ is not a scale parameter, because does not satisfy the condition

$$f(x;\xi,\mu,\sigma,\delta) = \frac{1}{\sigma} f\left(\frac{x}{\sigma};\xi,\mu,1,\delta\right),$$

since $T(x)/\sigma \neq T(x/\sigma)$.

On the other hand, the parameter μ in (5) is a location parameter. To prove this, it suffices to observe that $T(x;\mu,\delta) = T(x-\mu;0,\delta)$ and

$$\begin{aligned} F(x;\xi,\mu,\sigma,\delta) &= F(T(x-\mu);\xi,0,\sigma) \\ &= F(T(x);\xi,\mu,\sigma). \end{aligned}$$

Thus, $f(x;\xi,\mu,\sigma,\delta) = f(x-\mu;\xi,0,\sigma,\delta)$.

The model (5) is a generalization of the GEV distribution, because when $\delta = 0$ the BGEV distribution returns to GEV distribution. That is, $X \sim \mathrm{BGEV}(\xi,\mu,\sigma,0) = \mathrm{GEV}(\xi,\mu,\sigma)$.

From the expressions in (5) and (9), it is simple to obtain the survival and hazard functions. These functions are useful in the area of reliability and for calculating risk measures in other areas. The survival and hazard functions are given by the expressions:

$$S(y) = \begin{cases} 1 - \exp\left[-\left[1 + \xi \left(\frac{T(y)}{\sigma} \right) \right]^{-1/\xi} \right], & \xi \neq 0 \\ 1 - \exp\left[-\exp\left[-\frac{T(y)}{\sigma} \right] \right], & \xi = 0 \end{cases}$$

and

$$h(y) = \begin{cases} \dfrac{\frac{1}{\sigma}\left[1 + \xi \left(\frac{T(y)}{\sigma} \right) \right]^{(-1/\xi)-1} \exp\left[-\left[1 + \xi \left(\frac{T(y)}{\sigma} \right) \right]^{-1/\xi} \right] T'(y)}{1 - \exp\left[-\left[1 + \xi \left(\frac{T(y)}{\sigma} \right) \right]^{-1/\xi} \right]}, & \xi \neq 0 \\ \dfrac{\exp\left(-\frac{T(y)}{\sigma} \right) \exp\left[-\exp\left(-\frac{T(y)}{\sigma} \right) \right] T'(y)}{1 - \exp\left[-\exp\left[-\frac{T(y)}{\sigma} \right] \right]}, & \xi = 0, \end{cases}$$

respectively.

An important property of the new BGEV model is that its quantile function has a simple closed-form expression. This feature is extremely useful for simulation procedures and the calculation of risk measures in various applied fields.

From (1), (5), and (7) we have that the quantile function of BGEV model is given by

$$Q(y) = \begin{cases} \operatorname{sgn}\left\{ \frac{\sigma}{\xi} \left[(-\log(y))^{-\xi} - 1 \right] \right\} \left| \frac{\sigma}{\xi} \left[(-\log(y))^{-\xi} - 1 \right] \right|^{1/(\delta+1)} + \mu, & \text{if } \xi \neq 0 \\ \operatorname{sgn}[-\sigma \log(-\log(y))] \left| -\sigma \log(-\log(y)) \right|^{1/(\delta+1)} + \mu, & \text{if } \xi = 0. \end{cases}$$

2.2. Graphic illustrations of new BGEV distribution

The versatility of the new PDF, defined in (9), is illustrated in Figures 1 - 4. Depending on the combination of parameters, the PDF can be unimodal or bimodal, symmetric or asymmetric, and have a heavy or light tail. To better understand the role of each of the four parameters in the PDF, we consider four scenarios. In each scenario, we fix three parameters and let the fourth parameter vary to understand its effect on the curves. In each of the Figures 1 - 4 are the graphs of the PDF (f), CDF (F), survival (S), and hazard functions (h) of $X \sim \text{BGEV}(\xi, \mu, \sigma, \delta)$.

In Figures 1 - 3, the graphs of the four functions change as the values of ξ, δ, and σ vary. This illustrates our comment above that the parameters ξ, δ, and σ are shape parameters. Figure 1 shows the effect of the parameter δ on the curves. When $\delta = 0$ the PDF is unimodal and is bimodal for $\delta > 0$. Furthermore, the larger the value of δ, the further apart and larger the modes are and the heavier the tails. The effect of the parameter ξ on the curves is shown in Figure 2. As ξ increases, the density tails are heavier and the asymmetry becomes more evident. In Figure 3 one can see that the parameter σ also modifies the PDF. The parameter σ is not a scale parameter, since the PDF remains fixed at the local minimum. This confirms our proof above that σ is not a scale parameter. In Figure 4, the PDF only moves with the variation of μ. This also confirms that μ is a location parameter. Regarding the hazard function h, by depending on the combination of model parameters, the h function is increasing, decreasing, unimodal, N-shaped, or M-shaped. In other words, the BGEV distribution is quite flexible for modeling data in the survival/reliability.

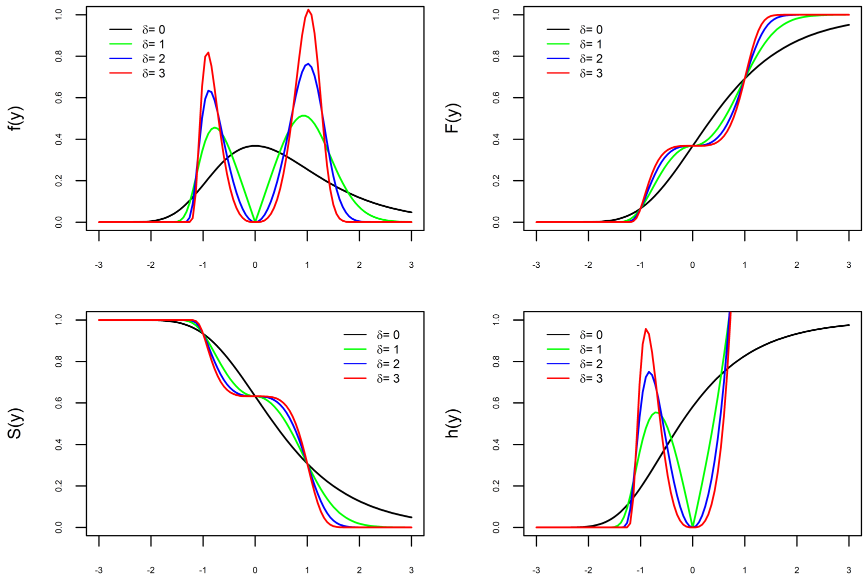

Figure 1. Graphs of $X \sim \text{BGEV}(0,0,1,\delta)$ with δ **varying**: PDF (top left), CDF (top right), Survival (bottom left), and Hazard (bottom right).

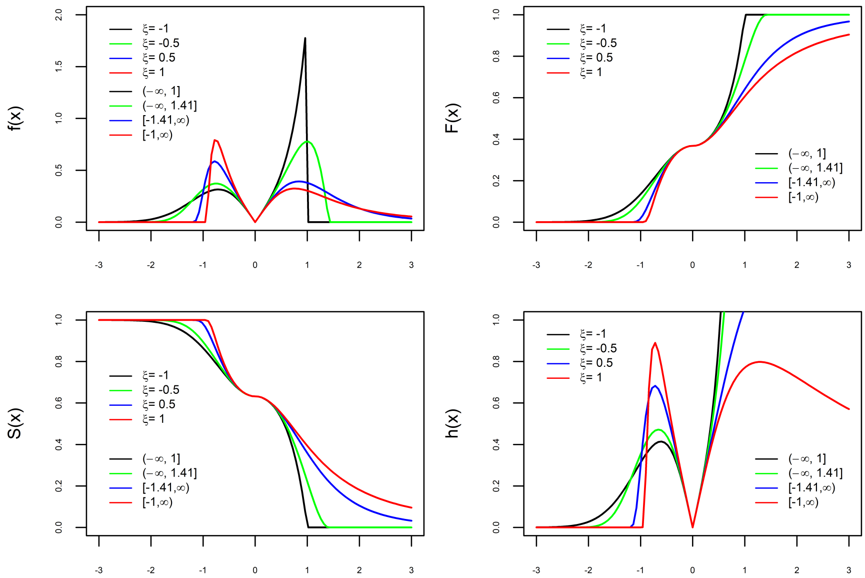

Figure 2. Graphs of $X \sim \text{BGEV}(\xi,0,1,1)$ with ξ **varying**: PDF (top left), CDF (top right), Survival (bottom left), and Hazard (bottom right).

2.3. Properties

2.3.1. Identifiability

In statistics, identifiability is an important property that a family of distributions must satisfy for accurate inference. A distribution from a family is identifiable if different values of the parameters should produce different probability distributions. In other words, the parameter of the distribution is unique. The following shows that the family of distributions of $F(\cdot;\xi,\mu,\sigma,\delta)$ in (5) is identifiable. In addition, we obtain other properties, including formulas for the moments and quantile functions.

Let $\mathcal{F} = \{F = F(\cdot;\theta)\}$ be a family of CDFs. This class $\mathcal{F}$ is identifiable if and only if for any $F_1 = F(\cdot;\theta_1), F_2 = F(\cdot;\theta_2) \in \mathcal{F}$ the equality $F_1 = F_2$ implies $\theta_1 = \theta_2$.

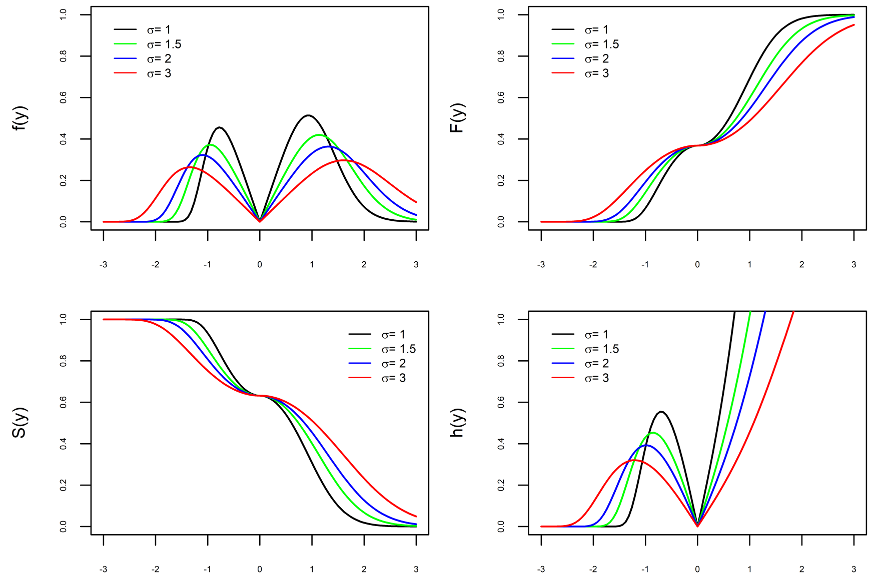

Figure 3. Graphs of $X \sim \mathrm{BGEV}(0, 0, \sigma, 1)$ with $\boldsymbol{\sigma}$ **varying**. PDF (top left), CDF (top right), Survival (bottom left), and Hazard (bottom right).

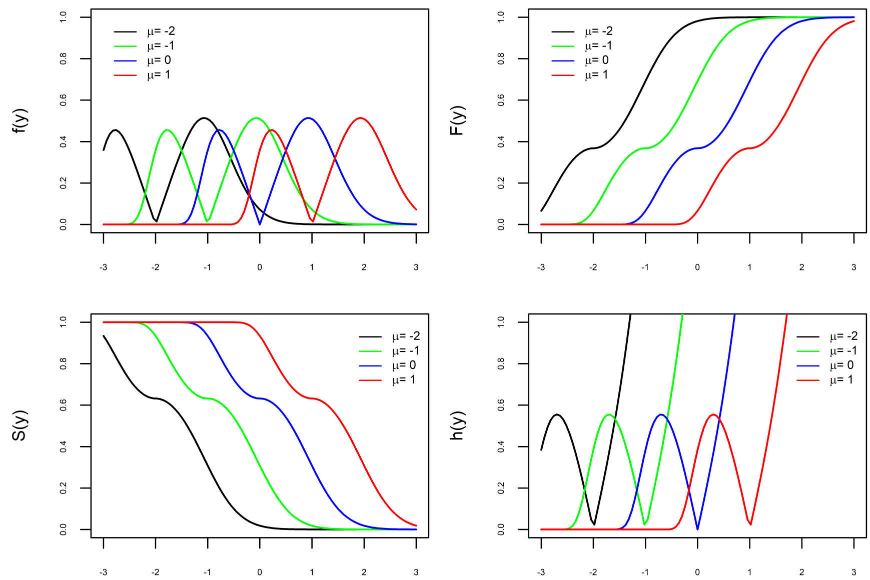

Figure 4. Graphs of $X \sim \mathrm{BGEV}(0, \mu, 1, 1)$ with $\boldsymbol{\mu}$ **varying**: PDF (top left), CDF (top right), Survival (bottom left), and Hazard (bottom right).

Proposition 1. *The family of BGEV distributions* $\mathcal{F}_{BGEV} = \{F(\cdot; \xi, \mu, \sigma, \delta) : F(\cdot; \xi, \mu, \sigma, \delta)$ *as* (5)$\}$ *is identifiable.*

Proof. The authors of [7] demonstrated the identifiability of the finite mixture of GEV distributions, particularly that the family of a GEV component is identifiable. That is, the family $\mathcal{F}_G = \{F(\cdot; \xi, \mu, \sigma) : F(\cdot; \xi, \mu, \sigma)$ as (1)$\}$ is identifiable. Thus, to prove that for any $F(\cdot; \xi_1, \mu_1, \sigma_1, \delta_1)), \ F(\cdot; \xi_2, \mu_2, \sigma_2, \delta_2) \in \mathcal{F}_{BGEV}$ equality

$$F(\cdot; \xi_1, \mu_1, \sigma_1, \delta_1) = F(\cdot; \xi_2, \mu_2, \sigma_2, \delta_2)$$

implies $\xi_1 = \xi_2, \mu_1 = \mu_2, \sigma_1 = \sigma_2, \delta_1 = \delta_2$. This occurs if

$$\exp\left\{-\left[1+\xi_1\left(\frac{(x-\mu_1)|x-\mu_1|^{\delta_1}}{\sigma_1}\right)\right]^{-1/\xi_1}\right\} = \exp\left\{-\left[1+\xi_2\left(\frac{(x-\mu_2)|x-\mu_2|^{\delta_2}}{\sigma_2}\right)\right]^{-1/\xi_2}\right\}, \tag{10}$$

for $\xi \neq 0$, and

$$\exp\left\{-\exp\left[-\left(\frac{(x-\mu_1)|x-\mu_1|^{\delta_1}}{\sigma}\right)\right]\right\} = \exp\left\{-\exp\left[-\left(\frac{(x-\mu_2)|x-\mu_2|^{\delta_2}}{\sigma}\right)\right]\right\}, \tag{11}$$

for $\xi = 0$. Since $\mathcal{F}_G$ is identifiable, we have that $\xi_1 = \xi_2, \mu_1 = \mu_2$, and $\sigma_1 = \sigma_2$. Therefore equations (10) and (11) occur if and only if

$$|x|^{\delta_1} = |x|^{\delta_2}.$$

This is satisfied if and only if $\delta_1 = \delta_2$.

□

2.3.2. Moments and moment generating function

To calculate the moments of $X \sim \text{BGEV}(\xi, \mu, \sigma, \delta)$, first consider the gamma function, the upper incomplete gamma function, and the lower incomplete gamma function, defined according to [2], respectively, by:

$$\Gamma(a) := \int_0^\infty t^{a-1}e^{-t}dt \tag{12}$$

$$\Gamma(a,x) := \int_x^\infty t^{a-1}e^{-t}dt \tag{13}$$

and

$$\gamma(a,x) := \int_0^x t^{a-1}e^{-t}dt, \tag{14}$$

where $a \in \mathbb{R}^+$.

Proposition 2. *Let $X \sim \text{BGEV}(\xi, \mu, \sigma, \delta)$ with $\xi \neq 0$, then the k-th integer moment of X is given by*

$$\begin{aligned} E(X^k) &= \sum_{j=0}^{k}\binom{k}{j}(-1)^{\frac{(k-j)(\delta+2)}{\delta+1}}\left(\frac{\sigma}{\xi}\right)^{\frac{k-j}{\delta+1}}\left[\sum_{i=0}^{[|\frac{k-j}{\delta+1}|]}\binom{[|\frac{k-j}{\delta+1}|]}{i}(-1)^i\gamma\left(1-\xi\left([|\frac{k-j}{\delta+1}|]-i\right),1\right)\right] \\ &+ \sum_{j=0}^{k}\binom{k}{j}\left(\frac{\sigma}{\xi}\right)^{\frac{k-j}{\delta+1}}\left[\sum_{i=0}^{[|\frac{k-j}{\delta+1}|]}\binom{[|\frac{k-j}{\delta+1}|]}{i}(-1)^i\Gamma\left(1-\xi\left([|\frac{k-j}{\delta+1}|]-i\right),1\right)\right], \end{aligned} \tag{15}$$

when $\xi > 0$*, whenever* $k < \frac{\delta+1}{\xi}$ *and*

$$
\begin{aligned}
E(X^k) &= \sum_{j=0}^{k}\binom{k}{j}(-1)^{\frac{(k-j)(\delta+2)}{\delta+1}}\left(\frac{\sigma}{\xi}\right)^{\frac{k-j}{\delta+1}}\left[\sum_{i=0}^{[|\frac{k-j}{\delta+1}|]}\binom{[|\frac{k-j}{\delta+1}|]}{i}(-1)^i\Gamma\left(1-\xi\left([|\frac{k-j}{\delta+1}|]-i\right),1\right)\right] \\
&+ \sum_{j=0}^{k}\binom{k}{j}\left(\frac{\sigma}{\xi}\right)^{\frac{k-j}{\delta+1}}\left[\sum_{i=0}^{[|\frac{k-j}{\delta+1}|]}\binom{[|\frac{k-j}{\delta+1}|]}{i}(-1)^i\gamma\left(1-\xi\left([|\frac{k-j}{\delta+1}|]-i\right),1\right)\right],
\end{aligned}
$$

when $\xi < 0$*, whenever* $\xi < \frac{\delta+1}{k}$.

Proof. By definition

$$
E(X^k) = \int_{-\infty}^{+\infty} x^k f(T(x);\xi,0,\sigma)T'(x)dx, \tag{16}
$$

where $f(\cdot;\xi,0,\sigma)$ is defined in (2), T as in (6), and T' given in (8). By substituting $y = T(x)$ into (16), the moments are expressed as follows:

$$
E(X^k) = \int_{-\infty}^{+\infty} [\mathrm{sng}(y)|y|^{\frac{1}{\delta+1}} + \mu]^k f(y;\xi,0,\sigma)dy. \tag{17}
$$

As $k \in \mathrm{Z}^+$, the Newton Binomial formula is used, so (17) is updated by the integral

$$
\begin{aligned}
E(X^k) &= \sum_{j=0}^{k}\binom{k}{j}\mu^j\left[\int_{-\infty}^{+\infty}[\mathrm{sng}(y)]^{k-j}|y|^{\frac{k-j}{\delta+1}}f(y;\xi,0,\sigma)dy\right] \\
&= \sum_{j=0}^{k}\binom{k}{j}\mu^j(-1)^{\frac{(k-j)(\delta+2)}{\delta+1}}E\left(Y^{\frac{k-j}{\delta+1}}I_{[Y<0]}\right) + \sum_{j=0}^{k}\binom{k}{j}\mu^jE\left(Y^{\frac{k-j}{\delta+1}}I_{[Y\geq 0]}\right),
\end{aligned} \tag{18}
$$

where $Y \sim \mathrm{GEV}(\xi,0,\sigma)$ and I_A is the indicator function of the set A; $I_A(\omega) = 1$ if $\omega \in A$ and $I_A(\omega) = 0$ otherwise. Now we need to analyze the cases where $\xi > 0$ and $\xi < 0$.

Case $\xi > 0$: By replacing (2) in (18) and changing the variable $t = \left[1+\frac{\xi}{\sigma}y\right]^{-1/\xi}$, for $r = [|\frac{k-j}{\delta+1}|]$, it follows that

$$
\begin{aligned}
E(Y^kI_{[Y\geq 0]}) &= \int_0^{+\infty} y^r\frac{1}{\sigma}\left[1+\frac{\xi}{\sigma}y\right]^{-\frac{1}{\xi}-1}\exp\left\{-\left[1+\frac{\xi}{\sigma}y\right]^{-\frac{1}{\xi}}\right\}dy \\
&= \int_1^{+\infty}\left(\frac{\sigma}{\xi}t^{-\xi}-\frac{\sigma}{\xi}\right)^r e^{-t}dt.
\end{aligned} \tag{19}
$$

Newton's Binomial is used in (19), and we obtain

$$
E(Y^kI_{[Y\geq 0]}) = \sum_{i=0}^{r}\binom{r}{i}(-1)^i\left(\frac{\sigma}{\xi}\right)^r\int_1^{+\infty}t^{-\xi(r-i)}e^{-t}dt. \tag{20}
$$

In the same way, the k-th moment of Y truncated in the negative part is

$$\begin{aligned}
E(Y^k I_{[Y<0]}) &= \int_{-\frac{\sigma}{\xi}}^{0} y^r \frac{1}{\sigma}\left[1+\frac{\xi}{\sigma}y\right]^{-\frac{1}{\xi}-1} \exp\left\{-\left[1+\frac{\xi}{\sigma}y\right]^{-\frac{1}{\xi}}\right\}dy \\
&= \int_0^1 \left(\frac{\sigma}{\xi}t^{-\xi}-\frac{\sigma}{\xi}\right)^r e^{-t}dt \\
&= \sum_{i=0}^{r}\binom{r}{i}(-1)^i\left(\frac{\sigma}{\xi}\right)^r \int_0^1 t^{-\xi(r-i)}e^{-t}dt. \qquad (21)
\end{aligned}$$

The lower incomplete gamma function (14) and the upper incomplete gamma function (13) are used to represent the integrals of (20) and (21), respectively. Consequently, the proof of (15) follows by substituting these updates into equation (18).

Case $\xi < 0$: The same procedure as in the case where $\xi > 0$ is repeated respecting the support $\{y : y \in (-\infty - \frac{\sigma}{\xi}]\}$ of $Y \sim \text{GEV}(\xi, 0, \sigma)$. $\square$

Remark. From Proposition 2, we have that for $\xi > 0,\ E(X^k)$ is finite for $k < \frac{\delta+1}{\xi}$. That is, the two shape parameters ξ and δ influence the weight of the tail of the new distribution. Consequently, the tail index of the new BGEV distribution is $\frac{\delta+1}{\xi}$. That is, the right tail of the BGEV distribution can be heavier than the tail of the GEV distribution.

In the following corollary, from Proposition 2, we obtain a known result.

Corollary 1. *Let $X \sim \text{BGEV}(\xi, \mu, \sigma, 0) = \text{GEV}(\xi, \mu, \sigma)$ where $\xi \neq 0$. The k-th moment of X is given by*

$$E(X^k) = \sum_{j=0}^{k}\binom{k}{j}\mu^j\left(\frac{\sigma}{\xi}\right)^{k-j}\Gamma(1-\xi(k-j)).$$

Proof. From (18), when $\delta = 0$, we obtain

$$\begin{aligned}
E(X^k) &= \sum_{j=0}^{k}\binom{k}{j}\mu^j E\left(Y^{k-j}\right) \\
&= E\left(\sum_{j=0}^{k}\binom{k}{j}\mu^j Y^{k-j}\right) \\
&= E(Y+\mu)^k,
\end{aligned}$$

where $Y \sim \text{GEV}(\xi, 0, \sigma)$ and $Y + \mu \sim \text{GEV}(\xi, \mu, \sigma)$. The proof ends with the use of the expressions (20) and (21) and the fact that $\Gamma(x, s) + \gamma(x, s) = \Gamma(x)$. $\square$

The mean of a random variable $X \sim \text{BGEV}(\xi, \mu, \sigma, \delta)$ exists when $\xi < \delta + 1$. It is given in the following Corollary.

Corollary 2. *Let $X \sim \mathrm{BGEV}(\xi, \mu, \sigma, \delta)$ with $\xi \neq 0$. Then, for $\xi > 0$*

$$
\begin{aligned}
E(X) &= (-1)^{\frac{\delta+2}{\delta+1}} \left(\frac{\sigma}{\xi}\right)^{\frac{1}{\delta+1}} \left[\sum_{i=0}^{1} \binom{1}{i} (-1)^i \gamma(1-\xi(1-i), 1)\right] \\
&+ \left(\frac{\sigma}{\xi}\right)^{\frac{1}{\delta+1}} \left[\sum_{i=0}^{1} \binom{1}{i} (-1)^i \Gamma(1-\xi(1-i), 1)\right], \ \textit{for } \xi > 0,
\end{aligned}
$$

and

$$
\begin{aligned}
E(X) &= (-1)^{\frac{\delta+2}{\delta+1}} \left(\frac{\sigma}{\xi}\right)^{\frac{1}{\delta+1}} \left[\sum_{i=0}^{1} \binom{1}{i} (-1)^i \Gamma(1-\xi(1-i), 1)\right] \\
&+ \left(\frac{\sigma}{\xi}\right)^{\frac{1}{\delta+1}} \left[\sum_{i=0}^{1} \binom{1}{i} (-1)^i \gamma(1-\xi(1-i), 1)\right], \ \textit{for } \xi < 0.
\end{aligned}
$$

For $\xi = 0$ an expression of the moment generating function was obtained. It is given in the following proposition.

Proposition 3. *Let $X \sim \mathrm{BGEV}(\xi, \mu, \sigma, \delta)$ with $\xi = 0$. The moment generating function of X is given by*

$$
M_X(t) = e^{\mu t} \sum_{k=0}^{\infty} \frac{t^k}{k!} \left[(-1)^{\frac{k(\delta+2)}{\delta+1}} E\left(Y^{\frac{k}{\delta+1}} I_{[Y<0]}\right) + E\left(Y^{\frac{k}{\delta+1}} I_{[Y \geq 0]}\right)\right], \tag{22}
$$

where $Y \sim \mathrm{GEV}(0, 0, \sigma)$.

Proof. By definition, we have

$$
M_X(t) = \int_{-\infty}^{+\infty} \frac{1}{\sigma} \exp\{tx\} \exp\left\{-\frac{T(x)}{\sigma} - \exp\left[-\frac{T(x)}{\sigma}\right]\right\} dx. \tag{23}
$$

When using the substitution $y = T(x)$ in (23) and the fact that $x = \mathrm{sgn}(\ln y^{-œ}) |\ln y^{-œ}|^{\frac{1}{ffi+1}} +$ ¯ we have what

$$
M_X(t) = e^{\mu t} \int_0^{+\infty} \exp\left\{\mathrm{sgn}(\ln y^{-œ}) |\ln y^{-œ}|^{\frac{1}{ffi+1}} \mathrm{t}\right\} \exp\{-y\} dy. \tag{24}
$$

The new substitution $s = \ln(y^{-\sigma})$ allows you to update (24) by

$$
M_X(t) = \frac{e^{\mu t}}{\sigma} \int_{-\infty}^{+\infty} \exp\left\{\mathrm{sgn(s)} |\mathrm{s}|^{\frac{1}{ffi+1}} \mathrm{t}\right\} \exp\left\{-\frac{s}{\sigma} - \exp\left[-\frac{s}{\sigma}\right]\right\} dy. \tag{25}
$$

Finally, the series representation of the exponential function is used. Thus (25) is rewritten by the equation

$$
\begin{aligned}
M_X(t) &= e^{\mu t} \sum_{k=0}^{-\infty} \frac{t^k}{k!} (-1)^{\frac{k(\delta+2)}{\delta+1}} \int_{-\infty}^{0} \frac{1}{\sigma} s^{\frac{k}{\delta+1}} \exp\left\{-\frac{s}{\sigma} - \exp\left[-\frac{s}{\sigma}\right]\right\} ds \\
&+ e^{\mu t} \sum_{k=0}^{-\infty} \frac{t^k}{k!} \int_{0}^{+\infty} \frac{1}{\sigma} s^{\frac{k}{\delta+1}} \exp\left\{-\frac{s}{\sigma} - \exp\left[-\frac{s}{\sigma}\right]\right\} ds.
\end{aligned}
$$

□

The following result is a particular case of (23). It coincides with the moment-generating function of the Gumbel distribution ([12]).

Corollary 3. *Let $X \sim \text{BGEV}(\xi, \mu, \sigma, 0) = \text{GEV}(\xi, \mu, \sigma)$. The moment generating function of X is given by:*

$$M_X(t) = e^{\mu t}\Gamma(1 - \sigma t). \tag{26}$$

Proof. When $\delta = 0$ the expression (22) reduces to

$$\begin{aligned} M_X(t) &= e^{\mu t} \sum_{k=0}^{\infty} \frac{t^k}{k!} E\left(Y^{\frac{k}{\delta+1}}\right) \\ &= e^{\mu t} E\left(\sum_{k=0}^{\infty} \frac{(tY)^k}{k!}\right) \\ &= e^{\mu t} E\left(e^{tY}\right) \\ &= e^{\mu t}\Gamma(1 - \sigma t), \end{aligned}$$

where $Y \sim \text{GEV}(0, 0, \sigma)$. □

The mean of $X \sim \text{BGEV}(0, \mu, \sigma, \delta)$ always exists. It is given in the following Corollary.

Corollary 4. *Let $X \sim \text{BGEV}(0, \mu, \sigma, \delta)$. The expectation of X is given by:*

$$E(X) = \mu E\left(Y^{1/(\delta+1)}\right) + (-1)^{\frac{\delta+2}{\delta+1}} E\left(Y^{1/(\delta+1)} I_{[Y<0]}\right) + E\left(Y^{1/(\delta+1)} I_{Y \geq 0}\right),$$

where $Y \sim \text{GEV}(0, 0, \sigma)$.

Proof. The proof follows from the derivative of (22) at $t = 0$. □

2.3.3. Entropy

The differential entropy of the BGEV distribution is given in the following proposition.

Proposition 4. *Let $X \sim \text{BGEV}(\xi, \mu, \sigma, \delta)$ with $\xi \geq 0$, then the entropy of X is given by*

$$H(X) = 1 + (1 + \xi)\gamma + \ln\left(\frac{\sigma}{\delta + 1}\right) - \frac{1}{\delta + 1} E[\ln |Y|], \tag{27}$$

where γ is the Euler constant and $Y \sim \text{GEV}(\xi, 0, \sigma)$.

Proof. By definition

$$H(X) = -\int_{-\infty}^{\infty} f(x; \xi, \mu, \sigma, \delta) \ln[f(x; \xi, \mu, \sigma, \delta)] dx. \tag{28}$$

Case $\xi > 0$. From (9), the equation (28) becomes

$$
\begin{aligned}
H(X) &= \int_{-\infty}^{\infty} \ln(\sigma) f(x;\xi,\mu,\sigma,\delta)dx + \int_{-\infty}^{\infty} \left(\frac{1}{\xi}+1\right) \ln\left[1+\xi\frac{T(x)}{\sigma}\right] f(x;\xi,\mu,\sigma,\delta)dx \\
&+ \int_{-\infty}^{\infty} \left[1+\xi\frac{T(x)}{\sigma}\right]^{-1/\xi} f(x;\xi,\mu,\sigma,\delta)dx - \int_{-\infty}^{\infty} \ln(T'(x)) f(x;\xi,\mu,\sigma,\delta)dx.
\end{aligned}
\tag{29}
$$

With the substitution $y = T(x)$ in (29), it follows that

$$
H(X) = \ln\left(\frac{\sigma}{\delta+1}\right) + E(1+\xi Y)^{-1/\xi} + \left(\frac{1}{\xi}+1\right) E[\ln(1+\xi Y)] - \frac{1}{\delta+1} E[\ln|Y|],
$$

where $Y \sim \text{GEV}(\xi, 0, 1)$ as (1). Due to the fact that $E(1+\xi Y)^{-1/\xi} = 1$ and $E[\ln(1+\xi Y)] = \xi\gamma$, the proof of (27) is complete.

Case $\xi = 0$. Again, from (9), the equation (28) becomes

$$
\begin{aligned}
H(X) &= \int_{-\infty}^{\infty} \ln(\sigma) f(x;0,\mu,\sigma,\delta)dx + \int_{-\infty}^{\infty} \frac{T(x)}{\sigma} f(x;0,\mu,\sigma,\delta)dx \\
&= \int_{-\infty}^{\infty} e^{-\frac{T(x)}{\sigma}} f(x;0,\mu,\sigma,\delta)dx - \int_{-\infty}^{\infty} \ln(T'(x)) f(x;0,\mu,\sigma,\delta)dx.
\end{aligned}
\tag{30}
$$

With the substitution $y = T(x)$, the equation (30) is updated by

$$
H(X) = \ln\left(\frac{\sigma}{\delta+1}\right) + E(Y) + E(e^{-Y}) - \frac{1}{\delta+1} E[\ln|Y|]. \tag{31}
$$

where $Y \sim \text{GEV}(0, 0, 1)$. Since $E(Y) = \gamma$ and $E(e^{-Y}) = 1$, the equation (31) proves (27). $\square$

3. Application

In this section, to demonstrate the applicability of the bimodal GEV model, $F(\cdot;\xi,\mu,\sigma,\delta)$, with PDF (9), we use data on the minimum humidity of Goiânia. It is the second most populous city in the Central-West region of Brazil, surpassed only by Brasilia, the capital of Brazil. The city is an important economic hub in the region and is considered a strategic center for areas such as industry, medicine, fashion, and agriculture. In Goiânia, the climate is tropical with a dry season, with two well-defined seasons: rainy (from October to April) and dry (from May to September). In the dry season, relative humidity reaches critical levels, and can be close to 10%, characterizing a state of emergency.

The data used here correspond to the period from January 1, 2011 to December 31, 2022 and come from the automatic weather station A002 in Goiânia. Data recording is hourly. They are available by the National Institute of Meteorology on the website https://portal.inmet.gov.br/. The relative humidity (HUM) is calculated as the percentage of water vapor in the atmosphere.

Table 1 shows the descriptive statistics of HUM. In the period corresponding to the data used here, the minimum humidity recorded was 21.83%.

Table 1. Descriptive statistics of HUM.

Mean	Standard deviation	Median	Maximum	Minimum
64.00	14.54	66.56	91.12	21.83

Since the original data of HUM exhibit temporal dependence and the model $F(\cdot;\xi,\mu,\sigma,\delta)$ is for independent and identically distributed (i.i.d.) data, we first applied the minimum block technique ([11]) to obtain a subsample of minimum values of HUM. At the 5% level, the Ljung-Box test ([4]) verified the serial independence of the subsample of minima for blocks of size $N = 1440$ hours (60 days).

The left panel in Figure 5 shows the histogram of the initial data, and the right panel shows the histogram of the subsample of the minima. This right panel clearly shows that the minima of the data have a bimodal shape. On the other hand, since the bimodal GEV distribution is the limit in the distribution of the extremes, it makes perfect sense to continue fitting these minima by the $F_{\mathrm{BG}}(\cdot;\xi,\mu,\sigma,\delta)$ distribution.

To estimate the parameters of the GEV and bimodal GEV distributions; $F(\cdot;\xi,\mu,\sigma)$ and $F(\cdot;\xi,\mu,\sigma,\delta)$, we use the maximum likelihood technique that is implemented in the EVD [22] and bgev package [17] in the R Project for Statistical Computing [18]. Table 2 shows the estimates and standard errors for the minimum HUM data.

Table 2. Estimates and standard errors under GEV and BGEV distributions for the minimum HUM data.

	BGEV		GEV	
Parameters	Estimates	Std. error	Estimates	Std. error
ξ	-0.37	0.09	-0.58	0.08
μ	41.88	0.49	43.07	1.85
σ	64.31	2.92	14.66	1.53
δ	0.54	0.18	-	-

We note that the estimate of δ is 0.54 which indicates that there is inherent bimodality in this data. The parameter $\delta > 0$ in the BGEV distribution is associated with the presence of bimodality, as previously discussed in Subsection 2.2. This indicates that the bimodal BGEV distribution performs better in fitting this data. To assess the goodness of fit of the minimum HUM data by the BGEV and GEV distributions, we used the Akaike Information Criterion (AIC) ([3]) for both models. The AIC results of the BGEV and GEV distributions were 611.2 and 2550.6, respectively. These results indicate that the BGEV model is more suitable than the GEV model for fitting the minimum relative humidity data of Goiânia. The better performance of the BGEV distribution is further illustrate in the right panel of Figure 5 which presents the histogram of the adjusted GEV and BGEV densities for the minimum relative humidity data.

4. Conclusion

The GEV distribution is a crucial tool for modeling extreme data. However, this distribution is not well-suited for datasets that exhibit bimodal behavior. In this work, we examine a recent extension of the GEV distribution that accommodates bimodal data, known as the BGEV distribution introduced by [16] and later redefined by [17].

In short, the main contributions of this paper are as follows. First, it presents a detailed explanation of the redefinition of the BGEV model. Second, the versatility of the BGEV distribution is illustrated through graphical representations of its PDF, which can be highly flexible, exhibiting unimodal or bimodal characteristics, as well as being symmetric or asymmetric, and possessing either heavy or light tails. Third, it provides a comprehensive proof of key properties of the new BGEV distribution, including identifiability, moments, the moment-generating function, and differential entropy. Fourth, it illustrates the usefulness of the new BGEV distribution, through the application of climate data. Overall, the BGEV distribution is more effective than the GEV distribution when bimodality is inherent in the data.

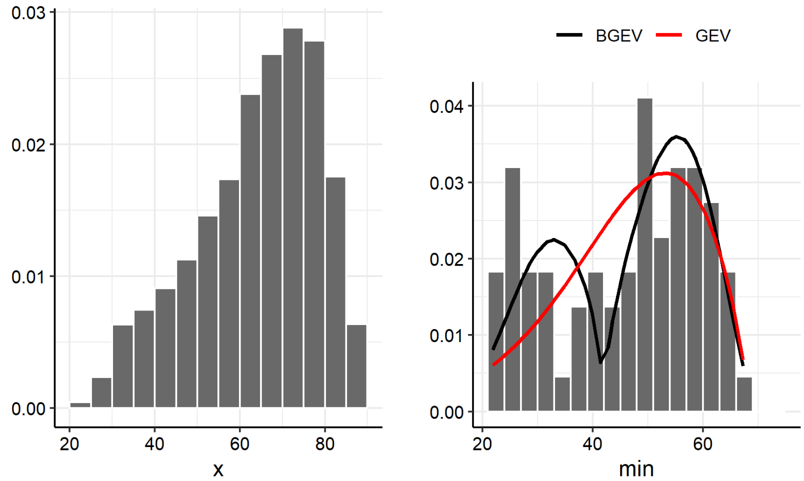

Figure 5. Histogram of the original relative humidity data (left panel) and histogram versus adjusted GEV and BGEV densities for the minimum relative humidity data (right panel).

A natural extension of the present work is the development of a regression model based on the BGEV distribution. The authors of this paper are currently developing a new class of regression models based on a median reparameterization of the redefined BGEV distribution discussed here. This work is in progress and the results will be reported elsewhere. Another promising extension of this work consists in developing time series models with innovations following the BGEV distribution.

Funding: This research was funded by DPI/BCE/UnB through Call for Proposals No. 001/2025 DPI/BCE/UnB.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AIC	Akaike Information Criterion
BGEV	Bimodal Generalized Extreme Value
CDF	Cumulative Distribution Function
GEV	Generalized Extreme Value
HUM	Relative Humidity
PDF	Probability Density Function

References

1. Aryal, G. R.; Tsokos, C. P. On the transmuted extreme value distribution with application. *Nonlinear Anal.* **2009**, 71, 1401–1407.
2. Abramowitz, M.; Stegun, I. A. *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*; Dover Publications: New York, 1965.
3. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control.* **1974**, 19, 716–723.
4. Box, G. E. P.; Pierce, D. A. Distribution of residual correlations in autoregressive-integrated moving average time series models. *J. Am. Stat. Assoc.* **1970**, 65, 1509–1526.
5. Embrechts, P.; Klüppelberg, C.; Mikosch, T. *Modelling Extremal Events: for Insurance and Finance*; Springer-Verlag: Berlin, 1997.
6. Fisher, R.A.; Tippett, L. H. C. Limiting forms of the frequency distribution of the largest and smallest member of a sample. *Proc. Cambridge Philos. Soc.* **1928**, 24, 180–190.
7. Gonçalves, C. R.; Otiniano, C. E. G.; Crivinel, E. C. Estimation of a nonlinear discriminant function from a mixture of two GEV distributions. *J Stat Comput Sim.* **2018**, 88, 1147–1171.
8. Gnedenko, B. V. Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals. Math.* **1943**, 44, 423–453.
9. Haan, L.; Ferreira, A. *Extreme Value Theory: An Introduction*; Springer: New York, 2006.

10. Jenkinson, A. F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Q. J. R. Meteorol. Soc.* **1955**, 81, 158–171.
11. Jondeau, E.; Poon, S. H.; Rockinger, M. *Financial Modeling Under Non-Gaussian Distributions*; Springer: New York, 2007.
12. Kotz, S.; Nadarajah, S. *Extreme Value Distributions: Theory and Applications*; Imperial College Press: London, 2000.
13. Krakauer, N. Y. Extending the blended generalized extreme value distribution. *Discovery. civ. eng.* **2024**, 1, 97.
14. Nascimento, F.; Bourguignon, M.; Leão, J. Extended generalized extreme value distribution with applications in environmental data. *Hacettepe. J. Math. Stat.*. **2016**, 45, 1847–1864.
15. Otiniano, C. E. G.; Paiva, B. S.; Neto, D. S. B. M. The transmuted GEV distribution: properties and application. *Commun. Stat. Appl. Methods*. **2019**, 26, 239–259.
16. Otiniano, C. E. G.; Paiva, B. S.; Vila, R.; Bourguignon, M. A bimodal model for extremes data. *Environ. Ecol. Stat.*. **2023**, 30, 261–288.
17. Otiniano, C. E. G.; Oliveira, Y. L. S.; Sousa, T. R. Bimodal GEV distribution with location parameter, 2024. https://CRAN.R-project.org/package.bgev.
18. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2022.
19. Reiss, R-D.; Thomas, M. *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*; Birkhäuser: Switzerland, 2007.
20. Resnik, S. *Extreme Values, Regular Variation and Weak Convergence*; Springer-Verlag: New York, 1987.
21. Rudd, E. M.; Jain, L. P.; Scheirer, W. J.; Boult, T. E. The extreme value machine. *IEEE Trans Pattern Anal Mach Intell*. **2017**, 40, 762–768.
22. Stephenson, A. G. evd: Extreme value distributions. *R News*. **2002**, 2, 31–32.