A bimodal model for extremes data

Cira E. G. Otiniano1 · Bianca S. Paiva1 · Roberto Vila1 · Marcelo Bourguignon2

Received: 11 February 2022 / Accepted: 21 May 2023 / Published online: 9 June 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract
In extreme values theory, for a sufficiently large block size, the maxima distribution is approximated by the generalized extreme value (GEV) distribution. The GEV distribution is a family of continuous probability distributions, which has wide applicability in several areas, including hydrology, engineering, science, ecology, and finance. However, the GEV distribution is not suitable for modeling extreme bimodal data. In this paper, we propose an extension of the GEV distribution that incorporates an additional parameter. The additional parameter introduces bimodality and arises tail weight, i.e., this proposed extension is more flexible than the GEV distribution. Inference for the proposed distribution was performed under the likelihood paradigm. A Monte Carlo experiment is conducted to evaluate the performances of these estimators in finite samples with a discussion of the results. Finally, the proposed distribution is applied to environmental data sets, illustrating their capabilities in challenging cases in extreme value theory.

Keywords Bimodality · Environmental data · Generalized extreme value distribution · Maximum likelihood

Bianca S. Paiva, Roberto Vila and Marcelo Bourguignon have contributed equally to this work.

Handling Editor: Luiz Duczmal.

Cira E. G. Otiniano
ciragotiniano@gmail.com
Bianca S. Paiva
bianca.unb@gmail.com
Roberto Vila
rovig161@gmail.com
Marcelo Bourguignon
m.p.bourguignon@gmail.com

1 Department of Statistics, University of Brasília, Darcy Ribeiro, Brasília, DF 70910-900, Brazil
2 Department of Statistics, Federal University of Rio Grande do Norte, Natal, RN 59078-970, Brazil
1 Introduction

The generalized extreme value (GEV) distribution is widely used to model extreme events in several areas, such as finance, insurance, hydrology, and bioinformatics, among others. In journals of statistics and applied areas, a huge amount of articles with applications of the GEV distribution can be found. Theory and applications of extreme value distributions can be found in the books by Kotz and Nadarajah (2000), Coles (2001), Beirlant et al. (2004), Haan and Ferreira (2010), Embrechts et al. (2013), Longin (2016), Scheirer (2017), among others. In journals of statistics and applied areas, a huge amount of articles of extreme value distributions can be found. Theory and applications of extreme value distributions can be found in the books by Kotz and Nadarajah (2000), Coles (2001), Beirlant et al. (2004), Haan and Ferreira (2010), Embrechts et al. (2013), Longin (2016), Scheirer (2017), among others.

A random variable $Y$ has a Generalized Extreme Value (GEV) distribution with shape parameter $\xi$, location parameter $\mu \in \mathbb{R}$, and scale parameter $\sigma > 0$ denoted by $Y \sim F(\cdot; \xi, \mu, \sigma)$, if its probability density function (PDF) is given by

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

valid for $y > \mu - \sigma/\xi$ in the case $\xi > 0$, and $y < \mu - \sigma/\xi$ in the case $\xi < 0$.

The cumulative distribution function (CDF) of the GEV distribution is given by

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

When $\xi = 0$ the pdf Eq. (1) is also known as the Gumbel distribution.

Here a GEV random variable with shape parameter $\xi \in \mathbb{R}$ and standard scale we denote by $Y \sim F_G(\cdot; \xi, \mu) := F(\cdot; \xi, \mu, 1)$. In this case, the PDF and CDF are given respectively by

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

and

$$F_G(y; \xi, \mu) = \begin{cases} \exp \left\{ - \left[ 1 + \xi (y - \mu) \right]^{-1/\xi} \right\}, & \xi \neq 0, \\ - \exp \left[ - (y - \mu) \right], & \xi = 0. \end{cases}$$

In extreme value modeling, in particular, environmental data (see Sect. 6) with bimodality has appeared more and more. The statistical modeling of this type of data requires distributions that capture bimodality. The usual GEV distribution is not suitable for modeling extreme bimodal data. In this context, Nascimento et al. (2016) provided some new extended models to the GEV distribution as a baseline function and the transmuted GEV distribution introduced by Aryal and Tsokos (2011), showing advantages compared with the standard GEV distribution. Eljabri and Nadarajah (2017) studied the Kumaraswamy GEV distribution. However, all the
models cited above are not suitable for capturing bimodality. In this sense, the mixture of GEV distributions and the mixture of extreme value distributions has been an alternative (Otiniano et al. 2016). The difficulty with these models is in the estimation process, as six sub-models must be considered.

This paper presents a new model for bimodal extremes, based on a transformation of the standard GEV, with the hope it yields a “better fit” in certain extremes analysis (see Sect. 6). The inclusion of an additional parameter in the GEV role is to introduce bimodality, that is, the new distribution is more flexible than the GEV distribution, and without computational complications for the estimation of its parameters.

In Sect. 2, we define the BGEV model and derive the main properties of the bimodal GEV distribution. Several graphic illustrations of the BGEV model are shown in Sect. 3. In Sect. 4, the inference procedure is carried out under the likelihood paradigm. Already, in Sect. 5 discusses some simulation results for the estimation method. Two illustrative applications in environmental datasets are investigated and shown in Sect. 6. Conclusions are addressed in Sect. 7.

### 2 Results

In this section, inspired by the reference Swamee and Rathie (2007), we introduce a bimodal distribution of extreme value (BGEV). All results in this section are new in the literature.

#### 2.1 The bimodal GEV distribution

We define a random variable with BGEV distribution and parameters \( \xi, \mu, \sigma, \delta \), denoted \( X \sim F_{\text{BGEV}}(\cdot; \xi, \mu, \sigma, \delta) \), if its PDF is given by

\[
 f_{\text{BGEV}}(x; \xi, \mu, \sigma, \delta) = f_G(T_{\sigma, \delta}(x); \xi, \mu) \cdot T'_{\sigma, \delta}(x) 
 = \begin{cases} 
 [1 + \xi(T_{\sigma, \delta}(x) - \mu)]^{(-1/\xi) - 1} 
 \times \exp \left\{ - \left[ 1 + \xi(T_{\sigma, \delta}(x) - \mu) \right]^{1/\xi} \right\} T'_{\sigma, \delta}(x), & \xi \neq 0, \\
 \exp \left\{ -(T_{\sigma, \delta}(x) - \mu) - \exp \left\{ -(T_{\sigma, \delta}(x) - \mu) \right\} \right\} T'_{\sigma, \delta}(x), & \xi = 0,
\end{cases} 
\]

whose support depends on its parameters. It is

\[
 \text{Support} (f_{\text{BGEV}}) = \begin{cases} 
 \text{sgn}(\mu - \frac{1}{\xi}) \left( \frac{|\mu - \frac{1}{\xi}|}{\sigma} \right)^{1\frac{1}{\xi + 1}}, & \text{if } \xi > 0, \\
 -\infty, \text{sgn}(\mu - \frac{1}{\xi}) \left( \frac{|\mu - \frac{1}{\xi}|}{\sigma} \right)^{1\frac{1}{\xi + 1}}, & \text{if } \xi < 0, \\
 (-\infty, +\infty), & \text{if } \xi = 0.
\end{cases}
\]

The transformation
\[ T_{\sigma, \delta}(x) = \sigma |x|^\delta, \quad x \in \mathbb{R}, \quad \delta > -1, \quad \sigma > 0 \]

is an odd and invertible transformation with derivatives

\[ T'_{\sigma, \delta}(x) = \sigma(\delta + 1)|x|^\delta, \quad T''_{\sigma, \delta}(x) = \text{sgn}(x) \sigma(\delta + 1)\delta|x|^{\delta-1}, \]

and

\[ T^{(k)}_{\sigma, \delta}(x) = \left[ \text{sgn}(x) \right]^{k-1} \sigma \left[ \prod_{i=1}^{k-2} (\delta - i) \right] |x|^{\delta-(k-1)}, \quad k \geq 2. \]

Here, \( \text{sgn}(\cdot) \) denotes the sign function. The transformation \( T_{\sigma, \delta} \) in Eq. (5) is not new in the literature.

Daly et al. (2009) introduced a theoretical derivation of bimodality for soil moisture in humid-land environments, arriving at a density of the form \( r(x) \exp(-\gamma x + T(x)) \), where \( r(\cdot) \) and \( T(\cdot) \) are nonlinear functions and \( \gamma \) is a positive parameter. Swamee and Rathie (2007) used \( T_\delta(x) = x|x|^\delta \) to obtain functions closely approximating the normal and lognormal distributions. We emphasize that, for generating bimodal distributions, \( T_{\sigma, \delta} \) is applicable not only to GEV but also to other unimodal distributions, for example, we can take the Gaussian and the Maxwell-Boltzmann laws. In general, we cannot guarantee that the application of \( T_{\sigma, \delta} \) will always generate bimodality. Recently, several authors have used adequate transformations \( T \) in the original density to generate bimodality in distributions such as: Normal (Elal-Olivero 2010), Birnbaum-Saunders (Fonseca and Cribari-Neto 2018; Vila and Çankaya 2021), Weibull (Vila et al. 2020a) and Gamma (Vila et al. 2020b).

Since \( T_{\sigma, \delta} \) is non-decreasing for \( \delta > -1 \), then, its inverse function, denoted by \( T_{\sigma, \delta}^{-1} \), exists and is given by

\[ T_{\sigma, \delta}^{-1}(x) = \text{sgn}(x) \left( \frac{|x|}{\sigma} \right)^{1/(\delta+1)}. \]

The PDF Eq. (3) is valid for \( x > T_{\sigma, \delta}^{-1}(\mu - 1/\xi) \) in the case \( \xi > 0 \), for \( x < T_{\sigma, \delta}^{-1}(\mu - 1/\xi) \) in the case \( \xi < 0 \), and for \( x \in \mathbb{R} \) in the case \( \xi = 0 \).

Note that, for \( \xi > 0 \), the function \( f_{\text{BEV}} \) is a PDF, because

\[ \int_{T_{\sigma, \delta}^{-1}(\mu - 1/\xi)}^{\infty} f_{\text{BEV}}(x; \xi, \mu, \sigma, \delta) \, dx = \int_{\mu - 1/\xi}^{\infty} f_G(T_{\sigma, \delta}(x); \xi, \mu) \, dT_{\sigma, \delta}(x) = 1. \]

Similarly, it is verified that, for \( \xi \leq 0 \), \( f_{\text{BEV}} \) is a PDF.

The expression Eq. (3) is equivalent to

\[ f_{\text{BEV}}(x; \xi, \mu, \sigma, \delta) = \frac{d}{dx} F_G(T_{\sigma, \delta}(x); \xi, \mu), \]

where
\begin{equation}
F_{G}(T_{\sigma,\delta}(x);\xi, \mu) = F_{\text{BGEV}}(x; \xi, \mu, \sigma, \delta) 
\end{equation}

is the CDF of \( X \).

**Remark 1** Since the GEV distribution \((\xi \neq 0)\) approaches the Gumbel distribution \((\xi = 0)\) as \( \xi \to 0 \), in this work, unless mentioned, we will consider our study of the BGEV distribution only for the general case \( \xi \neq 0 \).

When \( \xi = 0 \), a different approach to bimodality can be found in the reference Otiniano et al. (2023).

When \( \delta = 0 \) the function Eq. (3) is of type Eq. (1); \( F_{\text{BGEV}}(x; \xi, \mu, \sigma, 0) = F_{G}(x; \xi, \mu/\sigma, 1/\sigma) \).

Some mathematical properties as monotonicity, bimodality property, stochastic representation, moments, quantiles and tail behavior of the BGEV distribution are discussed in the next subsection.

### 2.2 Monotonicity

Let \( m_\xi \) be the unique mode for the GEV distribution Eq. (2). To state the next result we define the following quantities: \( x_{\xi}^{\min} = \min\{0, T_{\sigma,\delta}^{-1}(m_\xi)\} \) and \( x_{\xi}^{\max} = \max\{0, T_{\sigma,\delta}^{-1}(m_\xi)\} \).

**Proposition 1** The following monotonicity properties it hold:

A) If \( \xi > 0 \) and \( \delta < 0 \), then the BGEV PDF is increasing for each \( x < x_{\xi}^{\min} \).

B) If \( \xi > 0 \) and \( \delta > 0 \), then the BGEV PDF is decreasing for each \( x < x_{\xi}^{\min} \).

C) If \( \xi < 0 \) and \( \delta < 0 \), then the BGEV PDF is decreasing for each \( x > x_{\xi}^{\max} \).

D) If \( \xi < 0 \) and \( \delta > 0 \), then the BGEV PDF is increasing for each \( x > x_{\xi}^{\max} \).

**Proof** Since \( T_{\sigma,\delta}^\prime(x) = \sigma(\delta + 1)|x|^{\delta} \), for \( \delta < 0 \), \( T_{\sigma,\delta}^\prime(x) \) is decreasing (resp. increasing) for each \( x > 0 \) (resp. \( x < 0 \)). For \( \delta > 0 \), \( T_{\sigma,\delta}^\prime(x) \) is increasing (resp. decreasing) for each \( x > 0 \) (resp. \( x < 0 \)).

On the other hand, since \( m_\xi \) is the mode of the GEV distribution Eq. (2), \( f_{G}(T_{\sigma,\delta}(x);\xi, \mu) \) is increasing (resp. decreasing) for each \( x < x_{\xi}^{\min} \) (resp. \( x > x_{\xi}^{\max} \)) and \( \xi > 0 \) (resp. \( \xi < 0 \)).

Since \( T_{\sigma,\delta}^\prime(x) \) is increasing and nonnegative for all \( x < 0 \), and \( f_{G}(T_{\sigma,\delta}(x);\xi, \mu) \) is increasing for each \( x < x_{\xi}^{\min} \) and \( \xi > 0 \) and \( \delta < 0 \), from definition Eq. (3) of BGEV density, we have that \( f_{\text{BGEV}}(x;\xi, \mu, \sigma, \delta) = f_{G}(T_{\sigma,\delta}(x);\xi, \mu) T_{\sigma,\delta}^\prime(x) \) is the product of two increasing and nonnegative functions. Thus, the BGEV PDF is increasing for each \( x < x_{\xi}^{\min} \). This completes the proof of first item.
The proof of the other items follows the same reasoning as the one of Item 1). \hfill \Box

2.3 Bimodality property

**Proposition 2** The point \( x \) is a critical point of BGEV density Eq. (3) if it is a solution of the following equation:

\[
T''_{\sigma,\delta}(x) - \frac{1 + \xi - \left[ 1 + \xi(T_{\sigma,\delta}(x) - \mu) \right]^{-1/\xi}}{1 + \xi(T_{\sigma,\delta}(x) - \mu)} \left[ T'_{\sigma,\delta}(x) \right]^2 = 0,
\]

where \( T_{\sigma,\delta}(x), T'_{\sigma,\delta}(x) \) and \( T''_{\sigma,\delta}(x) \) are given in Eqs. (5) and (6).

**Proof** The proof is trivial and omitted. \hfill \Box

**Theorem 1** (Uni- or bimodality, case \( \xi > 0 \)) Let \( \xi = 1/k \) be a rational number where \( k > 0 \) is an integer, \( \delta \neq 0, \delta > -1/(1 + \xi) \) and \( \mu < 1/\xi \). The PDF of the BGEV distribution is uni- or bimodal.

**Proof** By replacing the identities \( T_{\sigma,\delta}(0) = T'_{\sigma,\delta}(0) = T''_{\sigma,\delta}(0) = 0 \) in the equation of Proposition 2, we have that \( x = 0 \) is a critical point of BGEV density Eq. (3).

From now on, we assume that \( x \neq 0 \). By using the identities

\[
T'_{\sigma,\delta}(x) = (\delta + 1) \frac{T_{\sigma,\delta}(x)}{x} \quad \text{and} \quad T''_{\sigma,\delta}(x) = (\delta + 1) \frac{T_{\sigma,\delta}(x)}{x^2},
\]

a simple algebraic manipulation shows that the equation of Proposition 2 is equivalent to

\[
1 + \xi[T_{\sigma,\delta}(y) - \mu] - (1 + \xi - \left[ 1 + \xi[T_{\sigma,\delta}(x) - \mu] \right]^{-1/\xi})\delta T_{\sigma,\delta}(x) = 0.
\]

Let \( y = 1 + \xi(T_{\sigma,\delta}(x) - \mu) \). Since \( \xi > 0; y > 0 \). Then the above identity equivalently can be written as

\[
p(y) := ay^{1+1/\xi} - by^{1/\xi} + cy + d = 0 \quad \text{with } y > 0,
\]

where \( a = \xi - (1 + \xi)(\delta + 1), \quad b = (\mu\xi - 1)(1 + \xi)(\delta + 1), \quad c = \delta + 1 \) and \( d = (\mu\xi - 1)(\delta + 1) \). The conditions \( \delta \neq 0, \delta > -1/(1 + \xi) \) and \( \mu < 1/\xi \), ensure that \( a < 0, b < 0, c > 0 \) and \( d < 0 \). So the number of positive roots of the polynomial \( p(y) \) determines the number of remaining roots of the BGEV PDF.

By Descartes’ rule of signs (see Griffiths 1947; Xue 2012), the polynomial \( p(y) \) has two sign changes (the sequence signs is \(-, +, +, -\)), meaning that this polynomial has two or zero positive roots.

Assume that \( p(y) \) has two positive roots, denoted by \( y_1 \) and \( y_2 \). This implies that, in addition to \( x = 0 \), the BGEV PDF has other two critical points given by
Let $A$ be the set formed for all $(\mu, \delta) \in \mathbb{R} \times (-1, \infty)$ such that the following inequalities hold:

$$0 < \sqrt{(\delta + 1)^2[1 - 2(\mu - 1)]^2 - 4[1 - 2(\delta + 1)](\mu - 1)(\delta + 1) < (\delta + 1)[1 - 2(\mu - 1)].$$

By considering $\mu = 1/2$ and $\delta = 1$, we have that $(\mu, \delta) \in A$. That is, the set $A$ is non-empty.

**Corollary 2 (Bimodality, case $\xi = 1$)** Let $\xi = 1$, $\delta \neq 0$, $\delta > -1/2$ and $\mu < 1$. If $(\mu, \delta) \in A$ then the PDF of the BGEV distribution is bimodal.

**Proof** By applying Theorem 1, with $k = 1$, we obtain the uni- or bimodality of the BGEV distribution. In what follows, we prove that only bimodality can occur.

Indeed, as a sub-product of the proof of Theorem 1, it is sufficient to verify that the polynomial $p(y)$, $y > 0$, has exactly two different positive real roots. But, when $\xi = 1$, this polynomial is written as follows

$$p(y) = [1 - 2(\delta + 1)]y^2 + (\delta + 1)[1 - 2(\mu - 1)]y + (\mu - 1)(\delta + 1).$$

The conditions imposed on the corollary ensure that $1 - 2(\delta + 1) < 0$, $(\delta + 1)[1 - 2(\mu - 1)] > 0$ and $(\mu - 1)(\delta + 1) < 0$. Then, by Descartes’ rule of signs, this polynomial has either two (different) positive real roots or zero real roots (in this case, complex). Denote by $y_-$ and $y_+$ to these roots. Bhaskara’s formula provides the following roots:

$$y_{\pm} = \frac{-(\delta + 1)[1 - 2(\mu - 1)] \pm \sqrt{(\delta + 1)^2[1 - 2(\mu - 1)]^2 - 4[1 - 2(\delta + 1)](\mu - 1)(\delta + 1)}}{2[1 - 2(\delta + 1)]}.$$

Since $(\mu, \delta) \in A$, $y_-$ and $y_+$ are two (different) positive real roots. This completes the proof. □

Notice that, in Corollary 4, under the conditions imposed: $\xi = 1$, $\delta \neq 0$, $\delta > -1/2$, $\mu < 1$, by Descartes’ rule of signs, the quadratic polynomial

$$p(y) = [1 - 2(\delta + 1)]y^2 + (\delta + 1)[1 - 2(\mu - 1)]y + (\mu - 1)(\delta + 1)$$

has either two positive real roots or zero real roots (in this case complex). Denote by $y_-$ and $y_+$ to these roots. Since $(\mu, \delta) \in A$ we have that $y_-$ and $y_+$ are two positive roots.

**Theorem 3 (Uni- or bimodality, case $\xi < 0$)** Let $\xi = -1/k$ be a rational number where $k > 0$ is an integer, $\delta \neq 0$, $\delta > -1$ and $\mu > 1/\xi$. The PDF of the BGEV distribution is uni- or bimodal.

[Springer]
Proof Analogously to the proof of Theorem 1, see that $x = 0$ is a critical point of BGEV density.

Let $y = 1 + \xi(T_{\sigma,\delta}(x) - \mu)$ with $x \neq 0$. Since $\xi < 0; \gamma > 0$. Then, in this case, the polynomial $p(y), y > 0$, defined in Theorem 1, can be written as

$$p(y) = ay^{k+1} - by^k + cy + d = 0,$$

or equivalently

$$p(y) = cy^{k+1} + dy^k + ay - b = 0.$$  

Note that the conditions provided in theorem ensure that $a < 0, b < 0, c > 0$ and $d < 0$. Following the same steps as the proof of Theorem 1 we have that $p(y)$ has two or zero positive roots. This guarantees that the BGEV PDF has at most three real critical points (including $x = 0$). Since $\lim_{x \to \pm \infty} f_{BGEV}(x; \xi, \mu, \sigma, \delta) = 0$, the uni- or bimodality of the BGEV PDF follows. \hfill $\Box$

2.4 Stochastic representation of the BGEV distribution

Proposition 3 (Related distributions) Let $X \sim F_{BGEV}(\cdot; \xi, \mu, \sigma, \delta)$.

A) If $X \sim F_{BGEV}(\cdot; \xi, \mu, \sigma, \delta)$ then $Y = T_{\sigma,\delta}(X) \sim F_G(\cdot; \xi, \mu)$.

B) If $Y \sim F_G(\cdot; \xi, \mu)$ then $X = T^{-1}_{\sigma,\delta}(Y) \sim F_{BGEV}(\cdot; \xi, \mu, \sigma, \delta)$.

C) If $X \sim \text{Weibull}(1, \mu)$ then $T^{-1}_{\sigma,\delta}(\mu(1 - \ln X)) \sim F_{BGEV}(\cdot; 0, \mu, \sigma, \delta)$.

D) If $X \sim F_{BGEV}(\cdot; 0, \mu, \sigma, \delta)$ then $\exp\left( - (T_{\sigma,\delta}(X) - \mu) / \mu \right) \sim \text{Weibull}(1, \mu)$.

E) If $X_1 \sim F_{BGEV}(\cdot; 0, \mu_1, \sigma, \delta)$ and $X_2 \sim F_{BGEV}(\cdot; 0, \mu_2, \sigma, \delta)$ are independent random variables, then $T_{\sigma,\delta}(X_1) - T_{\sigma,\delta}(X_2) \sim \text{Logistic}(\mu_1 - \mu_2, 1)$.

F) If $X \sim F_{BGEV}(\cdot; \xi, \mu, \sigma, \delta)$ and $c \neq 0$ a constant, then $cX \sim F_{BGEV}(\cdot; \xi, |c| \mu, c^2 \sigma, \delta)$.

Proof If $X \sim F_{BGEV}(\cdot; \xi, \mu, \sigma, \delta)$, by Eq. (8), follows that

$$P(Y \leq y) = P(X \leq T^{-1}_{\sigma,\delta}(y)) = F_{BGEV}(T^{-1}_{\sigma,\delta}(y); \xi, \mu, \sigma, \delta) = F_G(y; \xi, \mu).$$

This proves the Item (1). Analogously, the proof of the second item follows.

For $X \sim \text{Weibull}(1, \mu)$ (Weibull distribution) it is known that $\mu(1 - \ln X) \sim F_G(\cdot; 0, \mu)$. Then, by using the stochastic representation of Item (2), the proof of Item (3) follows.

For $X \sim F_{BGEV}(\cdot; 0, \mu, \sigma, \delta)$, by Item (1), $T_{\sigma,\delta}(X) \sim F_G(\cdot; 0, \mu)$. By combining this with the well-known result: If $Y \sim F_G(\cdot; 0, \mu)$ then $\exp\left( - (Y - \mu) / \mu \right) \sim \text{Weibull}(1, \mu)$, the proof of Item (4) follows.

Let assume that $X_1 \sim F_{BGEV}(\cdot; 0, \mu_1, \sigma, \delta)$ and $X_2 \sim F_{BGEV}(\cdot; 0, \mu_2, \sigma, \delta)$ are independent. By Item (1), $T_{\sigma,\delta}(X_1) \sim F_G(\cdot; 0, \mu_1)$ and $T_{\sigma,\delta}(X_2) \sim F_G(\cdot; 0, \mu_2)$. By
combining this result with the known fact: If $Y_1 \sim F_G(\cdot;0, \mu_1)$ and $Y_2 \sim F_G(\cdot;0, \mu_2)$ are independent then $Y_1 - Y_2 \sim \text{Logistic}(\mu_1 - \mu_2, 1)$ (Logistic distribution), the proof of Item (5) follows.

Already, the proof of Item (6) follows by combining Item (1), the relation $|c|T_{\sigma,\delta}(x) = T_{|c|\sigma,\delta}(x)$, and the following well-known fact: If $Y \sim F_G(\cdot;\xi, \mu) = F_G(\cdot;\xi, \mu, 1)$ then $cY \sim F_G(\cdot;\xi, c\mu, c)$ for $c \neq 0$; with Item (2).

### 2.5 The BGEV model arising as a distribution limit of a normalized maximum

Let $Y_1, \ldots, Y_n$ be a sequence of independent and identically distributed (iid) random variables with CDF $F$ and let $M_n := \max\{Y_1, \ldots, Y_n\}$ denote the maximum. By Fisher-Tippett-Gnedenko theorem (Gnedenko 1943) there exist sequences of constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that

$$
\frac{M_n - b_n}{a_n} \xrightarrow{\mathcal{D}} Y \quad \text{with} \quad Y \sim F_G(\cdot;\xi, \mu),
$$

where “$\xrightarrow{\mathcal{D}}$” denotes convergence in distribution. Since $T_{\sigma,\delta}^{-1}$ is a continuous map, by applying the continuous mapping theorem, we have

$$
T_{\sigma,\delta}^{-1}\left(\frac{M_n - b_n}{a_n}\right) \xrightarrow{\mathcal{D}} X := T_{\sigma,\delta}^{-1}(Y) \quad \text{with} \quad X \sim F_{\text{BGEV}}(\cdot;\xi, \mu, \sigma, \delta),
$$

where we used the Item (2) of stochastic representation given in Proposition 3. Since $T_{\sigma,\delta}^{-1}$ is monotonically increasing, we have the identity

$$
T_{\sigma,\delta}^{-1}\left(\frac{M_n - b_n}{a_n}\right) = \max \left\{ T_{\sigma,\delta}^{-1}\left(\frac{Y_1 - b_n}{a_n}\right), \ldots, T_{\sigma,\delta}^{-1}\left(\frac{Y_n - b_n}{a_n}\right) \right\}.
$$

Therefore,

$$
\max \left\{ T_{\sigma,\delta}^{-1}\left(\frac{Y_1 - b_n}{a_n}\right), \ldots, T_{\sigma,\delta}^{-1}\left(\frac{Y_n - b_n}{a_n}\right) \right\} \xrightarrow{\mathcal{D}} X,
$$

with $X \sim F_{\text{BGEV}}(\cdot;\xi, \mu, \sigma, \delta)$. That is, the maximum of a sample of iid random variables converges in distribution to BGEV. In other words, $X$ is an extremal random variable (in the sense of Definition 6.1 of Gut 2013, p. 451). Since the classes of extremal distributions and max-stable distributions coincide (see Theorem 6.1 of Gut 2013, p. 452) it follows that $X$ is max-stable (see Definition 6.1 of Gut 2013, p. 451).

In the special case $b_n = 0$, we have

$$
\frac{\max\{Z_1, \ldots, Z_n\}}{a_n} \xrightarrow{\mathcal{D}} X,
$$

(10)
where \( Z_k = T_{-1, \delta}^{-1}(Y_k) = \text{sgn}(Y_k) Y_k^{1/(\delta+1)} \) for \( k = 1, \ldots, n \), and \( \tilde{a}_n = (\sigma a_n)^{1/(\delta+1)} \).

**Example 1** If we take the sequence \( Y_1, \ldots, Y_n \) of iid random variables distributed according to the standard Cauchy law:

\[
F(y) = \frac{1}{2} + \frac{1}{\pi} \arctan(y), \quad y \in \mathbb{R},
\]

then, by Fisher-Tippett-Gnedenko theorem, we have

\[
\frac{M_n - b_n}{a_n} \xrightarrow{d} Y \quad \text{with} \quad Y \sim F_G(\cdot; \xi, \mu),
\]

where \( \xi = 1 \) (Fréchet law), \( a_n = F^{-1}(1 - 1/n) = \tan \left( \pi(n - 2)/(2n) \right) \) and \( b_n = 0 \).

Letting \( Z_k = T_{-1, \delta}^{-1}(Y_k) \), note that \( Z_k \) has CDF \( F \circ T_{1, \delta} \), for \( k = 1, \ldots, n \). By convergence in Eq. (10), we get

\[
\max\{Z_1, \ldots, Z_n\} \xrightarrow{d} X \quad \text{with} \quad X \sim F_{BGEV}(\cdot; 1, \mu, \sigma, \delta).
\]

Therefore, under conditions of Corollary 2: \( \xi = 1, \delta \neq 0, \delta > -1/2, \mu < 1 \) and \( (\mu, \delta) \in A \), where the set \( A \) was defined lines before introducing Corollary 2, we have found an example of a sequence \( Z_1, \ldots, Z_n \) of iid random variables, with univariate distribution \( F \circ T_{1, \delta} \), whose maxima converges to some bimodal distribution (see Corollary 2).

### 2.6 Moments

Here we give a closed analytical formula for the \( k \)th moment of a random variable with BGEV distribution.

**Proposition 4** If \( X \sim F_{BGEV}(\cdot; \xi, \mu, \sigma, \delta) \), \( \xi < 1/k \) and \( \delta > -1 \), then

\[
\mathbb{E}(X^{k(\delta+1)}) = \begin{cases} 
\frac{1}{\xi^k \sigma^k} \sum_{i=0}^{k} \binom{k}{i} (\xi \mu - 1)^k \Gamma(1 - k \delta), & \mu - \frac{1}{\xi} > 0, \\
\frac{(-1)^i k(\delta+1)}{\xi^k \sigma^k} \sum_{i=0}^{k} \binom{k}{i} (\xi \mu - 1)^{k-i} \gamma(1 - \xi i; (1 - \xi \mu)^{-1/\xi}) & + \frac{1}{\xi^k \sigma^k} \sum_{i=0}^{k} \binom{k}{i} (\xi \mu - 1)^{k-i} \Gamma(1 - \xi i; (1 - \xi \mu)^{-1/\xi}), & \mu - \frac{1}{\xi} < 0,
\end{cases}
\]

where \( \gamma(a;x) = \int_x^\infty t^{a-1} e^{-t} \, dt \) and \( \Gamma(a;x) = \int_0^x t^{a-1} e^{-t} \, dt \) are the incomplete gamma functions.
Proof For $\xi > 0$, by using the stochastic representation of the BGEV distribution (see Item 2 of Proposition 3), we have
\[
\mathbb{E}(X^{k(\delta+1)}) = \frac{1}{\sigma^k} \int_{\mu-1/\xi}^{\infty} y^k f_G(y; \xi, \mu) \, dy = \frac{1}{\sigma^k} \mathbb{E}(y^k), \quad Y \sim F_G(\cdot; \xi, \mu),
\]
where
\[
\mathbb{E}(y^k) = \sum_{i=1}^{k} \binom{k}{i} \left( \frac{\mu - 1/\xi}{\xi} \right)^{k-i} \frac{1}{\xi} \Gamma(1 - \xi i) \quad \text{for} \quad \xi < 1/k.
\]

(ii) If $\mu - 1/\xi > 0$,
\[
\mathbb{E}(X^{k(\delta+1)}) = \frac{(-1)^{k(\delta+1)}}{\sigma^k} \int_{\mu-1/\xi}^{0} y^k f_G(y; \xi, \mu) \, dy + \frac{1}{\sigma^k} \int_{0}^{\infty} y^k f_G(y; \xi, \mu) \, dy
\]
\[
= \frac{(-1)^{k(\delta+1)}}{\sigma^k} I_1 + \frac{1}{\sigma^k} I_2.
\]
Using the PDF Eq. (2) and the substitution $w^{-1} = 1 + \xi(y - \mu)$, we obtain
\[
I_1 = \frac{1}{\xi^{k+1}} \int_{(1-\xi\mu)^{-1}}^{\infty} [w^{-1} + (\xi \mu - 1)]^k w^{1-1/\xi} e^{-w^{1/\xi}} \, dw.
\]
To solve the integral in Eq. (13) we use the Newton’s binomial formula, thus
\[
I_1 = \frac{1}{\xi^{k+1}} \sum_{i=1}^{k} \binom{k}{i} (\xi \mu - 1)^{k-i} \int_{(1-\xi\mu)^{-1}}^{\infty} w^{1-1-i} e^{-w^{1/\xi}} \, dw
\]
\[
= \frac{1}{\xi^{k}} \sum_{i=1}^{k} \binom{k}{i} (\xi \mu - 1)^{k-i} \int_{(1-\xi\mu)^{-1/\xi}}^{\infty} z^{-\xi i} e^{-z} \, dz,
\]
where in the last line we used the new substitution $z = w^{1/\xi}$.

Similarly we obtain
\[
I_2 = \frac{1}{\xi^{k}} \sum_{i=1}^{k} \binom{k}{i} (\xi \mu - 1)^{k-i} \int_{0}^{(1-\xi\mu)^{-1/\xi}} z^{-\xi i} e^{-z} \, dz.
\]
The proof is completed by expressing the integrals Eqs. (14) and (15) in terms of the incomplete gamma functions and then updating the equation Eq. (12).
For $\xi < 0$ the proof is similar. $\square$

As a sub-product of the proof of Proposition 4, we have the following result.

**Corollary 4** If $X \sim F_{BGEV}(\xi; \mu, \sigma, \delta)$, $\delta > -1$ and $\mu - 1/\xi > 0$, then the moments of $X^{\delta+1}$ of order $k$ satisfying the condition $\xi > 1/k$ don’t exist.

By taking $k(\delta + 1) = 1$ in Proposition 4 we obtain the following formula for the expected value.

**Corollary 5** Let $X \sim F_{BGEV}(\xi; \mu, \sigma, \delta)$ and $-1 < \delta \leq 0$, then

$$E(X) = \frac{1}{(\xi \sigma)^{\frac{1}{\delta+1}}} \sum_{i=0}^{\left\lfloor 1/\xi \right\rfloor} \left( \left\lfloor \frac{1}{\delta+1} \right\rfloor \right) (\xi \mu - 1)^{\left\lfloor \frac{1}{\delta+1} \right\rfloor - i} \Gamma(1 - \xi i)$$

• when $\mu - \frac{1}{\xi} > 0$, and

$$E(X) = \frac{1}{(\xi \sigma)^{\frac{1}{\delta+1}}} \sum_{i=0}^{\left\lfloor 1/\xi \right\rfloor} \left( \left\lfloor \frac{1}{\delta+1} \right\rfloor \right) (\xi \mu - 1)^{\left\lfloor \frac{1}{\delta+1} \right\rfloor - i} \theta\left(1 - \xi i; (1 - \xi \mu)^{-1/\xi}\right)$$

• when $\mu - \frac{1}{\xi} < 0$. Here, $\theta\left(1 - \xi i; (1 - \xi \mu)^{-1/\xi}\right) := \Gamma\left(1 - \xi i; (1 - \xi \mu)^{-1/\xi}\right) - \gamma\left(1 - \xi i; (1 - \xi \mu)^{-1/\xi}\right)$ and the value of $\|x\|$ is the largest integer that is less than or equal to $x$.

### 2.7 Quantiles

In the next section, through Monte Carlo simulation studies, we test the behavior of the parameter estimates of the BGEV model. The random samples of the population are simulated using the inverse transformation method based on quantiles. The quantile function of a random variable $X \sim F_{BGEV}(\xi; \mu, \sigma, \delta)$ is defined by

$$x_Q = \inf\{x \in \mathbb{R} : Q \leq F_{BGEV}(x; \xi, \mu, \sigma, \delta)\}$$

or equivalent by

$$x_Q = T_{\sigma, \delta}^{-1}(F_G^{-1}(Q)).$$

From Eq. (8), we have $F_G^{-1}(Q) = \mu + [(-\ln Q)^{-\xi} - 1]/\xi$ for $\xi > 0$ and $Q \in [0, 1)$ or $\xi < 0$ and $Q \in (0, 1]$. Now, by using the expression Eq. (7) of $T_{\sigma, \delta}^{-1}$, we get
2.8 Tail behavior of BGEV

According to Embrechts et al. (2013), asymptotic estimates of probability models are omnipresent in statistics, insurance mathematics, mathematical finance, and hydrology. In this sense, the theory of regular variation plays a crucial role. Concepts and properties of regularly varying can be found in the book of Bingham et al. (1987). In order to study the asymptotic behavior of the tail of the BGEV distribution, here we show some of the results of regular variation.

**Definition 1** A positive Lebesgue measurable function $h$ on $(0, \infty)$ is regularly varying at $\infty$ of index $\alpha \in \mathbb{R}$ ($h \in \mathcal{R}_\alpha$) if

$$
\lim_{t \to \infty} \frac{h(tx)}{h(t)} = x^\alpha.
$$

Suppose $F$ is a distribution function such that $F(x) < 1$ for all $x > 0$. Let $\overline{F}(x) = 1 - F(x)$ and $\overline{F} \in \mathcal{R}_{-\alpha}$ for some $\alpha > 0$, then

$$
\overline{F}(x) = o(x^{-\alpha}L(x))
$$

for some $L \in \mathcal{R}_0$, that is, $L$ is a slowly varying function.

**Definition 2** Let $F$ be a continuous univariate distribution on $\mathbb{R}$.

- The distribution $F$ has the least light-tail distribution if, for any $t > 0$,

$$
\lim_{x \to -\infty} \frac{\exp(tx)}{F(x)} = \infty.
$$

- The distribution $F$ has upper light-tail distribution if, for any $t > 0$,

$$
\lim_{x \to \infty} \frac{\exp(-tx)}{F(x)} = \infty.
$$

- The distribution $F$ has a simple exponential tail if it has the least light-tailed and upper light-tail distribution.
- The distribution $F$ has upper heavy-tail distribution if, for any $t > 0$,  

\[
x_Q = \begin{cases} 
\left[ \frac{\mu}{\sigma} + \frac{(-\ln Q)^{-\xi} - 1}{\sigma \xi} \right]^{1/(\delta+1)}, & \mu + \frac{(-\ln Q)^{-\xi} - 1}{\xi} > 0, \\
(-1)^{2\alpha} \left[ \frac{\mu}{\sigma} + \frac{(-\ln Q)^{-\xi} - 1}{\sigma \xi} \right]^{1/(\delta+1)}, & \mu + \frac{(-\ln Q)^{-\xi} - 1}{\xi} < 0.
\end{cases}
\]
The expression Eq. (17) indicates that the tail of the distribution $F$ slowly decays, as $x^{-\alpha}$, that is $F$ has upper heavy-tail with index $\alpha$. Thus, $F$ is Pareto-type distribution.

**Proposition 5** Let $X \sim F_{BGEV}(\cdot; \xi, \mu, \sigma, \delta)$ with $\xi > 0$ and $\delta > -1$. Then

$$F_{BGEV}(x; \xi, \mu, \sigma, \delta) = o\left(x^{-\left(\frac{(\delta+1)}{\xi}\right)}\right), \quad x > 0,$$

where $F_{BGEV}(x; \xi, \mu, \sigma, \delta)$ is given in Eq. (8).

**Proof** For all $\delta > -1$ and $x > 0$: $r := xT_{\sigma, \delta}(t) \to \infty$ and $s := T_{\sigma, \delta}(t) \to \infty$ as $t \to \infty$. Then, a straightforward calculation shows that

$$\lim_{t \to \infty} \frac{F_{BGEV}(tx; \xi, \mu, \sigma, \delta)}{F_{BGEV}(t; \xi, \mu, \sigma, \delta)} = \lim_{t \to \infty} \frac{F_G(r|x^{\delta}; \xi, \mu)}{F_G(r; \xi, \mu)} \cdot \lim_{s \to \infty} \frac{F_G(ss^{\delta}; \xi, \mu)}{F_G(s; \xi, \mu)} = x^{-\left(\frac{(\delta+1)}{\xi}\right)},$$

because, for $\xi > 0$, the GEV distribution is regularly varying with index $1/\xi$.

From Eq. (18) we have that $F_{BGEV}$ is Pareto-type distribution with tail index $(\delta + 1)/\xi$. The new parameter $\delta$ affects the weight of the tail.

It is well-known that the ordinary GEV distribution (when $\delta = 0$) can be divided into three types of tail behavior, depending on the sign of the shape parameter $\xi$. More precisely, if $\xi = 0$ (Gumbel law) the GEV has simple exponential tail; if $\xi < 0$ (reversed Weibull law) the GEV distribution has the least light-tail and finite right-end-point; while for $\xi > 0$ (Fréchet law) the GEV has upper heavy-tail, regularly varying with index $1/\xi$ and infinite right-end-point. In what follows we prove that the type of tail behavior of BGEV has similar properties.

**Theorem 6** The shape parameter $\xi$ and the bimodality parameter $\delta$ govern the tail behavior of the BGEV distribution type of the following form:

(a) If $\xi = 0$, for any $\delta > 0$ the BGEV has simple exponential tail.
(b) If $\xi < 0$, for any $\delta > 0$ the BGEV has the least light-tail distribution and finite right-end-point.
(c) If $\xi > 0$, for any $-1 < \delta < 0$ the BGEV has upper heavy-tail distribution, regularly varying with tail index $(\delta + 1)/\xi$ and infinite right-end-point.

**Proof** Firstly, see that
\[ \forall \delta > -1 : y := T_{\sigma, \delta}(x) \longrightarrow \pm \infty \text{ as } x \rightarrow \pm \infty \] (18)

and

\[
L := \lim_{x \to \pm \infty} \frac{\exp(-tx)}{\exp(-tT_{\sigma, \delta}(x))} = \lim_{x \to \pm \infty} \exp \left( -tx(1 - \sigma|x|^{\delta}) \right) = \begin{cases} 0, & -1 < \delta < 0, \\ \infty, & \delta > 0. \end{cases} \] (19)

In order to prove the Item (a), we assume that \( \xi = 0 \) and \( \delta > 0 \). Since \( F_{BGEV}(x; \xi, \mu, \sigma, \delta) = F_{G}(y; \xi, \mu) \), for any \( t > 0 \) we get

\[
\lim_{x \to \pm \infty} \frac{\exp(tx)}{F_{BGEV}(x; \xi, \mu, \sigma, \delta)} = L \cdot \lim_{y \to \pm \infty} \frac{\exp(ty)}{F_{G}(y; \xi, \mu)} = : L \cdot L_{1},
\]

\[
\lim_{x \to \pm \infty} \frac{\exp(-tx)}{F_{BGEV}(x; \xi, \mu, \sigma, \delta)} = L \cdot \lim_{y \to \pm \infty} \frac{\exp(-ty)}{F_{G}(y; \xi, \mu)} = : L \cdot L_{2},
\]

where \( L \) is as in Eq. (20). Furthermore, for \( \xi = 0 \) the GEV has simple exponential tail, that is, \( L_{1} = L_{2} = \infty \). Consequently, for \( \xi = 0, \delta > 0 \) and for any \( t > 0 \),

\[
\lim_{x \to \pm \infty} \frac{\exp(tx)}{F_{BGEV}(x; \xi, \mu, \sigma, \delta)} = \lim_{x \to \pm \infty} \frac{\exp(-tx)}{F_{BGEV}(x; \xi, \mu, \sigma, \delta)} = 0.
\]

Then the simple exponential tailedness of BGEV distribution follows. This proves the first item.

Now, let \( \xi < 0 \) and \( \delta > 0 \). For any \( t > 0 \) we obtain

\[
\lim_{x \to \pm \infty} \frac{\exp(tx)}{F_{BGEV}(x; \xi, \mu, \sigma, \delta)} = L \cdot \lim_{y \to \pm \infty} \frac{\exp(ty)}{F_{G}(y; \xi, \mu)} = : L \cdot L_{3},
\]

where \( L \) is as in Eq. (20). Since for \( \xi < 0 \) the GEV is the least light-tailed distribution, \( L_{3} = \infty \). This implies, for \( \xi < 0, \delta > 0 \) and for any \( t > 0 \), that

\[
\lim_{x \to \pm \infty} \frac{\exp(tx)}{F_{BGEV}(x; \xi, \mu, \sigma, \delta)} = 0.
\]

That is, the BGEV is the least light-tailed distribution. Furthermore, by definition of BGEV distribution, it is clear that the BGEV has a finite right-end-point given by \( T^{-1}_{\sigma, \delta}(\mu - 1/\xi) \). This proves the Item (b).

Finally, suppose that \( \xi > 0 \) and \(-1 < \delta < 0 \). For any \( t > 0 \) we have

\[
\lim_{x \to \pm \infty} \frac{\exp(-tx)}{F_{BGEV}(x; \xi, \mu, \sigma, \delta)} = L \cdot \lim_{y \to \pm \infty} \frac{\exp(-ty)}{F_{G}(y; \xi, \mu)} = : L \cdot L_{4},
\]

where \( L_{1} \) is given in Eq. (20). But, for \( \xi > 0 \) the GEV is upper heavy-tailed distribution (see Beirlant et al. 2001), that is, \( L_{4} = 0 \). Therefore, for \( \xi > 0, -1 < \delta < 0 \) and for any \( t > 0 \),
This guarantees the upper heavy-tailedness of BGEV distribution. From Proposition 5 it follows that the BGEV is regularly varying with tail index \((\delta + 1)/\xi\). Moreover, since \(\lim_{x \to \infty} f_{BGEV}(x; \xi, \mu, \delta) = 0\), we have that the BGEV has infinite right-end-point. Thus, we complete the proof of Item (c).

\[ \lim_{x \to \infty} \frac{\exp(-tx)}{F_{BGEV}(x; \xi, \mu, \delta)} = 0. \]

\(\square\)

Corollary 7 The BGEV distribution has a transition from light-tailed distributions to heavy-tailed.
3 Graphical study

The role that the parameters $\sigma$, $\mu$, $\xi$, $\delta$ play in the BGEV distribution was investigated graphically by generating different densities with variations in each parameter. Figure 1 shows that $\sigma$ is a scale parameter. That is, $\sigma$ plays the same role in GEV and BGEV. Positive and negative values of $\mu$ and $\xi$ were considered to evaluate the role of $\mu$ and $\xi$ in the BGEV density. In Fig. 2 are the BGEV densities for $\sigma = 1$, $\delta = 1$ with $\xi \in \{0.45, -0.45\}$ (upper panels), and $\xi \in \{1, -1\}$ (bottom panels). In this figure, it can be seen that the density is unimodal or bimodal for some combinations of parameters. In each of the four panels the value of $\xi$ is fixed, but note that the shape of the density changes as the value of $\mu$ varies. Thus, it can be concluded that the parameter $\mu$ is a parameter of shape and not of location as in the GEV density. On the other hand, when varying $\xi$, with the other fixed parameters, the shape of $f_{\text{BGEV}}$ changes. For example, for $\mu < 0$, $f_{\text{BGEV}}(\cdot; 1, -2, 1, 1)$ is unimodal and $f_{\text{BGEV}}(\cdot; -1, -2, 1, 1)$ is bimodal, both curves concentrated on negative values. For $\mu > 0$, the opposite occurs, the density $f_{\text{BGEV}}(\cdot; 1, 1, 1, 1)$ is bimodal and
The log-likelihood function for \( f_{\text{BGEV}}(\cdot; \xi, \mu, \sigma, \delta) \) is given by

\[
\ell_i(\Theta; x) = n \ln \sigma + n \ln(\delta + 1) \\
+ \sum_{i=1}^{n} \left[ \delta \ln |x_i| - \left(1 + \frac{1}{\xi}\right) \ln \Psi_i(\Theta) - \Psi_i^{-1/\xi}(\Theta) \right],
\]

where \( \Psi_i(\Theta) = 1 + \xi \left( \sigma x_i |x_i|^\delta - \mu \right), \; i = 1, \ldots, n \). Note that, for all \( i = 1, \ldots, n \),

\[
\frac{\partial \Psi_i(\Theta)}{\partial \mu} = -\xi, \quad \frac{\partial \Psi_i(\Theta)}{\partial \sigma} = \xi \sigma x_i |x_i|^\delta \ln |x_i|, \quad \frac{\partial \Psi_i(\Theta)}{\partial \xi} = \sigma x_i |x_i|^\delta - \mu.
\]

Then, the ML estimates of \( \mu, \sigma, \delta, \xi \) are the solutions of the following system of equations

\[
\begin{align*}
\frac{\partial \ell_i(\Theta; x)}{\partial \mu} &= \mu \sum_{i=1}^{n} \Omega_i(\Theta) = 0, \\
\frac{\partial \ell_i(\Theta; x)}{\partial \sigma} &= \frac{n}{\sigma} - \sum_{i=1}^{n} x_i |x_i|^\delta \Omega_i(\Theta) = 0, \\
\frac{\partial \ell_i(\Theta; x)}{\partial \delta} &= \frac{1}{\delta + 1} + \sum_{i=1}^{n} \left[ \ln |x_i| + \sigma x_i |x_i|^\delta \ln |x_i| \right] \Omega_i(\Theta) = 0, \\
\frac{\partial \ell_i(\Theta; x)}{\partial \xi} &= \xi^{-2} \sum_{i=1}^{n} (\sigma x_i |x_i|^\delta - \mu) \left\{ \ln \Psi_i(\Theta) - \xi \Omega_i(\Theta) \right\} = 0,
\end{align*}
\]

where \( \Omega_i(\Theta) = \Psi_i^{-1}(\Theta) [1 + \xi - \Psi_i^{-1/\xi}(\Theta)] \), \( i = 1, \ldots, n \).

Likelihood maximization in bimodal models should be done with care since local optima are often present. In this paper, we have used the Nelder–Mead method (Nelder and Mead 1965). Nelder–Mead is a derivative-free method [Nocedal and Wright 2006, chap. 9] requiring only the likelihood as input. We have had no problem maximizing the loglikelihood in those cases with the optimization routines.
available through the `optim` function from R (R-Team 2020), even from far-away starting values (see Sect. 5). To verify that the trial converges to a local or a global maximum we have detected by inspection of profile likelihoods (Andrade et al. 2016). For a good discussion about likelihood maximization in bimodal models, see Andrade and Rathie (2016).

A potential difficulty with the use of likelihood methods for the GEV concerns the regularity conditions that are required for the usual asymptotic properties associated with the maximum likelihood estimator to be valid. Such conditions are not satisfied by the GEV model because the end-points of the GEV distribution are functions of the parameter values: \( \mu - \sigma / \xi \) is an upper end-point of the distribution when \( \xi < 0 \), and a lower end-point when \( \xi > 0 \). This violation of the usual regularity conditions means that the standard asymptotic likelihood results are not automatically applicable. The same goes for the BGEV distribution, because the support of the BGEV is also a function of its parameters, according to Eq. (4). Thus, it is not possible to calculate confidence intervals and coverage probabilities using the maximum likelihood methods. For the construction of confidence intervals for the model parameters, we suggest employing the parametric bootstrap (Efron 1979).

| \( n \) | \( \delta \) | \( \hat{\xi} \) | \( \hat{\mu} \) | \( \hat{\sigma} \) | \( \hat{\delta} \) |
|-------|-----|-----|-----|-----|-----|
| 50    | \(-0.5\) | \(-0.0942 (0.2041)\) | 0.1626 (1.0605) | \(-0.0016 (0.1639)\) | 0.0460 (0.0799) |
| 0.5   | \(-0.1317 (0.2133)\) | 0.1763 (1.0693) | \(-0.0023 (0.1615)\) | 0.1727 (0.2585) |
| 1.0   | \(-0.0879 (0.1987)\) | 0.1853 (1.1360) | \(-0.0029 (0.1504)\) | 0.0685 (0.3052) |
| 2.0   | \(-0.0976 (0.1997)\) | 0.2007 (1.1417) | \(-0.0036 (0.1567)\) | 0.0589 (0.4826) |
| 100   | \(-0.5\) | \(-0.0381 (0.1329)\) | 0.0522 (0.8038) | 0.0005 (1.086) | 0.0218 (0.0513) |
| 0.5   | \(-0.0578 (0.1275)\) | 0.0508 (0.8111) | 0.0002 (1.1151) | 0.0801 (0.1598) |
| 1.0   | \(-0.0631 (0.1335)\) | 0.1096 (0.7765) | 0.0010 (1.1115) | 0.0611 (0.2135) |
| 2.0   | \(-0.0941 (0.1343)\) | 0.0039 (0.8075) | 0.0003 (1.1112) | 0.0459 (0.3124) |
| 200   | \(-0.5\) | \(-0.0218 (0.0808)\) | 0.0857 (0.5429) | 0.0013 (0.0790) | 0.0098 (0.0352) |
| 0.5   | \(-0.0360 (0.0854)\) | 0.0116 (0.5366) | 0.0002 (0.0809) | 0.0355 (0.1037) |
| 1.0   | \(-0.0265 (0.0856)\) | 0.0081 (0.5321) | 0.0001 (0.0805) | 0.0185 (0.1282) |
| 2.0   | \(-0.0304 (0.0844)\) | 0.0418 (0.5534) | 0.0013 (0.0819) | 0.0244 (0.2157) |
| 400   | \(-0.5\) | \(-0.0147 (0.0572)\) | 0.0392 (0.3942) | 0.0031 (0.1051) | 0.0037 (0.0236) |
| 0.5   | \(-0.0167 (0.0554)\) | \(-0.0071 (0.3762)\) | 0.0019 (0.0809) | 0.0140 (0.0690) |
| 1.0   | \(-0.0119 (0.0616)\) | 0.0685 (0.4099) | 0.0014 (0.0947) | 0.0131 (0.0966) |
| 2.0   | \(-0.0089 (0.0567)\) | 0.0465 (0.3801) | 0.0013 (0.0720) | 0.0093 (0.1389) |
Table 2 Relative biases and RMSE (in parentheses) of the estimates of the parameters for \((\xi, \mu, \sigma) = (0.3, 1, 10)\) and some values of \(\delta\) and \(n\)

| \(n\) | \(\delta\) | \(\hat{\xi}\) | \(\hat{\mu}\) | \(\hat{\sigma}\) | \(\hat{\delta}\) |
|-------|-----------|---------------|---------------|---------------|---------------|
| 50    | -0.5      | 0.0666 (0.1252) | 0.0731 (1.5491) | -0.0043 (0.1735) | 0.0215 (0.0705) |
|       | 0.5       | 0.0995 (0.1309) | 0.2292 (1.6648) | -0.0057 (0.1801) | 0.0844 (0.2320) |
|       | 1.0       | 0.0824 (0.1205) | 0.1529 (1.5431) | -0.0050 (0.1661) | 0.0497 (0.2965) |
|       | 2.0       | 0.1037 (0.1297) | 0.1518 (1.6236) | -0.0053 (0.1730) | 0.0371 (0.4403) |
| 100   | -0.5      | 0.0571 (0.0784) | 0.1067 (1.1245) | -0.0014 (0.1208) | 0.0068 (0.0465) |
|       | 0.5       | 0.0437 (0.0758) | 0.0624 (1.0449) | -0.0010 (0.1188) | 0.0468 (0.1506) |
|       | 1.0       | 0.0364 (0.0757) | 0.0255 (1.1573) | -0.0003 (0.1373) | 0.0274 (0.1902) |
|       | 2.0       | 0.0476 (0.0803) | 0.0590 (1.0858) | -0.0009 (0.1259) | 0.0159 (0.3008) |
| 200   | -0.5      | 0.0249 (0.0522) | 0.0216 (0.7847) | 0.0049 (0.1421) | 0.0058 (0.0341) |
|       | 0.5       | 0.0190 (0.0515) | 0.0052 (0.7976) | 0.0037 (0.1393) | 0.0205 (0.1044) |
|       | 1.0       | 0.0346 (0.0479) | 0.0914 (0.7987) | 0.0019 (0.1299) | 0.0172 (0.1217) |
|       | 2.0       | 0.0324 (0.0509) | 0.0955 (0.7718) | 0.0002 (0.1196) | 0.0007 (0.1858) |
| 400   | -0.5      | 0.0184 (0.0343) | 0.0349 (0.5137) | -0.0021 (0.0828) | 0.0003 (0.0226) |
|       | 0.5       | 0.0106 (0.0348) | 0.0259 (0.5315) | -0.0014 (0.0689) | 0.0099 (0.0718) |
|       | 1.0       | 0.0186 (0.0340) | 0.0518 (0.5555) | -0.0016 (0.0761) | 0.0021 (0.0906) |
|       | 2.0       | 0.0244 (0.0337) | 0.0610 (0.5430) | -0.0005 (0.0764) | 0.0008 (0.1306) |

Fig. 5 Box plots from 500 simulated estimates of \(\xi, \mu, \sigma\) and \(\delta\) for some values of \(n\). (a)-(d) \(\xi = -0.5, \mu = 1, \sigma = 10\) and \(\delta = 0.5\), (e)-(h) \(\xi = -0.5, \mu = 1, \sigma = 10\) and \(\delta = -0.5\)

5 Numerical illustrations

In this Section, we present two simulation studies related to the BGEV model. The first study is devoted to the recovery of parameters under different scenarios, and the
second study assesses the performance of the BGEV model with a misspecification distribution.

5.1 Recovery parameters

In this Section, we study the properties of the ML estimators in finite samples. The simulation was performed using the R programming language; see http://www.r-project.org. The number of Monte Carlo replications was $R = 500$. The sample sizes considered are $n = 50, 100, 200, 400$ and the values of the parameters are the combinations $\{(\xi, \mu, \sigma) = (-0.5, 1, 10), (\xi, \mu, \sigma) = (0.3, 1, 10)\} \times \{\delta = -0.5, \delta = 0.5, \delta = 1.0, \delta = 2.0\}$. Tables 1 and 2 present the relative biases and root mean squared error (RMSE) of the estimates of the parameters of the model. From these tables, notice that the relative biases and the RMSE decrease as the size of the sample increases, evidencing that the estimators are asymptotically unbiased.

The previous findings are confirmed by the box plots shown in Figs. 5 and 6. Therefore, we recommend the use $\hat{\xi}, \hat{\mu}, \hat{\sigma}$ and $\hat{\delta}$ as the estimators for the parameters $\xi, \mu, \sigma$ and $\delta$ of a BGEV model, having a good performance in terms of relative bias and RMSE.

5.2 Robustness study

In this second study, we consider a misspecification study for the bimodal distribution. In each case, 1000 replicates were drawn and for each was used the
Fig. 7 Probability density function when $X \sim IBW(1.0, -1.0, 1.0, 2.0)$ (red) and when $X \sim BGEV(-0.9, -1.8, 9.0, 2.7)$ (blue)

Table 3 Empirical mean of the estimates of the parameters for $(\xi, \mu, \sigma, \delta) = (-0.9, -1.8, 9.0, 2.7)$ for BGEV the distribution, $(\xi, \mu, \sigma, \delta) = (1.0, -1.0, 1.0, 2.0)$ for the IBW distribution and some values of $n$. Columns indicate the real model and rows represent the fitted model

| Parameter | $n$ | Model | IBW data | BGEV data |
|-----------|-----|-------|----------|-----------|
| $\xi$     | 50  | IBW   | 0.9653   | 8.0175    |
|           |     |       | BGEV     | -0.9816   |
|           | 100 | IBW   | 0.9823   | 8.7244    |
|           |     |       | BGEV     | -0.0464   |
|           | 200 | IBW   | 0.9927   | 8.6365    |
|           |     |       | BGEV     | -0.3891   |
| $\mu$     | 50  | IBW   | -0.9819  | -4.8777   |
|           |     |       | BGEV     | -2.8518   |
|           | 100 | IBW   | -0.9911  | -4.8371   |
|           |     |       | BGEV     | -2.8138   |
|           | 200 | IBW   | -0.9957  | -4.9300   |
|           |     |       | BGEV     | -2.7184   |
| $\sigma$  | 50  | IBW   | 0.9864   | 4.9507    |
|           |     |       | BGEV     | 12.767    |
|           | 100 | IBW   | 0.9884   | 5.1892    |
|           |     |       | BGEV     | 12.484    |
|           | 200 | IBW   | 0.9947   | 5.14306   |
|           |     |       | BGEV     | 12.502    |
| $\delta$  | 50  | IBW   | 2.1442   | 2.6563    |
|           |     |       | BGEV     | 2.4893    |
|           | 100 | IBW   | 2.0661   | 3.1006    |
|           |     |       | BGEV     | 2.3706    |
|           | 200 | IBW   | 2.0229   | 2.9039    |
|           |     |       | BGEV     | 2.4876    |
BGEV and invertible bimodal Weibull (IBW) (Silva 2022) bimodal models to perform the parameter estimation.

The CDF of the IBW distribution is given by

\[ G(x; \xi, \mu, \sigma, \delta) = G(W_\delta(x)\xi; \mu, \sigma), \]

where \( W_\delta(x) = x|x|^\delta, \quad \delta > 0, \)

\[
G(x; \xi, \mu, \sigma) = \begin{cases} 
1 - \exp \left\{-\frac{x-\mu}{\sigma}\right\} & x \geq \mu, \\
0, & x < \mu,
\end{cases}
\]

\( \xi > 0 \) is shape parameter, \( \sigma > 0 \) scale parameter, and \( \mu \in \mathbb{R} \) location parameter. When \( \mu < 0 \) the density of the IBW distribution is bimodal. For more details of the IBW distribution, see [28].

In each replication, a random sample of size \( n \) is drawn from the BGEV\((-0.9, -1.8, 9.0, 2.7)\) distribution or IBW\((1.0, -1.0, 1.0, 2.0)\) distribution. Figure 7 suggested that the BGEV and IBW models are competitive for this scenario. Results are presented in Table 3. The ML estimates are close to the true values of the model parameters, which indicates the ‘robustness’ of each model when estimating the parameters under model misspecification.

6 Applications to environmental data

In this section, to illustrate the applicability of the model proposed and its advantages, the BGEV distribution was fitted to two data sets of the climate of the Federal District in Brazil.

Like the GEV distribution, the BGEV distribution is also a limit distribution of normalized maximums, as shown in the limit Eq. (9) in Sect. 2.5. The stochastic representation of the BGEV distribution and its relationship to the GEV distribution [Proposition 3, Item (1)] is illustrated here by fitting maximum values. To estimate the BGEV model parameters, the MLE method (as discussed in Sect. 4) was implemented in the R software (R-Team 2020). As for the estimates of the parameters of the GEV distribution, the “fExtremes” package from Rmetrics was used (Wuertz et al. 2017).

The data series employed were taken from INMET (National Institute of Meteorology) at www.inmet.gov.br.Data Stations-AutomaticStations. The data here utilized correspond to the wind speed during the period 10/12/2018 to 10/12/2019 (365 days) and the temperature at the oval point during the period 24/10/2018 to 20/10/2019 (361 days). Each day has 24 observations (one per hour). According to Andrade and Rathie (2016), working with small samples makes the assessment of bimodality even harder, i.e., it is very hard to quantify bimodality in small samples. Andrade and Rathie (2016) suggested that \( n \geq 100 \) is highly desirable. We used the maximum block technique with blocks of size 24 based in the daily maximum value. Furthermore, the transformation \((x – \text{Mean})/SD\), where SD is the standard deviation
Fig. 8 Histogram of temperature data set (first panel) and histogram of wind speed data set (second panel)

Table 4 Descriptive statistics for the standardized data sets

|                  | Minimum | 1st Quartile | Median | Mean  | 3rd Quartile | Maximum |
|------------------|---------|--------------|--------|-------|--------------|---------|
| Wind speed       | −2.26   | −1.40        | −1.08  | −0.75 | −0.38        | 1.54    |
| Temperature at   | −3.44   | −0.43        | 0.71   | 0.32  | 1.12         | 1.88    |
| the oval point   |         |              |        |       |              |         |

Table 5 MLEs of the model parameters and the statistics KS, AD and \(-2\ell(\Theta)\) for the two data sets

| Estimates        | \(\hat{\mu}\) | \(\hat{\sigma}\) | \(\hat{\xi}\) | \(\hat{\delta}\) | KS      | AD       | \(-2\ell(\Theta)\) |
|------------------|----------------|------------------|--------------|---------------|---------|----------|-------------------|
| Wind speed       | −1.0430        | 0.6516           | −0.3042      | 0.9434        | 0.05396 | 47.802   | 22204.9           |
| GEV              | −0.4957        | 0.6853           | 0.1330       | 0              | 0.24109 | 45.179   | 22214.6           |
| Temperature      | \(\hat{\mu}\) | \(\hat{\sigma}\) | \(\hat{\xi}\) | \(\hat{\delta}\) | KS      | AD       | \(-2\ell(\Theta)\) |
| BGEV             | 0.1804         | 0.8197           | −0.4839      | 0.4201        | 0.0738  | 66.297   | 20346.1           |
| GEV              | −0.1955        | 1.0599           | −0.6045      | 0              | 0.3261  | 70.706   | 22111.0           |

 Springer
of the data, is used to adjust it in the BGEV distribution. We applied Ljung-Box’s test (Trapletti 2016) to verify the null hypothesis of serial independence of the data sets. The test statistics did not reject the null hypothesis at the significance level of
Finally, from Fig. 8, we can see that the data set has bimodality. The descriptive statistics are given in Table 4.

The MLEs of the parameters, the value of $-2\ell(\Theta)$, the KS and AD statistics for the BGEV and GEV models are listed in Table 5. In general, the smaller the values of these statistics, the better the fit. The graphical study performed in Sect. 4 helped in choosing of the initial values of the estimates. For the wind speed, the
initial values were \( \mu = 0, \sigma = 1, \xi = -0.5 \) and \( \delta = 0.5 \) and for the temperature at the oval point were \( \mu = 0, \sigma = 1, \xi = -0.25 \) and \( \delta = 0.5 \). Since the values are smaller for the BGEV distribution compared with those values of the other models, the new distribution seems to be a very competitive model for these data sets. These results illustrate the potentiality of the BGEV model and the importance of the additional parameter (Figs. 9 and 10).

Here we illustrate the relationship between the BGEV distribution and the GEV distribution, according to Item 1 of Proposition 3. Then, by using the estimates of \( \sigma \) and \( \delta \), shown in Table 5, we define the functions

\[
y_{\text{top}} = 0.8197x|x|^{0.4201} \quad \text{and} \quad y_{\text{ws}} = 0.6516x|x|^{0.9434}
\]

for the oval point temperature data and for wind speed data, respectively. The histograms of the new data from Eq. (23) are shown in Fig. 11.

Now by using the \texttt{fExtremes} package from Rmetrics (Wuertz et al. 2017) we fit these data by the GEV distribution. The results of the GEV estimates and the statistics functions are in Table 6, for the temperature data at the oval point and wind speed (Figs. 12 and 13).

### 7 Conclusion

When extreme value data show bimodality, despite its broad sense of applicability in many fields, the GEV distribution is not suitable. In this article, we propose a generalization of the GEV distribution, called BGEV distribution, with an additional parameter, which modifies the behavior of the distribution, composing as an alternative model for single maxima events. The GEV distribution appears as a particular case. We present relevant properties of the model, and through graphical studies we have shown its wide flexibility. The good performance of the MLEs of the parameters was tested via Monte Carlo simulation. Applications of the BGEV distribution for two extreme data sets show that the new distribution can be used to provide better adjustments than the GEV to model when the data show bimodality. We hope this new distribution may attract wider applications for extreme values analysis.

### References

Andrade BB, Rathie PN (2016) Fitting asymmetric bimodal data with selected distributions. J Stat Comput Simul 86(16):3205–3224
Andrade BB, Bolfarine H, Siroky AN (2016) Random number generation and estimation with the bimodal asymmetric power-normal distribution. J Stat Comput Simul 86(3):460–476
Aryal RG, Tsokos PC (2011) Transmuted weibull distribution: a generalization of the weibull probability distribution. Eur J Pure Appl Math 4(2):89–102
Beirlant J, Matthys G, Dierckx G (2001) Heavy-tailed distributions and rating. ASTIN Bull. 31(1):37–58
Beirlant J, Goegebeur Y, Segers J, Teugels J (2004) Statistics of extreme theory and applications. Wiley, Hoboken
Bingham NH, Goldie CM, Teugels JL (1987) Regular variation. Cambridge University Press, Cambridge
Coles S (2001) An introduction to statistical modeling of extreme Values. Springer, London
Daly E, Zinger Y, Deletic A, Fletcher TD (2009) A possible mechanism for soil moisture bimodality in humid-land environments. Geophys Res Lett. 36(7):L07402
Efron B (1979) Bootstrap methods: another look at the jackknife. Ann Stat 7:1–26
Elal-Olivero D (2010) Alpha-skew-normal distribution. Proyecc J Math 29:224–240
Eljabri S, Nadarajah S (2017) The Kumaraswamy GEV distribution. Commun Stat Theory Methods 46:10203–10235
Embrechts P, Klüppelberg C, Mikosch T (2013) Modelling extremal events: for insurance and finance. Springer Science & Business Media, Berlin
Fonseca R, Cribari-Neto F (2018) Inference in a bimodal Birnbaum-Saunders model. Math Comput Simul 146:134–159
Gnedenko B (1943) Sur la distribution limite du terme maximum D’Une serie aleatoire. Ann Math 44(3):423–453
Griffiths L (1947) Introduction to the theory of equations. Wiley, New York
Gut A (2013) Probability: a graduate course. Springer, New York
Haan L, Ferreira A (2010) Extreme value theory: an introduction. Springer Science & Business Media, Berlin
Kotz SL, Nadarajah S (2000) Extreme value distributions: theory and applications. Imperial College Press, London
Longin F (2016) Extreme events in finance: a handbook of extreme value theory and its applications. Wiley, Hoboken
Nascimento F, Bourguignon M, Leão J (2016) Extended generalized extreme value distribution with applications in environmental data. Hacettepe J Math Stat 45:1847–1864
Nelder JA, Mead R (1965) A simplex algorithm for function minimization. Comput J 7:308–31
Nocedal J, Wright S (2006) Numerical optimization, 2d edn. Springer, New York
Otiniano CEG, Vila R, Brom PC, Bourguignon M (2023) On the bimodal Gumbel model with application to environmental data. Austrian J Stat 52:45–65
R-Team (2020) A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
Scheirer WJ (2017) Extreme value theory—based methods for visual recognition. Morgan & Claypool Publishers, California
Silva BLS (2022) A new invertible bimodal Weibull model. Dissertation, Universidade de Brasília, Brasília
Swamee PK, Rathie PN (2007) Invertible alternatives to normal and lognormal distributions. J Hifrol Eng 12(2):218–221
Trapletti, A (2016) R Box-Pierce and Ljung-Box tests. stat.ethz.ch. https://stat.ethz.ch/R-manual/R-devel/library/stats/html/box.test.html Accessed 05 Jun 2016
Wuertz, D., Setz, T., Chalabi, Y. (2017) Package fextremes package: rmetrics—modelling extreme events in finance. https://cran.r-project.org/web/packages/fExtremes/fExtremes.pdf . Accessed 12 Nov 2017
Vila R, Çankaya MN (2021) A bimodal weibull distribution: properties and inference. J Appl Stat 10:1–19
Vila R, Ferreira L, Saulo H, Prataviera F, Ortega E (2020a) A bimodal gamma distribution: properties, regression model and applications. Statistics. 54(3):469–493
Vila R, Leao J, Saulo H, Shahzad MN, Santos-Neto M (2020b) On a bimodal Birnbaum-Saunders distribution with applications to lifetime data. Braz J Prob Stat 34:495–518
Xue J (2012) Loop tiling for parallelism. Springer, US

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.