Brazilian Journal of Probability and Statistics
2017, Vol. 31, No. 4, 801–820
https://doi.org/10.1214/17-BJPS378

A Bayesian approach to extended models for exceedance

Fernando Ferraz do Nascimento[a] **and Marcelo Bourguignon Pereira**[b]

[a]*Universidade Federal do Piauí, Brazil*
[b]*Universidade Federal do Rio Grande do Norte, Brazil*

Abstract. In extreme value theory, the generalized Pareto distribution (GPD) is a family of continuous distribution used to model the tail of the distribution to values higher than a threshold u. Several works have used this method to approximate the tail of distribution. In this paper, we propose two extensions of GPD, including an additional shape parameter, to provide a more flexible distribution for exceedance. Some properties of these approximations are presented. Inference for these extensions were performed under the Bayesian paradigm, and the results showed fit improvement when compared with the standard GPD in applications to environmental and financial data.

1 Introduction

1.1 Extreme value theory

Extreme value analysis has become of fundamental importance in recent decades and it can basically be divided into two areas; maxima and exceedance analysis. The first consists in observing the maximum observation of each block of size n, where the limit distribution when $n \to \infty$ is known as Generalized Extreme Value (GEV) distribution (von Mises, 1954; Jenkinson, 1955). The other approach consists in modeling the tail of the data, from a threshold u, using the limit distribution known as GPD. Coles (2001), Embrechts, Kluppelberg and Mikosch (1997) and Ferreira and de Haan (2006) are some important references about the main models of extreme values as well as its properties.

In both approaches, the distribution is a limit case, and what is done in practice with data analysis is to choose a measure which is high enough (n to the maximum or u for excesses), such as the resulting observations can be well modeled by the boundary case. However, measures that are too high can result in quite a few observations, resulting in large variance estimates. Because of this limitation, it is necessary to search for more flexible models for the distribution of extreme values.

The cumulative distribution function (c.d.f.) and probability density function (p.d.f.) of the GPD are given by (Pickands, 1975)

$$G(x; u, \sigma, \xi) = \begin{cases} 1 - \left[1 + \xi(x-u)/\sigma\right]^{-1/\xi}, & \xi \neq 0, \\ 1 - \exp\left[-(x-u)/\sigma\right], & \xi \to 0, \end{cases} \tag{1.1}$$

Key words and phrases. Extreme value theory, exceedance analysis, generalized Pareto distribution, generalized classes of distributions, MCMC.
Received September 2016; accepted September 2017.

and

$$g(x; u, \sigma, \xi) = \begin{cases} \sigma^{-1}[1 + \xi(x - u)/\sigma]^{-(1+\xi)/\xi}, & \xi \neq 0, \\ \sigma^{-1} \exp[-(x - u)/\sigma], & \xi \to 0. \end{cases} \tag{1.2}$$

Functions (1.1) and (1.2) are defined in $x > u$ for $\xi > 0$ and $0 \leq (x - u) \leq -\sigma/\xi$ for $\xi < 0$. Pickands (1975) and Davison and Smith (1990) show properties that justify the use of GPD, such as the threshold stability, that is, if X follows the GPD, then the conditional distribution of $(X - u|X > u)$ has also GPD.

In extreme values it is of utmost importance to have the knowledge of extreme measures for the data. An extreme quantile of distribution is a value z_p such that $P(X < z_p) = p$, for values of p close to 1. For the GPD, the quantile function can be obtained inverting equation (1.1), obtaining

$$z_p = \begin{cases} u + \dfrac{[(1 - p)^{-\xi} - 1]\sigma}{\xi}, & \xi \neq 0, \\ u - \sigma \log(1 - p), & \xi \to 0, \end{cases} \tag{1.3}$$

where $p \in (0, 1)$.

This formula is valid only if we consider all the observations above the threshold. However, in real data, most of the data is below the threshold, and the value of u is estimated as the quantile $1 - N_u/N$, where N is the total number of observations, and N_u is the number of observations above the threshold. In this case, the quantile function z_p represented by all observations, including values below the threshold, is given by replacing p in equation (1.3) for $p^* = 1 - (1 - p)N/N_u$.

Consider a random variable with twice differentiable distribution function $F_X(x)$ as well as density function $f_X(x)$. Pickands (1986) showed that the scaled excess random variable $(X - u)/h(u)|X > u$ converges in distribution to the GP distribution if $\lim_{u \to x^{F_X}} h'_X(u) = \xi$, where $x^{F_X} = \sup\{x : F_X(x) < 1\}$, $h(x)$ is the reciprocal hazard function of X given by $h_X(x) = [1 - F_X(x)]/f_X(x)$ and $h'_X(x) = dh_X(x)/dx$. Thus, for the GPD, we have

$$\lim_{u \to x^{F_X}} h'_X(u) = \begin{cases} \xi, & \xi \neq 0, \\ 0, & \xi \to 0. \end{cases}$$

1.2 Generalization of distributions

There has been an increased interest in defining new classes of univariate continuous distribution introducing additional shape parameters to the baseline model, as lifetime (Gupta and Kundu, 2001), Environmental (Ristić and Balakrishnan, 2012), Medical (Ortega, Cordeiro and Kattan, 2012) and Economical (McDonald and Xu, 1993), there is a need for more flexible models for real data, as well as increasing some levels of skewness and kurtosis.

Marshall and Olkin (1997) derived an important method of including an extra parameter to a given baseline model thus defining extended distribution. The Marshall and Olkin ("MO" for short) transformation furnishes a wide range of behaviors with respect to the baseline distributions. The geometrical and inferential properties associated with the generated distribution depend on the values of the extra parameter. These characteristics provide more flexibility to the MO generated distributions. Let $\overline{G}(x) = 1 - G(x)$ and $g(x) = dG(x)/dx$ be the survival and density functions of a continuous random variable X with baseline c.d.f. $G(x)$. Then, the MO extended distribution has survival function given by

$$\overline{F}(x;\delta) = \frac{\delta\overline{G}(x)}{1-\overline{\delta}\overline{G}(x)} = \frac{\delta\overline{G}(x)}{G(x)+\delta\overline{G}(x)}, \qquad x \in \mathcal{X} \subseteq \mathbb{R}, \delta > 0, \tag{1.4}$$

where $\overline{\delta} = 1 - \delta$. Note that, when $\delta = 1$, $\overline{F}(x) = \overline{G}(x)$. The family (1.4) has p.d.f. given by

$$f(x;\delta) = \frac{\delta g(x)}{[1-\overline{\delta}\overline{G}(x)]^2}, \qquad x \in \mathcal{X} \subseteq \mathbb{R}, \delta > 0.$$

Recently, Cordeiro, Alizadeh and Marinho (2016) proposed and studied a broad family of univariate distributions through a particular case of half-logistic distribution. This new family stems from the general class: if $G(x)$ denotes the baseline c.d.f. of a random variable, then a generalized class of distributions can be defined by

$$F(x;\lambda) = \frac{1-[1-G(x)]^{\lambda}}{1+[1-G(x)]^{\lambda}}, \qquad x \in \mathcal{X} \subseteq \mathbb{R}, \tag{1.5}$$

where $\lambda > 0$ is an additional shape parameter. This family of distributions has p.d.f. given by

$$f(x;\lambda) = \frac{2\lambda g(x)[1-G(x)]^{\lambda-1}}{\{1+[1-G(x)]^{\lambda}\}^2}. \tag{1.6}$$

Equation (1.6) will be most tractable when $G(x)$ and $g(x)$ have simple expressions.

In extreme value modeling, Nascimento, Bourguignon and Leão (2016) provided new extended models to the GEV distribution as a baseline function, showing advantages compared with the standard GEV distribution. In the exceedance analysis, the most recent generalizations of the GP distribution were proposed by Papastathopoulos and Tawn (2013), which extended the GDP by defining the beta GDP, gamma GDP and exponentiated Pareto distributions, respectively, based on the class of generalized distributions introduced by Eugene, Lee and Famoye (2002), Zografos and Balakrishnan (2009) and Mudholkar and Hutson (1996). Papastathopoulos and Tawn (2013) referred to their generalizations as EGP1, EGP2 and EGP3 distributions, respectively. The generalized distributions are obtained by taking any parent $G(x)$ distribution in the c.d.f. of beta and gamma distributions with one and two additional shape parameters, which role is to introduce skewness and to vary tail weight. In fact, Papastathopoulos and Tawn (2013) wrote: "The inclusion of this parameter offers additional structure for the main body of the distribution, improves the stability of the modified scale, tail index and return level estimates to threshold choice and allows a lower threshold to be selected". However, the distributions given by Papastathopoulos and Tawn (2013) are complicated. The c.d.f. and p.d.f. of the generalizations involve the incomplete beta function and the incomplete gamma function.

This paper presents two new models for exceedance, based on extensions of the standard GPD, with the hope it yields a "better fit" in certain exceedance analysis (see Section 5). The inclusion of additional parameters in the GDP role is to introduce skewness, generate a distribution with heavy tails, offer additional structure for the main body of the distribution and improve the stability of the modified scale (Papastathopoulos and Tawn, 2013), that is, the new models are more flexible than the GPD. Initially, we show cumulative, density, quantile function, and some of the properties. These new generalizations have some advantages in terms of tractability, since it does not involve any special function such as do the beta and gamma functions. Furthermore, the proposed distributions have only two and three parameters, in contrast with some generalizations of the GPD which have four parameters (Papastathopoulos and Tawn, 2013). The parameter estimation of these new models are done under the Bayesian paradigm.

The paper will unfold as follows. We defined two generalizations of the GDP: the Marshall–Olkin GPD and the half-logistic GPD. Some statistical properties of the new distributions are derived. In Section 3, inference procedure is carried out under the Bayesian paradigm. In Section 4, a simulation study is performed in order to assess the accuracy of the estimators. In Section 5, four illustrative applications in environmental and financial data sets are investigated. Finally, concluding remarks are presented in Section 6.

2 Construction of extension for exceedance

In this section, we present two density functions that are generalizations of the GPD density.

2.1 The Marshall–Olkin generalized Pareto distribution (MOGPD)

Taking the GPD as the baseline model in (1.1), and applying the Marshall–Olkin generalization as in (1.4), we have

$$F(x;u,\sigma,\xi,\delta)=\begin{cases}\dfrac{1-[1+\xi(x-u)/\sigma]^{-1/\xi}}{1-\overline{\delta}[1+\xi(x-u)/\sigma]^{-1/\xi}}, & \xi\neq 0,\\[2ex] \dfrac{1-\exp[-(x-u)/\sigma]}{1-\overline{\delta}\exp(-x/\sigma)}, & \xi\to 0,\end{cases}$$

where $x > u$ for $\xi > 0$ and $0 \le (x-u) \le -\sigma/\xi$ for $\xi < 0$ with $\overline{\delta} = 1-\delta, \delta > 0$. The respective p.d.f. is

$$f(x; u, \sigma, \xi, \delta) = \begin{cases} \dfrac{\delta[1+\xi(x-u)/\sigma]^{-(1+\xi)/\xi}}{\sigma\{1-\overline{\delta}[1+\xi(x-u)/\sigma]^{-1/\xi}\}^2}, & \xi \neq 0, \\ \dfrac{\delta \exp[-(x-u)/\sigma]}{\sigma\{1-\overline{\delta}\exp[-(x-u)/\sigma]\}^2}, & \xi \to 0. \end{cases}$$

The MOGPD includes the GDP when $\delta = 1$. When $\delta \to 0^+$, the MOGPD converges to a distribution degenerated at zero. Hence, the parameter δ can be interpreted as a degeneration parameter.

The quantile function of the MOGPD is

$$z_p = \begin{cases} u + \dfrac{[(\frac{1-p}{1-\overline{\delta}p})^{-\xi} - 1]\sigma}{\xi}, & \xi \neq 0, \\ u - \sigma \log\left(\dfrac{1-p}{1-\overline{\delta}p}\right), & \xi \to 0, \end{cases} \tag{2.1}$$

where $p \in (0, 1)$. There are some advantages of the MOGPD over the EGP1 and EGP2 distributions (Papastathopoulos and Tawn, 2013), especially in terms of tractability since its distribution and quantile functions do not involve any special functions.

Proposition 2.1. *If X is a MOGPD, then*

$$\lim_{u \to x^{F_X}} h'_X(u) = \begin{cases} \xi, & \xi \neq 0, \\ 0, & \xi \to 0. \end{cases}$$

The proof is shown in the Appendix.

2.2 The half-logistic generalized Pareto distribution (HLGPD)

Taking the GPD as the baseline model in (1.1), and applying the half-logistic generalization as in (1.5), we have

$$F(x; u, \sigma, \xi, \lambda) = \begin{cases} \dfrac{1-[1+\xi(x-u)/\sigma]^{-\lambda/\xi}}{1+[1+\xi(x-u)/\sigma]^{-\lambda/\xi}}, & \xi \neq 0, \\ \dfrac{1-\exp[-\lambda^*(x-u)]}{1+\exp[-\lambda^*(x-u)]}, & \xi \to 0, \end{cases}$$

where $x > u$ for $\xi > 0$ and $0 \le (x-u) \le -\sigma/\xi$ for $\xi < 0$, with $\lambda > 0$ is an additional shape parameter (Cordeiro, Alizadeh and Marinho, 2016), and $\lambda^* = \lambda/\sigma$.

The p.d.f. of the HLGPD is

$$f(x;u,\sigma,\xi,\lambda)=\begin{cases}\dfrac{2\lambda[1+\xi(x-u)/\sigma]^{-(\lambda+\xi)/\xi}}{\sigma\{1+[1+\xi(x-u)/\sigma]^{-\lambda/\xi}\}^2} & \xi\neq 0,\\ \dfrac{2\lambda^*\exp[-\lambda^*(x-u)]}{\{1+\exp[-\lambda^*(x-u)]\}^2}, & \xi\to 0.\end{cases}$$

Note that the p.d.f. above is very simple, since it does not involve any special function such as does the beta function.

The quantile function of HLGPD is given by

$$z_p=\begin{cases}u+\dfrac{[(\frac{1-p}{1+p})^{-\xi/\lambda}-1]\sigma}{\xi}, & \xi\neq 0,\\ u-\dfrac{1}{\lambda^*}\log\Big(\dfrac{1-p}{1+p}\Big), & \xi\to 0,\end{cases}\tag{2.2}$$

where $p\in(0,1)$.

Remark 2.1. If we consider $\lambda^*=\lambda/\sigma$ and $\xi^*=\xi/\sigma$, the HLGPD can be written in function of just two parameters λ^* and ξ^*, and the density can be rewritten as, for $\xi\neq 0$

$$f(x;u,\xi^*,\lambda^*)=\frac{2\lambda^*[1+\xi^*(x-u)]^{-(\lambda^*+\xi^*)/\xi^*}}{\{1+[1+\xi^*(x-u)]^{-\lambda^*/\xi^*}\}^2}.\tag{2.3}$$

For inference purposes in the next sections, it would be preferable to use the re-parametrized density in (2.3) for the HLGPD.

Remark 2.2. The HLPGD with $\lambda=1$ is the MOGPD with $\delta=2$.

Proposition 2.2. *If X is a HLGPD, then*

$$\lim_{u\to x^{F_x}} h'_X(u)=\begin{cases}\dfrac{\xi}{\lambda}, & \xi\neq 0,\\ 0, & \xi\to 0.\end{cases}$$

The proof is shown in the Appendix.

2.3 Return levels

In extreme value analysis, something that is of much more importance than the knowledge about central position and dispersion of the distribution is how to predict the frequency of rare events, as they usually cause catastrophes and damage to society as a whole. With the fact that these events are in tail of the distribution, they are represented by a high quantile. So it is through the high quantiles of these distributions that we take a measure called return levels.

The $t-th$ return level is the event expected once every t periods of time, given by the quantile $p = 1 - 1/t$ of distribution. As an example, the return level in $t = 10$ is such a rare event that it is expected to occur once every 10 time periods, represented by the quantile $p = 1 - 1/10 = 0.90$. In exceedance distribution, a heavy tail represents higher frequencies of extreme events, and these values represent a lower quantile than a distribution with a light tail. In this section, we show the results of standard GPD quantiles compared with the generalizations.

Proposition 2.3. *Let $z_{M,p}$ be the quantile p of MOGPD with parameters (σ, ξ, δ) and let $z_{G,p}$ be the quantile p for the GPD with parameters (σ, ξ). Then,*

1. *If $\delta < 1$, $z_{G,p} > z_{M,p}$.*
2. *If $\delta > 1$, $z_{G,p} < z_{M,p}$.*

The proof is presented in the Appendix.

Proposition 2.4. *Let $z_{H,p}$ be the quantile p of HLGPD with parameters (σ, ξ, λ) and let $z_{G,p}$ be the quantile p for the GPD with parameters (σ, ξ). Then,*

1. *If $\xi = 0$, $z_{G,p} > z_{H,p}$ for $\lambda > \lambda^{**}$ and $z_{G,p} < z_{H,p}$ for $\lambda < \lambda^{**}$.*
2. *If $\xi < 0$, then $z_{G,p} > z_{H,p}$ and if $\xi > 0$, then $z_{G,p} < z_{H,p}$,*

where $\lambda^{**} = -\log(1 + p)$.

The proof is presented in the Appendix.

3 Estimation and inference

3.1 Likelihood function

Inference for the proposed model can be performed using classical methods based on the maximum likelihood estimators, or under the Bayesian approach. In both cases the knowledge of likelihood function is necessary. In exceedance models, we perform the likelihood function in the upper points of a threshold.

Let $x_1, \dots, x_n$ be observed values from the MOGPD with parameters σ, ξ and δ. Let $\boldsymbol{\theta}_1 = (\sigma, \xi, \delta)^\top$. The total log-likelihood function for $\boldsymbol{\theta}_1$ is given by ($\xi \neq 0$)

$$\ell^{\text{MOGPD}}(\boldsymbol{\theta}_1) = -n \log(\sigma/\delta) - (1/\xi + 1) \sum_{i=1}^{n_u} \log\bigl[1 + \xi(x_i - u)/\sigma\bigr]$$
$$- 2 \sum_{i=1}^{n_u} \log\bigl\{1 - \bar{\delta}\bigl[1 + \xi(x_i - u)/\sigma\bigr]^{-1/\xi}\bigr\},$$

where n_u is the number of points above the threshold, provided that $1 + \xi(x_i - u)/\sigma > 0$, for $i = 1, \dots, n$.

Now, let $x_1, \dots, x_n$ be observed values from the HLGPD with parameters σ, ξ and δ. Let $\boldsymbol{\theta}_2 = (\sigma, \xi, \lambda)^\top$. Then, the total log-likelihood function for $\boldsymbol{\theta}_2$ is given by ($\xi \neq 0$)

$$\ell^{\mathrm{HLGPD}}(\boldsymbol{\theta}_2) = -n \log[\sigma/(2\lambda)] - (\lambda/\xi + 1) \sum_{i=1}^{n_u} \log[1 + \xi(x_i - u)/\sigma] - 2 \sum_{i=1}^{n_u} \log\{1 + [1 + \xi(x_i - u)/\sigma]^{-\lambda/\xi}\},$$

provided that $1 + \xi(x_i - u)/\sigma > 0$, for $i = 1, \dots, n$.

The maximum likelihood estimators are defined as the values that maximize the log-likelihood function. However, the classical regularity conditions for the asymptotic properties of maximum likelihood estimators are not satisfied. When $\xi < -1$, the maximum likelihood estimators do not exist. For $-1 < \xi < -0.5$, they can have problems (Castillo and Hadi, 1997). For more details, see Smith (1985). Thus, in this paper, the parameter estimation of these new distributions was done under the Bayesian paradigm. The Bayesian approach directly allows the use of the likelihood function, while other works, for example Papastathopoulos and Tawn (2013), perform the inference based on quantiles of the exceedance distribution. They use asymptotic properties to find the distribution of the parameters, whereas the posterior distribution is exact.

3.2 Bayesian approach and MCMC

The Bayesian paradigm is used to make the inference for the models proposed in this work. For the GPD distribution, works as Do Nascimento, Gamerman and Lopes (2012) show how to proceed with inference for this model, and as their work, we used the Jeffreys prior proposed in Eugenia Castellanos and Cabras (2007) that obtained the non-informative prior for (σ, ξ) as

$$\pi(\sigma, \xi) \propto \sigma^{-1}(1 + \xi)^{-1}(1 + 2\xi)^{-1/2}, \qquad \xi > -0.5, \sigma > 0. \tag{3.1}$$

They also showed that this prior leads to proper posterior distributions.

Gamma prior with large variance is used for the additional shape parameter for each generalization, of the δ for MOGPD and λ for HLGPD. We choose prior Gamma(a, b) with values of $a = 0.001$ and $b = 0.001$ in a non-informative prior scenario with mean 1 and variance 1000. Combining the prior with the likelihood information, we obtained the posterior distribution for each generalization.

Let $\mathbf{x} = (x_1, \dots, x_n)^\top$ be observed values from the MOGPD with parameters σ, ξ and δ. Let $\boldsymbol{\theta}_1 = (\sigma, \xi, \delta)^\top$. The posterior distribution of $\boldsymbol{\theta}_1$ is given by ($\xi \neq 0$)

$$\pi_M(\boldsymbol{\theta}_1|\mathbf{x}) \propto \frac{\delta^{a-1} e^{-b\delta}}{\sigma(1 + \xi)(1 + 2\xi)^{1/2}} \left(\frac{\delta}{\sigma}\right)^{n_u} \prod_{i=1}^{n_u} \left\{ \frac{[1 + \frac{\xi}{\sigma}(x_i - u)]^{-\frac{(1+\xi)}{\xi}}}{[1 - \bar{\delta}(1 + \frac{\xi}{\sigma}(x_i - u))^{-\frac{1}{\xi}}]^2} \right\},$$

where n_u is the number of points above the threshold, provided that $1 + \xi(x_i - u)/\sigma > 0$, for $i = 1, \dots, n$.

Let $\mathbf{x} = (x_1, \dots, x_n)^\top$ be observed values from the HLGPD with parameters σ, ξ and δ. Let $\boldsymbol{\theta}_2 = (\sigma, \xi, \lambda)^\top$. The proportional of posterior distribution of $\boldsymbol{\theta}_2$ is given by ($\xi \neq 0$)

$$\pi_H(\boldsymbol{\theta}_2|\mathbf{x}) \propto \frac{\lambda^{a-1}e^{-b\lambda}}{\sigma(1+\xi)(1+2\xi)^{1/2}}\left(\frac{2\lambda}{\sigma}\right)^{n_u}\prod_{i=1}^{n_u}\left\{\frac{[1+\frac{\xi}{\sigma}(x_i-u)]^{-\frac{(\lambda+\xi)}{\xi}}}{[1+(1+\frac{\xi}{\sigma}(x_i-u))^{-\frac{\lambda}{\xi}}]^2}\right\}$$

provided that $1 + \xi(x_i - u)/\sigma > 0$, for $i = 1, \dots, n$.

Inference cannot be performed analytically and, because of that, approximating MCMC methods were used (Gamerman and Lopes, 2006), notwithstanding, other methods for estimation should be used as quadrature rules. That can be seen, for example, in Santos, Gamerman and Franco (2017). Convergence was assessed by running two parallel chains with different starting values. The algorithm consists in sample points for (σ, ξ) and the additional parameter accepts these points according to the Metropolis rule. The sampling of (σ, ξ) is similar to Do Nascimento, Gamerman and Lopes (2012). With the MOGPD, for the δ parameter, we performed the Metropolis step using the Gamma distribution as proposed. For the HLGPD, the parameter λ^* is also proposed from the Gamma distribution. In all the situations, the proposal distribution has the mean located at the previous step of the chain, and the variance is chosen to result in a reasonable convergence rate, usually around 0.44 according to the optimal rate proposed by Roberts and Rosenthal (2009).

4 Simulation study

Simulations were performed in different types of configuration with the parameters, from the following two extensions, the HLGPD and MOGPD. The objective of the simulation study is to verify some aspects of Bayesian inference, such as parameter identifiability and algorithm efficiency in retrieving the true values of the parameters.

In total $K = 500$ replications were simulated with sample size $n = 1000$ points in a configuration where $(\sigma, u) = (20, 50)$. We simulated points for $\xi = (-0.4, 0.5)$, with the purpose of verifying the behavior of the extension for the short and heavy tail. With the additional shape parameter in MOGPD, it was simulated points with parameters $\delta = (0.3, 0.8, 2)$ and for HLGPD, the points were simulated with additional parameters $\lambda = (0.3, 0.8, 2)$. By consequence of Remark 2.1, the HLGPD can reduce to a two parameter problem, and estimation and inference were performed to the parameters $\lambda^* = \lambda/\sigma$ and $\xi^* = \xi/\sigma$, since the joint estimation of the 3 parameters of the HLGPD model could cause non-identifiability problems or higher variance of the estimates.

For each one of the $K = 500$ replications of sample size $n = 1000$, the inference was performed using both new generalized model, proposed in this work, and standard GPD and EGPD's distributions of Papastathopoulos and Tawn (2013) in order to make comparisons i.e., we generated observations of the proposed distributions and fitted it with the true distribution (proposed distributions) and the GPD and EGPD's for each replicate. The Deviance Information Criterion (DIC, Spiegelhalter et al. (2002)) and Root Mean Square Error of each replicate was collected to compare this fit measure of the proposed model against GPD model. A criteria to show if the estimation of the proposed distribution fits better than the GPD was the Wilcoxon test that compares the median of 500 DIC's produced by the proposed model against the median of 500 DIC's of the GPD and EGPD's models.

Table 1 shows the main results for the 500 replications in each parameter configuration. We can see that in most of the configurations, the average in posterior means of the parameters is near the true value, with a low average of the Root Mean Square Error. The model is less precise in MOGPD when $\xi = 0.5$. In all replication scenarios, the DIC average to the proposed model is lower than the DIC of the GPD and EGPD's fit, showing the advantage of using an extension. For the MOGPD model, the difference between the DIC of the GPD is less significant when $\delta = 0.8$ and $\xi = -0.4$, due to this case, it is the most similar to the GPD ($\delta = 1$). For all the other cases, the MOGPD fit showed being significantly different and better than the other extensions. For the HLGPD model, the difference between the DIC is more significant in cases where $\lambda = 0.3$. For $\xi > 0$, the results of DIC suggest a similar adjustment between the HLGPD, GPD, and EGPD fits for $\lambda = 0.8$ and $\lambda = 2$. Looking at the coverage probability of 95% credibility interval, we can note that in almost all scenarios, more than 90% of replicates contain the true value of the parameter inside the credibility interval. The exception was the case MOGPD when $\xi = 0.5$, which is less precise.

Figure 1 shows the return level plot estimated by the proposed model in one of the replications, compared with the true returns. We can see the efficiency of the estimation procedures in all of the configurations, the true return line is inside of the credibility interval in estimated returns.

5 Applications to real data

In order to show the utility of the new distributions, extreme data was analyzed in the environmental and financial areas. Environmental data consists of daily river flow and rainfall observations, which were previously analyzed by Do Nascimento, Gamerman and Lopes (2012) using the GPD distribution. Financial data consists of daily stock returns, used previously in Nascimento, Gamerman and Lopes (2016). In order to avoid excess clustering and decrease the dependence between consecutive days, maxima of sets of m days were considered. When analyzing SP500 and

Table 1 *Summary statistics for* 500 *replicates*

	APM	AMSE	CPI	APM	AMSE	CPI	APM	AMSE	CPI
MOGPD	$\delta = 0.3, \xi = -0.4$			$\delta = 0.8, \xi = -0.4$			$\delta = 2, \xi = -0.4$		
δ	0.307	0.059	0.896	0.801	0.150	0.926	2.008	0.362	0.924
σ	20.24	2.81	0.876	20.43	2.46	0.908	20.28	2.028	0.938
ξ	−0.402	0.081	0.916	−0.409	0.065	0.920	−0.405	0.051	0.940
DIC_{MOGPD}	6139.6	Wx.test	–	7052.0	Wx.test	–	7497.8	Wx.test	–
DIC_{GPD}	6170.4	<0.001	–	7053.9	0.320	–	7548.2	<0.001	–
DIC_{EGPD1}	6159.7	<0.001	–	7053.1	0.459	–	7505.9	<0.001	–
DIC_{EGPD2}	6159.4	<0.001	–	7053.5	0.458	–	7507.9	<0.001	–
DIC_{EGPD3}	6159.0	<0.001	–	7052.1	0.439	–	9999.7	<0.001	–
	$\delta = 0.3, \xi = 0.5$			$\delta = 0.8, \xi = 0.5$			$\delta = 2, \xi = 0.5$		
δ	0.171	0.219	0.226	0.455	0.657	0.374	1.319	1.741	0.574
σ	54.22	36.18	0.218	72.95	56.69	0.372	96.46	83.73	0.372
ξ	0.302	0.219	0.256	0.409	0.310	0.386	0.330	0.207	0.572
DIC_{MOGPD}	7015.8	Wx.test	–	8613.7	Wx.test	–	9967.1	Wx.test	–
DIC_{GPD}	7071.4	<0.001	–	8647.9	<0.001	–	9996.0	<0.001	–
DIC_{EGPD1}	7071.2	<0.001	–	8646.7	<0.001	–	9996.5	<0.001	–
DIC_{EGPD2}	7070.8	<0.001	–	8646.7	<0.001	–	9996.5	<0.001	–
DIC_{EGPD3}	7070.9	<0.001	–	8646.5	<0.001	–	9996.8	<0.001	–
HLGPD	$\lambda = 0.3, \xi = -0.4$			$\lambda = 0.8, \xi = -0.4$			$\lambda = 2, \xi = -0.4$		
λ	0.301	5.699	0.934	0.789	15.21	0.914	2.019	37.98	0.936
ξ	−0.399	7.500	0.952	−0.399	7.601	0.922	−0.389	7.610	0.932
DIC_{HLGPD}	7320.2	Wx.test	–	7667.24	Wx.test	–	6664.8	Wx.test	–
DIC_{GPD}	8580.1	<0.001	–	7788.1	<0.001	–	6688.4	<0.001	–
DIC_{EGPD1}	8006.6	<0.001	–	7694.1	<0.001	–	6674.8	0.002	–
DIC_{EGPD2}	7850.7	<0.001	–	7685.2	<0.001	–	6674.6	0.002	–
DIC_{EGPD3}	8106.2	<0.001	–	7703.7	<0.001	–	6674.9	<0.001	–
	$\lambda = 0.3, \xi = 0.5$			$\lambda = 0.8, \xi = 0.5$			$\lambda = 2, \xi = 0.5$		
λ	0.307	5.692	0.938	0.807	15.192	0.950	2.023	37.97	0.912
ξ	0.516	9.483	0.930	0.509	9.491	0.954	0.520	9.480	0.922
DIC_{HLGPD}	15,621.2	Wx.test	–	10,788.6	Wx.test	–	7918.0	Wx.test	–
DIC_{GPD}	15,621.8	0.462	–	10,790.6	0.853	–	7923.4	0.185	–
DIC_{EGPD1}	17,221.2	<0.001	–	10,790.1	0.380	–	7921.6	0.453	–
DIC_{EGPD2}	17,195.2	<0.001	–	10,788.1	0.640	–	7921.2	0.458	–
DIC_{EGPD3}	17,303.5	<0.001	–	10,791.6	0.500	–	7921.0	0.446	–

APM represents the average of posterior means, AMSE the average of the Root mean square error, and CPI is the coverage probability of 95% of the credibility intervals. DIC_m is the adjustment measure calculated for each simulation using the m model for estimation. Wx.test is the Wilcoxon-test and presents the p-value of test comparing the median the DIC to the 500 replicates from the proposed model compared with other useful models.

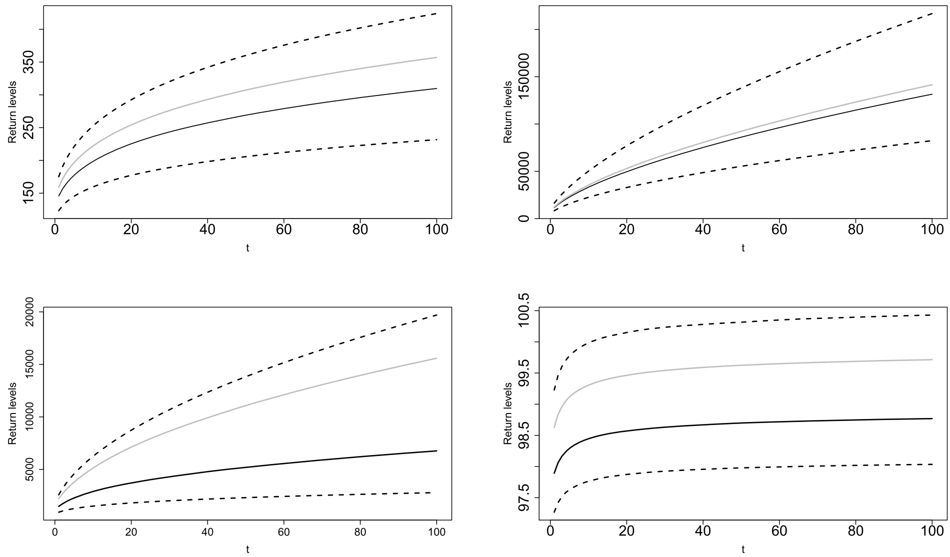

Figure 1 *Return Level plots for the simulations. Top: HLGPD with $\xi = 0.5$ with $\lambda = 2$ (left) and $\lambda = 0.8$ (right). Bottom: MOGPD with $\xi = 0.5$ and $\delta = 0.3$ (left) and $\xi = -0.4$ and $\delta = 0.8$ (right). Full lines: Posterior mean estimated by the proposed model. Dashed line: 95% C.I of the proposed model. Grey line: True value.*

Petrobrás data, we considered groups of $m = 5$ and $m = 3$ days, respectively. To the environmental data sets, which presents stronger daily dependence, we considered groups of $m = 15$. The previous work that used these applications treated the estimation of the threshold, and we will use these values which were obtained as fixed parameter for this work, as the main objective here is to compare the new distributions with the GPD. Figure 2 shows the histogram of the data sets to the values higher than to respective thresholds.

The proposed distributions were applied in several real data applications to environmental and financial data sets. We compared the two distributions of this work with standard GPD and the generalizations proposed in work by Papastathopoulos and Tawn (2013).

For each application, the best model was chosen using the DIC criteria. Table 2 shows the DIC fit measures for application data, where the GPD are the results for standard GPD; the MOGPD and HLGPD are the results for extensions of models described in Sections 2.1 and 2.2, respectively; the EGP1, EGP2 and EGP3 are the results of Examples 1, 2 and 3 of extensions proposed in Papastathopoulos and Tawn (2013), described on page 134.

Table 3 presents posterior mean and 95% credibility intervals for the parameters to each model. In the MOGPD case, values of $\delta \approx 1$ show the proximity of the estimation with the standard GPD. As for the HLGPD, the estimation we do not directly do for λ, but $\lambda^* = \lambda/\sigma$ and $\xi^* = \xi/\sigma$. We are not able to compare this

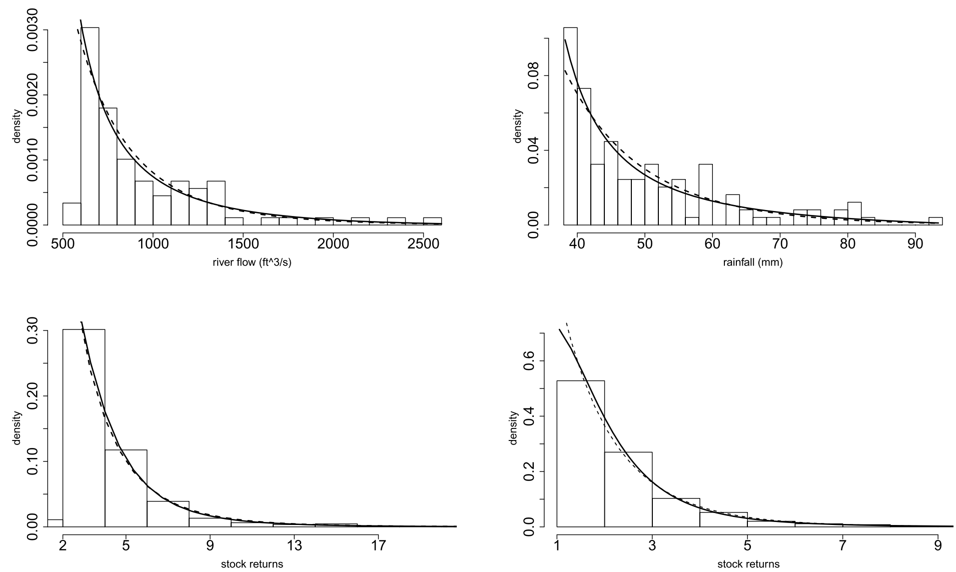

Figure 2 *Posterior predictive distribution for data. First line Left*: *Espirito Santo*; *Right*: *Grandola.Second line Left*: *Petrobrás*; *Right*: *SP*500. *Full line*: *Estimation using the best model according to DIC. Dashed line*: *Estimation using standard GPD.*

Table 2 *DIC measures for applications*

Model	GPD	MOGPD	HLGPD	EGP1	EGP2	EGP3
Rainfall–Grandola	864.76	**859.48**	868.93	865.70	865.59	865.75
River–Espirito Santo	1232.98	**1232.42**	1234.95	1234.08	1234.44	1234.51
Financial–Petrobrás	3151.17	**3127.71**	3133.54	3143.79	3143.89	3143.96
Financial–SP500	4697.55	4694.88	**4693.25**	4698.42	4697.80	4698.90

extension with the GPD ($\lambda = 1$). A method to perform this comparison in the case for HLGPD is to use the estimation of ξ/σ of the standard GPD and compare with ξ^* from HLGPD. We can see that for the two environmental applications, which the MOGPD model presented lower DIC, the value of the additional parameter δ is significantly lower than 1, while for Petrobrás this parameter is higher than 1. In the SP500 financial application, the credibility interval of parameter δ contains the particular case of GPD in the MOGPD model, while in the HLGPD model, the estimation of ξ^* was very close to ξ/σ from the GPD, indicating that this application is the one that the extensions have pointed out to be closer to the standard GPD.

Table 3 *Posterior means and 95% CI of the parameters in each model*

	MOGPD			HLGPD		GPD		
Model	δ	ξ	σ	λ^*	ξ^*	ξ	σ	ξ/σ
Grandola	0.39 (0.23; 0.69)	−0.37 (−0.49; −0.03)	27.05 (15.42; 36.77)	0.17 (0.12; 0.23)	0.05 (0.01; 0.12)	0.03 (−0.18; 0.29)	12.15 (8.88; 16.03)	0.01 (−0.01; 0.03)
Esp. Santo	0.50 (0.29; 0.98)	−0.12 (−0.34; 0.27)	643.1 (300.8; 873.6)	0.006 (0.004; 0.009)	0.002 (0.0006; 0.006)	0.11 (−0.09; 0.40)	339.6 (240.3; 457.4)	0.0004 (−0.001; 0.002)
Petrobrás	1.61 (1.16; 2.31)	0.22 (0.15; 0.32)	1.36 (1.01; 1.79)	0.88 (0.80; 0.98)	0.23 (0.16; 0.31)	0.16 (0.10; 0.23)	1.90 (1.73; 2.08)	0.09 (0.05; 0.12)
SP500	1.17 (0.87; 1.71)	0.11 (0.02; 0.18)	1.07 (0.83; 1.38)	1.43 (1.33; 1.53)	0.31 (0.23; 0.39)	0.07 (0.03; 0.12)	1.19 (1.12; 1.27)	0.06 (0.02; 0.11)

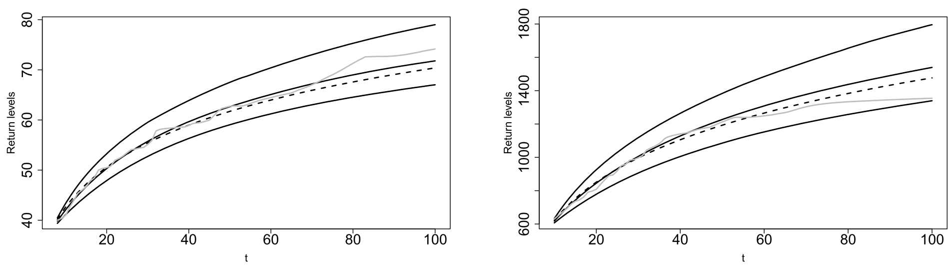

Figure 3 *Return Level plots for environmental data. Left: Grandola; Right: Espirito Santo. Full line: Estimation using the best model according to DIC, with respective 95% CI. Dashed line: Estimation using standard GPD. Grey line: Empirical returns.*

5.1 Environmental data sets: River flood and rainfall

The first analysis consists in the measurement of the amount of rain in Grandola monitoring station in the South of Portugal. The data was recorded daily from 1931 to 2008. A sample size of 925 fortnightly maxima was analysed. The second environmental analysis consists in the measurement of the levels of the Espirito Santo river flow located in Northeast of Puerto Rico in ft^3/s. The data was recorded on a daily basis from April 1967 to September 2002 and it was also analyzed in other works such as Do Nascimento, Gamerman and Lopes (2012). A size of 864 fortnightly maxima data was analysed. The threshold parameter here is fixed, and was chosen according to the work by Do Nascimento, Gamerman and Lopes (2012), who proposed a Bayesian approach to this parameter.

Table 2 shows that using an extension of GPD proposed in this work is the one which obtained better results. For Grandola and Espirito Santo data sets, the MOGPD presented a better fit. It is possible to observe this in the examples of Papastathopoulos and Tawn (2013), which produced very similar measurements, as this example comes from the same family of generalizations. Figure 2 shows the predictive lines of MOGPD and GPD curves fitting the datasets. The MOGPD curve appears to be in a different shape as a consequence of the additional parameter, and the datasets would be better fitted comparing with the standard GPD fitting.

Figure 3 shows the return level plot of Grandola and Espirito Santo river, using MOGPD (the best according to DIC) and standard GPD. It is possible to verify that the proposed model fits near the empirical quantile, and the return estimation by the MOGPD is usually a bit higher than the standard GPD. A rainfall of 65 mm is expected to occur in the Grandola station once every $t = 55$ periods of time, while a river quota of $1400 ft^3/s$ is expected to occur once every $t = 73$ periods of time.

5.2 Financial data sets: Petrobrás and SP500 index

The proposed models were applied in a study of two financial time series: Petrobrás and SP500 returns. The first is the biggest (non-banking) company in Brazil,

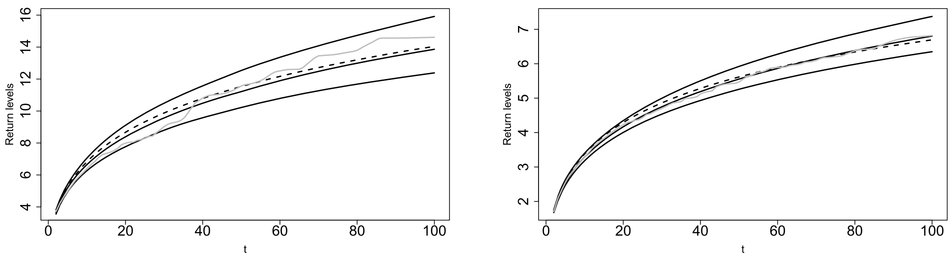

Figure 4 *Return Level plots for financial data. Left: Petrobrás; Right: SP*500. *Full line: Estimation using the best model according to DIC, with respective* 95% *CI. Dashed line: Estimation using standard GPD. Grey line: Empirical returns.*

while the second is a stock market index based on the markets capitalizations of 500 large companies which have common stock listed on the NYSE or NASDAQ. We collected daily closing prices p_t, and converted to returns $x_t = p_t/p_{t-1} - 1$, and then transformed to absolute returns $y_t = 100 \times |x_t - \bar{x}|$, where $\bar{x}$ is the sample average of x_t. The subtraction of $\bar{x}$ from each x_t is used to avoid zeros and the multiplication of 100 is introduced for convenience of the presentation. The threshold parameter here is fixed, and chosen according to the work of Nascimento, Gamerman and Lopes (2016), that analysed this same data set, proposing a Bayesian approach to this parameter. For SP500 data we analyzed data from January 2nd 1950 to February 18th 2014 or 3345 5-day maxima. For Petrobrás data we analysed data from October 8th 2000 to February 18th 2014 or 1131 3-day maxima.

Table 2 shows that the HLGPD of Section 2.2 presented the best fit measures for the SP500, while MOGPD had the best fit to Petrobrás. The second line of Figure 2 shows the predictive distribution of the financial examples of the best model (according of DIC) compared with standard GPD. We can verify that, although there is some difference between the curves, in both applications the predictive curve fits well the true histogram of the data.

Figure 4 shows the return level of the best extension model (according to DIC) against GPD. We can confirm that the proposed extension model and empirical returns present similar results. The return level of 10% is expected to occur once every 33 periods of times to Petrobrás, and as for SP500 data this same number of periods represents a return of approximately 5%.

6 Concluding remarks

In this paper, we proposed two extensions to the GPD, as a flexible alternative to standard GPD for exceedance analysis. Although GPD is considered a limit distribution of exceedance, as in practice we do not work in the limiting situation, the flexibility of the GPD extensions proposed can present better results. Therefore, we proposed a Bayesian approach to the estimation of the extensions, as well as the extensions of Papastathopoulos and Tawn (2013).

We showed important properties of each generalization, and performed an inference for each case. The inference procedure showed efficiency to recover value of parameters and quantile, with low variability, based on results of 500 replications. Application in real data sets showed an advantage in considering a new parameter in GPD. This work can be extended as well as proposing new generalizations to GPD, and to propose a method which chooses the best generalization for each data set. We hope these new models may attract wider applications for exceedance analysis.

Appendix

Proof of the Proposition 2.1

Let f and F the density and cumulative distribution of MOGPD, respectively. Then, for $\xi \neq 0$, we have

$$\begin{aligned}\frac{1-F(x;u,\sigma,\xi,\delta)}{f(x;u,\sigma,\xi,\delta)} &= \frac{[1+\xi(x-u)/\sigma]\{1-\overline{\delta}[1+\xi(x-u)/\sigma]^{-1/\xi}\}}{\sigma} \\ &= \sigma^{-1}\big[1+\xi(x-u)/\sigma\big]-\sigma^{-1}\overline{\delta}\big[1+\xi(x-u)/\sigma\big]^{-1/\xi+1}.\end{aligned}$$

Thus,

$$h'_X(x) = \xi - \overline{\delta}(\xi-1)\big[1+\xi(x-u)/\sigma\big]^{-1/\xi}.$$

For $\xi \to 0$,

$$\frac{1-F(x;u,\sigma,\xi,\delta)}{f(x;u,\sigma,\xi,\delta)} = \frac{1}{\sigma} - \frac{\overline{\delta}\exp[-(x-u)/\sigma]}{\sigma}.$$

Thus,

$$h'_X(x) = \overline{\delta}\frac{\exp[-(x-u)/\sigma]}{\sigma^2}.$$

Therefore,

$$\lim_{u \to x^{F_X}} h'_X(u) = \begin{cases} \xi, & \xi \neq 0, \\ 0, & \xi \to 0. \end{cases}$$

Proof of the Proposition 2.2

Let f and F the density and cumulative distribution of HLGPD, respectively. Then, for $\xi \neq 0$, we have

$$\frac{1-F(x;u,\sigma,\xi,\lambda)}{f(x;u,\sigma,\xi,\lambda)} = \frac{1+[1+\xi(x-u)/\sigma]}{\lambda^*} + \frac{1+[1+\xi(x-u)/\sigma]^{-\lambda^*/\xi^*+1}}{\lambda^*}.$$

Thus,

$$h'_X(x) = \frac{\xi}{\lambda} + (\xi^* - \lambda^*)\big[1 + \xi(x-u)/\sigma\big]^{-\lambda^*/\xi^*}.$$

For $\xi \to 0$,

$$\frac{1 - F(x; u, \sigma, \xi, \lambda)}{f(x; u, \sigma, \xi, \lambda)} = \frac{1}{\lambda^*} + \frac{\exp[-\lambda^*(x-u)]}{\lambda^*}.$$

Thus,

$$h'_X(x) = -\exp\big[-\lambda^*(x-u)\big].$$

Therefore,

$$\lim_{u \to x^{F_X}} h'_X(u) = \begin{cases} \dfrac{\xi}{\lambda}, & \xi \neq 0, \\ 0, & \xi \to 0. \end{cases}$$

Proof of the Proposition 2.3

Let $Z_{G,p}$ the quantile of GPD given in (1.3) and let $Z_{M,p}$ the quantile of MOGPD given in (2.1). Then, for $\xi = 0$,

$$\begin{aligned} Z_{G,p} - Z_{M,p} &= \sigma\big\{-\log(1-p) + \log(1-p) - \log\big[1 - (1-\delta)p\big]\big\} \\ &= \sigma\big\{-\log\big[1 - (1-\delta)p\big]\big\}. \end{aligned}$$

If $\delta < 1$, $[1-(1-\delta)p < 1]$, and then $Z_{G,p} - Z_{M,p} > 0$. If $\delta > 1$, $[1-(1-\delta)p > 1]$ and $Z_{G,p} - Z_{M,p} < 0$.

For $\xi \neq 0$,

$$\begin{aligned} Z_{G,p} - Z_{M,p} &= \frac{\sigma}{\xi}\left\{(1-p)^{-\xi} - \frac{(1-p)^{-\xi}}{[1-(1-\delta)p]}\right\} \\ &= \frac{\sigma}{\xi}(1-p)^{-\xi}\big\{1 - \big[1 - (1-\delta)p\big]^{\xi}\big\}. \end{aligned}$$

If $\delta < 1$, can be verified that for any value of ξ, $Z_{G,p} - Z_{M,p} > 0$.

Proof of the Proposition 2.4

Let $Z_{G,p}$ the quantile of GPD given in (1.3) and let $Z_{H,p}$ the quantile of HLGPD given in (2.2). Then, for $\xi = 0$,

$$\begin{aligned} Z_{G,p} - Z_{H,p} &= \sigma\big[-\log(1-p) + \log(1-p)/\lambda - \log(1+p)\big] \\ &= \sigma\big[(1/\lambda)\log(1-p) + \log\big(1-p^2\big)\big]. \end{aligned}$$

If $\lambda > \lambda^{**}$, $Z_{G,p} - Z_{H,p} > 0$. If $\lambda < \lambda^{**}$, $Z_{G,p} - Z_{H,p} < 0$.

For $\xi \neq 0$,

$$
\begin{aligned}
z_{G,p} - z_{H,p} &= \frac{\sigma}{\xi}\left[(1-p)^{-\xi} - \left(\frac{1-p}{1+p}\right)^{-\xi/\lambda}\right] \\
&= \frac{\sigma}{\xi}(1-p)^{-\xi}\left[1 - \left(\frac{1-p}{1+p}\right)^{-1/\lambda}\right].
\end{aligned}
$$

As $\lambda > 0$ and $(1-p)/(1+p) < 1$, if $\xi < 0$, $z_{G,p} - z_{H,p} > 0$ and if $\xi > 0$, $z_{G,p} - z_{H,p} < 0$.

Acknowledgments

We would like to thank the two anonymous referees as well as the two guest editors for the constructive suggestions and comments. This research work was supported in part by CNPq, grant number 472727/2013-3.

References

Castillo, E. and Hadi, A. S. (1997). Fitting the generalized Pareto distribution to data. *J. Amer. Statist. Assoc.* **92**, 1609–1620. MR1615270

Coles, S. (2001). *An Introduction to Statistical Modelling of Extreme Values*. *Springer Series in Statistics*. London: Springer-Verlag London, Ltd. MR1932132

Cordeiro, G. M., Alizadeh, M. and Marinho, R. D. M. (2016). The type I half-logistic family of distributions. *Journal of Statistical Computation and Simulation* **86**, 707–728.

Davison, A. C. and Smith, R. L. (1990). Models for exceedances over high thresholds (with discussion). *Journal of the Royal Statistical Society, Series B* **52**, 393–442.

Do Nascimento, F. F., Gamerman, D. and Lopes, H. F. (2012). A semiparametric Bayesian approach to extreme value estimation. *Stat. Comput.* **22**, 661–675. MR2865043

Embrechts, P., Kluppelberg, C. and Mikosch (1997). *Modelling Extremal Events for Insurance and Finance*. New York: Springer.

Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics Theory and Methods* **31**, 497–512.

Eugenia Castellanos, M. and Cabras, S. (2007). A default Bayesian procedure for the generalized Pareto distribution. *J. Statisti. Plann. Inference* **137**, 373–483. MR2298951

Ferreira, A. and de Haan, L. (2006). *Extreme Value Theory. And Introduction*. New York: Springer.

Gamerman, D. and Lopes, H. F. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, 2nd ed. Baton Rouge: Chapman and Hall/CRC.

Gupta, R. D. and Kundu, D. (2001). Exponentiated exponential family: An alternative to gamma and Weibull distributions. *Biometrical Journal* **43**, 117–130.

Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological events. *Quarterly Journal of the Royal Meteorology Society* **81**, 158–171.

Marshall, A. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika* **84**, 641–652.

McDonald, J. B. and Xu, Y. J. (1993). A generalization of the beta distribution with applications. *Journal of Econometrics* **66**, 133–152.

Mudholkar, G. S. and Hutson, A. D. (1996). The exponentiated Weibull family: Some properties and a flood data application. *Communications in Statistics Theory and Methods* **25**, 3059–3083.

Nascimento, F. F., Bourguignon, M. and Leão, J. (2016). Extended generalized extreme value distribution with applications in environmental data. *Hacettepe Journal of Mathematics and Statistics* **45**, 1847–1864.

Nascimento, F. F., Gamerman, D. and Lopes, H. F. (2016). Time-varying extreme pattern with dynamic models. *Test* **25**, 131–149.

Ortega, E. M. M., Cordeiro, G. M. and Kattan, M. W. (2012). The negative binomial beta Weibull regression model to predict the cure of prostate cancer. *Journal of Applied Statistics* **39**, 1191–1210.

Papastathopoulos, I. and Tawn, J. A. (2013). Extended generalized Pareto models for tail estimation. *Journal of Statistical Planning and Inference* **143**, 131–143.

Pickands, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics* **3**, 119–131.

Pickands, J. (1986). The continuous and differentiable domains of attraction in extreme value theory. *Annals of Probability* **14**, 996–1004.

Ristić, M. M. and Balakrishnan, N. (2012). The gamma exponentiated exponential distribution. *Journal of Statistical Computation and Simulation* **8**, 1191–1206.

Roberts, G. O. and Rosenthal, J. S. (2009). Examples of adaptive MCMC. *Journal of Computational and Graphical Statistics* **18**, 349–367.

Santos, T. R., Gamerman, D. and Franco, G. C. (2017). Reliability analysis via non-Gaussian state-space models. *IEEE Transactions on Reliability* **66**, 309–318.

Smith, R. L. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika* **72**, 67–90.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B, Methodological* **64**, 583–639.

von Mises, R. (1954). La distribution de la plus grande de nvaleurs. *American Mathematical Society* **2**, 271–294.

Zografos, K. and Balakrishnan, N. (2009). On families of beta and generalized gamma-generated distributions and associated inference. *Statistical Methodology* **6**, 344–362.

Departamento de Estatística
Universidade Federal do Piauí
Campus Ministro Petrônio Portela
Ininga, Teresina, PI 64049-550
Brazil
E-mail: fernandofn@ufpi.edu.br

Departamento de Estatística
Universidade Federal do Rio Grande do Norte
Natal-RN 59064-741
Brazil
E-mail: m.p.bourguignon@gmail.com