Modelling jointly low, moderate and heavy rainfall intensities without a threshold selection

Philippe Naveau,[1] Raphael Huser,[2] Pierre Ribereau,[3], and Alexis Hannart[4]

Corresponding author: Philippe Naveau, Laboratoire des Sciences du Climat et de l'Environnement, IPSL-CNRS, France. (naveau@lsce.ipsl.fr)

[1]Laboratoire des Sciences du Climat et de l'Environnement, IPSL-CNRS, France

[2]King Abdullah University of Science and Technology, Saudi Arabia

[3]Claude Bernard University Lyon 1, France

[4]CNRS, Buenos-Aires, Argentina

Abstract. In statistics, extreme events are classically defined as maxima over a block length (e.g., annual maxima of daily precipitation) or as excesses above a given large threshold. These definitions allow hydrologists and flood planners to apply Extreme-Value Theory (EVT) to their time series of interest. Even in the stationary univariate context, this approach has at least two main drawbacks. First, working with maxima or excesses implies that a lot of observations (those below the chosen threshold or smaller than the block maximum) are completely disregarded. The range of precipitation is artificially shopped down into two pieces, namely large intensities and the rest which, necessarily imposes different statistical models for each piece. Second, this strategy raises a non-trivial and very practical difficultly: how to choose the threshold, or equivalently the block size, which correctly discriminates between low and heavy rainfall intensities.

To address these issues, we propose a statistical model in which EVT results apply not only to heavy, but also to low precipitation amounts. Our model is in compliance with EVT at both ends of the spectrum and allows a smooth transition between the two tails, while keeping a low number of parameters. In terms of inference, we have implemented and tested two classical methods of estimation, likelihood maximisation and probability weighed moments. Last but not least, there is no need to choose a threshold to define low and high excesses. The performance and flexibility of this approach are illustrated on simulated and hourly precipitation recorded in Lyon, France.

1. Introduction

There exists a wide range of distribution families to statistically model rainfall intensities. For example, *Katz* [1977], *Vrac et al.* [2007], and *Wilks* [2006] argued that most of the precipitation variability can be approximated by Gamma distributions. It is, however, also well known [see, e.g., *Katz et al.*, 2002] [*Katz et al.*, 2002] that the tail of the Gamma distribution can be too light to capture heavy rainfall intensities. This leads to underestimate return levels and other quantities linked to high quantiles of precipitation amounts. Consequently, hydrological, societal and economical impacts associated with heavy rains (e.g., floods) can be misestimated. To solve this issue, a popular approach in hydrology [e.g. *Katz et al.*, 2002] is to disregard small and moderate precipitation values and to focus only on the largest rainfall amounts. The advantage of this strategy is that an elegant mathematical framework called *Extreme-Value theory* (EVT), originating from the pioneering work of *Fisher and Tippett* [1928] and regularly adapted during the last decades [e.g. *de Haan and Ferreira*, 2006], dictates the distribution of heavy precipitation. Specifically, EVT states that rainfall excesses, i.e., amounts of rain greater than a given threshold u, may be approximated by a Generalized Pareto (GP) distribution, provided the threshold and the number of observations are large enough and some mild conditions are satisfied (see Section 2 for the GP definition).

Numerous studies [see, e.g., *Katz et al.*, 2002; *Cooley et al.*, 2007] have illustrated how the GP distribution can be applied to climate and hydrology sciences. An obvious drawback is that the GP only models data exceeding a given high threshold [e.g., *Dupuis*, 1999], and one can wonder how to model the remaining observations (i.e., lower than the threshold)

or equivalently how to deal with the entire range of the data. Recently, there have been a few attempts at modelling the full range of the observations. Carreau and her co-authors (*Carreau and Bengio* [2006]; *Carreau et al.* [2009]; *Carreau and Vrac* [2011]) investigated a semi-parametric mixture model that combines hybrid densities built by stitching a Gaussian density with a heavy-tailed GP density. The estimation was performed with a neural network approach and applied in a regression context. This so-called hybrid Pareto model has many advantages, but also two drawbacks. First, it can produce negative values because the low part of the distribution is based on a Gaussian variable, an unwelcome feature for rainfall data. Second, the stitching between the Gaussian and the Pareto densities is obtained by imposing a strong constraint on the GP and Gaussian parameters. This automatically links the GP shape parameter with the bulk of the distribution. In the i.i.d. case, *Frigessi et al.* [2002] proposed another approach based on a mixture model of two components. The first one represents the bulk of the distribution and the second one focuses on the upper tail, with a weight function smoothly connecting the two parts. The Frigessi model may be defined as

$$c\Big[(1 - p_{\mu,\tau}(x))\; g_\gamma(x) + p_{\mu,\tau}(x)\; h_\xi(x/\sigma)/\sigma\Big],$$

where $x > 0$, c is a normalizing constant, g_γ corresponds to a light-tailed density with parameters γ, and the function h_ξ represents a heavy-tailed GP density with shape parameter $\xi > 0$. One of the most interesting aspect of this density mixture is the weight function $p_{\mu,\tau}(\cdot)$ defined by

$$p_{\mu,\tau}(x) = \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{x-\mu}{\tau}\right).$$

Since this weight function is non-decreasing, takes values in $(0, 1]$ and tends to unity as $x \to \infty$, *Frigessi et al.* [2002] argued that it can play the role of an unsupervised threshold selection algorithm. While *Frigessi et al.* [2002] chose to parametrize the light density g_γ as a Weibull density in their fire loss application, g_γ was a Gamma density in the precipitation data studied by *Vrac and Naveau* [2007]. Overall, Frigessi's model has conceptually a lot of advantages; in particular, it removes the delicate choice of a pre-determined threshold. In practice, there are important drawbacks. Frigessi's model has a lot of parameters (six) in the simple i.i.d. context and inference is not straightforward. In their simulation study under the true model, *Frigessi et al.* [2002] wrote on page 227 that "for all parameters, the estimates are rather spread. Especially, τ and the Weibull parameters are difficult to estimate, and the estimators are clearly dependent". In addition, their tables 2 and 3 clearly showed that the GP shape parameter ξ was strongly underestimated (the true value of ξ was above the 75% quantile estimate). Concerning the weight function $p_{\mu,\tau}(x)$, *Vrac and Naveau* [2007] observed those difficulties with rainfall data: the estimates of τ were poor and very close to zero, meaning that the weight function for their rainfall data converged to a step function with a jump at μ. Hence, a discontinuity around the value μ was re-introduced, which is an undesirable feature. Another drawback resides in the constraint of having a strictly positive GP shape parameter, owing to identifiability problems when $\xi = 0$. In hydrology, the GP shape parameter can tend to zero when the time scale increases, e.g., say from hourly to weekly rainfall amounts.

The research developed below attempts to address the same issues treated by authors like *Frigessi et al.* [2002], *Carreau and Bengio* [2006], *Tancredi et al.* [2011] and *Li et al.* [2012], but we would like to avoid the use of mixtures that can quickly inflate the number

of parameters [see *MacDonald et al.*, 2011, for more details on non-parametric approaches]. A popular road in statistics to increase the flexibility of a given density, here the GP, is to simply multiply it by a simple nonnegative function and renormalizing the product to make it a valid pdf. This simple idea is the cornerstone of the so-called *skewed distributions* research field [e.g., *Genton*, 2004]. The main difficulty resides in choosing a multiplicative function that has be to simple enough to keep computational issues at bay, and complex enough to bring a real added value in terms of flexibility. For example, one can carefully choose an appropriate multiplicative function by taking advantage of second order rates of convergence for a variety of tails [see *Falk et al.*, 2010], see Section 2. It is also important to emphasize that this skew-based approach is not the unique way to construct GP distribution extensions.

Recently, *Papastathopoulos and Tawn* [2013] proposed a very interesting and general alternative to build a variety of GP distribution extensions. As any continuous random variable can be generated by applying its inverse cdf to uniform draws (see also the CDF-t transform used in hydrology, e.g., *Kallache et al.*, 2011), one can generate GP-like random variables by replacing the uniform draws by something richer like Beta distributed draws. Again, the main difficulty is to find the right balance between computational simplicity, added flexibility and desirable features, such as retaining the upper tail behaviour. *Papastathopoulos and Tawn* [2013] proposed mainly three types of extensions. Besides establishing the link with skewed distributions, one major difference of the present paper with respect to *Papastathopoulos and Tawn* [2013] is that we take advantage of EVT to model also low rainfall intensities. Although low precipitation amounts are bounded by zero, the lower tail should, in principle, also comply with EVT.

To finish this brief overview about extended GP distributions, we would also like to mention the work of *Beirlant et al.* [2009] who introduced and studied another type of extended GP distribution. As in *Papastathopoulos and Tawn* [2013], their ultimate goal was to improve the estimation of the upper tail shape parameter, and they carefully studied how to select a suitable threshold. Our aim is different in the sense that we want to model adequately the full range of rainfall, and not just to improve inference for high quantiles. For example, most crop computer models require simulating low, moderate and large precipitation to explore their sensitivity. In summary, our main goal here is to offer a practical model and fast estimation procedures to describe the full precipitation range while bypassing a threshold selection, and not only to improve, per say, the estimation of high quantiles like in *Papastathopoulos and Tawn* [2013] and *Beirlant et al.* [2009]. In particular, low rainfall will also be modelled using the EVT paradigm.

The paper is organised as follows. After recalling a few basic concepts used in EVT, Section 2 presents several types of extended GP models and describes a simple sampling scheme. Section 3 discusses inference procedures based on probability weighted moments and maximum likelihood, the performance of which is assessed by an extensive simulation study in Section 4. Section 5 discusses an application to hourly precipitation in Lyon, France. Finally, we summarize our results and discuss some future research directions in Section 6.

2. A rainfall intensity model

2.1. Heavy rainfall modelling

According to basic univariate EVT [e.g. *Coles*, 2001; *Embrechts et al.*, 1997], the probability that large rainfall amount exceeding a well-chosen high threshold u are larger than

x can be approximated by a Generalized Pareto (GP) tail defined as

$$\overline{H}_\xi \left(\frac{x-u}{\sigma} \right),$$

where the survival function $\overline{H}_\xi$ corresponds to

$$\overline{H}_\xi(x) = \begin{cases} (1+\xi x)_+^{-1/\xi}, & \text{if } \xi \neq 0, \\ \exp(-x), & \text{if } \xi = 0, \end{cases} \quad (1)$$

with $a_+ = \max(a, 0)$. The scalar $\sigma > 0$ represents the scale parameter. The shape parameter ξ describes the GP tail behavior. If ξ is negative, the upper tail is bounded. If ξ is zero, this corresponds to the case of an exponential distribution, where all moments are finite. If ξ is positive, the upper tail is unbounded but higher moments eventually become infinite. These three cases are termed "short-tailed", "light-tailed", and "heavy-tailed", respectively. The flexibility of the GP distribution to describe three different types of tail behavior makes it a universal tool for modelling excesses. In our case, we assume that heavy rainfall data have either exponential tails ($\xi = 0$) or heavy tails ($\xi > 0$). This condition appears to be satisfied for most heavy rainfall data [e.g., *Katz et al.*, 2002].

In practice, one difficulty in fitting the GP distribution to rainfall measurements resides in the choice of an appropriate threshold u. Despite a lot of research on this topic, finding a simple, fast and efficient threshold selection scheme remains an elusive task in the realm of hydrological applications [see *Dupuis*, 1999, for details].

Although rarely used in hydrology and climatology, a few approaches of GP distribution extensions have been studied in theoretical statistics. For example, the third edition of the book of *Falk et al.* [2010] describes different extensions of the GP density. For non-negative GP shape parameters, these authors studied the theoretical properties of densities (called

$Q_1(\delta)$ and $Q_3(\delta)$ in their book, see their Proposition 2.2.1) of the form

$$\text{cst} \times \frac{1}{\sigma} h_\xi\{(x-b)/\sigma\} \left[1 + \mathcal{O}\left\{\overline{H}_\xi^\delta(x/\sigma)\right\}\right], \quad (2)$$

where $\delta > 0$ and the notation $\mathcal{O}(v)$ means that the ratio $\mathcal{O}(v)/v$ is bounded as v converges to zero. This class encompasses a broad family of densities. The main idea of Equation (2), basically multiplying a density like $h_\xi(.)$ by another function, has been extensively studied in the so-called *skewed distributions* research field [e.g., *Genton*, 2004]. The archetypal example is the skew normal pdf [e.g., *Azzalini*, 1985] defined by the product

$$2\phi(x)\ \Phi(\lambda x)$$

of $\phi(x)$, the standard Gaussian pdf, with its cdf $\Phi(x)$. The parameter λ regulates the skewness, with $\lambda = 0$ yielding the normal pdf. Besides the Gaussian building block, one can extend other densities of interest. For example, *Ribereau et al.* [2014] proposed an extension of the Generalized Extreme-Value distribution that improves the fit of maxima over fixed or random block sizes.

2.2. Low rainfall modelling

Before explaining our approach to model the full precipitation range (zeros excluded), we need to address the often overlooked question of how to model low rainfall intensities. EVT can also be applied to the lower tail if we flip the sign of the rainfall amount X, i.e., defining $Y = -X$. The largest negative rainfall in a sample of observations Y may be fitted using a GP distribution with a negative shape parameter, say $-1/\kappa$ with $\kappa \geq 0$, and a positive scale parameter, say $\nu > 0$. Mathematically, this means that the upper tail of Y should tend to a GP distribution, i.e.,

$$\mathbb{P}(Y > -x | Y > -u) \approx \overline{H}_{-1/\kappa}\left(\frac{-x+u}{\nu}\right),$$

as u approaches zero for any x such that $0 < x < u$. Obviously, the upper limit of Y is zero, which explains why the shape parameter has to be negative in order to have a short-tailed GP. Furthermore, it also implies that the threshold u has to be chosen such that $\overline{H}_{-1/\kappa}(0) = 0$, leading to the constraint $u = \kappa\nu$. Consequently,

$$\mathbb{P}(Y > -x | Y > -u) \approx \text{cst} \times x^{\kappa}.$$

In other words, this suggests that low rainfall intensities might be adequately described by a power law

$$\mathbb{P}(X \leq x) \approx \text{cst} \times x^{\kappa}, \text{ for any small } x \geq 0.$$

Notice that this condition is satisfied by a gamma density $f(x) \propto x^{\kappa-1}e^{-x/\theta}$, $x \geq 0$, $\kappa, \theta > 0$.

2.3. Full range modelling

According to the two previous sections, we would like, in order to be in compliance with EVT on both sides of the rainfall spectrum, to have

$$\mathbb{P}(X \leq x) = \begin{cases} 1 - \text{cst} \times \overline{H}_{\xi}\left(\frac{x}{\sigma}\right), & \text{for any large } x, \\ \text{cst} \times x^{\kappa}, & \text{for any small } x \text{ near } 0. \end{cases} \quad (3)$$

The widely-used gamma density is in agreement with (3) for low values, but fails at representing correctly high values, while the contrary is true for the GP distribution. The exponentially decaying tail of the gamma density typically leads to a drastic underestimation of probabilities of extreme events. We therefore aim at creating a gamma-like density that ressembles a GP density on both tails, while bypassing the threshold selection problem that brings two unwelcome discontinuities between low and moderate, and moderate and heavy rainfall. One strategy could be to create a mixture based on $\overline{H}_{\xi}(x)$ and x^{κ}. This approach is certainly valuable but we do not pursue it here for the following reasons.

First, the function x^κ needs to be defined on the compact support $[0,1]$ to be a valid cdf, and the end point of this interval would create a discontinuity in the mixture. Second, the inference might be complex (using an EM algorithm). Third, the computation of return levels (high quantiles) is not explicit. Alternatively, we want to propose a single pdf with the appropriate lower and upper tails and no hidden states. To achieve this goal, we shall follow the footsteps of *Papastathopoulos and Tawn* [2013]. Although these authors focused on the upper tail only, their approach can be somehow adapted for modelling the full rainfall range.

Our main starting point is the classical scheme used to simulate GP distributed random draws via the formula

$$\sigma\, H_\xi^{-1}(U),$$

where U represents a random variable uniformly distributed on $[0,1]$ and H_ξ^{-1} corresponds to the inverse GP cdf. A simple way to add flexibility to this simulation scheme is to define

$$X = \sigma\, H_\xi^{-1}\{G^{-1}(U)\}, \tag{4}$$

where G is a continuous cdf on $[0,1]$. The main question is to find a class of distributions G such that the upper tail behaviour with shape parameter ξ is preserved and the cdf of X for value near zero behaves like x^κ. If we denote the tail of G by $\overline{G} = 1 - G$, these constraints can be fulfilled whenever

(A)

$$\lim_{u \downarrow 0} \frac{\overline{G}(1-u)}{u} = a, \text{ for some finite } a > 0,$$

(B)

$$\lim_{u \downarrow 0} \frac{G\{u\, v(u)\}}{G(u)} = b, \text{ for some finite } b > 0,$$

where $v(u)$ is any positive function such $v(u) = 1 + o(u)$ as $u \to 0$,

(C)

$$\lim_{u \downarrow 0} \frac{u^{\kappa}}{G(u)} = c, \text{ for some finite } c > 0.$$

To understand these three conditions, one has to notice that Equation (4) may be expressed in terms of the cdf $F(x) = \mathbb{P}(X \leq x)$ or the tail $\overline{F}(x) = \mathbb{P}(X > x)$ as

$$F(x) = G\left\{H_{\xi}\left(\frac{x}{\sigma}\right)\right\} \text{ and } \overline{F}(x) = \overline{G}\left\{H_{\xi}\left(\frac{x}{\sigma}\right)\right\}. \tag{5}$$

Constraint (A) implies that the upper tail of X is equivalent to a Pareto tail in the sense that the ratio $\overline{F}(x)/\overline{H}_{\xi}(x/\sigma) = \overline{G}(1-u)/u$ with $u = \overline{H}_{\xi}(x/\sigma)$ converges to a constant, as $x \to \infty$. Similarly, the ratio $F(u)/G(u) = G\{u\ v(u)\}/G(u)$ with $v(u) = H_{\xi}(u/\sigma)/u$ converges also to a constant, as $u \to 0$, because of the constraint (B) and the Taylor expansion $\{1-(1+\xi u)^{-1/\xi}\}/u = 1+o(u)$. In other terms, the constraint (B) ensures that low values are driven by G and constraint (C) forces G to behave as a GPD of Weibull type for this lower tail.

Before presenting parametric examples of the function G, we would like to emphasise two important practical advantages of (4). This equation provides a straightforward and fast way of simulate random draws from the cdf F whenever the G^{-1} is available. The same is true when return levels have to be computed using the explicit formula

$$x_p = F^{-1}(p) = \begin{cases} \frac{\sigma}{\xi}\left[\{1 - G^{-1}(p)\}^{-\xi} - 1\right], & \text{if } \xi > 0, \\ -\frac{\sigma}{\delta}\log\{1 - G^{-1}(p)\}, & \text{if } \xi = 0, \end{cases} \tag{6}$$

$0 < p < 1$. Practical reasons should drive our choice concerning particular parametric forms of the cdf G.

2.4. Parametric families

The simplest choice for the cdf $G(u)$ is the power law distribution u^κ, $\kappa > 0$, defined on the unit interval. The three constraints (A), (B) and (C), are satisfied and this family corresponds to the type III introduced by *Papastathopoulos and Tawn* [2013]. Figure 1 displays the corresponding density for different lower tail behaviours ($\kappa = 1, 2, 5$) and fixed upper tail decay ($\xi = 0.5$), compared with a gamma density. The GP model (putting mass $1/\sigma$ at zero) is recovered when $\kappa = 1$, and more flexibility for low values is achieved by varying this parameter. As expected, the gamma and extended GP densities behave similarly for small and moderate values, but the discrepancy increases further in the tail. A way to increase the flexibility of this fairly simple model is to consider a mixture of power laws, i.e., $G(u) = pu^{\kappa_1} + (1-p)u^{\kappa_2}$, with $p \in [0,1]$ and $\kappa_1, \kappa_2 > 0$. This model again satisfies the three conditions (A), (B) and (C), with the latter holding by setting $\kappa = \min(\kappa_1, \kappa_2)$. This means that the lower tail behaviour is controlled by $\min(\kappa_1, \kappa_2)$, whereas $\max(\kappa_1, \kappa_2)$ modifies the shape of the density in its central part. With this specification, model (5) has five parameters. Figure 2 illustrates the flexibility of this model with $p = 0.5$, $\kappa_1 = 2$ and different values of κ_2, by comparison with a Gamma density.

To describe another nontrivial and interesting choice for $G(u)$, we can make a link with the work of *Falk et al.* [2010], see (2), by setting for some $\delta > 0$

$$G(u) = \mathbb{P}\left(1 - B_\delta^{1/\delta} \leq u\right), \text{ for } u \in [0,1], \tag{7}$$

where B_δ corresponds to a Beta random variable with cdf

$$V_\delta(u) = \mathbb{P}(B_\delta \leq u) = (1 + 1/\delta)\, u^{1/\delta} \left(1 - \frac{u}{1+\delta}\right),$$

and pdf

$$\frac{(1+1/\delta)}{\delta}\, u^{1/\delta-1}(1-u). \tag{8}$$

This fairly complex choice of $G(u)$ corresponds to the very simple case where $\mathcal{O}(x) = -x$ and $b = 0$ in (2), i.e.,

$$f(x;\xi,\sigma,\delta) = (1+1/\delta)\,\frac{1}{\sigma}h_\xi(x/\sigma)\,\left\{1-\overline{H}_\xi^\delta(x/\sigma)\right\}, \tag{9}$$

which is illustrated in Figure 3. By noticing that $G(u) = \overline{V}_\delta\left\{(1-u)^\delta\right\}$, where $\overline{V}_\delta = 1-V_\delta$, it is possible to check that the constraints (A)–(C) are satisfied for $\kappa = 2$. As δ increases to infinity, $f(x;\xi,\sigma,\delta)$ becomes closer to the GP density. Moreover, for large x, the tail behaviour of both densities is equivalent and, consequently, very heavy rainfall is captured in the same fashion for both densities, i.e., through the shape parameter ξ. This can also be justified by noticing that

$$\overline{F}(x;\xi,\sigma,\delta) = (1+1/\delta)\overline{H}_\xi(x/\sigma)\left[1-\frac{1}{1+\delta}\overline{H}_\xi^\delta(x/\sigma)\right]. \tag{10}$$

This means that, since $\overline{H}_\xi(x)$ converges to zero as x tends to infinity, the asymptotic tail behavior of X is equivalent to that of a GP distribution, i.e., for large x,

$$\begin{aligned}\overline{F}(x;\xi,\sigma,\delta) &\sim (1+1/\delta)\overline{H}_\xi(x/\sigma),\\ &= \overline{H}_\xi\left(\frac{x-u_\delta}{\sigma_\delta}\right),\end{aligned}$$

where $\sigma_\delta = \sigma c_\delta^\xi$ and $u_\delta = \sigma(c_\delta^\xi - 1)/\xi$. In other words, the extra term $\{1-\overline{H}_\xi^\delta(x/\sigma)\}$ in (9) does not affect the extremal index ξ representing the main driver of very extreme events. The parameter δ rather increases modelling flexibility for the central part of the distribution. This parameter could be interpreted as a "threshold tuning parameter" that has to be estimated from the data at hand.

Regarding the behaviour of $f(x;\xi,\sigma,\delta)$ near zero, we can immediately notice from (9) that $f(0;\xi,\sigma,\delta)=0$. However, a drawback of (9) resides in the fact that the Taylor expansion $(1+x)^a \sim 1+ax$ (for small x) implies that the lower tail behavior of $f(x;\xi,\sigma,\delta)$ is

$$f(x;\xi,\sigma,\delta) \sim \frac{1+\delta}{\sigma^2}\, x, \text{ for } x \text{ near zero.}$$

This means that the lower tail $F(x;\xi,\sigma,\delta)$ is of type x^2 (i.e., $\kappa=2$ in condition (C)) and consequently, the lower tail behavior is not estimated from the data but imposed by the choice of model (9).

A solution to this problem might be to add an extra parameter $\kappa>0$ controlling the lower tail behaviour as in constraint (C) as follows:

$$G(u) = \left[\overline{V}_\delta\left\{(1-u)^\delta\right\}\right]^{\kappa/2}.$$

This leads to the cdf

$$F(x) = \left[\overline{V}_\delta\left\{\overline{H}_\xi^\delta\left(\frac{x}{\sigma}\right)\right\}\right]^{\kappa/2}, \tag{11}$$

where κ, δ, ξ describe the low, moderate and upper parts of the distribution, respectively. In particular, the lower and upper tails are, by construction, GP with shape parameters κ and ξ, respectively. Furthermore, our choice to build $G(u)$ from a Beta cdf makes the simulation and the computation of return levels straightforward. Specifically, the quantile function corresponding to (11) may be expressed as

$$x_p = F^{-1}(p) = \begin{cases} \frac{\sigma}{\xi}\left[\{V_\delta^{-1}(1-p^{2/\kappa})\}^{-\xi/\delta}-1\right], & \text{if } \xi>0, \\ -\frac{\sigma}{\delta}\log\{V_\delta^{-1}(1-p^{2/\kappa})\}, & \text{if } \xi=0, \end{cases} \tag{12}$$

where $V_\delta^{-1}(u)$ is the quantile function of the Beta random variable B_δ. From an hydrological point of view, Equation (12) offers a simple way to compute return levels x_p for

any return period $1/p$. To simulate from the model, one simply needs to randomly draw a uniform variable U in $[0, 1]$, and then to apply the quantile function (12) as $X = F^{-1}(U)$.

3. Inference

3.1. The two classical estimation approaches

A variety of inference methods exists and, for heavy rainfall analysis, two options are popular among hydrologists: maximizing the likelihood function (ML) and a method of moments based on Probability Weighted Moments (PWMs). In this paper, we investigate how these two classical inference techniques can be implemented within our framework. Concerning the ML approach, the likelihood based on (5) can be easily obtained whenever the function $G(u)$ is easily differentiable. This is the case for the parametric families introduced in Section 2.4.

The PWMs approach is based on the hypothesis that PWMs can be easily obtained and quickly computed. The following section deals with this aspect.

3.2. Probability Weighted Moments

The PWMs method has a long tradition in statistical hydrology [e.g., *Landwehr et al.*, 1979; *Hosking and Wallis*, 1987]. It has been recently revisited by statisticians [e.g., *Ferreira and de Haan*, 2014] and applied in various settings [e.g., *Naveau et al.*, 2014]. A recent study of *Caeiro and Gomes* [2011] theoretically compared different estimation methods for the shape parameter ξ. Besides its simplicity, the PWMs approach usually performs reasonably well compared to other estimation procedures. Additional arguments in favour of PWMs are that they are typically quickly computed, even in non-stationary contexts *Naveau et al.* [2014].

The idea of the PWMs approach is simple: for a pdf specified by s parameters (e.g., δ, κ, σ and ξ), we need to find the explicit expressions of s (weighted) moments that are function of these parameters. Having s empirical moments and s unknown parameters, we can pursue a method-of-moments by solving these equations [e.g., *Diebolt et al.*, 2008, 2007].

For model (5), it is convenient to work with PWMs of the form

$$\mu_s = \mathbb{E}\left(X\overline{F}^s(X)\right), \quad s = 0, 1, \ldots.$$

From (6), it can be shown that for $\xi \neq 0$,

$$\mu_s = \frac{\sigma}{\xi}\left(\mathbb{E}\left[\{1 - G^{-1}(U)\}^{-\xi}(1-U)^s\right] - \frac{1}{1+s}\right), \tag{13}$$

where U corresponds to a uniformly distributed random variable on $[0, 1]$. In the following subsections, we detail calculations for our extended GP densities introduced above.

The Appendix provides the explicit PWMs for the parametric families defined by $G(u) = u^\kappa$, $G(u) = pu^{\kappa_1} + (1-p)u^{\kappa_2}$ and $G(u) = \overline{V}_\delta\left\{(1-u)^\delta\right\}$, respectively. For these cases, it would be ideal to write down a simple system like in the GP case, see (B1) in the Appendix. But, the unknown parameters cannot always be explicitly expressed in function of PWMs. In practice, this limitation does not cause any particular problem because statistical softwares like the R package **gmm** [*Chaussé*, 2010] provides numerical solutions to such method-of-moments system of equations. Confidence intervals can also be obtained from this R package. The extension to the case $G(u) = [\overline{V}_\delta\left\{(1-u)^\delta\right\}]^{\kappa/2}$ is, however, more complicated and probability weighted moments have to be also computed by Monte Carlo simulations using a large number of replicates. The following section illustrates how the inference performs.

4. Simulation study

In this section, we assess the performance of the PWMs and ML estimators by simulation. All results presented here are based on the following setting. The scale parameter in model (5) is always set to one, $\sigma = 1$. The shape upper tail parameters can take three values $\xi = 0.1, 0.2, 0.3$, classical values for precipitation data. The sample size is fixed to $n = 300$ —a number which more or less corresponds to the data size in our application—, and 10^5 replicates are used to compute basic statistical metrics like root mean squared errors (RMSEs). We have also tried different settings (results available upon request), which provide similar conclusions. In this article, we will explore the four following setups:

(i) $G(u) = u^{\kappa}$, with lower tail parameter $\kappa \in \{1, 2, 5, 10\}$;

(ii) $G(u) = \overline{V}_{\delta}\left\{(1-u)^{\delta}\right\}$, with skewness parameter $\delta \in \{0.5, 1, 2, 5\}$,

(iii) $G(u) = pu^{\kappa_1} + (1-p)u^{\kappa_2}$, with $p = .4$, $\kappa_1 \in \{1, 2, 5, 10\}$ and $\kappa_2 \in \{2, 5, 10, 20\}$,

(iv) $G(u) = [\overline{V}_{\delta}\left\{(1-u)^{\delta}\right\}]^{\kappa/2}$, with $\delta \in \{0.5, 1, 2, 5\}$, and $\kappa \in \{1, 2, 5, 10\}$,

with starting values for the numerical solvers fixed arbitrarily to $\sigma = 2$, $\xi = 0.15$, $\kappa = 3$, $\delta = 1.5$, $p = 0.5$, $\kappa_1 = 3$ and $\kappa_2 = 4$. For model (iii), the estimation of the five parameters σ, ξ, p, κ_1 and κ_2 requires PWMs of orders $s = 0, 1, 2, 3, 4$, while for model (iv), only the first four moments can be used to estimate σ, ξ, κ and δ.

To summarize the performance of the ML and PWM estimators, Figure 4 displays boxplots of estimated parameters and 99%-quantiles for representative cases, see the caption for the exact parameters values (model (iv) gives similar and results available upon request). Skewness parameters like δ or κ_2 are difficult to estimate by ML, as it was noticed by *Sartori* [2006] for the skew normal case and *Ribereau et al.* [2014] for the extended GEV case. Figure 4 also shows generally higher variability of PWMs estimators compared to ML estimators, especially for tail decay parameters (ξ or κ) and high quantiles.

More importantly, both inferential methods provide reasonable estimates for a moderate sample size of $n = 300$, even with a five-parameter model like (iii) (although the ML estimator for $\kappa_2 = 5$ appears to have reached its limit).

To fine tune our comparison between these two classical estimation approaches, Table 1 reports the ratio of RMSEs of PWMs and ML estimators for the four models. The ratio of RMSEs for the estimated 99%-quantile is reported in bold. Values lower than unity indicate that PWMs estimators perform better, and vice versa.

Overall, PWMs and ML estimators have a similar performance, though the latter are generally slightly better. This is especially the case for model (i) with large $\kappa \geq 5$, although high quantiles may still be reasonably well estimated by PWMs (see the case $\kappa = 10$, $\xi = 0.1$). Model (ii) emphasises one more time that PWMs perform well for the skewness parameter δ, but interestingly, high quantiles seem to be better estimated using the ML approach. Similar conclusions apply for 98% and 99.5%-quantiles. The case of model (iii) exacerbates that five parameters imply a much higher instability, especially for $\kappa_2 = 2$ by ML and for $\kappa_1 = 10$ by PWMs. For model (iv), ML estimators slightly but constantly outperform PWMs in terms of RMSE for high quantiles. This is surprising because the skewness parameter δ has much higher variability with the ML approach; figures are available upon request.

One may wonder if imposing a parametric model on the whole dataset "deteriorates" the fit of the largest values that could be obtained by a classical GP approach based on a small fraction of extreme data. To assess this, we simulated data from model (i) with $\sigma = 1$, $\xi = 0.2$, $\kappa = 2$ (based on the same setting as before), and estimated the upper tail parameter ξ and the 0.99-quantile, either by fitting the true model to the whole dataset

by ML or by fitting the GP distribution to excesses above the 95%-quantile. In a nutshell, the full range modelling approach improves the estimate of ξ, the ratio of RMSEs being equals 3.22, while this ratio is only 1.12 for the 0.99-quantile, indicating that the gain is weaker for high (though not extremely high) quantiles. This simple experiment suggests that the bulk of the distribution may help to estimate upper (respectively lower) tail parameters or high (respectively low) quantiles, provided that the assumed model is, at best, the correct one or, at least, flexible enough to adapt itself to the data at hand.

5. Hourly rainfall in Lyon (France)

As an illustrative example, we analyze hourly precipitation from 1996 to 2011 recorded at the French weather station of Lyon. The different extended GP models described in Section 2 were separately fitted to rainfall intensities for each season using ML and PWMs, as outlined in Section 3. To reduce temporal short-term dependence, every third observation was retained for the analysis of each time series. After removing the dry events (i.e., zero values) and thresholding the observations at 1mm, the sample sizes are equal to 282 (Spring), 251 (Summer), 336 (Fall) and 184 (Winter); see the histograms in the left panels of Figure 5. According the the Akaike information criterion (AIC) and simple graphical diagnostics, model (i) with $G(u) = u^{\kappa}$ in (4) seems to perform quite well overall; the fits of models (ii) with $G(u) = \overline{V}_{\delta}\left\{(1-u)^{\delta}\right\}$ and (iv) with $G(u) = [\overline{V}_{\delta}\left\{(1-u)^{\delta}\right\}]^{\kappa/2}$ are comparable, although the skewness parameter δ is generally badly estimated, while model (iii) with $G(u) = pu^{\kappa_1} + (1-p)u^{\kappa_2}$ is over-parametrized and consistently appears to be the worst. For these reasons, Figure 5 reports the results of model (i) only. The other results are available upon request.

In the simulation study in Section 4, the ML approach was found to have generally lower RMSEs than PWMs. However, these results were based on the assumption that the models fitted were well-specified. When dealing with real data, however, this strong assumption is no longer valid, and we have found that PWMs are (much) more robust against model misspecification since they are based on useful summary statistics, rather than the exact values of observations. This is especially important when a model is fitted to the full range of hourly rainfall data, because the discretization (here, to the closest 0.1mm) strongly affects low values, and consequently the estimation of parameters if this feature is not properly taken into account. In particular, the upper tail shape parameter ξ is constantly over-estimated using the classical ML approach, and the effect of the discretization increases as the distribution becomes more concentrated around zero. To counteract this undesirable effect, one might consider two possibilities: One may view the data as being left-censored (here, no data were directly observed below 0.1mm), or one may assume that data observed in the interval $[x, x + 0.1)$ have been rounded to x and modify the likelihood function accordingly. Both approaches significantly improve the inference compared to the naive ML approach (especially for the Spring, Summer and Fall data), and correct the estimation of the shape parameter; the left-censored likelihood was found to be best censoring. Table 2 reports the estimated parameters obtained by maximizing the censored likelihood, with 95%-confidence intervals based on 500 non-parametric bootstrap replicates. Figure 5 displays the corresponding fitted densities, as well as quantile-quantile plots representing the quality of the fit.

Overall there is a good agreement between the censored likelihood and probability weighted moments approaches, except for the Winter season where the former perform

very badly compared to the latter, strongly overestimating the lower and upper tail shape parameters and large quantiles.

This study illustrates the convenience and robustness of PWMs, when it comes to fitting a distribution to the full range of the data in a misspecified setting.

6. Conclusions and perspectives

In this work, we attempt to show that it may be possible to model jointly low, moderate and heavy rainfall, without fixing a threshold. Such a strategy coupled with two classical inference approaches has many practical advantages. It is fast to implement, simple to interpret with only a few parameters and in compliance with EVT for both the upper and lower tails. One drawback could be that we impose a specific form for moderate precipitation. In cases for which this limitation is too stringent, it is always possible to couple our method with a weather-type approach, i.e., by assuming that daily rainfall has to belong to a weather type driven by some specific atmospheric pattern [see, e.g. *Ailliot et al.*, 2015]. Hence, our distributions could be fitted for each weather type and consequently, a mixture of pdfs will represent the whole rainfall spectrum. One can even imagine to impose the weather-type mixture on the function G; see our model (iii) in Section 4.

Concerning the inference, PWMs and ML approaches perform adequately on simulated data, with a slight advantage of the MLE for high quantiles inference. PWMs seem to be better for skewness. With observed rainfall recordings, the PWMs method appears to be more straightforward and robust, the MLE having to be fine-tuned to handle the censoring at 0.1 mm due to precision errors. In any case, both approaches are fast and we advice to use both and compare estimates. It would be of interest to implement a

Bayesian approach. Putting an informative prior on δ and ξ could certainly improve the analysis for these two parameters that are never simple to infer.

Our object of study was rainfall and consequently, we limit our GP distribution shape parameter to be non-negative. For other applications, it would be interesting to relax this hypothesis. Going back to the topic of precipitation, obvious extensions should be explored. In this work, we did not address the important issue of modelling rainfall occurrences (i.e., wet or dry events). This leads to the question on how to couple Bernoulli type events with continuous intensities [see, e.g. *Koch and Naveau*, 2015]. In the same vein, the modelling of rainfall amounts at multi-sites remains a statistical challenge and it would of interest to develop a multivariate version of our proposed distributions.

Acknowledgments. Part of this work has been supported by the ANR-DADA, LEFE-INSU-Multirisk, CHAVANA and Extremoscope projects. The authors acknowledge Meteo France for the Lyon precipitation time series that available to anyone upon request. The authors would also like very much to credit the contributors of the *R Core Team* [2013].

References

Ailliot, P., D. Allard, V. Monbet and P. Naveau, (2015). Stochastic weather generators: an overview of weather type models. *Journal de la Société Française de Statistiques*, 156: 101–113.

Azzalini, A. A.,(1985) A class of distributions which include the normal, *Scandinavian Journal of Statistix*, 12:171 178..

Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications*, Wiley Series in Probability and Statistics.

Beirlant J., E. Joossens, J. Segers,(2009) Second-order refined peaks-over-threshold modelling for heavy-tailed distributions, *Journal of Statistical Planning and Inference*, 2800 – 2815.

Caeiro F. and M. I. Gomes,(2011) Semi-parametric tail inference through probability-weighted moments, *Journal of Statistical Planning and Inference*, 937950.

Carreau, J., and Bengio, Y. (2006). A Hybrid Pareto Model for Asymmetric Fat-tailed Data: the Univariate Case, *Extremes*.

Carreau J., P. Naveau and E. Sauquet (2009). A Statistical Rainfall-Runoff Mixture Model with Heavy-Tailed Components, *Water Resources Research*.

Carreau J. and M. Vrac, (2011). Stochastic Downscaling of Precipitation with Neural Network Conditional Mixture Models, Water Resources Research, Vol. 47.

Chaussé, P. (2010). Computing Generalized Method of Moments and Generalized Empirical Likelihood with R. *Journal of Statistical Software*, **34-11**, 1-35.

Coles, S.G. (2001) *An introduction to statistical modelling of extreme values*. Springer Series in Statistics.

Cooley D., Nychka, D., and Naveau, P. (2007). Bayesian Spatial Modeling of Extreme Precipitation Return Levels, *Journal of the American Statistical Association*, 102(479):824840.

De Haan, L., and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer Series in Operations Research and Financial Engineering.

Dupuis, D.J.(1999), Exceedances over high thresholds: A guide to threshold selection," *Extremes*, 1(3), 251–261.

Diebolt, J., A. Guillou, P. Naveau, and P. Ribereau (2008), Improving probability-weighted moment methods for the generalized extreme value distribution, *REVSTAT - Statistical Journal*, *6*, 33–50.

Diebolt, J., A. Guillou, and I. Rached (2007), Approximation of the distribution of excesses through a generalized probability-weighted moments method, *J. Stat. Plan. Infer.*, *137*, 841–857.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*, volume 33 of *Applications of Mathematics*, Springer-Verlag, Berlin.

Falk M., Hüsler J., Reiss R.-D. (2010). *Laws of Small Numbers: Extremes and Rare Events.*. Birkhäuser Verlag, Basel 2010, 3rd Ed.

Fisher, R.A., and Tippett, L.H.C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, 24:180–190.

Ferreira Ana and Laurens de Haan (2014). On the block maxima method in extreme value theory. *http://arxiv.org/abs/1310.3222*.

Frigessi, A., Haug, O., and Rue, H. (2002). A dynamic mixture model for unsupervised tail estimation without threshold selection. *Extremes*, 5:219–235.

Furrer R. and P. Naveau (2007). Probability weighted moments properties for small samples. *Statistics & Probability Letters*, 77 (2007) 190195.

Genton, M. G. (2004), *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*. Edited Volume, Chapman & Hall / CRC, Boca Raton, FL, 416 pp..

514 Greenwood, J. A., J. M. Landwehr, N. C. Matalas, and J. R. Wallis (1979), Probability-
515 weighted moments: definition and relation to parameters of several distributions ex-
516 pressable in inverse form, *Water Resour. Res.*, *15*, 1049–1054.

517 Hosking, J. R. M., and J. R. Wallis, (1987), Parameter and quantile estimation for the
518 generalized Pareto distribution, *Technometrics*, *29*, 339–349.

519 Kallache, M., M. Vrac, P. Naveau, and P.A. Michelangeli (2011). Nonstation-
520 ary probabilistic downscaling of extreme precipitation, *J. Geophys. Res.*, **116**,
521 doi:10.1029/2010JD014892.

522 Katz, R., Parlange, M. and Naveau, P. (2002). Extremes in hydrology, *Advances in Water*
523 *Resources*, **25**, 1287–1304.

524 Katz, R.W. (1977). Precipitation as a chain-dependent process. *Journal of Applied Me-*
525 *teorology*, 16:671–676.

526 Koch, E. and Naveau, P. (2015). A frailty-contagion model for multi-site hourly precipi-
527 tation driven by atmospheric covariates. *Advances in Water Resources*, 78, 145–154.

528 Landwehr, J., N. Matalas, and J. R. Wallis (1979), Probability weighted moments com-
529 pared with some traditional techniques in estimating Gumbel parameters and quantiles,
530 *Water Resour. Res.*, *15*, 1055–1064.

531 Li, C., V. P. Singh, and A. K. Mishra (2012). Simulation of the entire range of daily
532 precipitation using a hybrid probability distribution. *Water Resources Research*, 48,
533 W03521, doi:10.1029/2011WR011446.

534 MacDonald, A., Scarrott, C., Lee, D., Darlow, B., Reale, M., Russell, G., (2011). A
535 flexible extreme value mixture model. *Computational Statistics and Data Analysis*, 55,
536 2137?2157.

Naveau P., A. Toreti, I. Smith and E. Xoplaki, (2014). A fast nonparametric spatio-temporal regression scheme for generalized Pareto distributed heavy precipitation. *Water Resources Research*, 50, doi:10.1002/2014WR015431.

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

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. `http://www.R-project.org/`.

Ribereau P., E. Masiello and P. Naveau (2014). Skew Generalized Extreme Value Distribution: Probability Weighted Moments Estimation and Application to Block Maxima Procedure. *Communications in Statistics Theory and Methods*, in press.

Sartori, N. (2006). Bias prevention of maximum likelihood estimates for scalar skew normal and skew t distributions. *Journal of Statistical Planning and Inference*, 136, 4259-4275.

Tancredi, A., Anderson, C., O'Hagan, A., (2006). Accounting for threshold uncertainty in extreme value estimation. *Extremes*, 9, 87-106.

Vrac, M., and Naveau, P. (2007). Stochastic downscaling of precipitation: From dry events to heavy rainfalls. *Water Resources Research*, 43, W07402, doi:10.1029/2006WR005308.

Vrac, M., Stein, M., and Hayhoe, K. (2007). Statistical downscaling of precipitation through a nonhomogeneous stochastic weather typing approach. *Climate Research*, 34: 169-184, doi: 10.3354/cr00696.

Wilks, D. (2006). *Statistical methods in the atmospheric sciences* (second edition). Elsevier, Oxford.

Appendix A: Appendix

A1. *Properties of the density defined by (9)*

To see that the random variable $\overline{H}_\xi^{-1}\left(B_\delta^{1/\delta}\right)$ follows (9), we can write

$$P\left[\overline{H}_\xi^{-1}\left(B_\delta^{1/\delta}\right) > x\right] = P\left[B_\delta < u(x)\right],$$

with $u(x) = \left(\overline{H}_\xi(x/\sigma)\right)^\delta$. Taking the derivative with respect to x of the l.h.s gives the density of interest (up to a negative factor). For the r.h.s., replacing u in (8) by $u(x)$ gives us

$$u'(x)\frac{1}{1+\delta}\, u(x)^{1/\delta-1}(1-u(x)) = -\frac{\delta}{1+\delta}\, h_\xi(x)(1-u(x))$$

because $u'(x) = -\delta u(x)^{1-1/\delta} h_\xi(x)$. The required result follows.

Besides making fast and simple simulations, the link between our extended GP and the beta density allows to compute moments quantities. The Beta random variable distributed as in (8) has the following moments

$$\mathbb{E}B_\delta^t = \frac{1+\delta}{(1+\delta\, t+\delta)(1+\delta\, t)}.$$

Suppose that X follows the density defined by (9) with $\xi < 1$. As we have

$$\overline{F}_\theta(x) = W_\delta(\overline{H}_\xi^\delta(x/\sigma)),$$

we can write from (4) that in distribution $X\overline{F}_\theta^s(X) = XW_\delta^s(B_\delta)$, and then

$$X\overline{F}_\theta^s(X) = \begin{cases} \frac{\sigma}{\xi}(B_\delta^{-\xi/\delta}-1)W_\delta^s(B_\delta), & \text{if } \xi > 0, \\ -\frac{\sigma}{\delta}(\log B_\delta)W_\delta^s(B_\delta) & \text{if } \xi = 0. \end{cases}$$

It follows $\mu_s = -\frac{1}{\delta}\mathbb{E}\left[(\log B_\delta)W_\delta^s(B_\delta)\right]$ for $\xi = 0$, and otherwise

$$\mu_s = \frac{\sigma}{\xi}\left(\mathbb{E}[B_\delta^{-\xi/\delta}W_\delta^s(B_\delta)] - \frac{1}{1+s}\right).$$

This expression can be compared to

$$\mu_s^* = \frac{\sigma}{\xi}\left[\frac{1}{1-\xi} - \frac{1}{1+s}\right].$$

In particular, we deduce for $s = 0$ that

$$\mu_0 = \frac{\sigma}{\xi}\left(\frac{1+\delta}{(1-\xi+\delta)(1-\xi)} - 1\right)$$

For $s = 1$, the definition of W_δ implies that $\mathbb{E}B_\delta^{-\xi/\delta} W_\delta(B_\delta)$ is equal to

$$\begin{aligned} & (1+1/\delta)\left(\mathbb{E}B_\delta^{-\xi/\delta+1/\delta} - \frac{1}{1+\delta}\mathbb{E}B_\delta^{-\xi/\delta+1/\delta+1}\right), \\ = \; & \frac{(1+\delta)(4-\xi+2\delta)}{(2-\xi)(2-\xi+\delta)(2-\xi+2\delta)}. \end{aligned}$$

Appendix B: *PWMs of the different parametric families defined in §2.4*

For the special case of the GP model with $G(u) = u$, the PWMs are known explicitly for $0 \le \xi < 1$ and we denote them by

$$\mu_s^* = \frac{\sigma}{(1+s)(1+s-\xi)}, \quad s = 0, 1, \ldots.$$

In particular, the two GP parameters may be expressed as

$$\xi = \frac{\mu_0^* - 4\mu_1^*}{\mu_0^* - 2\mu_1^*} \text{ and } \sigma = \mu_0^*\,(1-\xi). \tag{B1}$$

When $G(u) = u^\kappa$, Equation (13) may be rewritten for any non-negative integer s as

$$\begin{aligned} \mu_s &= \frac{\sigma}{\xi}\left[\sum_{j=0}^{s}\binom{s}{j}(-1)^j \mathbb{E}\left\{(1-U^{1/\kappa})^{-\xi}U^j\right\} - \frac{1}{1+s}\right]. \\ &= \frac{\sigma}{\xi}\left[\kappa\sum_{j=0}^{s}\binom{s}{j}(-1)^j\, B\{(j+1)\kappa, 1-\xi\} - \frac{1}{1+s}\right], \end{aligned}$$

where $B(\cdot,\cdot)$ represents the classical Beta function. It follows that

$$\begin{aligned} \mu_0 &= \frac{\sigma}{\xi}\left\{\kappa B(\kappa, 1-\xi) - 1\right\}, \\ \mu_1 &= \frac{\sigma}{\xi}\left[\kappa\left\{B(\kappa, 1-\xi) - B(2\kappa, 1-\xi)\right\} - \frac{1}{2}\right], \\ \mu_2 &= \frac{\sigma}{\xi}\left[\kappa\left\{B(\kappa, 1-\xi) - 2B(2\kappa, 1-\xi) + B(3\kappa, 1-\xi)\right\} - \frac{1}{3}\right]. \end{aligned}$$

Similarly, for $G(u) = pu^{\kappa_1} + (1-p)u^{\kappa_2}$, Equation (13) may be rewritten for any non-negative integer s as

$$\mu_s = \frac{\sigma}{\xi}\left[\sum_{j=0}^{s}\sum_{k=0}^{j}\binom{s}{j}\binom{j}{k}(-1)^j p^k (1-p)^{j-k} A_{j,k} - \frac{1}{1+s}\right],$$

where $A_{j,k} = p\kappa_1 B\{\kappa_1(k+1) + \kappa_2(j-k), 1-\xi\} + (1-p)\kappa_2 B\{\kappa_1 k + \kappa_2(j-k+1), 1-\xi\}$ and $B(\cdot,\cdot)$ is the Beta function.

PWMs can also be explicitly computed for the pdf defined by (9). For $s = 0, 1$ and 2, one has

$$\begin{aligned}
\mu_0 &= \mu_0^* \frac{2+\delta}{1+\delta} \times \frac{1-\xi/(2+\delta)}{1-\xi/(1+\delta)}, \\
\mu_1 &= \mu_1^* \frac{2}{\xi}\left[\left\{1 - \frac{\xi}{2(2+\delta)}\right\} \times \right. \\
&\quad \left. \left\{1 - \frac{\xi}{(2+\delta)}\right\}^{-1} \times \left\{1 - \frac{\xi}{2(1+\delta)}\right\}^{-1} - \frac{2-\xi}{2}\right], \\
\mu_2 &= \frac{\sigma}{\xi}\left\{\frac{(1+\delta)^3}{(3+\delta-\xi)(3-\xi)\delta^2} - \frac{2(1+\delta)^2}{\delta^2(3-\xi+2\delta)(3-\xi+\delta)}\right. \\
&\quad \left. + \frac{1+\delta}{\delta^2(3-\xi+3\delta)(3-\xi+2\delta)} - \frac{1}{3}\right\},
\end{aligned}$$

where μ_0^* and μ_1^* are defined in (B1).

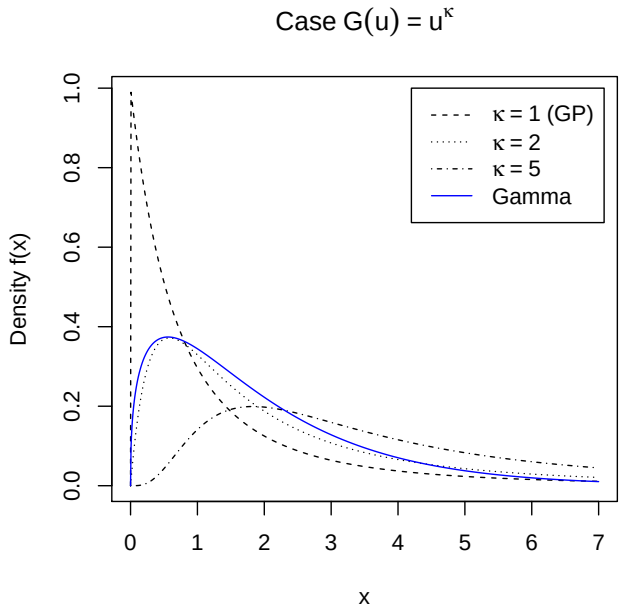

Figure 1. Density function corresponding to model (5) with $G(u) = u^{\kappa}$, for $\sigma = 1$, $\xi = 0.5$ and lower tail shape parameters $\kappa = 1, 2, 5$ (dashed, dotted, dashed-dotted black curves, respectively). The case $\kappa = 1$ corresponds to the exact GP density. The solid blue curve represents a gamma density with parameters $(1.4, 1.4)$.

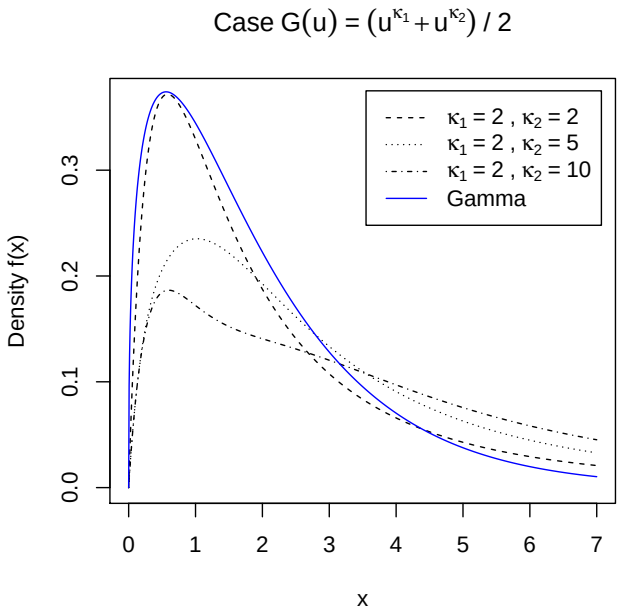

Figure 2. Density function corresponding to model (5) combined with $G(u) = (u^{\kappa_1} + u^{\kappa_2})/2$, for $\sigma = 1$, $\xi = 0.5$ and parameters $\kappa_1 = 2$, and $\kappa_2 = 2, 5, 10$ (dashed-dotted, dotted, dashed black curves, respectively). The solid blue curve represents a gamma density with parameters $(1.4, 1.4)$.

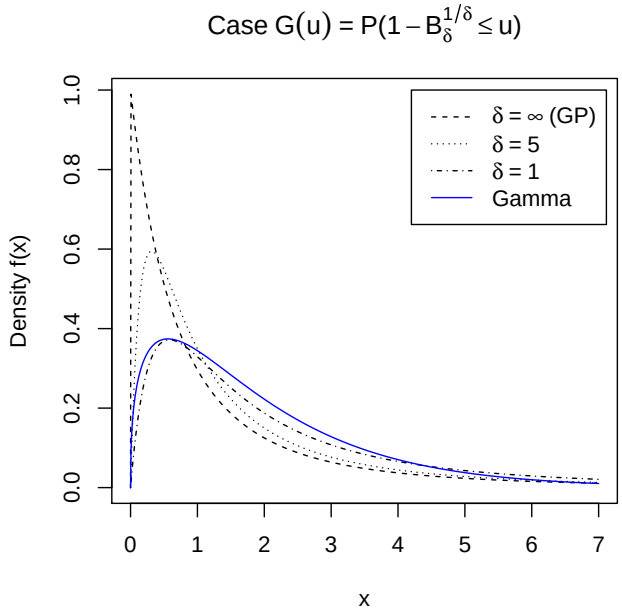

Figure 3. Density function corresponding to model (5) combined with (7), for $\sigma = 1$, $\xi = 0.5$ and parameters $\delta = 1, 5, \infty$ (dashed-dotted, dotted, dashed black curves, respectively). The limiting case $\delta = \infty$ corresponds to the exact GP density. The solid blue curve represents a gamma density with parameters $(1.4, 1.4)$.

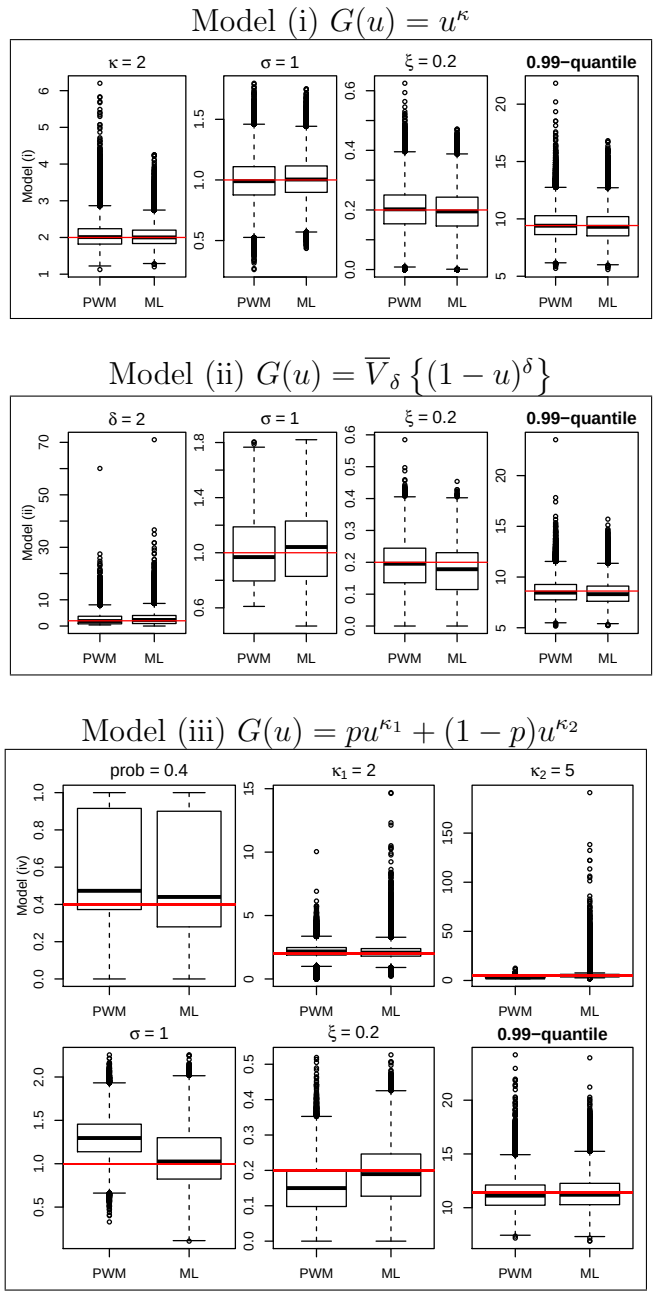

Figure 4. Boxplots of estimated parameters and 0.99-quantiles using PWMs (left) and ML (right), for model (i) with $\sigma = 1$, $\xi = 0.2$, $\kappa = 2$, model (ii) with $\sigma = 1$, $\xi = 0.2$, $\delta = 2$, and model (iii) with $p = 0.4$, $\kappa_1 = 2$, $\kappa_2 = 5$, $\sigma = 1$ and $\xi = 0.2$. Boxplots are based on 10^5 independent replicates, and true values are represented by horizontal red lines.

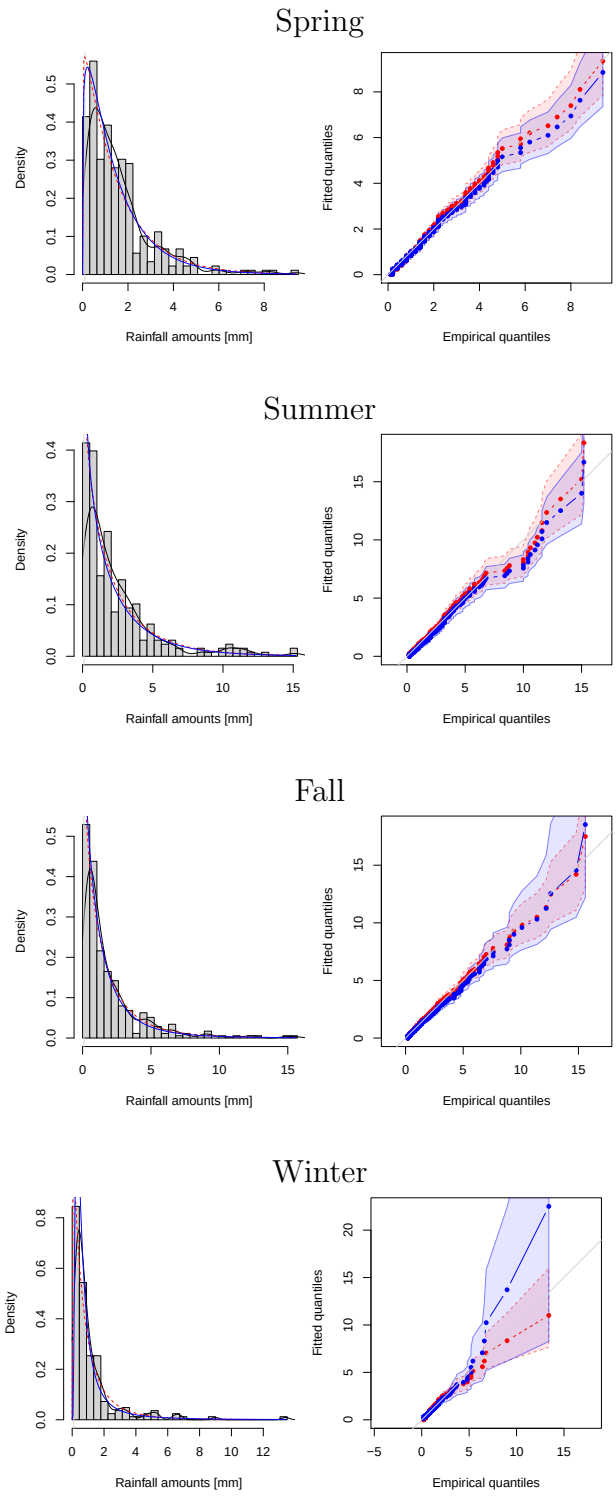

Figure 5. Hourly precipitation (1996-2011, Lyon, France). Left panels represent empirical histograms (grey) for the Spring (a), Summer (b), Fall (c) and Winter (d) data. A kernel-based density (black) and fitted EGP densities (by censored ML in solid blue, by PWMs in dashed red) are superimposed. Right panels correspond to the corresponding quantile-quantile plots with associated 95%-pointwise confidence intervals based on 500 bootstrap replicates.

Table 1. Ratio of root mean squared errors (RMSEs) of PWMs and MLE based on estimates obtained from 10^5 independent datasets of size $n = 300$. Each **cell in bold** represents the ratio of RMSEs for the 99%**-quantile**. Non-bold cells correspond to ratios for each parameter, i.e. $\sigma/\xi/\kappa$ for model (i), and $\sigma/\xi/\delta$ for model (ii).

(i) $G(u) = u^\kappa$			
κ	ξ		
	0.1	0.2	0.3
1	1.06/1.02/1.17	1.06/0.98/1.19	1.11/1.00/1.27
	1.01	**0.98**	**0.98**
2	1.05/1.02/1.13	1.07/1.00/1.18	1.15/1.05/1.33
	1.01	**0.99**	**1.02**
5	1.74/1.39/1.29	1.68/1.49/1.31	1.19/1.09/1.34
	1.00	**1.14**	**1.04**
10	2.40/1.70/1.41	4.21/2.82/1.32	1.39/1.26/0.69
	0.76	**1.26**	**1.06**

(ii) $G(u) = \overline{V}_\delta\left\{(1-u)^\delta\right\}$			
δ	ξ		
	0.1	0.2	0.3
0.5	1.00/1.03/0.98	0.91/0.90/0.91	0.92/0.92/0.91
	1.02	**1.02**	**1.05**
1	0.95/0.99/0.94	0.89/0.90/0.90	0.90/0.92/0.90
	1.03	**1.02**	**1.05**
2	1.02/1.00/0.95	0.95/0.93/0.91	0.93/0.94/0.92
	1.03	**1.01**	**1.04**
5	1.06/1.05/0.95	0.99/0.97/0.98	0.98/0.97/1.07
	1.02	**0.98**	**0.99**

(iii) $G(u) = pu^{\kappa_1} + (1-p)u^{\kappa_2}$				
κ_1	κ_2			
	2	5	10	20
1	**0.57**	**0.96**	**0.68**	**0.74**
2	–	**0.95**	**0.98**	**0.71**
5	–	–	**1.29**	**1.04**
10	–	–	–	**1.38**

(iv) $G(u) = [\overline{V}_\delta\left\{(1-u)^\delta\right\}]^{\kappa/2}$				
κ	δ			
	0.5	1	2	5
1	**1.13**	**1.13**	**1.09**	**1.02**
2	**1.09**	**1.09**	**1.06**	**1.01**
5	**1.10**	**1.09**	**1.07**	**1.04**
10	**1.02**	**1.13**	**1.03**	**1.07**

Table 2. Estimated parameters from fitting model (i) with $G(u) = u^{\kappa}$ in (4) to the hourly rainfall data by maximizing the censored likelihood function (ML-c) and by probability weighted moments (PWMs). 95%-confidence intervals (subscripts) are obtained from 500 non-parametric bootstrap replicates.

Spring data			
Method	Estimated parameters		
	κ	σ	ξ
ML-c	$1.19_{(0.89,1.66)}$	$1.28_{(0.92,1.63)}$	$0.06_{(0.00,0.20)}$
PWMs	$1.05_{(0.92,1.25)}$	$1.48_{(1.18,1.73)}$	$0.03_{(0.00,0.16)}$

Summer data			
Method	Estimated parameters		
	κ	σ	ξ
ML-c	$0.74_{(0.49,1.15)}$	$2.51_{(1.70,3.71)}$	$0.09_{(0.00,0.24)}$
PWMs	$0.86_{(0.74,1.08)}$	$2.37_{(1.76,2.93)}$	$0.13_{(0.00,0.26)}$

Fall data			
Method	Estimated parameters		
	κ	σ	ξ
ML-c	$0.97_{(0.60,1.79)}$	$1.34_{(0.67,2.18)}$	$0.27_{(0.06,0.49)}$
PWMs	$0.88_{(0.75,1.04)}$	$1.69_{(1.29,2.16)}$	$0.19_{(0.06,0.30)}$

Winter data			
Method	Estimated parameters		
	κ	σ	ξ
ML-c	$100_{(1.72,600)}$	$0.08_{(0.00,0.54)}$	$0.71_{(0.32,0.85)}$
PWMs	$1.01_{(0.85,1.26)}$	$0.83_{(0.60,1.08)}$	$0.31_{(0.16,0.46)}$