Uncertainty estimation of Intensity–Duration–Frequency relationships: A regional analysis

Victor Mélèse, Juliette Blanchet, Gilles Molinié

Uncertainty estimation of Intensity-Duration-Frequency relationships: a regional analysis

Victor Mélèse, Juliette Blanchet, Gilles Molinié

PII: S0022-1694(17)30519-X
DOI: http://dx.doi.org/10.1016/j.jhydrol.2017.07.054
Reference: HYDROL 22156

To appear in: *Journal of Hydrology*

Received Date: 3 May 2017
Revised Date: 19 July 2017
Accepted Date: 26 July 2017

Please cite this article as: Mélèse, V., Blanchet, J., Molinié, G., Uncertainty estimation of Intensity-Duration-Frequency relationships: a regional analysis, *Journal of Hydrology* (2017), doi: http://dx.doi.org/10.1016/j.jhydrol.2017.07.054

Uncertainty estimation of Intensity-Duration-Frequency relationships : a regional analysis.

Victor Mélèse[a,*], Juliette Blanchet[a], Gilles Molinié[a]

[a] *UGA - IGE CS 40700 38 058 Grenoble Cedex 9, France*

Abstract

We propose in this article a regional study of uncertainties in IDF curves derived from point-rainfall maxima. We develop two generalized extreme value models based on the simple scaling assumption, first in the frequentist framework and second in the Bayesian framework. Within the frequentist framework, uncertainties are obtained i) from the Gaussian density stemming from the asymptotic normality theorem of the maximum likelihood and ii) with a bootstrap procedure. Within the Bayesian framework, uncertainties are obtained from the posterior densities. We confront these two frameworks on the same database covering a large region of $100,000$ km^2 in southern France with contrasted rainfall regime, in order to be able to draw conclusion that are not specific to the data. The two frameworks are applied to 405 hourly stations with data back to the 1980's, accumulated in the range 3h-120h. We show that i) the Bayesian framework is more robust than the frequentist one to the starting point of the estimation procedure, ii) the posterior and the bootstrap densities are able to better adjust uncertainty estimation to the data than the Gaussian density, and iii) the bootstrap density give unreasonable confidence intervals, in particular for return levels associated to large return period. Therefore our recommendation goes towards the use of the Bayesian framework to compute uncertainty.

1. Introduction

Determining how often a storm of a given intensity is expected to occur requires an evaluation of its probability of occurrence, i.e. its return period. However extremeness of a rainfall event depends at which duration rainfall is considered. For this reason, Intensity-Duration-Frequency (IDF) curves are extensively used in water resources engineering for planning and design (Rantz, 1971; Cheng and AghaKouchak, 2014; Sarhadi and Soulis, 2017; Te Chow, 1988, chapter 14). They provide estimates of return levels for the continuum of durations and return periods. However a difficulty in producing IDF curves is that return periods of interest for risk mitigation amount usually to several hundreds of years, whereas series at disposal are most of the time much shorter. Estimating the 100-year return level, for example, relies then on extrapolating using some statistical model. Uncertainty is inherent to this estimation because no model is perfect. This is

∗. Corresponding Author
Email address: `victor.melese@univ-grenoble-alpes.fr` (Victor Mélèse)

particularly true for extreme value estimation –such as the 100-year return level– because it is based on few data, so a subsequent variability is induced by sampling. Risk evaluation should account for this uncertainty to avoid over-optimistic results (Coles and Pericchi, 2003). Since current infrastructure dealing with flooding and precipitation (e.g. dams or dikes) are based on IDF curves, ignoring uncertainty may result in sharp underestimation of flood risk and failure risk of critical infrastructures.

Few studies have explicitly examined uncertainty in IDF curves. They rely on two distinct theoretical frameworks making different modeling assumptions. The first one is a frequentist framework in which the IDF model parameters are treated as unknown real values. Estimation is usually made by moment- or likelihood-based methods and uncertainty is mainly obtained by a bootstrap resampling scheme to account for the influence of sampling on IDF estimation (Overeem et al., 2008; Hailegeorgis et al., 2013; Tung and Wong, 2014). The second one is a Bayesian framework. It differs from the frequentist framework in that the IDF model parameters are treated as random variables. Its estimation allows by nature uncertainty quantification by providing the most likely distribution for the parameters based on the data (Huard et al., 2010; Cheng and AghaKouchak, 2013; Chandra et al., 2015; Van de Vyver, 2015). The influence of the chosen framework on IDF uncertainty estimation has, to the best of our knowledge, never been addressed in the literature.

In this paper, we propose to confront the frequentist and Bayesian frameworks on the same database covering a large region with contrasted rainfall regimes, in order to be able to draw conclusion that are not specific to the data. The studied region covers $100,000$ km^2 of the southern part of France that is under mediterranean climatic influence and is notably well-instrumented with 563 hourly raingages since the mid-80s, from which we select the 405 stations featuring at least 10 years of observations. The IDF relationships used in this works rely on the simple scaling assumption (Gupta and Waymire, 1990), associated with a Generalized Extreme Value (GEV) distribution representing the frequency of annual maximum rainfall intensity. This model has been validated in the frequentist case in Blanchet et al. (2016a) for the same region. Here we mainly extend this work by assessing uncertainty in IDF relationships, which was missing in Blanchet et al. (2016a). We develop in Section 2 the Bayesian and frequentist frameworks of GEV-simple scaling IDF relationships. We present the data in Section 3 and give evidence of simple scaling in the range 3h-120h in the region in Section 4. We describe the workflow of analysis in Section 5. Finally, we confront the results of the two frameworks, with a particular focus on uncertainty estimation in Section 6.

2. Two frameworks of IDF relationships

2.1. Introduction

Return levels computation requires estimating the occurrence probability of annual maximum rainfall intensity, i.e. their probability density function (PDF). The founding theorem of extreme value theory (see Coles et al., 2001, for a full review) states that if independent and identically distributed data are blocked into sequences of observations and if each block is long enough, then the PDF of block maxima is approximately the Generalized Extreme Value (GEV) distribution. The combination of strict sense simple scaling and GEV theory for annual maximum rainfall intensity leads to the family of GEV-simple scaling models (Blanchet et al., 2016a). In the next sections, we develop two GEV-simple scaling models, respectively in the frequentist and the Bayesian frameworks. The main difference between the two is that model parameters are scalars under the frequentist framework and random variables under the Bayesian framework. In the following, we write random variables with bold symbols to distinguish them from scalars.

2.2. Frequentist framework

2.2.1. Model

The frequentist framework is the one considered in Blanchet et al. (2016a) in the same region and used in Borga et al. (2005) and Bougadis and Adamowski (2006). It relies on two assumptions. First, on the strict sense simple scaling assumption of Gupta and Waymire (1990) setting that

$$\mathrm{pr}(\boldsymbol{M}_D < x) = \mathrm{pr}\left\{\left(\frac{D}{D_{ref}}\right)^{-H} \boldsymbol{M}_{D_{ref}} < x\right\}, \tag{1}$$

where $\boldsymbol{M}_D$ is the random variable of annual maximum rainfall intensity for a duration D, $\boldsymbol{M}_{D_{ref}}$ is the random variable of annual maximum rainfall intensity for a duration of reference D_{ref} (D_{ref} = 3h in the application of Section 6), and H is a non-negative scalar called the scaling exponent. In terms of moments, Eq. 1 leads to the wide sense simple scaling assumption of Gupta and Waymire (1990)

$$\forall q \in \mathbb{R},\ \mathbb{E}(\boldsymbol{M}_D^q) = \left(\frac{D}{D_{ref}}\right)^{-Hq} \mathbb{E}(\boldsymbol{M}_{D_{ref}}^q), \tag{2}$$

which shows the advantage over (1) of being easily checked empirically on data, at least for moderate q, by computing the empirical moments and regressing them against the duration in log-log scale (see Section 4 for more details in our application).

The second assumption of our model is founded by extreme value theory and asserts that annual maximum rainfall intensity at reference duration, $\boldsymbol{M}_{D_{ref}}$, follows a Generalized Extreme Value (GEV), i.e. that

$$\mathrm{pr}(\boldsymbol{M}_{D_{ref}} < x) = \exp\left[-\left(1 + \xi\frac{x - \mu_{D_{ref}}}{\sigma_{D_{ref}}}\right)^{-\frac{1}{\xi}}\right], \tag{3}$$

provided $1 + \xi\frac{x-\mu_{D_{ref}}}{\sigma_{D_{ref}}} > 0$, where $\mu_{D_{ref}}$, $\sigma_{D_{ref}} > 0$, ξ are scalars, called respectively the location, scale and shape parameters. Case $\xi = 0$ corresponds to the Gumbel distribution

$$\mathrm{pr}(\boldsymbol{M}_{D_{ref}} < x) = \exp\left[-\exp\left(-\frac{x - \mu_{D_{ref}}}{\sigma_{D_{ref}}}\right)\right]. \tag{4}$$

(3) associated with (1) implies that annual maximum rainfall intensity $\boldsymbol{M}_D$ of any duration D follows a GEV distribution (Blanchet et al., 2016a) and that the GEV parameters at duration D and D_{ref} are linked through $\mu_D = \left(\frac{D}{D_{ref}}\right)^{-H} \mu_{D_{ref}}$, $\sigma_D = \left(\frac{D}{D_{ref}}\right)^{-H} \sigma_{D_{ref}}$, while the shape parameter ξ does not depend on the time scale. As a consequence, the IDF relationships relating the duration D, the return period T_R and the return level (i.e. the quantile of order $1 - 1/T_R$ of the corresponding GEV distribution) is given by

$$m_{D,T_R} = \left(\frac{D}{D_{ref}}\right)^{-H} \left\{ \mu_{D_{ref}} - \frac{\sigma_{D_{ref}}}{\xi} \left(1 - \left[-\log(1 - \frac{1}{T_R}) \right]^{-\xi} \right) \right\}. \tag{5}$$

2.2.2. Inference

The set of unknown parameters to be estimated is $\theta = (\mu_{D_{ref}}, \sigma_{D_{ref}}, \xi, H)$. As in Blanchet et al. (2016a), θ is estimated by maximizing the likelihood under the assumptions that i) annual maxima are independent from one year to another, and ii) annual maxima of a given year at different durations are independent. This later assumption is likely to be miss-specified. For instance a 4h annual maximum is likely to be correlated with a 3h annual maximum. However incorporating dependence among many durations complicates the modeling and its estimation (Davison et al., 2012; Cooley et al., 2012; Ribatet and Sedki, 2012; Davison and Huser, 2015), with little gain, if not loss, when only the marginal distributions are of interest (Sebille et al., 2017). We are in this case since IDF relationships relate to quantiles of marginal distributions. Under the assumption of independence, the model log-likelihood is given by

$$\begin{aligned} l(\theta) = &\sum_{D \in \mathcal{D}} n(D) \log \left(\frac{D}{D_{ref}}\right)^H - \log(\sigma_{D_{ref}}) \sum_{D \in \mathcal{D}} n(D) - \\ &\frac{\xi + 1}{\xi} \sum_{D \in \mathcal{D}} \sum_{i=1}^{n} \log \left(1 + \xi \frac{\left(\frac{D}{D_{ref}}\right)^H m_{D,i} - \mu_{D_{ref}}}{\sigma_{D_{ref}}} \right) - \\ &\sum_{D \in \mathcal{D}} \sum_{i=1}^{n} \left[1 + \xi \frac{\left(\frac{D}{D_{ref}}\right)^H m_{D,i} - \mu_{D_{ref}}}{\sigma_{D_{ref}}} \right]^{-\frac{1}{\xi}}, \end{aligned} \tag{6}$$

where $n(D)$ is the number of observed years at duration D, $m_{D,i}$ is the annual maximum rainfall intensity at the duration D for year number i and $\mathcal{D}$ is the set of considered durations. There is no analytical form for the maximum of l but maximization can be obtained numerically (e.g. quasi Newton method).

2.2.3. Uncertainty computation

We propose two ways of computing uncertainty in the frequentist framework. The first one relies on the asymptotic normality of the maximum likelihood estimator, but using the correction described in Davison (2008) and used in Van de Vyver (2012) to account for the fact that the likelihood (6) ignores dependence among maxima of the same year. Let $\tilde{\theta}$ denote the value maximizing the log likelihood function (6). It is function of the data m_D. Writing this in terms of random variables means that the maximum likelihood estimator

$\hat{\boldsymbol{\theta}}_{ML}$ is function of the random variable of annual maximum rainfall intensity $\boldsymbol{M}_D$. $\hat{\boldsymbol{\theta}}_{ML}$ is a random variable because it depends on the $\boldsymbol{M}_D$'s which are random, while $\tilde{\theta}$ is a scalar; it is a realization of $\hat{\boldsymbol{\theta}}_{ML}$. Being random, $\hat{\boldsymbol{\theta}}_{ML}$ has a distribution. The theorem of asymptotic normality of the maximum likelihood estimator provides an approximation for this distribution when the number of data is large. Under the correction of likelihood misspecification for dependence, it states that $\hat{\boldsymbol{\theta}}_{ML}$ can be considered as multivariate normal distributed, with mean approximated by $\tilde{\theta}$ and covariance matrix approximated by $\Sigma(\tilde{\theta}) = I(\tilde{\theta})^{-1}V(\tilde{\theta})I(\tilde{\theta})^{-1}$ where $I(\tilde{\theta})$ and $V(\tilde{\theta})$ are the 4×4 matrices

$$I(\theta) = -\sum_{i=1}^{n} \frac{\partial^2 l_i(\theta)}{\partial\theta\partial\theta^T}, \quad V(\theta) = \sum_{i=1}^{n} \frac{\partial l_i(\theta)}{\partial\theta} \frac{\partial l_i(\theta)}{\partial\theta^T},$$

evaluated in $\tilde{\theta}$. An approximate $(1-\alpha)$ confidence interval for θ_j, any of the four model parameters, is then given by

$$\tilde{\theta}_j \pm z_{\alpha/2}\sqrt{\Sigma_{jj}},$$

where $z_{\alpha/2}$ is the $(1-\alpha/2)$ quantile of the standard normal distribution and Σ_{jj} is the jth diagonal element of Σ.

Applying the delta method (Coles et al., 2001), the maximum likelihood estimator of the T_R-year return level at duration D can be considered as normal distributed with mean approximated by $g(\tilde{\theta})$ and variance approximated by $\tau^2(\tilde{\theta})$, where

$$\tau^2(\theta) = \frac{\partial g(\theta)}{\partial\theta^T}\Sigma(\tilde{\theta})\frac{\partial g(\theta)}{\partial\theta},$$

and g is the right-hand side function in (5). In particular, its $(1-\alpha)$ confidence interval is approximately

$$g(\tilde{\theta}) \pm z_{\alpha/2}\,\tau(\tilde{\theta}).$$

The second method to obtain uncertainties is based on bootstrap resampling. It allows to account for the influence of sampling on IDF estimation. It consists of resampling the data with replacement to obtain new samples. Let's assume that the annual maxima are stored in a matrix with one row per year and one column per duration. A bootstrap sample is constructed by drawing with replacement the lines of the matrix. The log likelihood function is maximized for each bootstrap sample, given a new estimate $\tilde{\theta}$, which is considered as a possible realization of the true estimator $\hat{\boldsymbol{\theta}}$. If R bootstrap samples are used, R realizations $\tilde{\theta}_1, \ldots, \tilde{\theta}_R$ are obtained. When R is large (e.g. $R = 1000$ in Section 6), usual density estimates (e.g. Kernel density) can be applied to $\tilde{\theta}_1, \ldots, \tilde{\theta}_R$ to obtained an approximate density for $\hat{\boldsymbol{\theta}}$. An approximate density for the T_R-year return level is obtained likewise by estimating the density of the $g(\tilde{\theta}_1), \ldots, g(\tilde{\theta}_R)$, where g is the right-hand side function in (5). Approximate $(1-\alpha)$ confidence intervals are obtained empirically as the interval bounded by the empirical quantiles of order $\alpha/2$ and $(1-\alpha/2)$.

2.3. Bayesian framework

2.3.1. Model and priors

As in the frequentist framework, the Bayesian framework relies on the strict sense simple scaling hypothesis combined with the GEV distribution. However in this case, the model parameters $\boldsymbol{\theta} = (\boldsymbol{\mu}_{D_{ref}}, \boldsymbol{\sigma}_{D_{ref}}, \boldsymbol{\xi}, \boldsymbol{H})$ are random variables. Thus the two above hypothesis, as all the equations derived in Section 2.2.1, still apply but conditionally on $\boldsymbol{\theta}$ equals to some $\theta = (\mu_{D_{ref}}, \sigma_{D_{ref}}, \xi, H)$. In particular, the strict sense simple scaling assumption of Gupta and Waymire (1990) turns into

$$\mathrm{pr}(\boldsymbol{M}_D < x | \boldsymbol{H} = H) = \mathrm{pr}\left\{ \left(\frac{D}{D_{ref}}\right)^{-H} \boldsymbol{M}_{D_{ref}} < x \right\}, \tag{7}$$

which leads, in terms of moments, to

$$\forall q \in \mathbb{R},\ \mathbb{E}(\boldsymbol{M}_D^q | \boldsymbol{H} = H) = \left(\frac{D}{D_{ref}}\right)^{-Hq} \mathbb{E}(\boldsymbol{M}_{D_{ref}}^q). \tag{8}$$

Likewise, conditional on $\boldsymbol{\theta} = \theta$, the annual maximum rainfall intensity $\boldsymbol{M}_D$ of any duration D, follows a GEV distribution, i.e.

$$\mathrm{pr}(\boldsymbol{M}_D < x | \boldsymbol{\theta} = \theta) = \exp\left[-\left(1 + \xi \frac{x - \mu_D}{\sigma_D}\right)^{-\frac{1}{\xi}} \right], \tag{9}$$

where $\mu_D = \left(\frac{D}{D_{ref}}\right)^{-H} \mu_{D_{ref}}$ and $\sigma_D = \left(\frac{D}{D_{ref}}\right)^{-H} \sigma_{D_{ref}}$.

Finally, the random variable of the T_R-year return level for duration D is given by

$$\boldsymbol{M}_{D,T_R} \overset{a.s.}{=} \left(\frac{D}{D_{ref}}\right)^{-\boldsymbol{H}} \left[\boldsymbol{\mu}_{D_{ref}} - \frac{\boldsymbol{\sigma}_{D_{ref}}}{\boldsymbol{\xi}} \left(1 - \left[-\log(1 - \frac{1}{T_R}) \right]^{-\boldsymbol{\xi}} \right) \right], \tag{10}$$

where $\overset{a.s.}{=}$ means equality almost surely.

Since (9) is conditional on $\boldsymbol{\theta}$, full modeling of $\boldsymbol{M}_D$ requires defining the density of $\boldsymbol{\theta}$, i.e. the prior density. Here we assume independence of the model parameters, i.e.

$$f(\theta) = f(\mu_{D_{ref}}) f(\sigma_{D_{ref}}) f(\xi) f(H). \tag{11}$$

We make this choice for the sake of simplicity but a separate analysis applied to the data of Section 3 revealed that actually choosing dependent or independent priors does does not affect the results.

In (11) univariate prior densities for $\mu_{D_{ref}}$, $\sigma_{D_{ref}}$, ξ and H have to be chosen. Choice of the prior density is crucial in Bayesian analysis and a whole field of research is devoted to this issue. Prior densities can be separated into two major classes, namely subjective (or informative) and objective (or uninformative) priors (Gelman et al., 2014; Beirlant et al., 2005, chapter 11). Subjective priors allow to bring prior knowledge to the analysis, based on expert information of different degrees. Objective priors (Berger, 2006; Kass and Wasserman, 1996) should be used when subjective analysis is not possible. Most common objective priors include the uniform density, Maximum Data Information prior (Zellner, 1998) and Jeffreys prior (Kass and Wasserman, 1996; Jeffreys, 1998). For what matters extreme rainfall and GEV distributions, there is no consensus on the choice of the priors. Coles and Tawn (1996) use expert information on extreme quantiles. Huard et al. (2010) and Chandra et al. (2015) use objective priors for the location (uniform) and scale (Jeffreys) but a weakly subjective prior for the shape (Beta). Coles and Pericchi (2003) uses objective priors for the three GEV parameters (Gaussian for the location and shape, log-Gaussian for the scale). For IDF relationships, Van de Vyver (2015) uses objective priors for the location, scale and scaling exponent (respectively Gaussian, log-Gaussian and uniform) and weakly subjective prior for the shape (Beta). Muller et al. (2008) also uses objective priors for the location, scale and scaling exponent (Gaussian for the first and log-Gaussian for the two latter) and weakly subjective prior for the shape (uniform).

In this work, we aim to use a model as general as possible in order to make a fair comparison of uncertainty with the frequentist framework, which does not include expert knowledge, so the four chosen priors are very weakly informative. For the location parameter at reference duration (3h), we choose an objective uniform density as in Huard et al. (2010) and Chandra et al. (2015). The bounds are chosen to span the worldwide values of $\boldsymbol{\mu}_{D_{ref}}$, from very arid to very humid regions, in order to use priors as little informative as possible for our data. In a study of more than 15,000 worldwide records, Papalexiou and Koutsoyiannis (2013) finds that the location parameter for annual maxima of daily rainfall ranges between 6 and 700mm/day. Since rainfall accumulation cannot be greater in 3h than in 24h, we can anticipate that the location parameter for annual maxima of 3h rainfall is worldwide no lower than 6mm/3h and no bigger that 700mm/3h, i.e. between 2 and 233mm/h at 3h duration. In order to be even less conservative, we set the lower and upper bounds of the uniform prior for $\boldsymbol{\mu}_{D_{ref}}$ to 0 and 250mm/h at 3h duration, respectively. Likewise, we use for the scale parameter at reference duration $\boldsymbol{\sigma}_{D_{ref}}$ a uniform prior with bounds 0.1 and 150mm/h at 3h duration, which extends over the range of values found in Papalexiou and Koutsoyiannis (2013) (2-400mm/day). For the shape parameter, we use the normal density, which tends to be less informative than the Beta prior used in Huard et al. (2010), Chandra et al. (2015) and Van de Vyver (2015), which has bounded tails. Papalexiou and Koutsoyiannis (2013) shows that the distribution of the shape parameter is approximately Gaussian with mean 0.1 and standard deviation 0.045. Here we consider a much less informative density by using a Gaussian prior with mean 0.1 but standard deviation 0.5. Finally, owing to the fact that the scaling parameter is non-negative and lower than 1, we choose for $\boldsymbol{H}$ a uniform density between 0 and 1, as in Van de Vyver (2015).

2.3.2. Inference

For shortness we denote $\boldsymbol{M}$ the set of annual maximum rainfall intensities, i.e. the set of $\boldsymbol{M}_{D,i}$, $D \in \mathcal{D}$, $i = 1, \ldots, n$. In the Bayesian framework, interest is in estimating the density of the parameters knowing the data, i.e. $f(\theta|\boldsymbol{M} = m)$, called the posterior density. The well known Bayes formula states that

$$f(\theta|m) = \frac{f(m|\theta)f(\theta)}{\int_\theta f(m|\theta)f(\theta)d\theta}, \tag{12}$$

where the prior density $f(\theta)$ is given by (11) with the aforementioned priors and $f(m|\theta)$ is the density associated to the data under (9), whose log expression is assumed to be given by (6). By doing this we assume that the maxima at different durations are independent conditional on the parameters. In a Bayesian framework, Van de Vyver (2015) and Muller et al. (2008) model dependence between two durations (namely 24h and 72h) with a logistic model, while Stephenson et al. (2016) uses max-stable processes to model dependence across several durations. However Sebille et al. (2017) shows by comparing different spatial models, including that of Stephenson et al. (2016), that when interest lies in the estimation of marginal quantities, such as return levels, the independence assumption is one of the most creditable one.

In our case, as often in Bayesian analysis, there is no analytical form for the posterior density (12) due to the presence of an integral in the normalizing constant. This problem can be overcome by using simulation based techniques such as Markov chain Monte Carlo (MCMC), which provides a way of simulating from complex distributions, such as $f(\theta|m)$, by simulating from Markov chains which have the target distributions as their stationary distributions. Estimates of the posterior distribution could then be obtained from the simulated sample at convergence of the Markov chains. There are many MCMC techniques, among which the most popular are the Gibbs sampler when it is possible to simulate from the full conditional distribution, or Metropolis-Hastings sampling otherwise. Here simulation from the full conditional distribution is not straightforward so we use Metropolis sampling, i.e. Metropolis-Hastings with symmetric jumping distributions (or proposal distribution). In our case, it proceeds as follows :

1. Draw a starting point $\theta^{(0)}$ for which $f(\theta^{(0)}|m)$ is defined and strictly superior to 0.
2. At each step t,
   - Draw a candidate θ^* from a symmetric jumping distribution $J_t(\theta^*|\theta^{(t-1)})$.
   - Derive the acceptance probability :

$$a = \min\left\{1, \frac{f(\theta^*|m)}{f(\theta^{(t-1)}|m)}\right\} = \min\left\{1, \frac{f(m|\theta^*)f(\theta^*)}{f(m|\theta^{(t-1)})f(\theta^{(t-1)})}\right\}, \tag{13}$$

— Accept or reject the candidate θ^*, i.e. set

$$\theta^{(t)} = \begin{cases} \theta^* \text{ with probability } a, \\ \theta^{(t-1)} \text{ otherwise.} \end{cases} \tag{14}$$

We use a Gaussian distribution for the jumping distribution $J_t(.|\theta^{(t-1)})$, with mean $\theta^{(t-1)}$ and diagonal covariance matrix with standard deviation set at initialization to $(0.3, 0.3, 0.025, 0.025)$, and then tuned during the first half iterations of the MCMC so that the acceptance rate of θ (i.e. the proportion of times θ^* is set to $\theta^{(t)}$) is between 30% and 50%. The resulting chain converges, after an initial burn-in period, to the posterior distribution. At the end of the algorithm, samples of the posterior density are obtained as $\theta^{(t)}$, for t exceeding the burn-in period. We will see in Section 5 how to monitor this convergence. Estimate of the posterior density can be obtained by usual (e.g. Kernel) density estimate based on an independent subsample of these $\theta^{(t)}$.

3. Data

The studied region corresponds to the southern part of France that is under Mediterranean climatic influence (see Fig. 1). It is limited to the south by the Mediterranean coast from Perpignan to Nice, to the west by the Pyrenees, to the north by the Massif Central and to the east by the southern Alps. Altitude ranges from 0 to more than 3000 m.a.s.l. The highest peaks are located in the the Alps and the Pyrenees while the Massif Central is mostly below 1500m. The mountain massifs design funnel-shaped domains that are known to experience severe storms generating flash-floods from various foothill rivers. Examples are provided by quite recent severe events causing numerous human losses and considerable damages that occurred in 1999 on the Aude River (Gaume et al., 2004), in 2002 on the Gard River (southern edge of the Massif Central - Delrieu et al., 2005), in 1992 on the Ouvèze River (eastern flank of the Alps - Sénési et al., 1996) and in 2010 on the Argens River (southern edge of the Alps - Ruin et al., 2014). Nevertheless a strong heterogeneity exists in terms of occurrence of such events in this area. The south-eastern edge of the Massif Central experiences most of the extreme storms and resulting flash-floods (Fig. 2 of Nuissier et al., 2008). The HyMeX field campaign (Ducrocq et al., 2013; Drobinski et al., 2014) illustrates a variety of meteorological situations blocking heavy rainfall systems over the region. The presence of the surrounding mountain massifs is critical in the positioning and stationarity of these systems (Nuissier et al., 2008).

The instrumented area covers a surface of about $100,000$ km^2, as displayed in Fig. 1. Hourly rainfall data are acquired by either Météo-France or Electricité de France since the mid 80's for the oldest. 563 hourly raingages with more than 10 years observations are available. We restrict the data to the three months of September-October-November (SON) since flash floods usually occur in Autumn in this region. Starting from hourly data, we create new databases by aggregating hourly rainfalls at 3h, 4h, 8h, 12h, 24h, 48h,

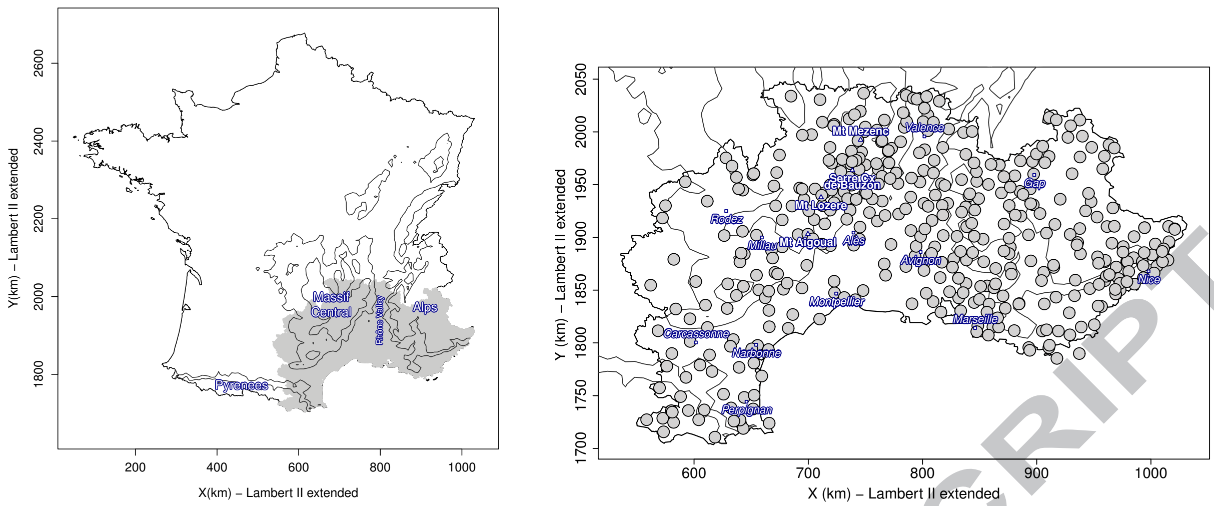

FIGURE 1: Map of studied region with main mountains peaks (triangle), main cities (square) and raingage locations (circle).

72h, 96h and 120h using 1h-length moving windows. We do not consider maxima at duration 1h and 2h because these maxima are likely to underestimate the true maxima when a sampling period of 1h is used. This underestimation is likely to decrease with duration. Then, SON maxima are extracted for each of these durations. Following Blanchet et al. (2016a), a given maximum is considered as missing if its rank is smaller than $pmiss \times N$ where $pmiss$ is the proportion of missing values for that season and duration, and N is the number of maxima for the considered duration. This allows us to consider maxima of very incomplete year (large $pmiss$), provided these maxima are large compared to the other maxima (i.e. their ranks are large). Finally a given SON season is considered as completely missing if at least four of the nine durations are missing and the whole station is considered as missing (i.e. excluded from the analysis) if less than 10 SON maxima are observed. Doing so, we end up with a set of 405 stations (see Fig. 1).

4. Evidence of simple scaling

We first give empirical evidence of simple scaling of rainfall in our region. It is not possible to check the strict sense simple scaling assumptions (1) and (7) directly on the data because they depend on H which is unknown. However, it is possible to check their counterpart versions (2) and (8) for the moments, which state in both frameworks that

- wide sense scaling hypothesis : the logarithm of moment of order q of annual maximum rainfall intensity is a linear function of the logarithm of duration,
- wide sense simple scaling hypothesis : the slope of the above linear functions is an affine function of q (i.e. of the form Hq).

We check wide sense scaling hypothesis for $q = 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2$ by computing, for each station, the empirical moment of order q of the maxima at each duration, and regressing the logarithm of

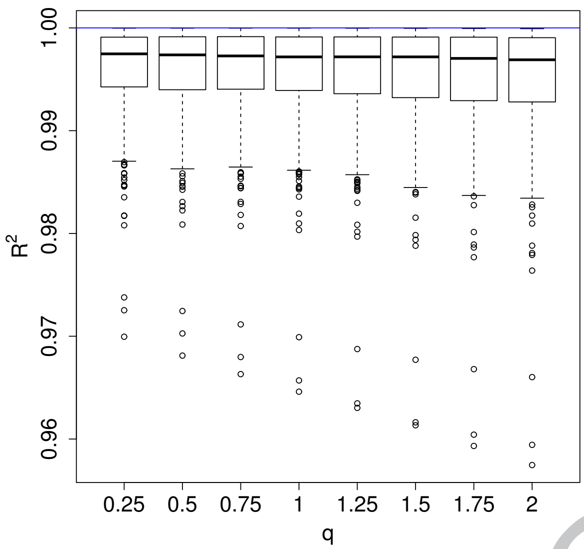

FIGURE 2: Boxplots of the correlation coefficients, R^2, of the empirical moments of order $q = 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2$ of maximum rainfall intensity versus duration in log-log scale. The blue horizontal line show the theoretical value under the wide sense scaling hypothesis.

these values with respect to the log duration. We show in Figure 2 the boxplots of the correlation coefficients, R^2, of these regressions for the 406 stations. We see that all R^2 are all close to one, as should be under the simple-scaling hypothesis. However, this gives only rough evidence of scaling because R^2 are computed over all durations from 3h to 120h, so it is not possible to assess whether specific durations tend to depart from the regressing lines, which would mean that the simple scaling hypothesis applies only on part of the considered durations. To check this, we consider the case $q = 1$ and compute the slope between averages of successive durations, i.e. between e_d and e_{d+1}, where e_d is the average of maximum rainfall intensity at the dth smallest duration, for a given station. Let call s_d this slope, $d = 1, \dots, 8$. Any ratio $s_d/s_{d'}$ should be one under the wide sense scaling hypothesis. We show in Fig. 3 the boxplots of the ratio s_d/s_{d+1}, $d = 1, \dots, 7$, for the 406 stations. We see that 95% of the ratio lie between 0.6 and 1.4, which can be considered as close to one given that each slope is computed on two points only. More importantly maybe, we do not see any break point in the 95% envelopes as d increases, so the wide sense scaling hypothesis seems to apply equally to all durations between 3h and 120h.

To check the wide sense simple scaling assumption, we consider the slopes of Fig. 2 for $q = 0.25, 0.5, 0.75, 1,$ $1.25, 1.5, 1.75, 2$, divide them by q, and denote c_k, $k = 1, \dots, 7$, these values. If the simple scaling holds, each c_k should equal H. Fig. 4 shows the ratio c_k/c_{k+1}, for $k = 1, \dots, 7$. We see that 95% of ratios lie between 1.011 and 0.988, with no value lower than 0.984 and larger than 1.021. This gives good evidence of wide sense simple scaling in the region.

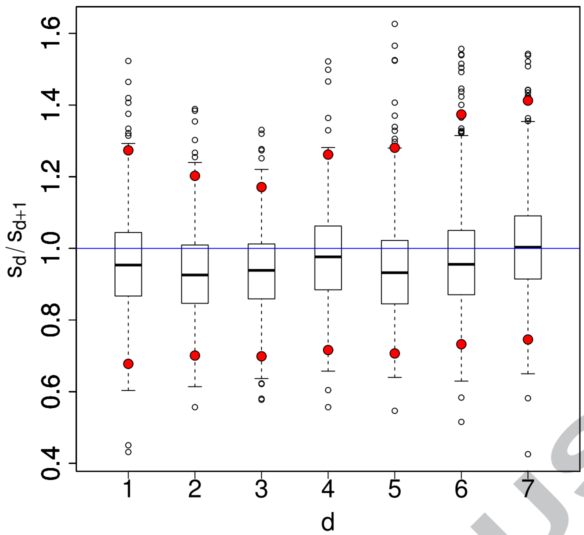

FIGURE 3: Boxplots of the ratio of the slopes s_d/s_{d+1}, for $d = 1, \ldots, 7$ and $q = 1$. The upper and lower red points show the quantiles of order 0.975 and 0.025, respectively. The blue horizontal line shows the theoretical value under the wide sense scaling hypothesis.

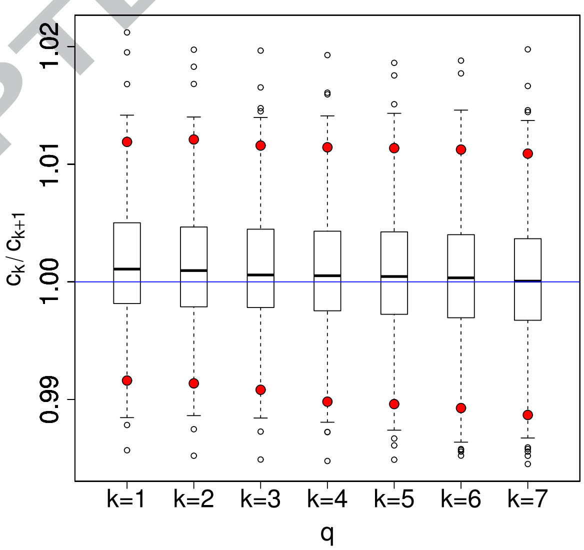

FIGURE 4: Boxplots of the ratio of the slopes c_k/c_{k+1}, for $k = 1, \ldots, 7$. The upper and lower red points show the quantiles of order 0.975 and 0.025, respectively. The blue horizontal line shows the theoretical value under the wide sense simple scaling hypothesis.

5. Workflow

5.1. *Frequentist framework*

The GEV simple scaling model in the frequentist framework (Section 2.2.1) is estimated at each station by maximizing the likelihood (6). Optimization is based on the gradient projection method of Byrd et al. (1995) allowing box constraints for the variables. Constraints are set on the scale parameter, which is restricted to strictly positive values, the shape parameter, which is restricted in the range $(-0.75, 0.75)$ and the scaling parameter H, which is constrained in the range $(0, 1)$. Optimization is initialized by $\theta_1 = (\mu_{D_{ref,1}}, \sigma_{D_{ref,1}}, \xi_1, H_1)$, which can be considered as a smart initialization in that it is built from that data of each station : ξ_1 is set to 0, corresponding to a Gumbel distribution. $\mu_{D_{ref,1}}$ and $\sigma_{D_{ref,1}}$ are estimated using the method of moments under the Gumbel assumption. Following (2) with $q = 1$, H_1 is set to the opposite of the regression slope of the log average maxima on the log duration (i.e. case $q = 1$ in Fig. 2). Starting from θ_1, the gradient projection algorithm stops in $\tilde{\theta}$, the maximum likelihood estimate, if it is unable to reduce the log likelihood (6) by a factor of $10^{-8} \times |l(\tilde{\theta})|$. Density estimates of the associated random variable are obtained i) from the theorem of asymptotic normality of the maximum likelihood estimator, and ii) by bootstrap resampling technique using 1000 bootstrap samples. Return level estimates and associated densities are derived from these estimations as detailed in Section 2.2.3.

5.2. *Bayesian framework*

The same starting points θ_1 is used to initialize Metropolis-Hastings algorithm in the Bayesian framework (Section 2.3.2). Convergence of the MCMC is monitored using the $\hat{R}$ convergence criteria of Gelman et al. (2014) chapter 6, based on five runs of the Metropolis-Hastings algorithm. Convergence is considered to be reached if $\hat{R} < 1.06$, which is obtained after $20,000$ iterations. The burn-in period is set to the first half iterations and every 10th iteration of the remaining 10,000 iterations is considered for the estimation of the posterior density, in order to reduce dependence within the sample. So, the posterior density estimation is based on 1000 samples. Posterior density estimates of return levels are obtained from (10), using these 1000 samples. To summarize any posterior density with one single value and, in particular, compare estimations with the frequentist framework, we decide to consider the posterior mean, i.e. the mean of the posterior density. Another common choice is to consider the mode of the posterior density (maximum a posteriori) but this is slightly less stable than the posterior mean.

6. Results and discussion

6.1. *IDF curves*

Although this is not the main focus of this study, we present below some results on IDF relationships because they are valuable from a climatological point of view by documenting the main hydrological processes

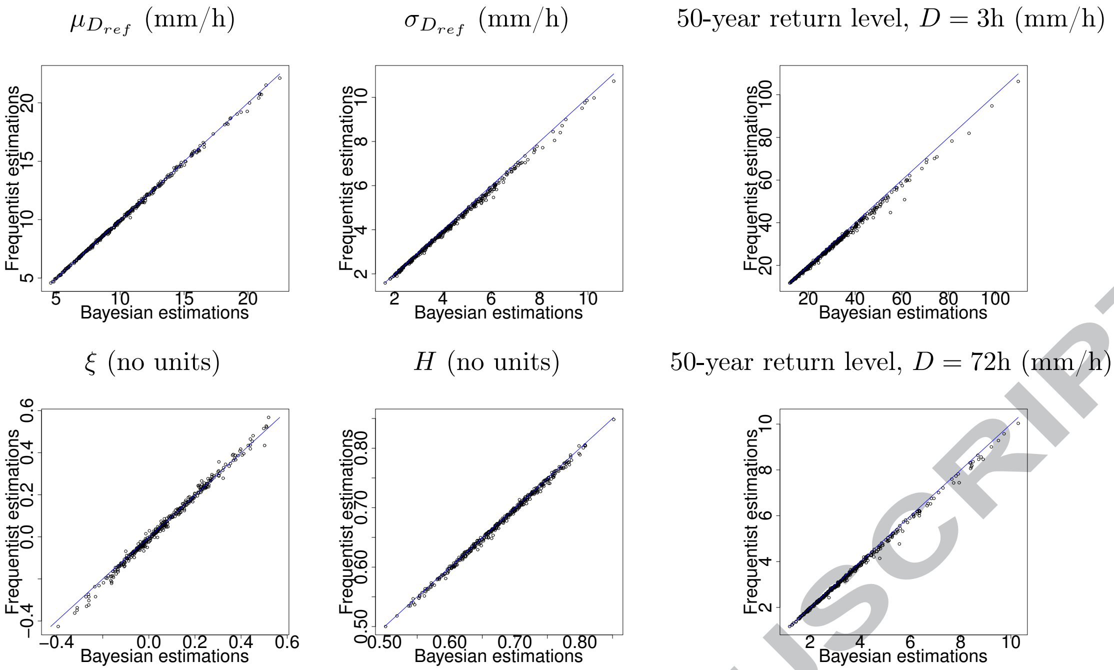

Figure 5: Scatter plot of the Bayesian (posterior mean) and frequentist (maximum likelihood), for $\mu_{D_{ref}}$, $\sigma_{D_{ref}}$, ξ, H and for the 2- and 50-year return levels at 3h and 72h durations.

leading to extreme rainfall in the region.

6.1.1. *Estimation and goodness-of-fit*

Fig. 5 compares the Bayesian (posterior mean) and frequentist (maximum likelihood) estimates. It shows that the framework has very little impact on these estimation with the chosen initialization. A separate analysis (not shown) revealed that actually the Bayesian framework is very little sensitive to initialization, whereas the frequentist framework requires a quite reasonable initialization. In order to assess goodness-of-fit of the estimated IDF curves, we consider two goodness-of-fit criteria proposed by Blanchet et al. (2016a) : the relative Root Mean Square Error (rRMSE) and the relative bias (rBIAS), respectively given by

$$\text{rRMSE}_i(D) = \left\{ n_i(D) \sum_{T_R} \left[\frac{m_{i,D,T_R} - \widehat{m_{i,D,T_R}}}{\sum_{T'_R} m_{i,D,T'_R}} \right]^2 \right\}^{1/2}, \tag{15}$$

and,

$$\text{rBIAS}_i(D) = \sum_{T_R} \left[\frac{m_{i,D,T_R} - \widehat{m_{i,D,T_R}}}{\sum_{T'_R} m_{i,D,T'_R}} \right], \tag{16}$$

where m_{i,D,T_R} is the empirical T_R-year return level for duration D and station i and $\widehat{m_{i,D,T_R}}$ is its estimation. The closer rBIAS and rRMSE to zero, the better the fit. We find that, under both frameworks, the absolute value of rBIAS is no bigger than 12% for 95% of the stations and rRMSE is no bigger that 26% for 95% of the data. This is of the same order as the values found in Blanchet et al. (2016a) on part of the region but using daily data on a much longer observation period (about 60 years).

6.1.2. Spatial variability of return level across durations

Fig. 6 displays the posterior mean estimations of the 2- and 50-year return levels at 3h and 72h durations. Fig. 6 shows that the 2- and 50-year return levels behave differently as the duration increases from 3h to 72h. Considering the 2-year return level, the largest values at 3h duration are found in the foothill around the town of Alès and along the overhanging Massif Central crest. Increasing the duration to 72h, the largest values are still found along the crest but, comparatively, the 2-year return level fade in the foothill.

Rainfall events featuring a 2-year return period are quite common as by definition they tend to occur regularly in one's life (every two years on average). Molinié et al. (2012) characterize the rainfall regimes in the Massif Central region. They show that the largest rainfalls at hourly duration usually occur both over the foothill and over the Massif Central crest. The rainfall characteristics are those of convective storms in terms of intermittency, diurnal cycle and spatial pattern. Increasing the duration to 72h, one may hypothesize that there is no stationary forcing of rainfall over the foothill, while the mountain crest or slope may continue to trigger rainfall if humidity remains sufficient. Molinié et al. (2012) shows that the spatial pattern of rainfall at daily duration over the mountain is similar to those of cellular storms.

Focussing on the 50-years return level, the largest values at 3h duration are found only in the foothill, while they extend over the mountain range at 72h duration. The persistence of large rainfall over the foothill during several hours requires an exceptional forcing in agreement with the exceptional characteristics of the 50-year return level event, which occur seldom in one's life (in average every 50 years). Example of such forcing is the cold pool thermal forcing described in Ducrocq et al. (2008). Other configurations producing severe and long lasting rainfall events have been observed during the HyMeX field campaign (Ducrocq et al., 2013; Drobinski et al., 2014). For example Bousquet et al. (2013) describes a mesoscale convective system impinging the Massif Central range from the west and producing a bow of heavy rainfall cells over the foothill. A different kind of precipitating system yielding large rainfall during periods of tens of hours over the Massif central crest is stationary shallow convective system (Miniscloux et al., 2001; Anquetin et al., 2003). This shallow convection may be combined with deep convection during several hours. Godart et al. (2011) shows that 40% of the largest daily rainfalls over the Massif central crest are produced by such systems.

6.1.3. Temporal variability of extreme rainfall

Eqs. (5) or (10) show that the T_R-year return levels at duration D is nothing else than the T_R-year return level at the reference duration D_{ref} multiplied by $(D/D')^{-H}$, for any D, D' and T_R. Note that the multiplying factor is independent on T_R, so it applies equally to any quantile. Case $H = 0$ corresponds to uniform rainfall with equal intensity whatever the duration. Case $H = 1$ corresponds to rainfall tending to concentrate in D_{ref} hours. Cases $0 < H < 1$ correspond to intermediate cases between uniform and concentrated rainfall. The closer H to one, the more rainfall tends to concentrate in few hours. So H informs on the temporal variability of extreme rainfall. Fig. 7 displays the posterior mean estimations of H in the

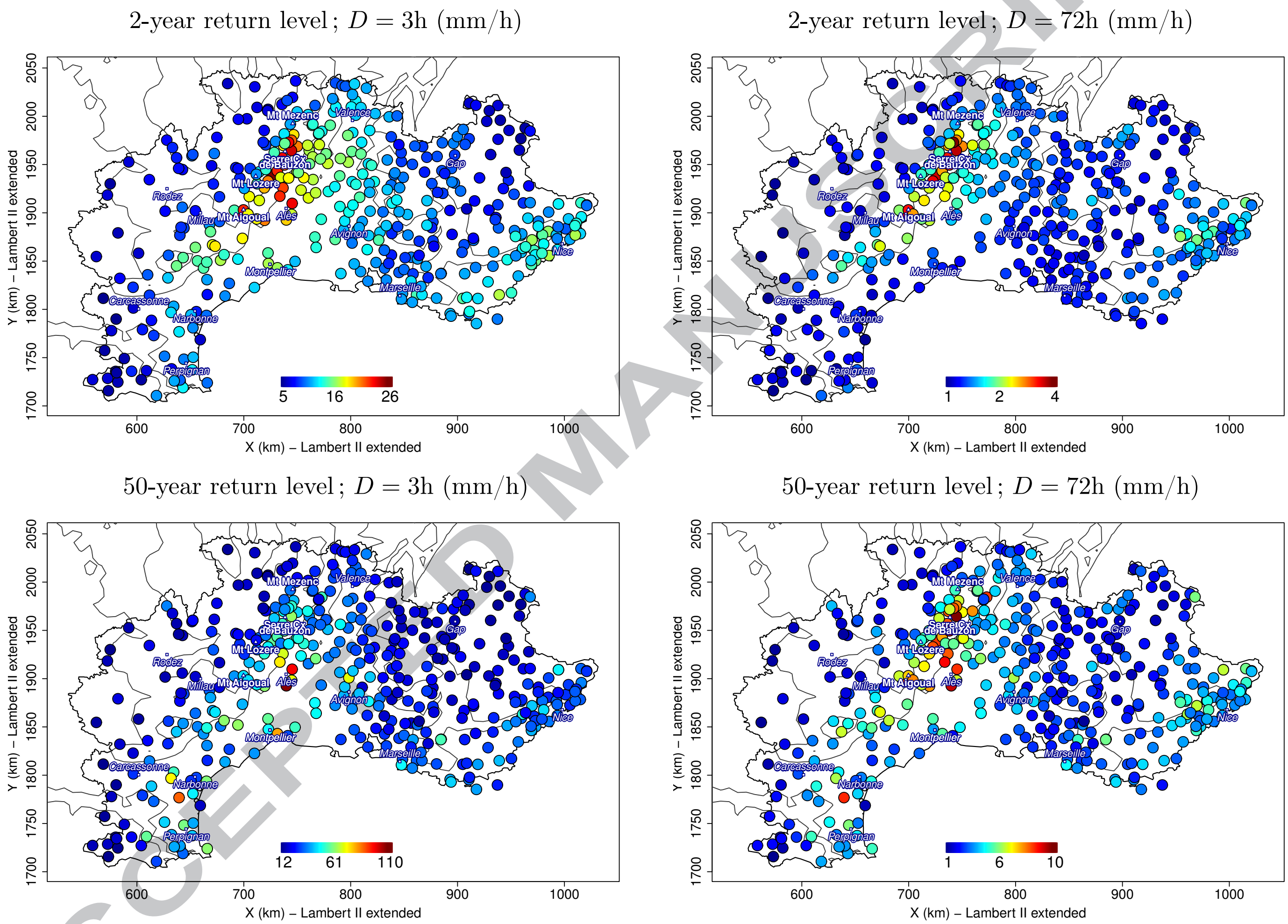

FIGURE 6: Posterior mean estimation of the 2- and 50-year return levels (mm/h) at 3h and 72h durations.

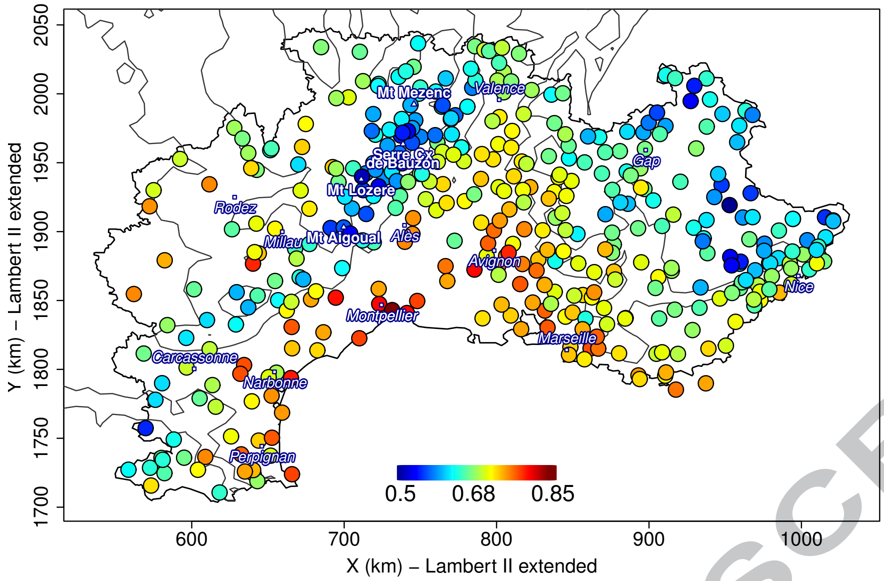

Figure 7: Posterior mean estimation of the scaling parameter H (no units).

region. The largest H are found along the Mediterranean coast between Perpignan and Marseille and along the Rhône valley ($0.7 - 0.85$). The lowest values are found along the Massif Central crest and in the south eastern Alps (H around 0.5). Thus two different extreme rainfall regimes are identified : i) mainly short and intense rainfall events along the Mediterranean shore and in the wide plain of the Rhône valley, which are likely to be controlled by deep convection, and ii) mainly long and regular rainfall events along the Massif Central crest and slope, which force stationary shallow or deep convection.

6.2. IDF uncertainty

6.2.1. The example of Montpellier

Before comparing the density estimates obtained with the different frameworks over the whole region, we start illustrating results on the station of Montpellier. This station is chosen because i) it shows among the largest values of 3h-rainfall intensity (84 mm/h at 3h duration, in autumn 2014), ii) Montpellier is a good illustration of the temporal variability of extreme rainfall : the median value of annual maximum 3h-rainfall intensity (15mm/h at 3h duration) is 50% bigger than the median value over the region (10mm/h at 3h duration), whereas at 72h duration it equals the regional median (1.25 mm/h at 72h duration), and iii) its population is among the biggest in the region (more than $250,000$ inhabitants in 2010), which make it a sensible case of risk analysis. Fig. 8 compares the density estimates of the parameters and 50-year return levels at 3h and 72h durations. In the frequentist framework, densities are obtained with either the theorem of asymptotic normality -in which case densities are Gaussian-, or the bootstrap resampling method. For the Bayesian framework, the posterior density is depicted. Fig. 8 illustrates that the posterior and bootstrap densities are able to better adjust to the data by being able to produce asymmetric densities with several modes. The posterior density of H departs particularly from the bell-like shape of a Gaussian with a flattened

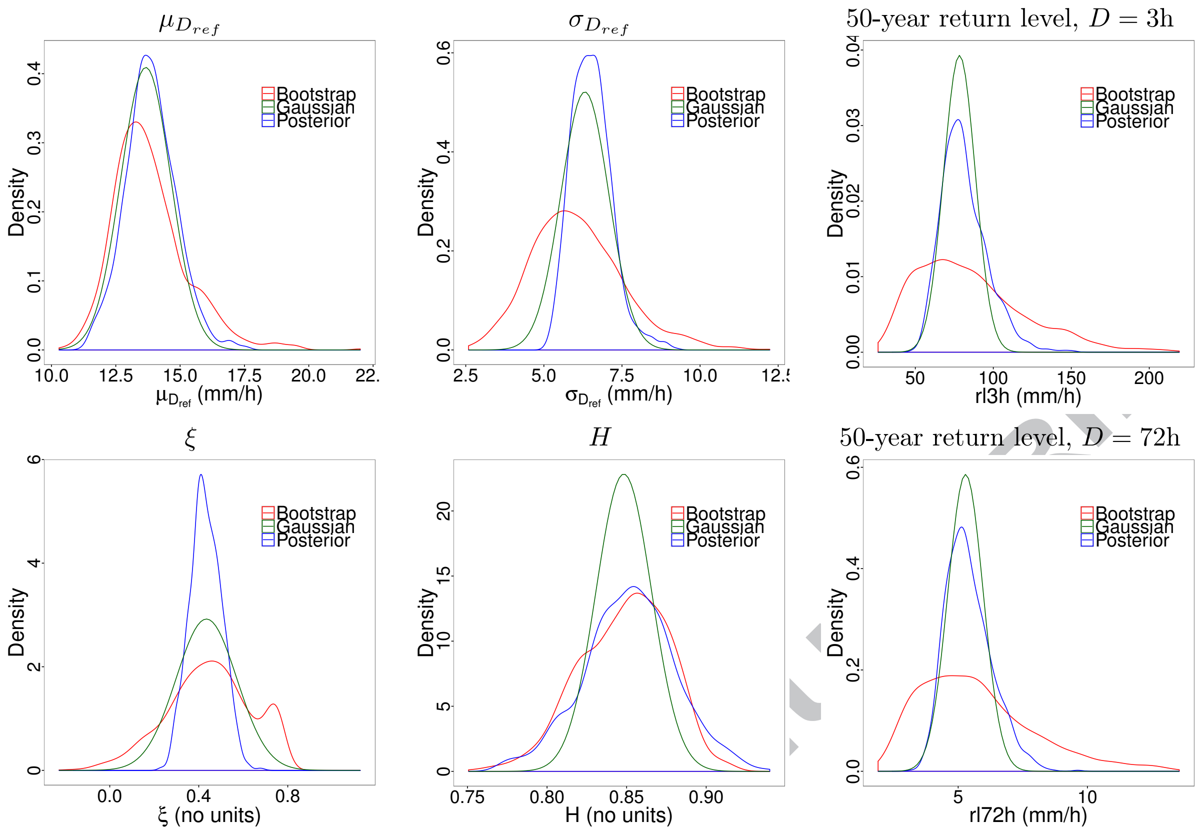

Figure 8: Density estimates of the model parameters and the 50-year return levels at 3h and 72h durations, for Montpellier station. Frequentist densities are obtained with the theorem of asymptotic normality (green) and the bootstrap resampling method (red). Bayesian densities are the posterior densities (blue).

peak between 0.83 and 0.87, which cannot be seen by application of the asymptotic normality theorem. The bootstrap method, on the opposite, produces similar density of H to the posterior density. Some asymmetry with respect to the mode is also found for ξ in the posterior density and even more in the bootstrap density. This produces asymmetry in return levels with a heavier right tails for the bootstrap and posterior densities than for the Gaussian density, whereas the left tails of the posterior and Gaussian densities are similar. Therefore the bootstrap and Bayesian methods are able to tell there is a greater likelihood for the 50-year return level to be over than under the estimated value, which is not possible when considering symmetric Gaussian densities.

The return level plot of Fig. 9 illustrates this asymmetry in the uncertainty of return levels for the bootstrap and posterior densities, particularly for large return periods. Whatever the return period, the {lower bound of the posterior and Gaussian confidence intervals are equal, whereas the upper bound differs significantly. We can thus postulate that, by imposing symmetry, the asymptotic normality theorem tends to underestimate the upper bound of the confidence interval. The bootstrap method allows asymmetry, however it gives much wider confidence intervals than the two other methods, even for the lower bound. Comparing the bootstrap and posterior densities in Fig. 8 shows that difference in the width of the confidence intervals is mainly due to differences in the scale $\sigma_{D_{ref}}$ and shape ξ parameters.

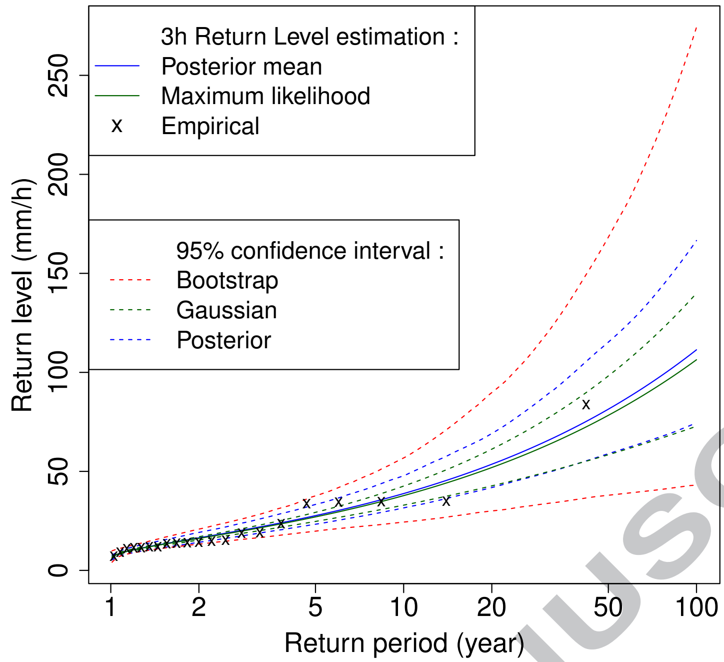

Figure 9: Return level plot at 3h duration. The crosses show the empirical values and the lines the predicted values in the frequentist (green) and Bayesian (blue) framework. The dotted lines are the 95% confidence intervals associated to the Gaussian (green), bootstrap (red) and posterior (blue) densities.

6.2.2. Regional study

The example of Montpellier showed asymmetry of the bootstrap and posterior densities, which is a good sign that these methods allow to better adjust uncertainty estimation to the data. To document this feature at the region scale, we compute the skewness s of the estimated densities at each station. If $s = 0$, the density is symmetric (as in the Gaussian case). If $s > 0$ the density is asymmetric and the right tail is heavier than the left tail. If $s < 0$, it is the opposite. The further s from zero, the greater the asymmetry. Fig. 10 shows the skewness of the bootstrap and Bayesian densities. For sake of readability, we represent the Kernel densities of the skewness and restrict the x-axis to comprise 95% of the values. For the GEV parameters, most skewness of the posterior densities are positive, meaning heavier right tails. This also applies for the bootstrap densities but to a lesser extent for ξ. For the scaling parameter, both left and right heavy tails are found with both methods. For the return levels, mainly positive skewness are found, corroborating what was found for the station of Montpellier in Section 6.2.1. For the great majority of the stations, there is a greater likelihood for the 50-year return level to be over than under its estimated value. This piece of information is of great importance for risk management and is missing when considering symmetric Gaussian densities according to the asymptotic normality theorem. Bootstrap skewness of all variables often largely exceed the Bayesian values. We can postulate that the bootstrap method tends to give too heavy right-tailed densities and are not recommended for the computation of uncertainty. The main reason is that the number of observed years

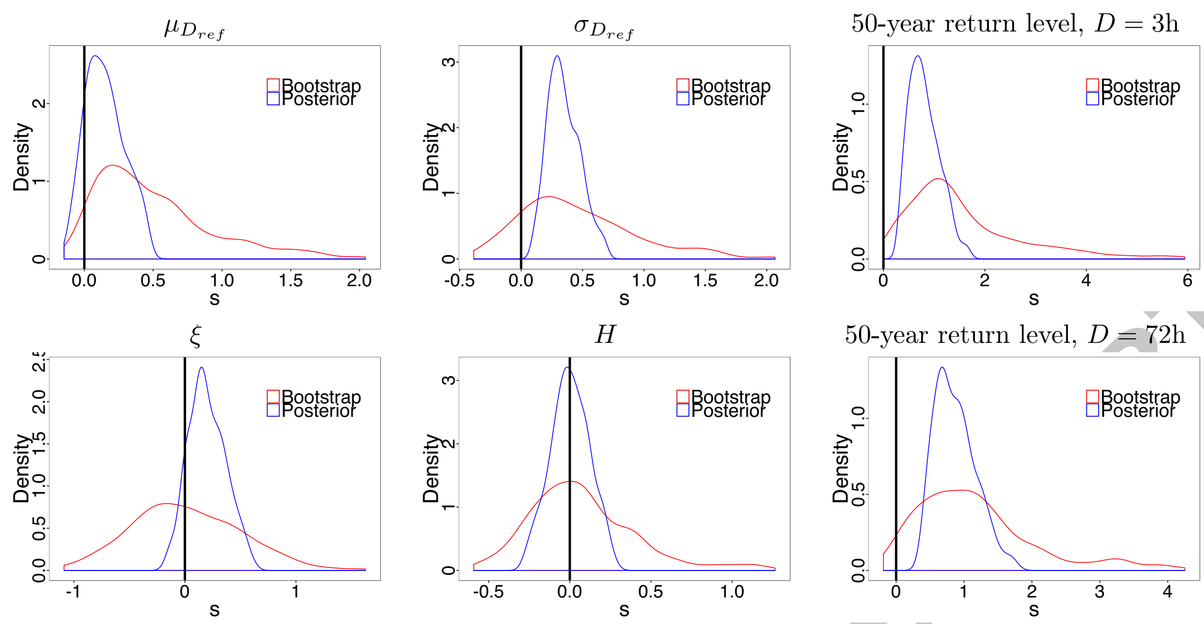

FIGURE 10: Skewness of the bootstrap (red) and posterior densities (blue) of the model parameters and the 50-year return levels at 3h and 72h durations. The black vertical line at 0 corresponds to symmetric density, as the Gaussian density.

per station is too small (20 years on average), while bootstrapping requires long series. To illustrate this, we compare in Fig. 11 the normalized range of 95% confidence interval of 50-year return level at 3h duration in the Bayesian and bootstrap cases. The normalized range is obtained by dividing the 95% confidence interval by either the maximum likelihood estimate (in the bootstrap case) or the posterior mean (in the Bayesian case). Fig. 11 illustrates that bootstrap uncertainty estimation is much more sensitive to the number of data than the Bayesian estimation, confirming that bootstrapping requires long series to work well, while the Bayesian estimation is much more robust. On the opposite there is no way of knowing whether the Bayesian confidence bands are too narrow but checking the return level plots of a large number of stations revealed that very few empirical estimates lie outside the 95% Bayesian confidence bands, which seems to confirm that Bayesian uncertainty estimation is reasonable.

We conclude this analysis by comparing uncertainty in 50-year return levels obtained from the Gaussian and posterior densities. We discard the bootstrap densities, which are often not reasonable. Fig. 12 compares the lower and upper bounds of the 95% confidence interval of the Gaussian and posterior densities at 3h duration. It shows that the lower bounds are usually similar in both cases whereas the upper bounds of the posterior density are always greater. This corroborates the results found for the station of Montpellier in Section 6.2.1 : the Bayesian framework allows to obtain asymmetric confidence bands extending further to large values. We conclude from Fig. 12 that the Gaussian density tends to underestimate uncertainty across the whole region.

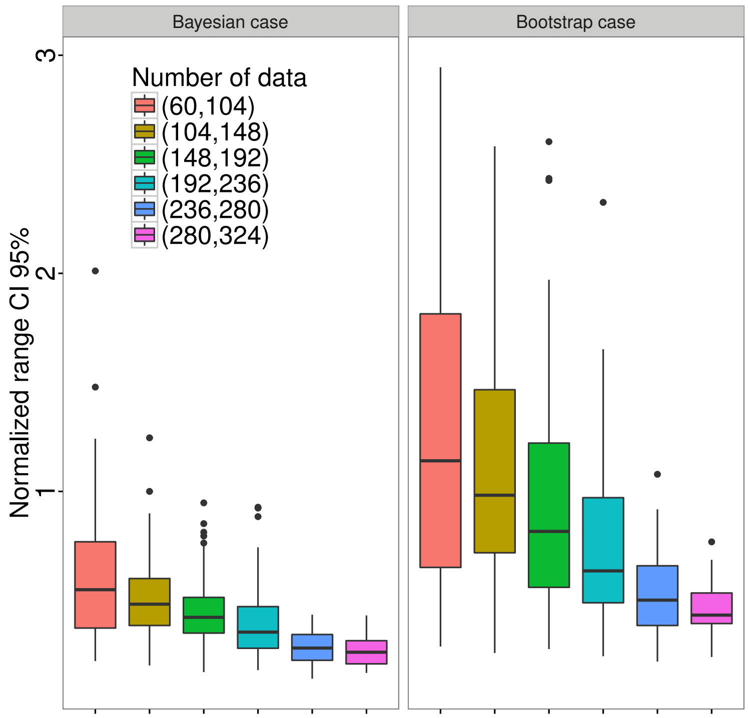

FIGURE 11: Boxplot of the normalized range of 95% confidence interval of 50-year return level at 3h duration obtained in the Bayesian case (left) and in the bootstrap case (right). Different boxplots are drawn depending on the number of data per station, summing the observed years of the nine durations.

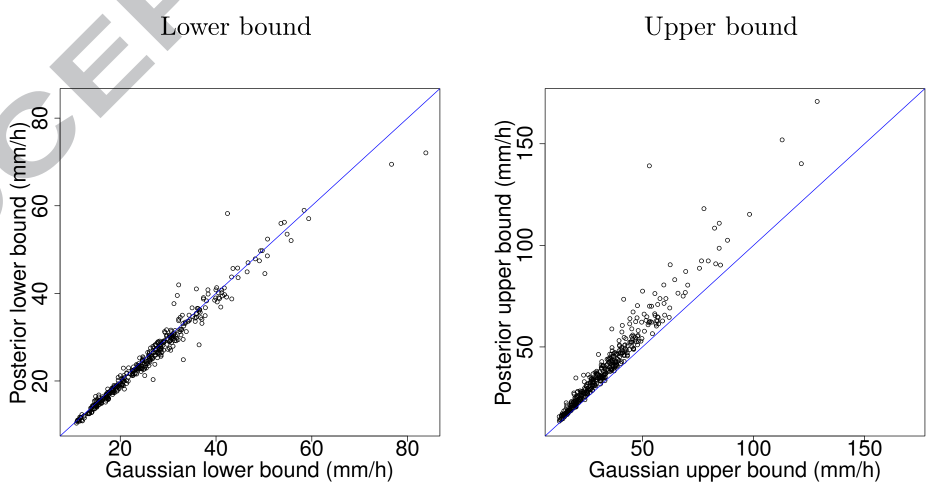

FIGURE 12: Bounds of 95% confidence interval of Gaussian density versus bounds of 95% confidence interval of posterior density of 50-year return level at 3h duration.

7. Conclusion

We conducted in this paper a regional study on the impact of using either a frequentist or Bayesian framework in the estimation of Intensity-Duration-Frequency relationships and subsequent uncertainty. Our analysis was applied to a large database covering a large Mediterranean region with contrasted rainfall regimes. It was shown that estimation is not very sensitive to the choice of framework if the starting point is chosen with care. Uncertainty estimation, however, depends on both framework and estimation method. It was shown that the posterior density (in the Bayesian framework) and the bootstrap density (in the frequentist framework) are able to better adjust uncertainty estimation to the data than the Gaussian density stemming for the asymptotic normality theorem (in the frequentist framework). They are in particular able to produce multi-modal asymmetric densities. However the bootstrap density tends to give unreasonable confidence intervals, in particular for return levels associated to large return period. The main reason is that the number of observed years per station is too small (20 years on average), while bootstrapping requires long series to work well. On the opposite there is no way of knowing whether the Bayesian confidence bands are too narrow but checking the return level plots of a large number of stations revealed that very few empirical estimates lie outside the 95% Bayesian confidence bands, which seems to confirm that Bayesian uncertainty estimation is pretty reasonable. By imposing symmetric confidence intervals, the Gaussian density tends to underestimate to upper bounds of the confidence intervals, which is an issue for risk management. The lack of objectivity of the Bayesian framework is the principal argument of those who rejects this framework (Efron, 2005), but this criticism does not apply to this work, which was conducted using very weakly subjective priors. Therefore our recommendation goes towards the use of the Bayesian framework to compute uncertainty because i) it better adjusts uncertainty computation to the data, and ii) it gives reasonable estimates of uncertainty. Our analysis further highlighted that uncertainty estimation is particularly important in IDF estimation in order to avoid over-optimistic results. For instance, in our case study, there is on average 95% chance for the 50-year return level to be between -20% and +30% of its estimation. Since current infrastructure dealing with flooding and precipitation (e.g. dams or dikes) are based on IDF curves, ignoring this uncertainty would result in large underestimation of flood risk and failure risk of critical infrastructures.

Although the Bayesian framework revealed to give reasonable estimates of IDF relationships and related uncertainties, estimation could be improved in two ways. First, relaxing the hypothesis of independence between durations assumed in this study. Although this hypothesis does not impact the estimation of IDF curves, it may have some impact on their uncertainty. However taking into account dependence between durations is not straightforward. Extreme value theory insures that dependence modelling between the continuum of durations should rely on max-stable processes, which are difficult to estimate in the frequentist framework (Davison et al., 2012), and even more in a Bayesian framework (Ribatet et al., 2012). To the best of our knowledge, max-stable processes have never been used in IDF estimation. This may be the subject of future work. Second improvement regards the consideration of nonstationarity of IDF curves in a context of global warming, for example by considering time-varying IDF relationships as in Cheng and AghaKouchak (2014) or Sarhadi and Soulis (2017), or, even better maybe, by considering covariations in temperature or other climate-related variable. A stationary assumption in a framework of nonstationarities may lead to underestimation of extreme precipitation, and therefore underestimation of flood risk or failure risk in infrastructure systems (Cheng and AghaKouchak, 2014; Sarhadi and Soulis, 2017). However nonstationarity in extreme precipitation seems not to be obvious for the studied region at daily time step (Blanchet et al., 2016b). Furthermore accounting for nonstationarity at subdaily scales would require much longer time series than those available so far for the region, which are most of the time less than 20-years long.

Anquetin, S., Minsicloux, F., Creutin, J.-D., and Cosma, S. (2003). Numerical simulation of orographic rainbands. *Journal of Geophysical Research : Atmospheres*, 108(D8). 8386.

Beirlant, J., Goegebeur, Y., Teugels, J., and Segers, J. (2005). *Statistics of Extremes*. John Wiley & Sons, Ltd.

Berger, J. (2006). The case for objective bayesian analysis. *Bayesian Anal.*, 1(3) :385–402.

Blanchet, J., Ceresetti, D., Molinié, G., and Creutin, J.-D. (2016a). A regional GEV scale-invariant framework for intensity duration frequency analysis. *Journal of Hydrology*, 540 :82 – 95.

Blanchet, J., Molinié, G., and Touati, J. (2016b). Spatial analysis of trend in extreme daily rainfall in southern France. *Climate Dynamics*, pages 1–14.

Borga, M., Vezzani, C., and Fontana, G. D. (2005). Regional rainfall depth–duration–frequency equations for an alpine region. *Natural Hazards*, 36(1) :221–235.

Bougadis, J. and Adamowski, K. (2006). Scaling model of a rainfall intensity-duration-frequency relationship. *Hydrological Processes*, 20(17) :3747–3757.

Bousquet, O., Ribaud, J.-F., Grazioli, J., Berne, A., and Delanoë, J. (2013). Interactions between airflow dynamics and cloud microphysics in complex terrain : The 24 september 2012 hpe observed during HyMeX IOP6. HyMeX Workshop, Cassis, France.

Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5) :1190–1208.

Chandra, R., Saha, U., and Mujumdar, P. (2015). Model and parameter uncertainty in {IDF} relationships under climate change. *Advances in Water Resources*, 79 :127 – 139.

Cheng, L. and AghaKouchak, A. (2013). Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate. *Scientific reports*, 4 :7093–7093.

Cheng, L. and AghaKouchak, A. (2014). Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate. *Scientific reports*, 4.

Coles, S., Bawa, J., Trenner, L., and Dorazio, P. (2001). *An introduction to statistical modeling of extreme values*, volume 208. Springer.

Coles, S. and Pericchi, L. (2003). Anticipating catastrophes through extreme value modelling. *Journal of the Royal Statistical Society : Series C (Applied Statistics)*, 52(4) :405–416.

Coles, S. G. and Tawn, J. A. (1996). A bayesian analysis of extreme rainfall data. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 45(4) :463–478.

Cooley, D., Cisewski, J., Erhardt, R. J., Jeon, S., Mannshardt, E., Omolo, B. O., and Sun, Y. (2012). A survey of spatial extremes : Measuring spatial dependence and modeling spatial effects. *REVSTAT*, 10(1) :135–165.

Davison, A. and Huser, R. (2015). Statistics of Extremes. *Annual Review of Statistics and Its Application*, 2(1) :203–235.

Davison, A. C. (2008). *Statistical Models*. Cambridge University Press, 1 edition.

Davison, A. C., Padoan, S. A., and Ribatet, M. (2012). Statistical modeling of spatial extremes. *Statistical science*, pages 161–186.

Delrieu, G., Nicol, J., Yates, E., Kirstetter, P.-E., Creutin, J.-D., Anquetin, S., Obled, C., Saulnier, G.-M., Ducrocq, V., Gaume, E., Payrastre, O., Andrieu, H., Ayral, P.-A., Bouvier, C., Neppel, L., Livet, M., Lang, M., du Châtelet, J. P., Walpersdorf, A., and Wobrock, W. (2005). The catastrophic flash-flood event of 8 and 9 september 2002 in the gard region, france : A first case study for the cévennes vivarais mediterranean hydrometeorological observatory. *Journal of Hydrometeorology*, 6(1) :34–52.

Drobinski, P., Ducrocq, V., Alpert, P., Anagnostou, E., Béranger, K., Borga, M., Braud, I., Chanzy, A., Davolio, S., Delrieu, G., Estournel, C., Boubrahmi, N. F., Font, J., Grubišić, V., Gualdi, S., Homar, V., Ivančan-Picek, B., Kottmeier, C., Kotroni, V., Lagouvardos, K., Lionello, P., Llasat, M. C., Ludwig, W., Lutoff, C., Mariotti, A., Richard, E., Romero, R., Rotunno, R., Roussot, O., Ruin, I., Somot, S., Taupier-Letage, I., Tintore, J., Uijlenhoet, R., and Wernli, H. (2014). HyMeX : A 10-Year Multidisciplinary Program on the Mediterranean Water Cycle. *Bulletin of the American Meteorological Society*, 95(7) :1063–1082.

Ducrocq, V., Braud, I., Davolio, S., Ferretti, R., Flamant, C., Jansa, A., Kalthoff, N., Richard, E., Taupier-Letage, I., Ayraland, P., Belamariand, S., Berneand, A., Borgaand, M., Boudevillain, B., Bock, O., Boichard, J.-L., Bouin, M.-N., Bousquet, O., Bouvier, C., Chiggiato, J., Cimini, D., Corsmeier, U., Coppola, L., Cocquerez, P., Defer, E., Delanoë, J., Di Girolamo, P., Doerenbecher, A., Drobinski, P., Dufournet, Y., Fourrié, N., Gourley, J., Labatut, L., Lambert, D., Le Coz, J., Marzano, F., Molinié, G., Montani, A., Nord, G., Nuret, M., Ramage, K., Rison, B., Roussot, O., Said, F., Schwarzenboeck, A., Testor, P., Van-Baelen, J., Vincendon, B., Aran, M., and Tamayo, J. (2013). Hymex-sop1, the field campaign dedicated to heavy precipitation and flash-flooding in the northwestern mediterranean. Available on line.

Ducrocq, V., Nuissier, O., Ricard, D., Lebeaupin, C., and Thouvenin, T. (2008). A numerical study of three catastrophic precipitating events over southern france. II : Mesoscale triggering and stationarity factors. *Quarterly Journal of the Royal Meteorological Society*, 134(630) :131–145.

Efron, B. (2005). Bayesians, frequentists, and scientists. *Journal of the American Statistical Association*, 100(469) :1–5.

Gaume, E., Livet, M., Desbordes, M., and Villeneuve, J.-P. (2004). Hydrological analysis of the river aude, france, flash flood on 12 and 13 november 1999. *Journal of Hydrology*, 286(1 4) :135 – 154.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2014). *Bayesian Data Analysis*. Chapman & Hall/CRC Texts in Statistical Science. Chapman and Hall/CRC, 3 edition.

Godart, A., Anquetin, S., Leblois, E., and Creutin, J.-D. (2011). The contribution of orographically driven banded precipitation to the rainfall climatology of a mediterranean region. *Journal of Applied Meteorology and Climatology*, 50(11) :2235–2246.

Gupta, V. K. and Waymire, E. (1990). Multiscaling properties of spatial rainfall and river flow distributions. *Journal of Geophysical Research : Atmospheres*, 95(D3) :1999–2009.

Hailegeorgis, T. T., Thorolfsson, S. T., and Alfredsen, K. (2013). Regional frequency analysis of extreme precipitation with consideration of uncertainties to update {IDF} curves for the city of trondheim. *Journal of Hydrology*, 498 :305 – 318.

Huard, D., Mailhot, A., and Duchesne, S. (2010). Bayesian estimation of intensity–duration–frequency curves and of the return period associated to a given rainfall event. *Stochastic Environmental Research and Risk Assessment*, 24(3) :337–347.

Jeffreys, H. (1998). *The theory of probability*. OUP Oxford.

Kass, R. E. and Wasserman, L. (1996). The selection of prior distributions by formal rules. *Journal of the American Statistical Association*, 91(435) :1343–1370.

Miniscloux, F., Creutin, J. D., and Anquetin, S. (2001). Geostatistical analysis of orographic rainbands. *Journal of Applied Meteorology*, 40(11) :1835–1854.

Molinié, G., Ceresetti, D., Anquetin, S., Creutin, J. D., and Boudevillain, B. (2012). Rainfall regime of a mountainous mediterranean region : Statistical analysis at short time steps. *Journal of Applied Meteorology and Climatology*, 51(3) :429–448.

Muller, A., Bacro, J.-N., and Lang, M. (2008). Bayesian comparison of different rainfall depth–duration–frequency relationships. *Stochastic Environmental Research and Risk Assessment*, 22(1) :33–46.

Nuissier, O., Ducrocq, V., Ricard, D., Lebeaupin, C., and Anquetin, S. (2008). A numerical study of three catastrophic precipitating events over southern france. i : Numerical framework and synoptic ingredients. *Quarterly Journal of the Royal Meteorological Society*, 134(630) :111–130.

Overeem, A., Buishand, A., and Holleman, I. (2008). Rainfall depth-duration-frequency curves and their uncertainties. *Journal of Hydrology*, 348(1) :124–134.

Papalexiou, S. M. and Koutsoyiannis, D. (2013). Battle of extreme value distributions : A global survey on extreme daily rainfall. *Water Resources Research*, 49(1) :187–201.

Rantz, S. (1971). Precipitation depth-duration-frequency relations for the san francisco bay region, california. *US Geological Survey, Prof. Paper*, pages 237–241.

Ribatet, M., Cooley, D., and Davison, A. C. (2012). Bayesian inference from composite likelihoods, with an application to spatial extremes. *Statistica Sinica*, 22(2) :813–845.

Ribatet, M. and Sedki, M. (2012). Extreme value copulas and max-stable processes. *Journal de la Société Française de Statistique*, 153(3) :138–150.

Ruin, I., Lutoff, C., Boudevillain, B., Creutin, J.-D., Anquetin, S., Rojo, M. B., Boissier, L., Bonnifait, L., Borga, M., Colbeau-Justin, L., Creton-Cazanave, L., Delrieu, G., Douvinet, J., Gaume, E., Gruntfest, E., Naulin, J.-P., Payrastre, O., , and Vannier, O. (2014). Social and Hydrological Responses to Extreme Precipitations : An Interdisciplinary Strategy for Postflood Investigation. *Weather, Climate, and Society*, 6(1) :135–153.

Sarhadi, A. and Soulis, E. D. (2017). Time-varying extreme rainfall intensity-duration-frequency curves in a changing climate. *Geophysical Research Letters*, 44(5) :2454–2463. 2016GL072201.

Sebille, Q., Fougères, A.-L., and Mercadier, C. (2017). Modeling extreme rainfall. *Spatial Statistics*.

Sénési, S., Bougeault, P., Chèze, J.-L., Cosentino, P., and Thepenier, R.-M. (1996). The vaison-la-romaine flash flood : Mesoscale analysis and predictability issues. *Weather and Forecasting*, 11(4) :417–442.

Stephenson, A. G., Lehmann, E. A., and Phatak, A. (2016). A max-stable process model for rainfall extremes at different accumulation durations. *Weather and Climate Extremes*, 13 :44 – 53.

Te Chow, V. (1988). *Applied hydrology.* Tata McGraw-Hill Education.

Tung, Y.-k. and Wong, C.-l. (2014). Assessment of design rainfall uncertainty for hydrologic engineering applications in hong kong. *Stochastic Environmental Research and Risk Assessment*, 28(3) :583–592.

Van de Vyver, H. (2012). Spatial regression models for extreme precipitation in belgium. *Water Resources Research*, 48(9). W09549.

Van de Vyver, H. (2015). Bayesian estimation of rainfall intensity–duration–frequency relationships. *Journal of Hydrology*, 529, Part 3 :1451 – 1463.

Zellner, A. (1998). Past and recent results on maximal data information priors. *Journal of Statistical Research*.

ACCEPTED MANUSCRIPT

We compare the Bayesian and frequentist frameworks for uncertainty estimation of IDF relationships.

We confront the two frameworks in a regional study.

The Bayesian framework allows to better adjust uncertainty estimation to the data.